Genetic Resources

Genetic Resources (2025)
Vol. 6, Issue 12
DOI: 10.46265/genresj.2025.12
www.genresj.org
ISSN: 2708-3764

Focus and Scope of *Genetic Resources*

Genetic Resources is an open access journal disseminating global knowledge and tools used by the community of practitioners of plant and animal genetic resources involved in monitoring, collecting, maintaining, conserving, characterizing and using genetic resources for food, agriculture and forestry. ***Genetic Resources*** publishes original research, methods, strategies, guidelines, case studies and reviews as well as opinion and other papers on a variety of topics of interest on the present and future use of genetic resources. These may include the acquisition, documentation, conservation, management, assessment, characterization and evaluation of genetic resources and their link to broader biodiversity, socioeconomic practices, policy guidelines or similar, serving stakeholders within and across sectors. Occasionally, ***Genetic Resources*** publishes special issues with a focus on selected topics of interest for the genetic resources community. The journal has a focus on the European region and also welcomes contributions of wider interest from all world regions.

Cover photos (from left to right):
Phureja potatoes, Peru. Credit: J. F. Seminario *et al*; Struma chicken, Bulgaria. Credit: Lukanov *et al*; *Phaseolus angustissimus* Asa Gray, USA. Credit: Sarah Dohle.

This journal is supported by the European Cooperative Programme for Plant Genetic Resources (ECPGR) and the European Regional Focal Point for Animal Genetic Resources (ERFP).

This journal was conceived as part of the GenRes Bridge project. This project received seed funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817580.

Editorial Office

ECPGR Secretariat
c/o Alliance of Bioversity International and CIAT
Via di San Domenico 1
000153 Rome, Italy

Submissions to www.genresj.org

Editorial Board

Genetic Resources

Genetic Resources (2025), 6 (12)
DOI: 10.46265/genresj.2025.12
www.genresj.org
ISSN: 2708-3764

TABLE OF CONTENTS

Original Articles

Cultivar loss and conservation of genetic resources of the phureja potato (*Solanum phureja* L., Phureja Group) in Peru
Juan F. Seminario, Leónidas S. Concha-Tupayachi, Alejandro Seminario-Cunya, Tulio Medina-Hinostrosa, Wilsser Silva
Pages 1–13
doi: 10.46265/genresj.JCDC4631

Agro-morphological and molecular characterization of Argentine maize (*Zea mays* L.) landraces of 'Cristalino Colorado' race
Raquel Alicia Defacio, Natalia Mercedes Paz, Ana Rosa Schlater, Marcelo Edmundo Ferrer, Sergio Jorge Bramardi
Pages 14–25
doi: 10.46265/genresj.TSJG3884

Phenotypic variability of *Smallanthus sonchifolius* germplasm of Peru
Angel Esteban Santa Cruz Padilla, Jorge Luis Vásquez Orrillo, Silvia Yanina Rodríguez López, Araceli Eugenio Leiva, Ricardo Manuel Bardales-Lozano, Juan F. Seminario, Hipolito Murga-Orrillo
Pages 26–38
doi: 10.46265/genresj.ZDEW3901

The promise of access and benefit-sharing is met through holistic policy reform: Insights from Colombia's genetic diversity and innovation landscape during COP16
Bob Kreiken, Lotte Asveld
Pages 39–56
doi: 10.46265/genresj.APNR6909

Ethnography of traditional healers and their indigenous medicinal plants in southern Philippines: Implications for conservation and sustainable use
Bhegie May G. Buay, Catherine Hazel M. Aguilar, Maria Celeste N. Banaticla-Hilario, Chenny Rose M. Rodriguez, Florence L. Zapico
Pages 57–72
doi: 10.46265/genresj.FQLF1923

Exploring the situation of transboundary breeds in Europe for their effective management and conservation
Eléonore Charvolin, Dimitrios Tsiokos, Grégoire Leroy, Christina Ligda
Pages 73–82
doi: 10.46265/genresj.BPGX5961

A fast and effective method to distinguish cultivated fonio species: conservation and evaluation perspectives

Sandrine Causse, Thomas Kaczmarek, Cécile Dubois, Enoch G. Achigan-Dako, Joseph Adjebeng-Danquah, Richard Y. Agyare, Louise Akanvou, Yacoubou Bakasso, Mamadou B. Barry, Baye M. Diop, Mame C. Gueye, Abdou R. Ibrahim Bio Yerima, Happiness O. Oselebe, Saidou Idi Sani, Edak A. Uyoh, Sylvie Vancoppenolle, Adeline Barnaud, Claire Billot, Jean-François Rami, Christian Leclerc

Pages 83–94
doi: 10.46265/genresj.JDUT8893

Solanum wild relative species indicate varying ecological resilience to climate change in Benin (West Africa)

Ahuéfa Mauricel Kégbé, Rodrigue Idohou, Birane Dieng, Gafarou Agounde, Anthony Egeru, Kandioura Noba, Achille Ephrem Assogbadjo

Pages 95–110
doi: 10.46265/genresj.ZRCI8675

Egg production characteristics of several Bulgarian chicken breeds

Hristo Lukanov, Ivelina Pavlova, Atanas Genchev, Todor Petrov

Pages 120–129
doi: 10.46265/genresj.QNIV2785

Novel germplasm of tepary and other *Phaseolus* bean wild relatives from dry areas of southwestern USA

Daniel G. Debouck, Richard C. Pratt, Sarah Dohle, Timothy Porch, Marcela Santaella, Luis Guillermo Santos, Milan O. Urban

Pages 131–152
doi: 10.46265/genresj.LSBJ4572

Genetic variation of Burgo chicken from Bengkulu, Indonesia, based on the ND1-mitochondrial DNA gene

Jarulis, Aceng Ruyani, Nurmeiliasari, Ahmat Fakhri Utama

Pages 153–159
doi: 10.46265/genresj.NZSL1498

Evaluation of red sanders (*Pterocarpus santalinus* L.f.) germplasm for conservation and breeding

Kumbarahally Murthigowda Shivaprasad, Pendela Surath Kumar, Avula Kishore, Dinesh Kumar, Swapnen Pattanaik

Pages 160–170
doi: 10.46265/genresj.UNBA8473

Biochemical characteristics of bread wheat genotypes related to SSR markers in moisture stress conditions

Fatemeh Bavandpouri, Ezatollah Farshadfar, Kianoosh Cheghamirza, Mohsen Farshadfar

Pages 171–193
doi: 10.46265/genresj.SLTA9371

Seed germinability of *Stylosanthes* spp. (Fabaceae) accessions under water stress conditions
Vitor Oliveira dos Santos, Aritana Alves da Silva, Robson de Jesus Santos, Uasley Caldas de Oliveira, Marilza Neves do Nascimento, Claudineia Regina Pelacani

Pages 194–204
doi: 10.46265/genresj.KWVH7038

What farmers value matters for the management of local breeds: A case study of the Pyrenean goat breed
Anne Lauvie, Marie-Odile Nozières-Petit, Fanny Thuault, Nathalie Couix

Pages 211–220
doi: 10.46265/genresj.PEGR2598

Short Communications

A public mid-density genotyping platform for North American Atlantic salmon (*Salmo salar* L.)
Dongyan Zhao, Josué Chinchilla-Vargas, Alexander M. Sandercock, Cristiane H. Taniguti, Roseanna Long, Yniv Palti, Guangtu Gao, Michael Pietrak, Samuel A. May, Kasia Heller-Uszynska, Craig T. Beil, Moira J. Sheehan

Pages 111–119
doi: 10.46265/genresj.SKPZ5680

A small-scale assessment of the availability of EURISCO accessions
Erik Wijnker, Theo van Hintum

Pages 205–210
doi: 10.46265/genresj.SRQL5271

Corrections

Correction: Organization of plant Biological Resource Centers for research in France: History, evolution and current status
Valérie Bergheaud, Jean-Marc Audergon, Arnaud Bellec, Anne Delaunay, Jerome Duminil, Stephane Dussert, Florence Esnault, Emmanuel Geoffriau, Brigitte Gouesnard, Christophe Jenny, Alain Label, Philippe Lashermes, Najate Maghnaoui, Cecile Marchal, Franciane Nuissier, Nilda Paulo-de-la-Reberdiere, Aurelia Priet, Valérie Rieucau, Paule Teres, Anne-Françoise Adam-Blondon

Page 130
doi: 10.46265/genresj.JZQN2033

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 1–13
DOI: 10.46265/genresj.JCDC4631
https://www.genresj.org
ISSN: 2708-3764

Cultivar loss and conservation of genetic resources of the phureja potato (*Solanum phureja* L., Phureja Group) in Peru

Juan F Seminario [*,a], **Leónidas S Concha-Tupayachi** [b], **Alejandro Seminario-Cunya** [c], **Tulio Medina-Hinostrosa** [d] **and Wilsser Silva** [*,e]

[a] *Facultad de Ciencias Agrarias, Programa de Raíces y Tubérculos Andinos, Universidad Nacional de Cajamarca, Peru*
[b] *Universidad Tecnológica de los Andes, Cusco, Peru*
[c] *Universidad Nacional Autónoma de Chota, Peru*
[d] *Dirección de Recursos Genéticos y Bioseguridad del Ministerio del Ambiente, Lima, Peru*
[e] *Programa de Doctorado en Ingeniería y Ciencias Ambientales, Universidad Nacional Agraria La Molina, Av. La Molina s/n, La Molina-Lima, Peru*

Abstract: *Solanum tuberosum* L. Phureja Group, known in Peru as 'phureja potato' or 'chaucha potato' and as 'criolla' in Colombia, is characterized by its earliness and the absence of dormancy in the tubers. It stands out for its nutritional value and its contribution to food security. However, it faces a high risk of disappearance in Peru. This study assessed its current status by collecting historical data, *ex situ* and *in situ* conservation analyses, and genetic erosion studies in local communities. Historical information suggests that phureja was relevant and abundant in the past. Currently, e*x situ* collections include 69 accessions, of which the International Potato Center conserves a significant portion. As for *in situ* conservation, 116 accessions have been identified. However, since 1992, genetic erosion has been documented in six departments of Peru. The main causes include: lack of time for continuous cultivation, prioritization of dairy farming, low seed quality, preference for more commercial modern or traditional cultivars, and the expansion of mining projects. The critical situation of the phureja potato requires urgent measures to collect new information and evaluate the remaining genetic variability. This assessment is essential to develop conservation and sustainability strategies to ensure its survival and its contribution to Peru's food and cultural well-being.

Keywords: Genetic erosion; potato diploid, *Solanum phureja*, yellow potato, chaucha

Citation: Seminario, J. F., Concha-Tupayachi, L. S., Seminario-Cunya, A., Medina-Hinostrosa, T., Silva, W. (2025). Cultivar loss and conservation of genetic resources of the phureja potato (*Solanum phureja* L., Phureja Group) in Peru. *Genetic Resources* 6 (12), 1–13. doi: 10.46265/genresj.JCDC4631.

Introduction

The understanding of global biodiversity remains limited, reaching only 20% of species. Especially in the centres of origin of cultivated plants, there are gaps in knowledge about how farmers preserve local cultivars. This, in turn, hinders the implementation of methodologies that could favour conservation, such as establishing baselines, conducting monitoring and collecting evidence on cultivar dynamics (losses and additions) and the likely genetic erosion involved (De Carvalho *et al*, 2016; Dawson *et al*, 2023).

Genetic erosion – the loss of crop genetic diversity in specific contexts of time and space – is a persistent concern in the agri-food field. Its dynamics, triggers, measurement methods and magnitude of losses are still not fully understood (van de Wouw *et al*, 2009; Khoury *et al*, 2022). This problem is especially critical

*Corresponding authors: Juan F Seminario (jfseminario@yahoo.es), Wilsser Silva (20240313@lamolina.edu.pe)

Received: 22.02.2025 | Accepted: 19.05.2025 | Published online: 07.07.2025

in key crops such as potato, which is key to food security, nutrition and sustainability of global agri-food systems. The reduction of the crop's genetic diversity limits its ability to adapt to environmental changes and phytosanitary challenges, threatening both agricultural productivity and the livelihoods of dependent communities (Burgos *et al*, 2020; Devaux *et al*, 2020).

The phureja potato is classified under the International Code of Botanical Nomenclature (ICBN) as the species *Solanum phureja* (Juz. & Buk.). Its taxonomic key indicators are: corolla lobes wider than half the petal length, short acuminata, leaves somewhat glossy, less densely pubescent (Ochoa, 2001). On the other hand, according to the International Code of Nomenclature for Cultivated Plants (ICNCP), this potato is located within the Phureja Group, as *Solanum tuberosum* L. Phureja Group, and its taxonomic key indicates that tubers sprouted at harvest (Huamán and Spooner, 2002). Its most notable features are its rapid maturation, from three to four months, and the lack of a resting period for its tubers (Ochoa, 2001; Huamán and Spooner, 2002).

The phureja potato stands out for its high nutritional and culinary value, and its richness in secondary metabolites with antioxidant properties. It is an important source of essential minerals such as potassium, iron and zinc, which reinforces its nutritional value (Cerón-Lasso *et al*, 2018; Beltrán-Penagos *et al*, 2020; Peña *et al*, 2021). Its tolerance to adverse conditions, such as poor soils and high altitudes, makes it a resilient option in the face of climate change. It also has great potential in the produce industry. Its role in genetic improvement is crucial, contributing valuable genes to develop more resistant and sustainable varieties (Gabriel and Franco, 2013; Ñústez and Rodríguez, 2020). It has been considered as one of the world's 50 foods of the future (Ñústez, 2021).

Despite its importance, the phureja potato has received little conservation attention in Peru. This situation is evident in the south and centre of the country, where several studies conducted in the last three decades (Zimmerer, 1991; De Haan and Thiele, 2004; CIP and FEDECH, 2006; Plasencia *et al*, 2018) have documented significant losses. Similarly, a report focusing on northern Peru (Seminario and Zarpán, 2011) suggest that this potato could be at risk of disappearing. Based on this background, the objective of this research was to gather evidence to evaluate its relative importance in the past and to analyze the loss of cultivars that occurred in the last decades in Peru.

Materials and methods

Records of the presence, importance and relative abundance of the phureja potato in Peru

Historical sources on ancient Peru were used, including the works of Varcárcel (1985), the report by J.B. Martinet at the end of the 19th century (Martinet, 1977), and the works of Herrera (1921) and Vargas (1936, 1946, 1955), among others. Historical information was also gathered on one of the pioneers in potato conservation and genetic improvement in the Paucartambo region (Cusco). This is Mr. Yabar, recognized for his valuable work between the 1930s and 1945, during which time he promoted early practices of conservation and selection of local potato cultivars (Yabar-Villagarcía, 2004). Another key source was the database of Ochoa (2003), which contains information on passport data from native potato collections carried out in Peru between 1947 and 1997.

Information about *ex situ* conservation of the phureja potato in Peru

In 2020, a contact was established with personnel responsible for the potato collections of the experimental stations at the National Institute for Agrarian Innovation (INIA), the institution in charge of preserving the genetic resources of cultivated plants in Peru, located in the localities of Puno, Cusco, Junín, Ayacucho and Cajamarca. During this period, detailed information was collected on the number of accessions (units of genetic material) of phureja potatoes conserved in each station. This work allowed us to understand the practices and criteria applied for the conservation of this germplasm in different regions of the country. In addition, in 2023, the database of the International Potato Center (CIP) was accessed to obtain the number of phureja potato entries and those corresponding to Peru.

In Cusco, access was granted to the germplasm database managed by a group of communities there known as the 'Parque de la Papa', managed by the NGO Andes del Cusco Association. This database is an exhaustive register of native potato cultivars, including phureja varieties, conserved *ex situ*. Among these materials are samples repatriated from CIP, preserved in collaboration with local communities associated with the park. In addition, information was collected from the Regional Center for Andean Biodiversity Research (CRIBA) of the Universidad Nacional de San Antonio Abad del Cusco (UNSAAC), specifically on the number of native potato and phureja varieties stored in its germplasm bank.

In the city of Paucartambo, the Salcedo Rojas family coordinates the activities of the Paucartambo Conservationists Association. After an interview with two family members, we were able to access (for systematization and analysis) the database on native potatoes maintained by 30 conservationists from the districts of Paucartambo, Challabamba, Colquepata, and Huancarani. In addition, an interview was conducted with Julio Hancco, an outstanding conservationist from the community of Pampacorral, located in the province of Lares, who is widely recognized for his work in the conservation of agricultural biodiversity.

In the Cajamarca region, interviews were conducted in 2011 and 2021 with 20 potato conservation farmers

in the provinces of Cajamarca, Celendín, Hualgayoc and San Marcos, previously identified in a study by Seminario and Zarpán (2011). The objective of these interviews was to determine the number of germplasm entries they maintain and to analyze the socioeconomic factors that affect their conservation work.

Information about *in situ* conservation of the phureja potato in Peru

The catalogues of native potatoes produced in Peru during the last two decades were compiled and analyzed. Those catalogues that specify the species to which the accessions belong or that, at least, record the duration of tuber dormancy were included. Although these documents do not provide direct information on cultivar erosion, the presence or absence of phureja potatoes in these collections is considered a relevant indicator of the level of conservation of this group on farmers' farms.

In Cajamarca region, information was collected on phureja potato collections carried out since 2005 in several provinces of this region. These data provided a comprehensive overview of the conservation efforts of potato cultivars and varieties by the communities. In addition, these activities made it possible to explore the challenges, strategies and traditional knowledge associated with *in situ* conservation, especially highlighting the crucial role of communities in the preservation of the genetic diversity of the phureja potato. Also in 2021, in the provinces of Chota and Cutervo, interviews were conducted with 45 potato farmers and information was collected on the phureja cultivars planted in the previous season and the varieties lost in the last two to three decades (Figure 1). In the locality of Waqanqa, located in the district of Paucartambo, Cusco, direct observation was carried out between 16 and 19 February 2020 with the objective of identifying potato cultivars. In addition, 16 farmers from the districts of Paucartambo and Challabamba, who participated in the XIX Agricultural Fair, held on 11-12 September 2023, were interviewed (Figure 1).

Empirical evidence on genetic erosion of the phureja potato in Peru

An exhaustive literature search for research related to the genetic erosion of the phureja potato was carried out, covering articles and other scientific publications produced between 1980 and 2023 that document the loss of native potato germplasm in Peru. For this purpose, physical libraries and bibliographic databases of wide coverage, such as Google Scholar, SciELO, Web of Science and Scopus, were consulted in order to gather and analyze information on this topic.

Results

Presence, importance and relative abundance of the phureja potato in Peru

According to information gathered from existing literature, catalogues, germplasm banks and farmers' knowledge, the phureja potato is grown or present in 11 of Peru's 25 regions. These regions are mainly located in the Andean zone, at altitudes ranging from 2,000 to 3,500 meters above sea level (Figure 1).

Since ancient times, the phureja potato has been an essential source of food and a key component in the food security of Andean communities (Forbes *et al*, 2020). In ancient Peru, five types of potato were identified, classified in Quechua language according to their morphological characteristics and agroecological adaptations: hatun papa (large potato), chaucha [= phureja] papa (early maturing potato), maguay papa (early planting potato), capo papa (possibly qjapo papa, associated with elevated areas and dark soils) and chiri papa (potato cultivated in cold regions) (Varcárcel, 1985).

Herrera (1921) described the diversity and quality of potatoes in Paucartambo (Cusco), classifying non-bitter potatoes into four groups. Within the group of elongated potatoes, he included the chaucha or phureja potato, characterized by its rapid growth, watery tubers and reddish buds, although there were also varieties with round tubers and white buds. These potatoes have a cultivation cycle of approximately three months. Vargas (1936) also highlighted the importance of phureja potatoes in Paucartambo. These, known locally as papa nueva, mosoc papa or misca mahuay, are sown early and are used both for immediate consumption and starch extraction. In addition, Vargas (1946) gathered a collection of 774 potato samples mainly from Puno, Apurímac and Cusco, of which 79% corresponded to *Solanum andigenum* and 10% to *Solanum stenotomum*. Later, in *Las Papas Sudperuanas, Parte II,* Vargas identified four clones of chaucha potatoes (1206, 1207, 1208 and 1209) belonging to *S. stenotomum*, and one clone (549) classified within *S. andigenum* (Vargas, 1946).

Luis Á. Yabar Ordóñez (1886-1965), a horticulturist originally from Paucartambo, stood out as a pioneer in the conservation of native potatoes in Peru during the period from 1930 to 1945. Yabar compiled and maintained a valuable collection of approximately 250 native potato cultivars, including the churumpis, chauchas or phurejas, miskillas, thumillas, tokolos and phasñachas groups (Yabar-Villagarcía, 2004). His work was widely recognized by specialists in the study of this tuber. In 1944, J. G. Hawkes acknowledged his contribution by naming a species in his honour: *S. yabari*, which included two varieties, *S. yabari* var. *pepino* and *S. yabari* var. *cuzcoense*. These species were later reclassified as part of *S. stenotomum* (Ochoa, 1999; Watanabe *et al*, 2008).

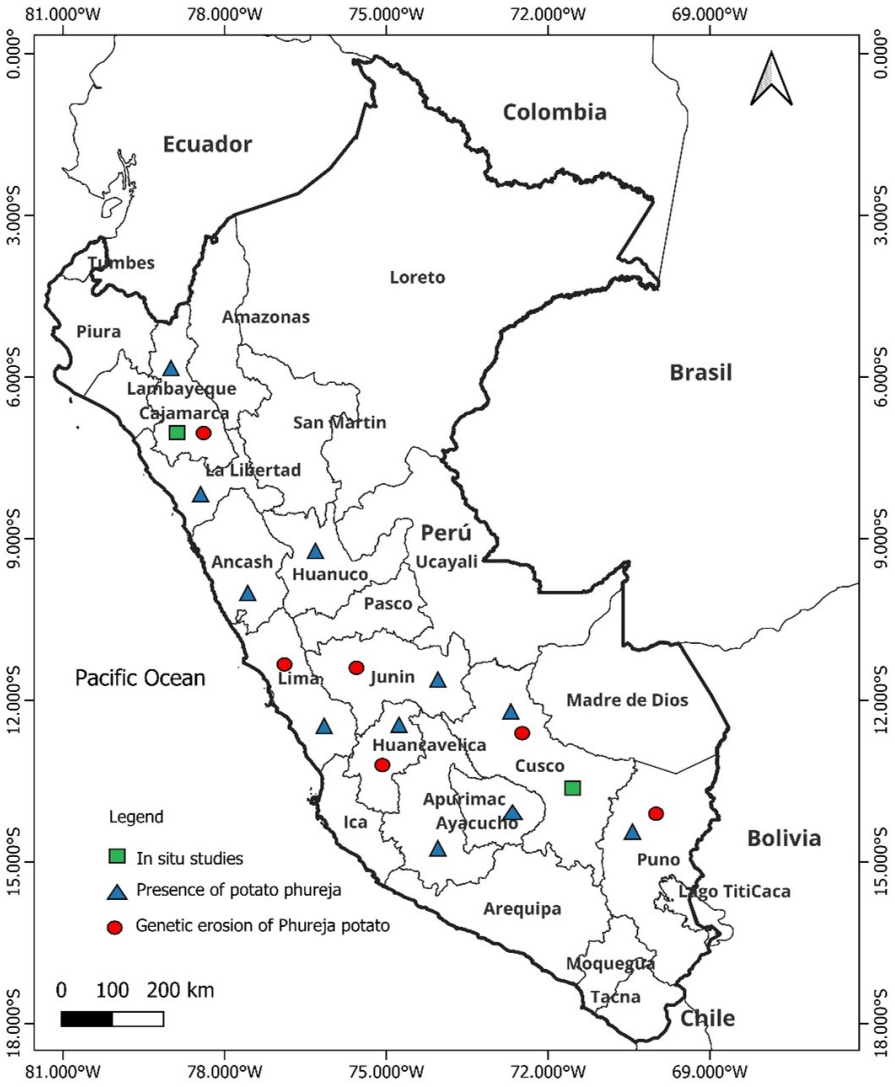


Figure 1. Map of the 25 regions of Peru showing the location of *in situ* studies, as well as the presence and erosion of the phureja potato.

A report from the mid-1940s noted that commercial potato varieties in Peru included chata blanca, chata rosada, chata negra, chata negra, amarilla, shitu, mauna and chaucha (Derteano, 1944). In the 1960s, C. Ochoa documented the presence of specimens of chaucha blanca, chaucha colorada and chaucha negra in the province of Sandia, Puno (Ochoa, 1964). Subsequently, Soukup (1970) described the chaucha potato or phureja potato as an early maturing variety characterized by its rapid cooking, even at the first boiling.

In Cajamarca, Iberico (1981) noted that the chaucha potato is distinguished by its characteristic colour, similar to that of egg yolk, as well as its extremely smooth texture and pleasant flavour. This tuber is commonly consumed parboiled and with skin. Leiva *et al* (1990) documented its use in traditional medicine, specifically to relieve headaches. In this context, it is applied sliced with salt as a poultice on the temples, or ground in combination with olluco (*Ullucus tuberosus*), white corn (*Zea mays*), salt and vinegar for therapeutic purposes.

C. Ochoa, noted for his extensive potato collections between 1947 and 1997, collected a total of 322 phureja potato samples from 15 departments of Peru. These samples were organized according to their relative abundance as follows: Amazonas (140), Puno (52), Cajamarca (30), Ayacucho (28), Piura (16), Huánuco (12), Junín (11), Lima (11), La Libertad (6), Apurímac (4), Huancavelica (4), Ancash (2), Cusco (2), Pasco (1) and Lambayeque (1). These data reflect the distribution and prevalence of the phureja potato during this period. However, Ochoa (2003) noted that these samples were preserved only as herbarium specimens in various museums in Peru and other parts of the world.

Ex situ conservation of the phureja potato in Peru

In Peru, 6,295 native potato accessions have been identified in eight databases of working collections,

including a university, a civil association and four experimental stations of INIA located in the Sierra (Table 1). Of these, only 0.3% correspond to the Phureja Group, which shows its minimal representation, even in the three main collections of the Cusco region, where its presence is insignificant. At the international level, CIP, which houses the world collection of potato germplasm, conserves only nine accessions of *Solanum tuberosum* L. Phureja Group from Peru, out of a total of 181 accessions of this species, most of which originate from Colombia and Ecuador (CIP, 2023).

The database of the Paucartambo Native Potato Growers' Association initially registered 2,500 cultivars. Each farmer maintains between 38 and 140 cultivars on average. Within this group, only two cultivars belong to the Phureja Group: Puka q'achirma and K'ello q'achirma. Both varieties are kept by a single farmer in the locality of Challabamba.

Julio Hancco, a well-known conservationist from Pampacorral, Cusco, reported that he had a collection of 350 potato cultivars, although none belong to the Phureja Group. He also reported having limited knowledge of this category of potato. Consistently, the analysis of the catalogue of a lot composed of 115 documented cultivars from his collection did not show the presence of cultivars of the Phureja Group either (Hancco *et al*, 2008).

In 2010, 20 conservationists of the Cajamarca Potato Producers Association dedicated to the conservation of native potatoes, including cultivars of the Phureja or Chaucha Group, were identified in Cajamarca. Of these, 17 maintained between 1 and 12 cultivars of the Phureja Group (Table 2). By 2021, the number of active conservationists decreased to 18, of which only 11 preserved between 1 and 10 cultivars of this group (Table 2). The total reduction in the number of native potato accessions in general and of the Phureja Group in particular was 50% and 52%, respectively.

The decrease in cultivars observed in 2021, together with the reduction in the number of active conservationists, is partly attributed to the advanced age of the latter, with an average age of 62 years, which limited their ability to continue conservation work. In addition, there were significant changes in the main economic activities of the conservationists: five focused primarily on livestock, four replaced agriculture with other occupations (such as carpentry, salaried work or temporary migration), three moved to urban areas while maintaining only partial agricultural activities, and one completely abandoned his activities due to chronic illness. Together, the remaining small collections total 35 accessions conserved *ex situ* in Cajamarca (Table 2).

In situ conservation of the phureja potato in Peru

Between 2006 and 2023, 27 regional catalogues of native potato were prepared, of which 19 met the established inclusion criteria. A total of 2,102 native potato accessions were identified in these catalogues, of which 4.3% correspond to the Phureja group (Table 3). Among the most outstanding catalogues are: Cuyo Cuyo (Puno), in the south, with 13 cultivars of the Phureja group (Wildlife Conservation Society , 2022); Huánuco, in the centre of the country, with 17 cultivars of phureja potato (Egúsquiza, 2015); and Cajamarca, in the north, which records 43 cultivars of Phureja. The Cajamarca collection was morphologically characterized, highlighting the first 15 accessions collected in three districts of Hualgayoc, which achieved a successful harvest (Figure 2) (Seminario and Zarpán, 2011; Seminario *et al*, 2019). In addition, in 2021, in the Cajamarca provinces of Chota and Cutervo, 24 cultivars of Phureja planted in local farms during the last agricultural seasons were identified, and identified by their traditional names.

The exploration conducted in the community of Waqanqa (Paucartambo, Cusco) indicated that potato cultivation has been reduced in area and cultivars. No phureja potato cultivars were found. On the contrary, corn, apple and pasture crops stood out. Also, 16 farmers participating in the XIX Agricultural Fair of Paucartambo (11-12 September 2023), who were interviewed, indicated that they maintained between 2 to 250 varieties of native potato, and nine of them planted one to two chaucha cultivars (puka chaucha or k'ello chaucha). They mentioned that these potatoes had been lost due to lack of time to attend to the crop, because of their rapid sprouting and the scarcity of seed.

Empirical evidence of genetic erosion of the phureja potato in Peru

Genetic erosion in potatoes of the Phureja Group had its first documented evidence in Waqanqa, the largest community in the Mapacho river valley, in Paucartambo. This group of potatoes, traditionally cultivated in the region, showed a progressive disappearance since the 1960s, being reduced to only four or five varieties in 1987. Factors such as labour shortages, migration to urban areas and low local appreciation of these cultivars led to their replacement by improved varieties (Zimmerer, 1991, 1992).

Canahua *et al* (2002) conducted a study in regions with extensive areas of potato and quinoa cultivation in six provinces of Puno. The results indicated that the cultivation of potatoes of the Phureja Group is restricted to small areas in the localities of Moho, Yunguyo and Viquechico. In addition, it was concluded that these cultivars are in danger of disappearing due to their low yields, despite being valued for their short maturity cycle and remarkable regeneration capacity.

De Haan and Thiele (2004) documented a decline in the frequency of cultivation of cultivars of the Phureja Group in the district of Yauyos, Lima. This phenomenon was attributed to factors such as limited seed availability, low capacity for germplasm exchange, labour shortages, and farmers' increasing preference for more commercial cultivars such as 'Huayro' and 'Peruanita'.

Table 1. Accessions of native potato and phureja potato, maintained *ex situ* in regional genebanks in Peru, 2020. [a], 778 collected within the Potato Park, 410 repatriated from CIP and 150 obtained through seed exchange; [b], Collected in Puno. N/A: Information not available.

Location	Native potato	Phureja potato	Source	Institutions
Cusco	2,500	12	L. Lizárraga, interview, 18 Feb 2020	RCABR -UNSAAC
Pisac, Cusco	1,345[a]	N/A	A. Argumedo, interview, 15 Feb 2020	Andes Association, Potato Park
Cusco	1,300	5[b]	L. Palomino, interview, 20 Apr 2020	INIA, Andenes
Ayacucho	400	N/A	M. Morote, interview, 2 May 2020	INIA, Canán
Junín	300	3	N. Zúñiga, interview, 2 Mar 2020	INIA, Santa Ana
Puno	450	1	R. Cahuana, interview, 15 May 2020	INIA, Illpa
Cajamarca	N/A	N/A	H. Roncal, interview, 25 May 2020	INIA, Baños del Inca
Total	**6,295**	**21**		

Table 2. Native potato accessions (total and phurejas) maintained by conservationist farmers in Cajamarca in 2010 and 2021. Source: Interviews conducted in 2010 (Seminario and Zarpán, 2011) and 2021, respectively. N/A: not available.

Conservationist name	Province/district	Age in 2021	2010		2021	
			Total acc.	Phureja acc.	Total acc.	Phureja acc.
Juan Huaccha	San Marcos/ Pedro Gálvez	62	200	2	90	2
Santos Abanto	San Marcos/Gregorio Pita	61	90	3	30	1
Pedro I. Abanto	San Marcos/Pedro Gálvez	49	45	3	25	0
Orestes Dávila	San Marcos/ José Sabogal	65	40	1	20	1
Termópilo Arévalo	Celendín/Sorochuco	59	90	4	100	10
Sergio Rodríguez	Celendín/Sorochuco	65	82	N/A	30	3
Armando Vergara	Celendín/Huazmín	54	100	10	N/A	N/A
Segundo D. Gil	Celendín/Huazmín	66	83	12	50	5
Idelso Garay	Celendín/Huazmín	71	65	2	80	0
Alindor Díaz	Cajamarca/ La Encañada	74	45	3	15	3
Miguel Riquelme	Cajamarca/La Encañada	73	45	4	50	2
Gumercindo Zelada	Cajamarca/Encañada	55	45	4	10	2
José I. Ayay Valdez	Cajamarca/ Cajamarca	70	75	0	70	1
Germán Sangay	Cajamarca/ Encañada	58	35	3	60	0
Emilio Huamán	Cajamarca/Namora	48	295	4	0	0
Abel Marín Ríos	Cajamarca/Namora	61	180	1	15	0
Juan E. Mendoza	Hualgayoc/Bambamarca	50	15	6	20	5
José Telmo Cabrera	San Marcos/Gregorio Pita	65	80	3	70	0
Luis Cabrera Ocas	San Marcos/Gregorio Pita	70	180	2	160	0
Wilson Pastor Marín	San Marcos/Gregorio Pita	67	50	0	20	0
Total			**1,840**	**67**	**915**	**35**

CIP and FEDECH (2006), together with De Haan *et al* (2010), conducted a study in Huancavelica on two potato groups, identifying the presence of 144 and 557 cultivars from four provinces and eight communities, but found no cultivars belonging to the Phureja Group. This absence was attributed to factors such as seed loss, temporary migration of inhabitants, substitution by modern cultivars and the lack of dormancy characteristic of this species. However, Brush *et al* (1981) documented the existence of a single cultivar of *Solanum phureja* called 'Pujuya' in the community of Aymará, Tayacaja district, Huancavelica, known for its frost resistance and also cultivated by farmers in nearby regions such as Chongos Alto, in Huancayo, Junín.

The Ministry of Agriculture conducted research in the southeastern Junín department on native potatoes, identifying that it was currently difficult to find cultivars of the Phureja Group (MINAGRI, Grupo Yanapai, INIA and CIP, 2017). Similarly, Plasencia *et al* (2018), in a study on the diversity of native potatoes in Challabamba (Paucartambo) and Quillcas (Junín), reported the presence of all the species studied, except *S. tuberosum* L. Phureja Group, although the causes of its absence were not determined.

Seminario and Zarpán (2011) and Seminario *et al* (2019) estimated a 17% reduction in cultivars of the Phureja Group in five provinces of Cajamarca over the previous two decades. The main causes of this decline include the lack of time to attend to the crop, due to its short sowing and harvesting cycles; the preference of local communities for livestock activities rather than agriculture; the low quality of seed, which in many cases

Table 3. Number of potato phureja cultivars in 19 Peruvian native potato catalogues in 2023.

Regions/Communities	Total cultivars	Phureja potato	Source
Huancavelica/ Huayta Corral, Tupac Amaru, Villa Hermosa, Pucara, Dos de Mayo	144	0	CIP and FEDECH (2006)
Cusco/ Huama, Huarqui, Poques, Patacancha, Willoc, Tauca	260	0	Cosio (2006)
Cusco/ Palccoyo, Acco Acco Phalla y Quisini (district of Sicuani)	141	0	Gutiérrez and Valencia (2010)
Cajamarca/ Chota and Lajas	23	5	INCAP Jorge Basadre (nd)
Cajamarca/ Three communities of Shitamalca	24	1	Programa Bioandes (nd)
Cajamarca	28	5	Cabrera and Pando (2011)
Cajamarca/22 communities	43	43	Seminario *et al* (2019)
Puno	86	0	Muñoz and Estaña (2012)
Cusco/ Quescay, Kcallacancha, Sipascancha Alta, Miscahuara	30	4	Revilla (2014)
Apurímac y Huancavelica.	24	0	Fonseca *et al* (2014)
La Libertad/ San Juan, La Soledad, Canucubamba, Macullida, Las Colpas, Arcopampa y Chugay	129	1	CIP, Asociación Pataz., and INIA (2015)
Huánuco/ 35 communities	296	17	Egúsquiza (2015)
Huánuco, Junín, Huancavelica, Ayacucho, Apurimac	12	0	Riveros and Peralta (2015)
Junín/ Seven communities and 14 families in the southeastern part of the department	146	1	MINAGRI, Grupo Yanapai, INIA and CIP (2017)
Apurímac/ Huayana y Pomacochas	119	0	PRODERIN (2018)
Apurimac, Cusco y Puno/113 communities of Apurimac, 58 de Cusco and 8 of Puno	200	0	Roldan *et al* (2019)
Huancavelica/ Castillapata, Paltamachay, Huachhua, Paccho Molinos, Santa Rosa de Pacchaccla y Pumaranra	184	0	CIP, Grupo Yanapai, Gobierno Regional de Huancavelica and AGUAPAN (2021)
Puno/ Cuyocuyo	91	13	Wildlife Conservation Society (2022)
La Libertad/ La Victoria	122	0	Asociación-Pataz, CIP, INIA and AGUAPAN (2023)
Total	**2,102**	**90**	

was depleted and with poor yields; the expansion of mining projects in the region; the trend towards the cultivation of modern varieties and more commercial native cultivars; labour shortages, caused by emigration and employment in non-agricultural activities; and the limited availability of seed for crop regeneration.

In 2021, a field study in the provinces of Chota and Cutervo (Cajamarca) revealed the loss of eight Phureja potato cultivars in the last two decades, reducing the total recorded from 32 to 24 (Table 2). The missing cultivars included Huevo de perdiz, Baya, Cemelina, Rosada, Morada, Negra, Chilopa and Amapola. The main causes identified are the rapid growth cycle of this group, which makes its management difficult; the preference for modern cultivars that are more competitive in the market; the prioritization of livestock; and the low quality of the available seeds, known as 'tired seeds'. Currently, the remaining germplasm of Phureja potato in Cajamarca comprises 67 cultivars, of which 43 have been morphologically characterized in five provinces, while 24 are only nominally registered in Chota and Cutervo. It is essential to extend studies to other provinces to evaluate diversity and promote its conservation.

Discussion

Historical evidence (Vargas, 1936, 1946, 1955; Herrera, 1921; Varcárcel, 1985) and Ochoa (2003) collections in 15 of the 19 departments where potatoes are grown in Peru highlight the historical relevance of the Phureja potato. Its distinctive characteristics, such as the absence of dormancy, precocity and adaptation to early harvests, underline its agricultural importance. However, it is necessary to thoroughly review the historical data (van de Wouw *et al*, 2009) and to explore again the sites visited by Ochoa to confirm the persistence of Phureja in these regions. This will allow updating knowledge about its persistence, as well as its agricultural and cultural value in the current context.

Ex situ collections are essential to preserve genetic diversity and prevent its loss, acting as a vital complement to *in situ* conservation. Both strategies

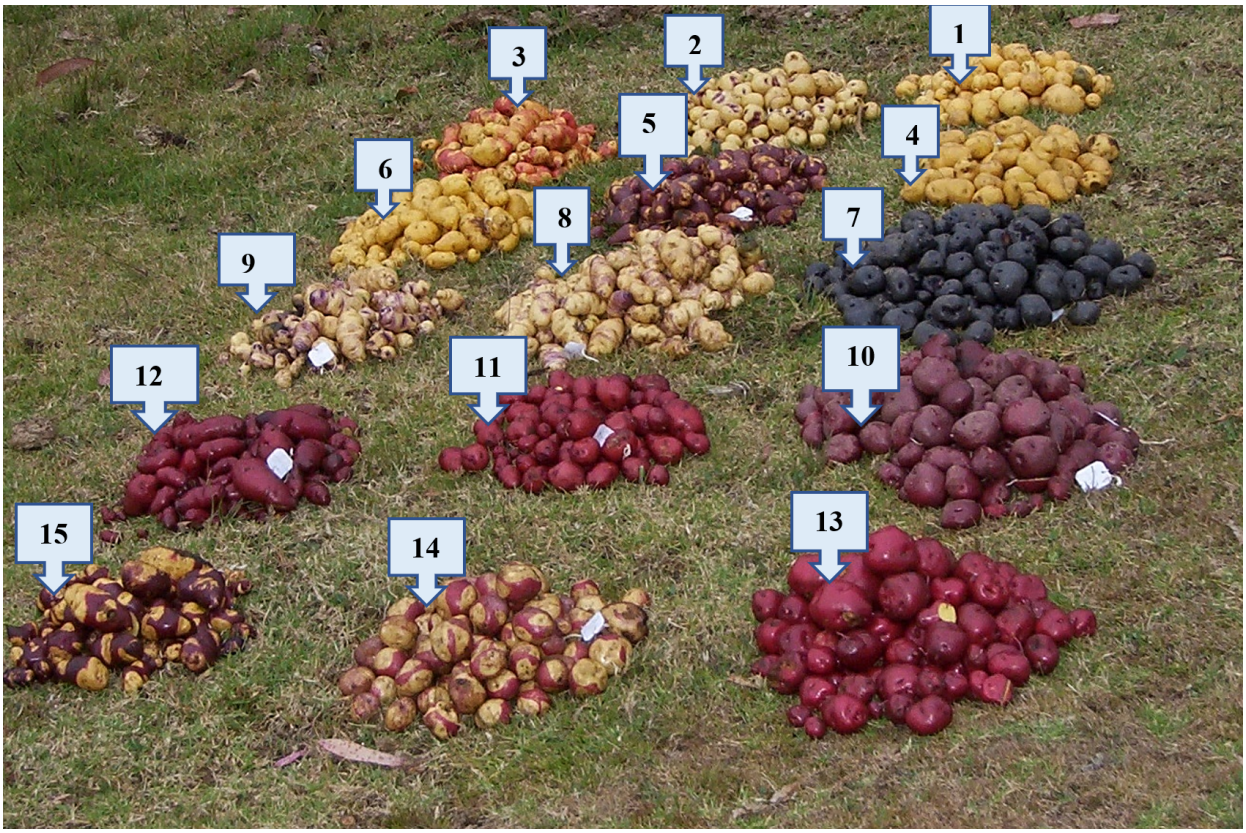


Figure 2. Phureja or chaucha potato cultivars harvested in the province of Hualgayoc, Cajamarca, northern Peru, in 2005: 1.Chaucha (Ch). Chachapoyana, 2. Ch. Blanca, 3. Ch. Huagalina, 4. Ch. Porpora, 5. Ch. Conda, 6. Ch. Amarilla, 7. Ch. Negra, 8. Ch. Colombiana, 9. Ch. Colombina negra, 10. Ch. Montañera, 11. Ch. Shoga, 12. Ch. Pabla, 13. Ch. Pimpinela, 14. Ch. Blanca peruanita, 15. Ch. Clavelina. In northern Peru, the names of phureja potato cultivars are usually composed of two elements: a primary common name, such as chaucha, and a secondary name which, in many cases, may coincide with the names of cultivars belonging to other potato groups. For example, Ch. Huagalina, whose secondary names may be related to specific physical characteristics.

are integrated to ensure the availability of these resources for future generations (Priyanka *et al*, 2021). A key advantage of *ex situ* collections is their accessibility to researchers and users, allowing their use in training, genetic improvement, research and repatriation to source communities (Fu, 2017; Ellis *et al*, 2020; Nagel *et al*, 2022). However, in Peru, *ex situ* germplasm of Phureja potato is limited, with only 65 accessions, distributed among state institutions (21), conservationists (35) and CIP (9). The reasons for this scarcity of *ex situ* samples need to be investigated. This could be due to the scarcity of these materials on farms or to a lack of interest among researchers and conservationists, due to the difficulty of preserving them because of their lack of dormancy.

This germplasm does not represent a significant complement for *in situ* conservation, nor does it constitute a solid base for the repatriation of cultivars (Joshi *et al*, 2020; Lüttringhaus *et al*, 2021). The most prominent collections of Phureja potato in the Andes are in Colombia: that of Nariño and northern Ecuador, together with the Colombian central collection, which together conserve 348 accessions (Rodríguez, 2010; Lasso *et al*, 2018). Surprisingly, the United States Potato Germplasm Bank (USPG) houses 144 Phureja potato accessions, all unique and without duplicates (Río and Bamberg, 2021). This highlights the importance of integrating international efforts for the conservation and study of this valuable genetic diversity.

Phureja potato cultivars maintained by conservationists or potato guardians in Cusco are scarce, while in Cajamarca they have experienced a drastic reduction of 48% between 2010 and 2021. Furthermore, the socioeconomic conditions of the 20 conservationists studied in Cajamarca are not favourable to guarantee efficient conservation. This evidences the need to fill a critical gap through studies on *in situ* conservation of phureja potatoes in other unexplored regions and to delve deeper into conservation dynamics in Cusco and Cajamarca. Addressing these areas will allow a more comprehensive understanding of the strategies needed to preserve this valuable genetic diversity in the context of the Peruvian Andes. The use of tools such as the four- or five-cell methodology offers an efficient way to obtain this information in a short period (Rana *et al*, 2006; Padulosi and Dulloo, 2012).

For *in situ* conservation, information from catalogues on native potatoes and specific reports in Cajamarca and Cusco were used. However, these catalogues present high variability in their content, influenced by the approach, purposes and descriptors used in their elaboration. Despite these differences, they proved to be

a valuable resource for the objectives of this research, as they provided an approximate view of the materials found *in situ*. Although they reflect information from a specific time, these documents also record cultural aspects and traditional knowledge, highlighting the work of potato conservationists. In addition, they can serve as an essential baseline for ongoing monitoring and evaluation of genetic diversity (MINAGRI, Grupo Yanapai, INIA and CIP, 2017).

The evidence gathered in this research suggests that, in Peru, *ex situ* and *in situ* conservation strategies for phureja potato are not operating in a complementary and efficient manner (Nagel *et al*, 2022). However, the data obtained may constitute a valuable reference for monitoring cultivar conservation, a priority aspect that has received little attention in cultivars in general (Padulosi and Dulloo, 2012).

The information available on the genetic erosion of the phureja potato in Peru is limited, with research conducted in only a few communities in six potato-producing departments. However, these studies provide an indication of the situation that may be occurring at the national level. Broader regional research that addresses remaining genetic variability and its relationship to the environment is essential.

The loss of cultivars is associated with factors such as the lack of time to plant and harvest crops in short periods, reflecting migration, and the prioritization of more profitable activities, such as dairy cattle ranching. In Cajamarca, areas previously dedicated to annual crops and potatoes are now used as pasture for dairy cattle. This trend, observed since the early 2000s (Winters *et al*, 2006), has been encouraged by the presence of three large milk collection companies and 1,052 artisanal dairy plants (INDECOPI, 2023), which guarantee investment security, attractive prices and immediate income for producers.

Cultivar loss is also attributed, in part, to poor seed quality (farmers say, “it no longer yields, it’s tired”), reflecting seed degeneration due to pathogen accumulation after prolonged vegetative propagation (Forbes *et al*, 2020; Sierra *et al*, 2020). The preference for modern cultivars and some native cultivars of greater commercial acceptance, observed in Waqanqa, Yauyos, Huancavelica and Cajamarca, is also a contributing factor, although studies are required to assess their impact. In addition, mining projects, especially in Cajamarca, where 53.6% of the territory of the Sierra provinces is concessioned to 33 mining companies (GPC, 2014), affect cultivars. This includes direct effects, such as employment (40% of workers in Yanacocha are local) (Yanacocha, 2018), and indirect effects, such as land sales and migration, which weaken agricultural sustainability.

In Peru, the possible occurrence of allelic or gene erosion, defined as the loss of alleles and their combinations, and genomic erosion, which implies the complete loss of the genome, has been observed in the phureja potato (Thormann and Engels, 2015). Allelic erosion occurs mainly due to the replacement of these cultivars with modern or traditional varieties of higher commercial value. Genomic erosion, on the other hand, is manifested through genetic displacement due to the elimination of the phureja potato from cropping systems. If this process continues, the implications would be serious for the country, as it would face the definitive loss of this species (*S. tuberosum* L. Phureja Group), which would have a significant impact on agricultural biodiversity and food security.

Conclusion

The phureja potato, is present in 11 of Peru’s 25 regions. Despite its historical relevance and former abundance, this valuable crop is now in danger of disappearing in the country. The germplasm of the phureja potato in Peru includes 90 accessions conserved *in situ*, 21 *ex situ* accessions maintained by state institutions, and 35 accessions safeguarded by farmers in the Cajamarca region. This research represents the first systematic effort to document the genetic diversity and erosion of this resource. It is crucial to complement this initial analysis with new regional data to validate its consistency and to develop a comprehensive national inventory to identify and preserve the remaining genetic variability.

The loss of genetic resources and the erosion of the phureja potato are influenced by several interrelated factors. Among them, farmers’ lack of time to attend to this crop, which requires continuous planting and harvesting, is aggravated by labour shortages due to the emigration of the most skilled members of rural families. In addition, the shift to more profitable activities, such as dairy farming, and the priority given to improved cultivars or commercial varieties displace the phureja potato. Other factors include the scarcity and poor quality of seed, and the presence of mining projects that affect cultivation areas.

Acknowledgements

The authors extend special thanks to: the native potato conservationists of Cajamarca and Paucartambo (Cusco); the professionals in charge of potato collections at INIA experimental stations; the farmers of Cajamarca and Paucartambo who participated in the interviews; to Alejandro Argumedo, Director of the Asociación Andes del Cusco, for facilitating access to the database of native potato from the Potato Park; to Luis Lizárraga, Director of the Biodiversity Research Center of UNSAAC; and to the Saucedo Rojas family of Paucartambo, for allowing us to examine the database of the collections that the native potato conservationists of this province maintain.

Author contributions

JFSC: conceptualization, methodology, writing original draft, revision. LSCT: data collection, formal analysis. ASC: writing original draft, review. TMH: data collection

and curation, WS: methodology, visualization, review and editing.

Ethics statement

The authors declare that this research did not require the approval of an ethics committee, as no clinical or experimental procedures requiring such approval were performed. However, the ethical principles applicable to research involving human subjects were strictly adhered to. Before each interview, participants were clearly informed about the objectives of the study, the voluntary nature of their participation, and their right to withdraw at any time without penalty. Informed consent was obtained orally and in language accessible to each interviewee. Each person was also given the option of authorizing the use of their name in the research. Their dignity, rights, and autonomy were respected at all times.

Conflict of interest statement

The authors declare no known conflicts of interest, financial or personal relationships that could influence the work or materials presented in this article.

References

Asociación-Pataz, CIP, INIA and AGUAPAN (2023). Catálogo de variedades de papa nativa de Tayabamba, La Libertad, Perú .

Beltrán-Penagos, M. A., Sánchez-Camargo, A. P., and Narváez-Cuenca, C. E. (2020). Proximal composition, bioactive compounds and biorefinery approach in potato tubers of Solanum tuberosum Group Phureja: a review. *International Journal of Food Science and Technology* 55(6), 2282–2295. doi: https://doi.org/10.1111/ijfs.14461

Brush, S. B., Carney, H. J., and Humán, Z. (1981). Dynamics of Andean potato agriculture. *Economic Botany* 35(1), 70–88. doi: https://doi.org/10.1007/BF02859217

Burgos, G., Felde, T. Z., Andre, C., and Kubow, S. (2020). The potato and its contribution to the human diet and health. In *The potato crop. Its agricultural, nutritional and social contribution to humankind,* ed. Campos, H. and Ortiz, O. 37-74. doi: https://link.springer.com/book/10.1007/978-3-030-28683-5.

Cabrera, H. and Pando, R. (2011). Catálogo de variedades mejoradas y nativas de papa en la Región de Cajamarca. Estación Experimental Baños del Inca, Instituto Nacional de Investigación e Innovación Agraria volume 75 (Martínez Compañón), 75p.

Canahua, A., Tapia, M., Ichuta, A., and Cutipa, Z. (2002). Gestión del espacio agrícola y agro diversidad en papa y quinua en comunidades campesinas de Puno. In *SEPIA IX,* ed. Pulgar-Vidal, M., Zegarra, E., and Urrutia, J. .

Cerón-Lasso, M., Alzate-Arbeláez, A. F., Rojano, B. A., and Ñuztez-Lopez, C. E. (2018). Composición fisicoquímica y propiedades antioxidantes de genotipos nativos de papa criolla (Solanum tuberosum GrupoPhureja). *Informacion Tecnologica* 29(3), 205–216. doi: https://doi.org/10.4067/S0718-07642018000300205

CIP (2023). CIP Genebank. url: https://genebank.cipotato.org/gringlobal/search.aspx.

CIP and FEDECH (2006). Catálogo de variedades de papa nativa de Huancavelica. Centro Internacional de la Papa; Federación Departamental de Comunidades Campesinas de Huancavelica (Metrocolor). url: https://hdl.handle.net/10568/101328.

CIP, Asociación Pataz., and INIA (2015). Catálogo de variedades de papa nativa de Chugay, La Libertad, Perú (Centro Internacional de la Papa).

CIP, Grupo Yanapai, Gobierno Regional de Huancavelica and AGUAPAN (2021). Catálogo Línea de Base de la Diversidad de Papa Nativa del Microcentro Yauli - Paucará, Huancavelica, Perú. Centro Internacional de la Papa (CIP), Lima, Perú, 300p.

Cosio, P. (2006). Variabilidad de papas nativas en seis comunidades de Calca y Urubamba, Cusco (Cusco, Perú: Asociacion Arariwa).

Dawson, T., Juarez, H., Maxted, N., and De Haan, S. (2023). Identifying priority sites for the on-farm conservation of landraces and systematic diversity monitoring through an integrated multi-level hotspot analysis: the case of potatoes in Peru. *Frontiers in Conservation Science* 4. doi: https://doi.org/10.3389/fcosc.2023.1130138

De Carvalho, M. A., Bebeli, P. J., Silva, A. M. B. D., Bettencourt, E., Slaski, J. J., and Dias, S. (2016). Agrobiodiversity: The importance of inventories in the assessment of crop diversity and its time and spatial changes. In *Genetic diversity and erosion in plants, sustainable development and biodiversity ,* ed. Ahuja, M. R. and Jain, S. M. (Springer), volume 8. doi: https://doi.org/10.1007/978-3-319-25954-3_9.

De Haan, S., Núñez, J., Bonierbale, M., and Ghislain, M. (2010). Multilevel agrobiodiversity and conservation of Andean potatoes in Central Peru: Species, morphological, genetic, and spatial diversity. *Mountain Research and Development* 30(3), 222–231. doi: https://doi.org/10.1659/MRD-JOURNAL-D-10-00020.1

De Haan, S. and Thiele, G. (2004). In situ conservation and potato seed systems in the Andes. In *Seed systems and crop genetic diversity on-farm,* ed. Jarvis, D. I., Sevilla-Panizo, R., Chávez-Servia, J. L., and Hodgkin, T. IPGRI (International Plant Genetic Resources Institute).

Derteano, C. (1944). Informe sobre el cultivo y costo de producción de papa. *Revista Nacional de Agricultura* 9(39).

Devaux, A., Goffart, J. P., Petsakos, A., Kromann, P., Gatto, M., Okello, J., Suarez, V., and Hareau, G. (2020). Global Food Security, Contributions from Sustainable Potato Agri-Food Systems. In *The potato crops. Its agricultural, nutritional and social contribution to humankind,* ed. Campos, H. and Ortiz,

O. volume 1, 3-34. doi: https://doi.org/10.1007/978-3-030-28683-5.

Egúsquiza, R. (2015). Catálogo de papas nativas cultivadas en Huánuco (Universidad Nacional Agraria La Molina).

Ellis, D., Salas, A., Chávez, O., Gomez, R., and Anglin, N. (2020). Ex situ conservation of potato [Solanum Section Petota (Solanaceae)] genetic resources in genebanks. In *The potato crop. Its agricultural, nutritional and social contribution to humankind,* ed. Campos, H. and Ortiz, O. volume 4, 109-138. doi: https://doi.org/10.1007/978-3-030-28683-5.

Fonseca, C., Burgos, G., Rodríguez, F., Muñoa, L., and Ordinola, M. (2014). Catálogo de variedades de papa nativa con potencial para la seguridad alimentaria y nutricional de Apurímac y Huancavelica. Lima: Centro Internacional de la Papa.

Forbes, G. A., Charkowski, A., Andrade-Piedra, J., Parker, M., and Schulte-Geldermann, E. (2020). The seed potato. In *The potato crop. Its agricultural, nutritional and social contribution to humankind,* ed. Campos, H. and Ortiz, O. volume 12, 431-447. doi: https://link.springer.com/book/10.1007/978-3-030-28683-5.

Fu, Y. B. (2017). The vulnerability of plant genetic resources conserved ex situ. *Crop Science* 57(5), 2314–2328. doi: https://doi.org/10.2135/cropsci2017.01.0014

Gabriel, J. and Franco, J. (2013). Solanum phureja Juz et Buk.: Valuable Source of Genetic Resistance to Potato Late Blight [Phytophthora infestans (Mont.) de Bary]. *Revista Latinoamericana de la Papa* 17(2), 131–142.

GPC (2014). Concesiones mineras en el Perú. Análisis y Propuestas de Política.

Gutiérrez, R. and Valencia, C. (2010). Las papas nativas de Canchis. Un catálogo de biodiversidad. Intermediate Technology Development Group (ITDG). Servicios Gráficos JMD, Lima, Perú.

Hancco, J., Blas, R., and Quispe, M. (2008). Pampacorral: catálogo de sus papas nativas (Universidad Nacional Agraria La Molina).

Herrera, F. (1921). Flora del departamento del Cuzco. Primera parte (Universidad del Cuzco), 2 edition.

Huamán, Z. and Spooner, D. M. (2002). Reclassification of landrace populations of cultivated potatoes (Solanum sect. Petota). *American Journal of Botany* 89(6), 947–965. doi: https://doi.org/10.3732/ajb.89.6.947

Iberico, L. (1981). El Folklore Agrario de Cajamarca (Universidad Nacional de Cajamarca).

INCAP Jorge Basadre (n.d.) Características del germoplasma nativo de Chota (Cajamarca, Perú) 27p.

INDECOPI (2023). Estudio de mercado sobre el sector lácteo en el Perú. Informe final (Lima, Perú: Instituto Nacional de Defensa de la Competencia y de la Protección de la Propiedad Intelectual).

Joshi, B. K., Gauchan, D., Bhandari, B., and Jarvis, D. (2020). Good practices for agrobiodiversity management (Kathmandu, Nepal: NAGRC, LI-BIRD and Bioversity International).

Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., De Haan, S., Engels, J. M. M., Guarino, L., Hoban, S., Mercer, K. L., Miller, A. J., Nabhan, G. P., Perales, H. R., Richards, C., Riggins, C., and Thormann, I. (2022). Crop genetic erosion: understanding and responding to loss of crop diversity. *New Phytologist* 233(1), 84–118. doi: https://doi.org/10.1111/nph.17733

Lasso, Z., Romero, Y., Coronel, B., Pérez-Cardona, O., and Valbuena, R. I. (2018). Dry matter and specific gravity content evaluation in the Central Colombian Collection Solanum tuberosum Group Andigena. World Potato Congress (Cusco, Perú) .

Leiva, S., Gonzalo, M., and Saenz, R. (1990). Medicina del campo (Departamento de Acción Social del Obispado de Cajamarca. Área de salud-medicina natural).

Lüttringhaus, S., Pradel, W., Suarez, V., Manrique-Carpintero, N. C., Anglin, N. L., Ellis, D., Hareau, G., Jamora, N., Smale, M., and Gómez, R. (2021). Dynamic guardianship of potato landraces by Andean communities and the genebank of the International Potato Center. *CABI Agriculture and Bioscience* 2(1). doi: https://doi.org/10.1186/s43170-021-00065-4

Martinet, J. B. (1977). La Agricultura en el Perú. Reimpreso del original de 1877 (Centro Peruano de Historia Económica, Universidad Nacional Mayor de San Marcos).

MINAGRI, Grupo Yanapai, INIA and CIP (2017). Catálogo de variedades de papa nativa del sureste del departamento de Junín - Perú . doi: https://doi.org/10.4160/9789290602088

Muñoz, C. and Estaña, W. (2012). Diversidad y variabilidad de papa nativa en Puno. Dirección Regional Agraria Puno (Lima, Perú) .

Nagel, M., Dulloo, M. E., Bissessur, P., Gravrilenko, P., Bamberg, J., Ellis, D., and Giovannini, P. (2022). Global strategy for the conservation of potato (Global Crop Diversity Trust). doi: https://doi.org/10.5447/ipk/2022/29

Ñústez, C. E. (2021). Papa chaucha (Solanum phureja). In *Los Andes y los alimentos del futuro. 50 Andean future foods,* ed. de Haan, S., Zeigler, M., and Guzmán, F. 157p.

Ñústez, C. E. and Rodríguez, L. E. (2020). Papa criolla (Solanum tuberosum Grupo Phureja (S. B. Universidad Nacional de Colombia).

Ochoa, C. (1999). Las papas de Sudamérica. Perú (Lima, Perú: Centro Internacional de la Papa).

Ochoa, C. M. (1964). Recuentos cromosómicos y determinación sistemática de papas nativas cultivadas en el Sur del Perú. *Anales Científicos. Perú* 2(1).

Ochoa, C. M. (2001). Las papas de Sudamérica. Bolivia (Lima, Perú: Centro Internacional de la Papa).

Ochoa, C. M. (2003). Las papas del Perú. Base de datos 1947-1997 (Gabriela Alcántara).

Padulosi, S. and Dulloo, E. (2012). Towards a viable system for monitoring agrobiodiversity on-farm: A

proposed new approach for Red Listing of cultivated plant species. In *On farm conservation of neglected and underutilized species: trends and novel approaches to cope with climate change. Proceedings of an International Conference held in Frankfurt, Germany, 14-16 June 2011,* ed. Padulosi, S., Bergamini, N., and Lawrence, T. (Bioversity International), 171-187. url: https://hdl.handle.net/10568/42046.

Peña, C., Palomeque, L., Restrepo-Sánchez, L. P., Kushalappa, A., Mosquera, T., and Narváez-Cuenca, C. E. (2021). Variation of mineral contents with nutritional interest in a collection of Solanum tuberosum group Phureja tubers. *International Journal of Food Science and Technology* 56(9), 4594–4603. doi: https://doi.org/10.1111/ijfs.15115

Plasencia, F., Juarez, H., Polreich, S., and De Haan, S. (2018). Evaluación de la distribución espacial de la biodiversidad de papa en los distritos de Challabamba en Cusco y Quilcas en Junín mediante el uso del mapeo participativo. *Rev. del Instituto de Investigación de La Facultad de Ingeniería Geológica, Minera, Metalúrgica y Geográfica* 21(41).

Priyanka, V., Kumar, R., Dhaliwal, I., and Kaushik, P. (2021). Germplasm conservation: Instrumental in agricultural biodiversity-A review. *Sustainability* 13(6743). doi: https://doi.org/10.3390/su13126743

PRODERIN (2018). La papa nativa en Apurímac. Identificación participativa de variedades en los distritos de Huayana y Pomacocha (Lima, Perú: Programa de desarrollo económico sostenible y de gestión estratégica de los recursos naturales en las regiones de Ayacucho, Apurímac, Huancavelica, Junín y Pasco).

Programa Bioandes (n.d.) Variedades de papas nativas y conocimientos campesinos. Microcuenca Shitamalca, San Marcos, Cajamarca (Cajamarca, Perú) 20-20.

Rana, R. B., Sthapit, B., Garforth, C., Subedi, A., and Javis, D. I. (2006). Four-cell analysis as a decision-making tool for conservation of agrobiodiversity on-farm. In *On-farm management of agricultural biodiversity in Nepal: Good practices,* ed. Sthapit, B., Shrestha, P., and Upadhyay, M., (Nepal: NARC/LI-BIRD/Bioversity International). Revised 2012 edition. url: https://hdl.handle.net/10568/104917.

Revilla, L. (2014). Costumbres de las papas nativas (Cusco, Perú: Centro de Servicios Agropecuarios), 187p.

Río, A. D. and Bamberg, J. (2021). An AFLP marker core subset for the cultivated potato species Solanum phureja (Solanum tuberosum L. subsp. andigenum). *American Journal of Potato Research* 98, 493–499. doi: https://doi.org/DOI:10.1007/s12230-021-09849-w

Riveros, C. and Peralta, J. (2015). Catálogo de 12 cultivares de papa nativa inscritas en el Registro de Cultivares Comerciales (Huancayo, Perú: FOVIDA (Fomento de la Vida)).

Rodríguez, L. (2010). Origen y evolución de la papa cultivada. Una revisión. *Agronomía Colombiana* 28(1), 9–17.

Roldan, A., Palomino, L., and Salas, A. R. (2019). Catálogo de variedades de papa nativa de las regiones de Apurímac, Cusco y Puno (Lima, Perú: Ministerio de Agricultura y Riego).

Seminario, J., Tapia, H., and Seminario, A. (2019). Los Solanum del Grupo Phureja de Cajamarca. Avances (Gráfica Bracamonte Heredia).

Seminario, J. and Zarpán, L. (2011). Conservación in situ on farm-ex situ de Solanum tuberosum L. grupo Phureja en la cuenca del Llaucano y áreas adyacentes. *Arnaldoa* 18(2), 103–114.

Sierra, A., Gallo, Y., Estrada, M., Gutiérrez, P. A., and Marín, M. (2020). Detección molecular de seis virus de ARN en brotes de tubérculos de papa criolla (Solanum phureja) en Antioquia, Colombia. *Bioagro* 32(2), 3–14.

Soukup, J. (1970). Vocabulario de los nombres vulgares de la flora peruana (Colegio Salesiano).

Thormann, I. and Engels, J. M. M. (2015). Genetic Diversity and Erosion—A Global Perspective. In *Genetic Diversity and Erosion in Plants. Sustainable Development and Biodiversity,* ed. Ahuja, M. and Jain, S., (Cham: Springer), volume 7.

van de Wouw, M., Kik, C., van Hintum, T., van Treuren, R., and Visser, B. (2009). Genetic erosion in crops: Concept, research results and challenges. *Plant Genetic Resources: Characterization and Utilization* 8(1), 1–15. doi: https://doi.org/10.1017/S1479262109990062

Varcárcel, L. E. (1985). Historia del Perú antiguo a través de la fuente escrita (Librería Editorial Juan Mejía Baca), 5 edition.

Vargas, C. (1936). El Solanum tubersoum a través del desenvolvimiento de las actividades humanas. *Revista Universitaria* 25(70), 138–223.

Vargas, C. (1946). Las papas sudperuanas. Parte I (Universidad Nacional del Cuzco).

Vargas, C. (1955). Las papas sudperuanas. Parte II. *Revista Universitaria* 44(108), 175–240.

Watanabe, L. K., Baigorria, M., and Olcese, O. (2008). Contribuciones al estudio de la papa en el Perú (San Marcos: San Marcos).

Wildlife Conservation Society (2022). Catálogo de variedades de papa nativa de la zona de agrobiodiversidad Andenes de Cuyocuyo (Press Off Graphics E.I.R.L.).

Winters, P., Hintze, L., and Ortiz, O. (2006). Rural development and the diversity of potatoes on farms in Cajamarca, Peru. In *Valuing Crop Biodiversity: On-Farm Genetic Resources and Economic Change,* ed. Smale, M. 161p.

Yabar-Villagarcía, A. A. (2004). El manicomio azul (Cusco, Perú: Paucartambo), 109p.

Yanacocha (2018). Reporte de Sostenibilidad (Yanacocha).

Zimmerer, K. S. (1991). Labor shortages and crop diversity in the Southern Peruvian Sierra. *Geographical Review* 81(4), 414–432. doi: https://doi.org/10.2307/215608

Zimmerer, K. S. (1992). The loss and maintenance of native crops in mountain agriculture. *GeoJournal* 27(1), 61–72. doi: https://doi.org/10.1007/BF00150635

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 14–25
DOI: 10.46265/genresj.TSJG3884
https://www.genresj.org
ISSN: 2708-3764

Agro-morphological and molecular characterization of Argentine maize (*Zea mays* L.) landraces of 'Cristalino Colorado' race

Raquel Alicia Defacio[a,*], Natalia Mercedes Paz[a], Ana Rosa Schlater[a], Marcelo Edmundo Ferrer[a], Sergio Jorge Bramardi[b,c]

[a] *Instituto Nacional de Tecnología Agropecuaria, Estación Experimental Agropecuaria Pergamino, CC31 Pergamino 2700, Buenos Aires, Argentina.*
[b] *Universidad Nacional del Comahue, Departamento de Estadística, Neuquén, Argentina*
[c] *Centro de Investigaciones en Toxicología Ambiental y Agrobiotecnología del Comahue (CONICET – UNCo), Neuquén, Argentina*

Abstract: Despite the high variability of Argentine maize (*Zea mays* L.) landraces, they are scarcely used by breeders due to the limited knowledge available about the genetic merit of these materials. In this study, we evaluated agro-morphological and molecular traits of 36 landraces of the 'Cristalino Colorado' race from Buenos Aires province, Argentina. Fifteen agro-morphological traits and five polymorphic microsatellite markers located on different chromosomes (48 alleles) were used. A principal component analysis was performed using average values of agro-morphological traits across two environments. Molecular markers were subjected to a principal coordinate analysis. A generalized procrustes analysis was used to evaluate agro-morphological and molecular traits together, showing seven groups. Distance between agro-morphological and molecular data had an average value of 0.24 and the range varied between 0.02 (ARZM01017) and 0.45 (ARZM01082). The results show that Argentine landraces of the 'Cristalino Colorado' race are a valuable source of new alleles for crop improvement. Studies of this type facilitate the selection of landraces for introduction in genetic breeding programmes and for the establishment of core collections.

Keywords: Genetic variability, Generalized Procrustes Analysis, SSR markers, agro-morphological traits, landraces.

Citation: Defacio, R.A., Paz, N.M., Schlater, A.R., Ferrer, M.E., Bramardi, S.J. (2025). Agro-morphological and molecular characterization of Argentine maize (*Zea mays* L.) landraces of 'Cristalino Colorado' race. *Genetic Resources* 6 (12), 14–25. doi: 10.46265/genresj.TSJG3884.

Introduction

Maize (*Zea mays L.*) landraces originated from long-term cultivation under natural and artificial selection in different environments and under different cultural management schemes (Xiang *et al*, 2010). These landraces maintain high genetic variation and good adaptation to the natural and anthropological environment where they have evolved (Lucchin *et al*, 2003; Di Pasquale *et al*, 2024). The mechanization of agriculture, the increase of urban areas, changes in consumption patterns and production systems has led to the replacement of landraces by improved varieties or hybrids (Pilling *et al*, 2020). Mechanization of agriculture and new market demands forced breeders to generate more uniform and productive crops with stable yield (Esquinas Alcázar, 2005). This homogeneity resulted in an irreversible loss of genetic variability, known as genetic erosion, with the consequent increase of the vulnerability of agricultural crops to future attack by biotic and abiotic stresses (Salhuana *et al*, 1998; Troyer *et al*, 1988; Esquinas Alcázar, 2005).

Genetic resources have long been an important source of new alleles for commercial plant breeding. However, high variability conserved in germplasm banks worldwide is poorly used because breeders prefer crosses among elite inbred lines for their improvement programmes (Vigouroux *et al*, 2008). Extensive exploitation of landraces is hampered by their high heterogeneity, low performance, seed underproduction and negative genetic load (Gorjanc *et al*, 2016).

* Corresponding author: Raquel Alicia Defacio (defacio.raquel@inta.gob.ar)

Received: 21.03.2025 | Accepted: 09.06.2025 | Published online: xx.xx.2025

Using conserved germplasm in future plant breeding requires systematic evaluation of desired traits (Xiang *et al*, 2010; Balconi *et al*, 2024). For this purpose, there are many descriptors, such as agro-morphological traits and molecular markers, which contribute different and equally important information.

Germplasm characterization and evaluation must be complemented with appropriate statistical analyses to obtain a more complete description of the landraces and establish relations among them. Properly studied and catalogued plant genetic resources can be available for plant breeding programmes (Bramardi, 2023).

In Argentina, as in other countries of the Americas, there is a large diversity of maize types. Argentine maize landraces are classified into 44 races mainly based on specific traits related to ear and grain descriptors, such as shape, colour and texture, and use (Cámara Hernández and Miante Alzogaray, 2003; Solari, 2007). Argentina has a leading role in the production of export maize 'Plata', typical of the 'Cristalino Colorado' race, being currently the only producer of this type of maize worldwide. The kernel of 'Cristalino Colorado' maize is of intense orange colour; the endosperm is mostly hard and glassy in the periphery and floury in the centre, lacking indentation (Secretaría de Agricultura, Ganadería, Pesca y Alimentación, 1997). This kind of grain is widely used in dry milling processes for human consumption and poultry feeding. This grain contains higher carotenoid concentrations than dent corns (Chandler *et al*, 2013); when included in the diet of chicken, it gives a desirable colour to the skin and egg yolks, without the addition of synthetic pigments. Besides, true metabolizable energy values of 'Cristalino Colorado' maize are higher than those of the dent maize due to the higher concentration of oil in the grain. 'Cristalino Colorado' maize can supply the calories required by cattle and pigs, with no need for additional oil in their diet. A 'Cristalino Colorado' hybrid is available in the Argentine market; however, it is of lower quality than traditional genotypes (Paz, 2009).

Previous studies of maize landraces conserved in the Active Germplasm Bank at 'Instituto Nacional de Tecnología Agropecuaria' (INTA) Pergamino (BAP) revealed a high degree of molecular and agro-morphological variability (Salhuana *et al*, 1998; Defacio *et al*, 2005; Paz *et al*, 2005; Defacio, 2009; Paz, 2009; Defacio, 2017; Heck *et al*, 2020; Rivas *et al*, 2022), as well as in disease resistance (Presello *et al*, 1996; Presello *et al*, 2006; Iglesias, 2008; Defacio *et al*, 2018) and grain quality traits (López *et al*, 2005; Heck *et al*, 2019). The aim of the present study was to characterize the variability of 36 maize landraces of 'Cristalino Colorado' race collected from Buenos Aires province, Argentina, based on agro-morphological traits and SSR markers, and the relationship among them.

Materials and methods

Plant material

Thirty-six maize landraces of 'Cristalino Colorado' race conserved at BAP were evaluated. These landraces were collected in Buenos Aires province between 1951 and 1963 (Luna and Safont Lis, 1978). Landrace passport descriptors and races are presented in Table 1. Four synthetic open-pollinated (OP) varieties developed by the INTA Pergamino corn breeding programme, Payagua INTA, Candelaria INTA, SP1234 and BS13p, were included as checks. Payagua INTA and Candelaria INTA have semi-dent endosperm and SP1234 belongs to the 'Cristalino Colorado' race. BS13p has dent endosperm and was developed through recurrent selection applied to BS13.

Agro-morphological characterization

Evaluations were performed in two environments, Pergamino (33°53′01″ S, 60°34′01″ W) and Ferré (34°07'30" S, 61°08'27" W), Buenos Aires province, Argentina, during the 2004/2005 growing season. Both are characterized by typical Argiudol soil, (INTA, 1972). The climate is classified as humid temperate, characterized by annual average rainfall of 1,000mm and average temperature of 16–18 ºC. Figure 1 presents meteorological data recorded as long-term averages (1982–2005) and during the 2004/2005 growing season for the two environments. Minimum and maximum temperatures and long-term precipitation data were obtained from the NASA POWER package in R (Sparks, 2018), while precipitation during the 2004/2005 growing season was manually recorded in the field.

Both trials were conducted using a randomized complete block design with two replications. Each plot was planted in two 5m rows with a spacing of 70cm and 30 plant hills. Standard agronomic practices were followed for successful crop growth.

Fifteen quantitative traits, based on maize descriptor (CIMMYT/IBPGR, 1991) were evaluated: days to anthesis (GDU, growing degree units), days to silk (GDU), anthesis-silking interval (GDU), ear length (cm), ear diameter (mm), number of kernel rows (number), kernels per row (number), kernel width (mm), kernel length (mm), plant height (cm), ear height (cm), plant height/ear height ratio (index), 1,000-kernel weight (g), yield (kg/ha), and prolificacy (index). Phenological traits, 1000-kernel weight, yield and prolificacy were measured on the complete plot. Morphological traits were collected on ten plants per plot, randomly selected in each plot, using the average of the 10 units for performing the analyses. A principal component analysis (PCA) was performed using the standardized data matrix obtained from the arithmetic means of the agro-morphological quantitative variables corresponding to both environments and replications, in order to obtain an average characterization throughout the environments (Zuliani *et al*, 2018). Pearson correlation was computed to assess the relationship among traits prior to PCA.

Molecular characterization

Landraces were evaluated using a set of five public SSR markers (phi080, phi072, phi034, bnlg439 and phi96100) located on different chromosomes, with a high degree of polymorphism. Oligonucleotide sequences are publicly available at the Maize Data Bank (www.agron.missouri.edu/Coop/SSR-Probes/SSR1.html). DNA was extracted from young leaves of 25 plants per landrace, according to Kleinhofs *et al* (1993). PCR reactions were carried out in a MJ Research PTC-100 thermocycler (USA). The amplification products were visualized in 6% polyacrylamide gels and were

Table 1. Landrace passport descriptors and races.

Identifier	Location	Department	Altitude (masl)	Latitude	Longitude	Race
ARZM01001	Acevedo	Pergamino	70	33°46' S	60°27' W	Cristalino Colorado
ARZM01002	Rancagua	Pergamino	69	34°02' S	60°30' W	C. Colorado – C. Amarillo
ARZM01003	Rancagua	Pergamino	69	34°02' S	60°30' W	Cristalino Colorado
ARZM01005	Arroyo Dulce	Salto	75	34°06' S	60°24' W	Cristalino Colorado
ARZM01006	Tacuarí	Salto	69	34°13' S	60°19' W	C. Colorado – C. Amarillo
ARZM01007	Salto	Salto	51	34°18' S	60°15' W	Cristalino Colorado
ARZM01008	Salto	Salto	51	34°18' S	60°15' W	Cristalino Colorado
ARZM01012	Arenales	General Arenales	84	34°19' S	61°18' W	C. Colorado – C. Amarillo
ARZM01013	Rojas	Rojas	69	34°12' S	60°44' W	C. Colorado – C. Amarillo
ARZM01014	Chacabuco	Chacabuco	69	34°38' S	60°29' W	Cristalino Colorado
ARZM01015	Salto	Salto	51	34°18' S	60°15' W	Cristalino Colorado
ARZM01016	Arroyo Burgos	Bartolomé Mitre		34°04' S	60°07' W	Cristalino Colorado
ARZM01017	San Pedro	San Pedro	27	33°42' S	59°41' W	C. Colorado – C. Amarillo
ARZM01022	Ortiz Basualdo	Pergamino	64	34°03' S	60°39' W	C. Colorado
ARZM01025	Hunter	Rojas	50	34°15' S	60°32' W	C. Colorado – C. Amarillo
ARZM01026	Ferré	General Arenales	88	34°08' S	61°08' W	Cristalino Colorado
ARZM01027	Carabelas	Rojas	83	34°03' S	60°52' W	Cristalino Colorado
ARZM01028	Colón	Colón	90	33°59' S	61°06' W	Cristalino Colorado
ARZM01030	Conesa	San Nicolás	58	33°36' S	60°22' W	Cristalino Colorado
ARZM01033	El Paraíso	Ramallo	33	33°34' S	59°59' W	C. Colorado – C. Amarillo
ARZM01036	Conesa	San Nicolás	58	33°36' S	60°22' W	Cristalino Colorado
ARZM01039	Rancagua	Pergamino	69	34°02' S	60°30' W	Cristalino Colorado
ARZM01044	Chivilcoy	Chivilcoy	55	34°54' S	60°01' W	Cristalino Colorado
ARZM01048	La Violeta	Pergamino	55	33°44' S	60°11' W	Cristalino Colorado
ARZM01058	Chivilcoy	Chivilcoy	55	34°54' S	60°01' W	C. Colorado – C. Amarillo
ARZM01062	Chacabuco	Chacabuco	69	34°38' S	60°29' W	C. Colorado – Amarillo Ocho Hileras
ARZM01066	Rojas	Rojas	69	34°12' S	60°44' W	Cristalino Colorado
ARZM01082	Nueva Roma	Torquinst	285	38°06' S	62°14' W	Cristalino Colorado
ARZM01086	Nueva Roma	Torquinst	285	38°06' S	62°14' W	Cristalino Colorado
ARZM01087	Nueva Roma	Torquinst	285	38°06' S	62°14' W	Cristalino Colorado
ARZM01092	Pigüé	Saavedra	298	37°41' S	62°24' W	Cristalino Colorado
ARZM01096	Coronel Suárez	Coronel Suárez	234	37°28' S	61°56' W	Cristalino Colorado
ARZM01102	Carhue	Adolfo Alsina	112	37°11' S	62°45' W	Cristalino Colorado
ARZM01124	Trenque Lauquen	Trenque Lauquen	96	35°58' S	62°44' W	Cristalino Colorado
ARZM01151	Mones Cazón	Pehuajó	88	35°48' S	61°53' W	Cristalino Colorado
ARZM01152	Carlos Tejedor	Carlos Tejedor	96	35°23' S	62°25' W	Cristalino Colorado

detected by silver nitrate staining (Benbouza *et al*, 2006). The relative allelic frequencies were calculated using the direct counting method from the individual genotypes found in each landrace.

The Prevosti distance (Prevosti, 1974) was calculated from the relative allelic frequencies of the molecular markers for each landrace to infer the relationship between landraces using a principal coordinate analysis (PCoA). The allele number was determined for each group.
Subsequently, expected heterozygosity (He) was calculated for each group identified in the PCoA, which was calculated using the following equation:

$$He = 1 - \sum_{i=1}^{k} p_i^2$$

where, p_i is the frequency of the $i^{\text{-th}}$ allele, and k is the number of alleles.

Joint analysis

For joint analysis, generalized procrustes analysis (GPA) (Gower, 1975) was performed. This method gets a better adjustment for the information provided by both PCA and PCoA. A consensus configuration was performed and represented in a two-dimensional space. A minimum spanning tree (MST) from the Euclidean matrix obtained from the first

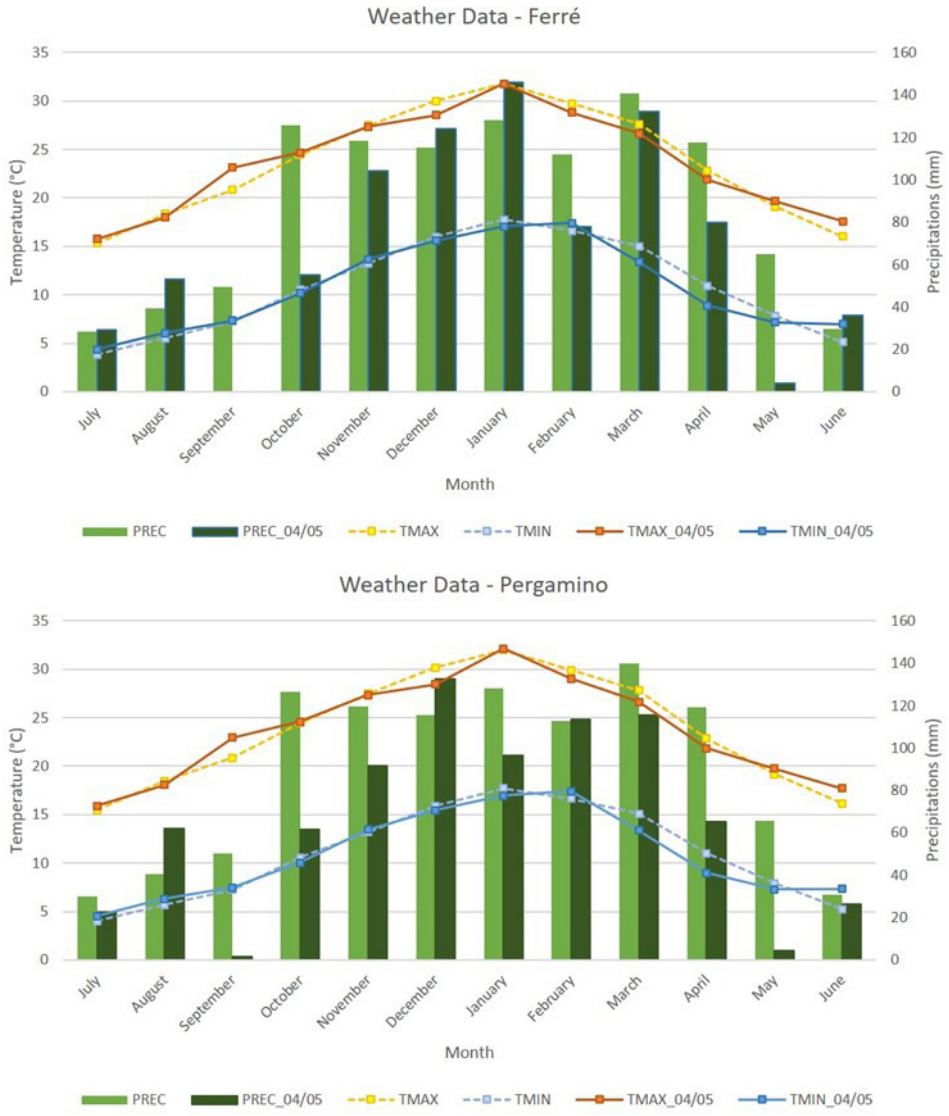


Figure 1. Meteorological data recorded as long-term averages and during the 2004/2005 growing season in Ferré (a) and Pergamino (b), Buenos Aires province, Argentina. PREC, long-term precipitations; PREC_04/05, precipitations in 2004/2005 growing season; TMAX, long-term maximum temperatures; TMAX_04/05, maximum temperatures in 2004/2005 growing season; TMIN, long-term minimum temperatures; TMIN_04/05, minimum temperatures in 2004/2005 growing season.

two GPA coordinates was added in the principal plane.

A Mantel test was performed to quantify the relationship between the molecular, agro-morphological and consensus distance matrix.

To determine concordance of molecular and agro-morphological characterizations at the landrace level, we calculated the Euclidean distance between two analogous points, i.e. points corresponding to a single landrace in the agro-morphological and molecular configurations.

All statistical analyses were carried out using the NTSYS programme (Numerical Taxonomic System ver. 2.11) (Rohlf, 2002).

Results and discussion

Agro-morphological characterization

Table 2 presents the mean values, standard errors, and ranges for each trait, calculated across landraces in each environment under evaluation.

In the Ferré 2004/2005 growing season, plants exhibited shorter height, a shorter anthesis-silking interval, and higher grain yields. These differences may be attributed to the higher precipitation levels recorded in Ferré during January, which match with the critical period for determining grain yield (30 days centred around flowering (Fischer and Palmer, 1984)). Maize is particularly sensitive during this stage, and any stress can increase the anthesis-silking interval, potentially leading to pollination failure and grain yield loss (Tao *et al*, 2023). The other traits did not show differences between the two environments.

The correlation matrix (Table 3) revealed strong and significant ($p < 0.01$) correlations between several traits: days to silking and days to anthesis ($r = 0.88$), plant height and ear height ($r = 0.87$), yield and ear diameter ($r = 0.79$), ear height and plant height/ear height ($r = -0.78$), yield and ear length ($r = 0.74$), and kernel length and ear diameter ($r = 0.73$). Similar correlations among these traits have been reported by other authors during the evaluation of maize landraces (Defacio, 2009; Javed *et al*, 2021; de Faria *et al*, 2022).

Table 2. Mean values, standard errors (S.E.), and ranges for each trait evaluated in two environments.

Environment	**Pergamino 2004/2005**			**Ferré 2004/2005**		
Trait	**Mean**	**S.E.**	**Range**	**Mean**	**S.E.**	**Range**
Ear length (cm)	15.99	0.14	13.10–19.50	16.56	0.11	14.00–18.70
Ear diameter (mm)	41.46	0.31	33.40–48.70	42.04	0.30	34.50–48.10
Number of kernel rows	13.16	0.17	10 .00–16.60	13.09	0.15	9.80–16.40
Kernel width (mm)	8.48	0.09	6.00–9.80	8.77	0.07	7.20–10.20
Kernel length (mm)	8.19	0.12	5.40–12.00	10.11	0.10	8.00–12.00
Kernels per row	31.25	0.40	21.60–38.40	34.90	0.28	29.90–40.00
Prolificacy (index)	0.94	0.02	0–1.36	1.01	0.01	0.66–1.33
1,000-kernel weight (g)	281.30	3.78	174.00–355.00	285.67	3.31	201.00–346.00
Yield (Kg/ ha)	5576.90	176.80	1341.10–9440.80	6838.30	172.90	2623.80–11561.30
Plant height (cm)	163.98	1.56	122.50–193.00	143.42	1.18	118.50–165.50
Ear height (cm)	96.88	1.22	68.00–117.50	84.20	1.10	65.00–106.50
Plant height/ear height ratio (index)	1.70	0.01	1.52–2.03	1.71	0.01	1.43–2.09
Days to anthesis (GDU)	1032.30	3.86	944.70–1100.15	903.23	3.49	837.60–1017.55
Days to silking (GDU)	1069.75	3.09	991.10–1143.75	974.92	5.10	895.90–1089.80
Anthesis-silking interval (GDU)	37.45	1.79	12.80–108.12	71.69	2.91	15.00–155.30

Table 3. Correlation matrix between evaluated traits. EL, ear length; ED, ear diameter; NKR, number of kernel rows; KW, kernel width; KL, kernel length; KR, kernels per row; HKW, 1,000-kernel weight; PH, plant height; EH, ear height; PH/EH, plant height/ear height ratio; PROL, prolificacy; DA, days to anthesis; DS, days to silking; ASI, anthesis–silking interval; NS, non-significant ($p > 0.05$); *, significant at $p < 0.05$; **, significant at $p < 0.01$.

	EL	**ED**	**NKR**	**KW**	**KL**	**KR**	**HKW**	**Yield**	**PH**	**EH**	**PH/ EH**	**PROL**	**DA**	**DS**	**ASI**
EL	1	**	NS	NS	**	**	**	**	**	NS	*	NS	NS	NS	NS
ED	0.56	1	**	NS	**	NS	**	**	NS	NS	NS	NS	NS	NS	NS
NKR	0.12	0.49	1	**	*	NS	NS	*	*	NS	NS	NS	NS	NS	NS
KW	0.16	-0.08	-0.46	1	NS	NS	*	NS	**	*	NS	*	NS	NS	NS
KL	0.51	0.73	0.36	0.10	1	NS	**	**	NS	NS	NS	NS	NS	NS	NS
KR	0.61	0.10	-0.17	0.18	0.26	1	NS	NS	*	NS	NS	NS	NS	NS	NS
HKW	0.62	0.64	0.09	0.38	0.66	0.11	1	**	*	NS	NS	*	NS	NS	NS
Yield	0.74	0.79	0.34	0.05	0.70	0.27	0.75	1	NS	NS	NS	**	NS	NS	NS
PH	0.46	0.15	-0.32	0.50	0.20	0.40	0.43	0.28	1	**	*	NS	NS	NS	NS
EH	0.23	0.02	-0.28	0.41	0.11	0.27	0.26	0.11	0.87	1	**	NS	*	*	NS
PH/EH	0.14	0.11	0.10	-0.13	0.03	-0.02	0.04	0.11	-0.37	-0.78	1	NS	**	**	NS
PROL	0.27	0.24	-0.16	0.33	0.27	0.10	0.35	0.44	0.22	0.07	0.14	1	NS	NS	NS
DA	-0.04	-0.03	-0.07	0.06	0.18	0.09	-0.11	-0.11	0.25	0.43	-0.50	-0.01	1	**	NS
DS	-0.02	0.04	-0.09	0.11	0.10	-0.02	-0.04	-0.09	0.29	0.43	-0.45	0.10	0.88	1	**
ASI	0.04	0.15	-0.06	0.12	-0.13	-0.21	0.13	0.03	0.14	0.10	-0.01	0.22	-0.03	0.45	1

Table 4. Axis loadings corresponding to PC1 and PC2

Trait	PC1	PC2
Ear lenght	0.8324	0.0588
Ear diameter	0.7851	0.3346
Number of kernel rows	0.2442	0.6234
Kernel width	0.2345	-0.5806
Kernel lenght	0.7884	0.1752
Kernels per row	0.4068	-0.2193
Prolificacy	0.3596	-0.1912
1,000-kernel weight	0.8402	0.0398
Yield	0.8861	0.2641
Plant height	0.5499	-0.6681
Ear height (cm)	0.3948	-0.7835
Plant height/ear height ratio	-0.0741	0.6170
Days to antesis	0.1024	-0.5616
Days to silking	0.1209	-0.6186
Anthesis-silking interval	0.0623	-0.2467

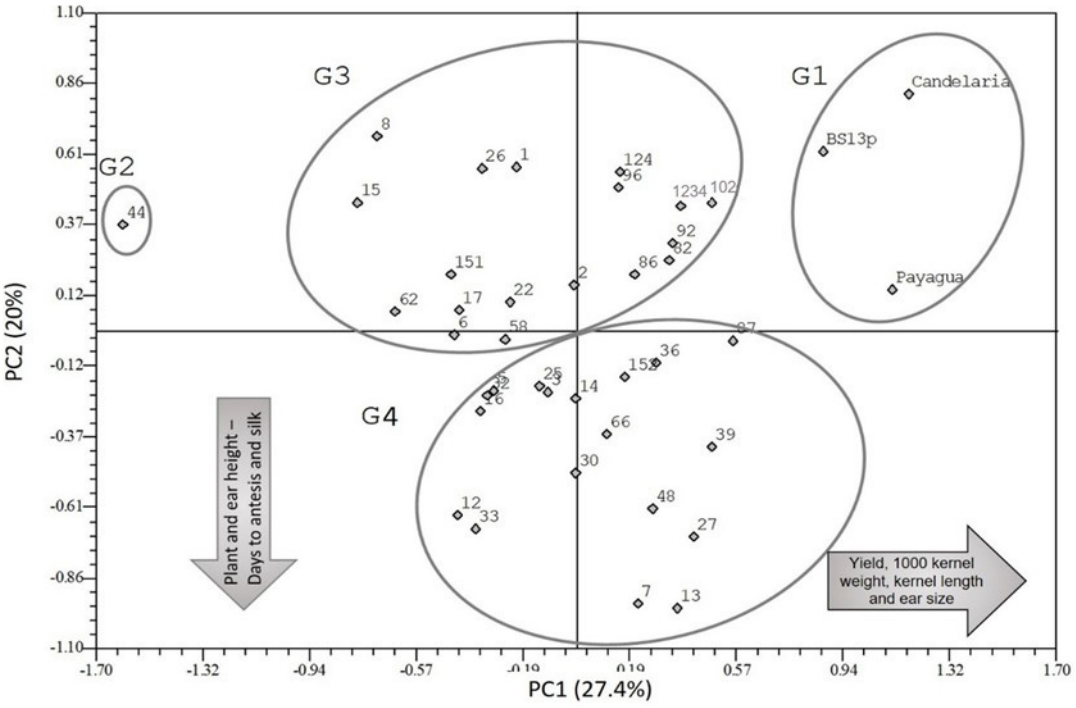


Figure 2. Principal Component Analysis of agro-morphological traits. Landraces are identified with the last numbers in their identifier (e.g. 13 corresponds to ARZM1013).

Results from the PCA (Figure 2) show that the first and second principal components (PC1 and PC2, respectively) accounted for 27.42% and 20.86% of the total variation, respectively. The axis loadings corresponding to PC1 and PC2 are shown in Table 4. PC1 was positively and strongly associated with yield, 1000-kernel weight, kernel length, ear diameter and ear length. PC2 was negatively and moderately associated with plant architecture traits (plant height and ear height), days to anthesis, and days to silking.

Landraces were classified in four groups by PCA, based on the distances observed between individuals in the direction of both established gradients.

G1. This group included BS13p, Candelaria INTA and Payagua INTA. These OP varieties were associated with the highest values of yield and its components, as well as shorter plant and ear height and fewer days to anthesis and silking than the rest of the evaluated landraces. This result agrees with the fact that these genotypes were selected for yield purposes.

G2. Represented by only one accession (ARZM01044) that showed the lowest yield and the smallest ear, grain size and 1,000-kernel weight of all landraces. This accession presented medium to low plant height and intermediate to fewer days to anthesis and silking.

G3. This group included landraces with average yield, 1,000-kernel weight, kernel length, ear diameter, and ear length, displaying shorter plants, lower ear height and fewer days to anthesis and silking than the G4 cluster.

G4. Represented by landraces with average yield, 1,000-kernel weight, kernel length, ear diameter and ear length, and highest values for days to anthesis and silking, high plant and ear height.

Molecular characterization

In this study, a set of five SSR markers was employed for the preliminary molecular characterization of maize landraces. Other authors (Di Pasquale *et al*, 2024; Joshi *et al*, 2020; Ignjatović Micić *et al*. 2013) have also used a low number of SSRs, ranging from 5 to 10, to evaluate maize landraces.

A total of 48 alleles were detected. The overall number of alleles per locus varied from 6 (phi034 and phi072) to 21 (bnlg439), with an average of 9.6 (Table 5). Six alleles were unique to landraces (unique or private alleles) while five other different alleles were present in two landraces (rare alleles).

The average number of alleles per locus obtained from landraces (9.6) was higher than the values reported by Reif *et al* (2003) (5.9), Warburton *et al* (2002) (6.3), Labate *et al* (2003) (6.5) and Di Pasquale *et al* (2024) (7.4), but lower than those reported by Barcaccia *et al* (2003) (20.75), Rivas *et al* (2022) (19.05) and Torres-Morales *et al* (2023) (25.39).

In the OP varieties, the assayed loci scored a mean number of alleles equal to 7.6, lower than landraces (9.6). This result is consistent with that obtained by Barcaccia *et al* (2003) of 10.25 vs. 19.75, showing that even though OP varieties have genetic variability, they originated from a narrow genetic base.

First and second principal axes of PCoA (Figure 3) accounted for 10.79 and 8.96% of the total variation, respectively. Landraces were distributed in four groups based on their relative distances on the principal plane, which differed from those obtained using PCA.

Table 5. Numbers of alleles per locus across landraces

SSR markers	No. of alleles
phi080	8
phi072	6
phi034	6
phi96100	7
bnlg439	21
Average	9.6

G1. Represented by ARZM01082 and ARZM01086 landraces.

G2. This group included ARZM01124, ARZM01102, ARZM01152, BS13p and Candelaria INTA.

G3 and G4. These groups included most of the evaluated landraces and were separated by the dispersion of the second principal coordinate (PCO2).

This clustering was not associated with the presence of rare or private alleles. Some landraces exhibited rare or private alleles (ARZM01003, ARZM01049, ARZM01102 and SP1234) but were clustered with other landraces.

Allele numbers for five SSR markers and expected heterozygosity were calculated for each group identified by the PCoA (Figure 4).

The number of alleles observed varied among groups, increasing from G1 to G4 in parallel with the number of landraces included in each group. This pattern suggests greater genetic variability in G4, in concordance with the broader dispersion of landraces observed for this group in the PCoA (Figure 3). The expected heterozygosity (He) also varied among groups, but no clear relationship was observed with the total number of alleles.

Joint analysis

The correlation between agro-morphological and molecular data matrices was very low and not significant (r = 0.07, p-value = 0.77). The different configurations obtained with both types of variables indicate that individual characterization offers additional information that can be used complementarily to know the genetic variability among landraces. Low correlation values between agronomic and molecular traits were found in 41 varieties of cucumber (*Cucumis sativus L.*) (Bramardi *et al*, 2005), in 37 Patagonian isolates of yeast (*Saccharomyces cerevisiae*) (Lopes *et al*, 2006), in 57 red clover landraces (*Trifolium pratense* L) (Dias *et al*, 2008), and in a set of banana (*Musa sp*.) clones (Ermini *et al*, 2016). For this reason, it is necessary to use a technique that gathers molecular and agro-morphological information.

According to the GPA results (Figure 5), seven groups of landraces were identified.

Some landraces denoted a high correspondence between molecular and agro-morphological characterizations. However, most landraces denoted a great discordance between agro-morphological and molecular data. Distance between both types of data presents a range between 0.02

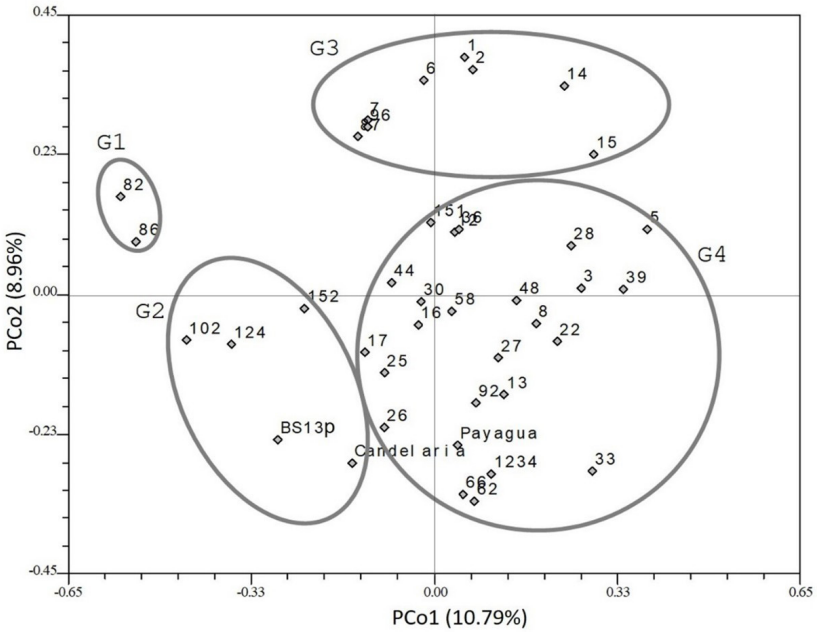


Figure 3. Principal coordinate analysis of molecular traits. Landraces are identified with the last numbers in their identifier (e.g. 13 corresponds to ARZM1013).

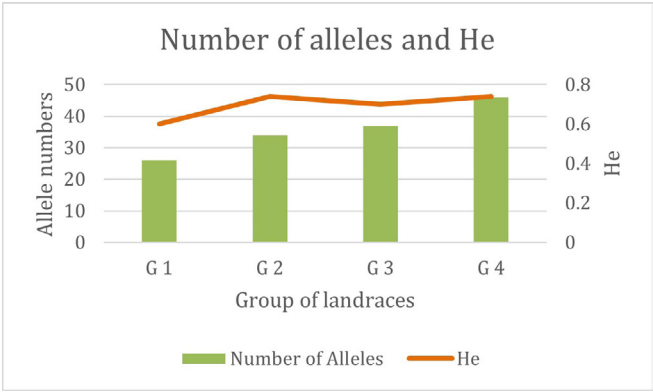


Figure 4. Allele numbers for five SSR markers and expected heterozygosity (He) for each group identified by the PCoA

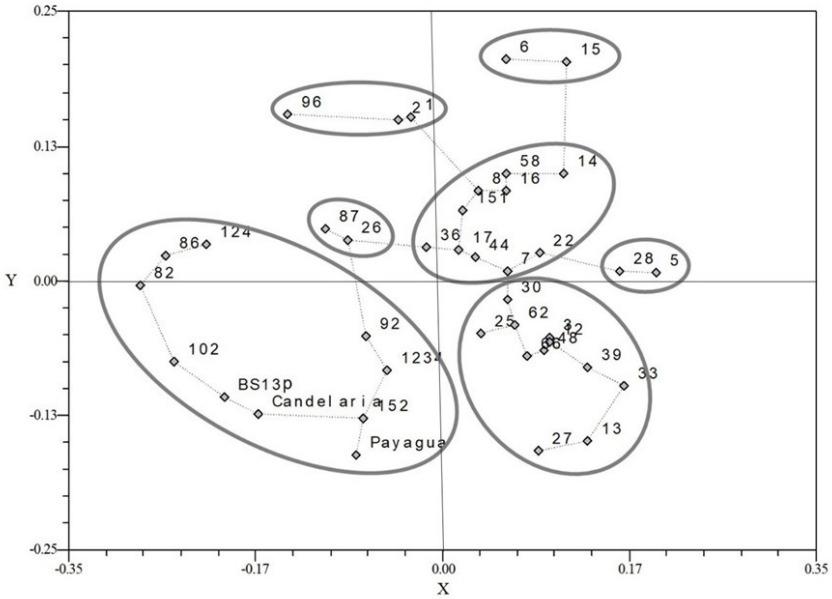


Figure 5. Generalized procrustes analysis of agro-morphological and molecular traits in the first two axis with Minimum Spanning Trees. Landraces are identified with the last numbers in their identifier (e.g. 13 corresponds to ARZM1013).

(ARZM01017) and 0.45 (ARZM01082), with an average of 0.24 (Table 6). The discrepancies observed between agro-morphological and molecular markers may be explained by the fact that the SSR used in this work are neutral and not linked to the agro-morphological traits under evaluation. Expression of these phenotypic traits is strongly influenced by environmental conditions and the selection performed by farmers according to local preferences. This selection contributed to phenotypic differentiation, which may not be reflected in neutral genomic regions, such as those assessed by SSR markers (Javed *et al*, 2021).

Table 6. Distance between two analogous points, i.e. points corresponding to a single landrace in the agro-morphological and molecular configurations

Identifier	Distance between molecular and agro-morphological traits
ARZM01001	0.26
ARZM01002	0.27
ARZM01003	0.21
ARZM01005	0.27
ARZM01006	0.23
ARZM01007	0.37
ARZM01008	0.24
ARZM01012	0.32
ARZM01013	0.42
ARZM01014	0.10
ARZM01015	0.19
ARZM01016	0.06
ARZM01017	0.02
ARZM01022	0.17
ARZM01025	0.14
ARZM01026	0.30
ARZM01027	0.37
ARZM01028	0.23
ARZM01030	0.15
ARZM01033	0.40
ARZM01036	0.06
ARZM01039	0.31
ARZM01044	0.08
ARZM01048	0.24
ARZM01058	0.11
ARZM01062	0.24
ARZM01066	0.24
ARZM01082	0.45
ARZM01086	0.42
ARZM01087	0.25
ARZM01092	0.21
ARZM01096	0.43
ARZM01102	0.26
ARZM01124	0.37
ARZM01151	0.08
ARZM01152	0.23
BS13P	0.16
Candelaria INTA	0.16
Payagua INTA	0.17
SP 1234	0.26

Consensus configuration grouped the four OP varieties and the six landraces. Interestingly, ARZM01044 formed a group by itself when evaluated by agro-morphological traits and was included in a group with a larger number of landraces when evaluated by molecular markers and in consensus analysis. ARZM01082 and ARZM01086 landraces were assigned to a separate group, according to molecular markers. However, according to agro-morphological traits, these landraces were grouped with other landraces. In GPA, these landraces were grouped with OP varieties. A similar situation was observed with checks Candelaria INTA, Payagua INTA and BS13p, which formed a distinct group according to agro-morphological analysis but grouped together with other landraces according to the molecular markers and GPA. The correlations among the three distance matrices (molecular, agro-morphological and consensus) were estimated using a Mantel test. A greater correlation was found between the consensus and the molecular and agro-morphological characterization (0.20 and 0.47, respectively) than between molecular and agro-morphological characterization (0.07). This result indicates that GPA allows the simultaneous characterization of a set of accessions with agro-morphological traits and SSR markers. There is no unique pattern of association among landraces, which emphasizes the importance of studying the different descriptors jointly to obtain the best description and interpretation of genetic diversity.

In conclusion, both agro-morphological and molecular variation were detected among the studied landraces, highlighting the importance of integrating both types of characterization to evaluate genetic diversity. GPA is a powerful statistical technique to align genetic and agro-morphological descriptors. Increasing the knowledge of the available genetic diversity in maize germplasm will facilitate the establishment of core collections. Furthermore, integrating agronomic performance with genetic data is critical to developing and optimizing future breeding strategies. Currently, maize breeding relies on a narrow genetic base. Incorporating landraces into crosses with elite varieties offers a promising approach to introduce novel alleles and broaden the genetic base of maize breeding. Moreover, the local adaptation exhibited by landraces represents a valuable source of germplasm for future needs in sustainable agriculture, particularly in the context of climate change.

To enhance the understanding of the genetic diversity of the conserved landraces, it is recommended to incorporate more molecular markers as well as landraces from other races and origins.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Raquel Defacio, Natalia Paz and Sergio Bramardi. The first draft of the manuscript was written by Raquel Defacio and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We are grateful to the staff of Germplasm Bank and Molecular Markers Laboratory at INTA Pergamino for their assistance in field trials and laboratory experiments, respectively. We also thank Dr. Juliana Iglesias for her precious collaboration and valuable suggestions.

This work was financed by the Instituto Nacional de Tecnologia Agropecuaria (INTA) and the Agencia PICTR2002-00109 'Conservación, valoración y desarrollo de recursos genéticos vegetales mediante el uso de nuevas tecnologías.'

Data availability statement

Accession-level data are available from the corresponding author upon reasonable request.

Conflict of interest statement

The authors have no competing interests to declare that are relevant to the content of this article.

References

Balconi, C., Galaretto, A., Malvar, R. A., Nicolas, S. D., Redaelli, R., Andjelkovic, V., Revilla, P., Bauland, C., Gouesnard, B., Butron, A., *et al (*2024) Genetic and Phenotypic Evaluation of European Maize Landraces as a Tool for Conservation and Valorization of Agrobiodiversity. Biology 13, 454. https://doi.org/ 10.3390/biology13060454

Barcaccia, G., Lucchin, M., Parrini, P. (2003) Characterization of a flint maize (*Zea mays* var. *indurata*) Italian landraces, II. Genetic diversity and relatedness assessed by SSR and Inter-SSR molecular markers. Genetic Resources and Crop Evolution 50:253-271. https://doi.org/10.1023/A:1023539901316

Benbouza, H., Jacquemin, J. M., Baudoin, J. P., Mergeai, G. (2006) Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. Biotechnologie, Agronomie, Société et Environnement, 10 (2), 77-81.: https://www.researchgate.net/publication/26433489

Bramardi, S. J. (2023) Three-Way Multivariate Analysis for the Characterization of Plant Genetic Resources. Modern Concepts & Developments in Agronomy (MCDA) 12(5): 1223-1227. https://doi.org/10.31031/MCDA.2023.12.000797

Bramardi, S.J., Bernet, G. P., Asíns, M. J., Carbonell, E. A. (2005) Simultaneous Agronomic and Molecular Characterization of Genotypes via the Generalised Procrustes Analysis: An Application to Cucumber. Crop Sci. 45:1603-1609. https://doi.org/10.2135/cropsci2004.0633

Cámara Hernández, J., Miante Alzogaray, A. M. (2003) Caracterización y clasificación, en razas, de maíces nativos de la Provincia de Misiones, Argentina. – IV Simposio de Recursos Genéticos para América Latina y el Caribe. Mar del Plata. Argentina.

Chandler, K., Lipka, A. E., Owens, B. F., Li, H., Buckler, E. S., Rocheford, T., Gore, M. A. (2013) Genetic Analysis of Visually Scored Orange Kernel Color in Maize. Crop Science 53(1):189-200. https://doi.org/10.2135/cropsci2012.02.0129

CIMMYT/IBPGR. 1991. Descriptores de maíz. México-Roma. 88pp.

de Faria, S. V., Zuffo, L. T., Rezende, W. M., Caixeta D. G., Pereira, H. D., Azevedo, C. F., DeLima R. O. (2022) Phenotypic and molecular characterization of a set of tropical maize inbred lines from a public breeding program in Brazil. BMC Genomics 23, 54. https://doi.org/10.1186/s12864-021-08127-7

Defacio, R. A., Hourquescos, M. J., Bramardi, S. J., Ferrer, M. E. (2005) Estudio de variabilidad en poblaciones nativas de maíz. Actas VIII Congreso Nacional de Maíz. Rosario Pp: 383-386.

Defacio, R. A. (2009) Caracterización y evaluación de la variabilidad genética en poblaciones nativas de maíz [*Zea mays L.*] de la Provincia de Buenos Aires en base a descriptores morfológicos y agronómicos. Tesis Maestría en Genética Vegetal; Área Mejoramiento Genético. UNR/INTA. Magister Scientiae. Marzo, 2009. 93p. https://rephip.unr.edu.ar/handle/2133/13927

Defacio, R. (2017) Evaluación comparativa de distintas estrategias de análisis de datos para la caracterización y ordenamiento de la variabilidad genética de poblaciones locales de maíz (Zea mays L.). Tesis de Doctorado en Ciencias Agrarias. UNR. Agosto 2017. 122p. https://rephip.unr.edu.ar/handle/2133/13925

Defacio, R. A., Iglesias, J., Kistner, M. B, Canteros, F. H., Parrado, J., Ferrer, M. E. (2018) Las poblaciones locales de maíz como fuente para la resistencia a enfermedades. Revista de Tecnología Agropecuaria: 10, 38: 18-21. http://hdl.handle.net/20.500.12123/4381

Di Pasquale, G. M., Stagnati, L., Lezzi, A., Lanubile, A., Marocco, A., Rossi, G., Busconi, M. (2024) Morphological and Genetic Characterization of Maize Landraces Adapted to Marginal Hills in North-West Italy. Plants 13, 1030. https://doi.org/10.3390/plants13071030

Esquinas Alcázar, J. (2005) Protecting crop genetic diversity for food security: political, ethical and technical challenges. Nat Rev Genet 6: 946-953. https://doi.org/10.1038/nrg1729

Ermini, J. L., Tenaglia, G., Pratta, G. R. (2016) Genetic diversity, ancestry relationships and consensus among phenotype and genotype in banana (Musa acuminata) clones from Formosa (Argentina) farmers. Plant Cell Biotechnology and Molecular Biology 17: 267-278. https://www.ikprress.org/index.php/PCBMB/article/view/1486.

Fischer, K. S., Palmer, F. E. (1984) Tropical Maize, In: P. R. Goldsworthy, P. R., Fischer,N. M. Eds., The Physiology of Tropical Field Crops, Wiley, Chichestor, pp. 213- 248.

Gorjanc, G., Jenko, J., Hearne, S. J., Hickey, J. M. (2016) Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genomics. 17-30. doi: https://doi.org/10.1186/s12864-015-2345-z.

Gower, J. C. (1975) Generalizad Procrustes Analisys. Physchometrika. 40: 33-51. https://doi.org/10.1007/BF02291478

Heck, M., Defacio, R., Ferrer, M., Cirilo, A., Fariza, S., De Lucia, A., Blaszchik, J. (2019) Evaluación de la calidad nutricional de variedades nativas de maíz de Misiones, Argentina". Revista de Investigaciones de la facultad de Ciencias Agrarias - UNR, 34. https://doi.org/10.35305/agro34.266

Heck, M., Defacio, R., Ferrer, M., Cirilo, A., Fariza, S., De Lucia, A.., Blaszchik, J. (2020). Evaluación de la variabilidad agromorfológica de poblaciones nativas de maíz de Misiones, Argentina. REevista de ciencia y tecnología 33: 6 - 12. https://www.fceqyn.unam.edu.ar/recyt/index.php/recyt/article/view/264

Iglesias, J. (2008) Potencial de germoplasma nativo de maíz como donante de genes de resistencia a Fusarium asociado a bajo contenido de micotoxinas. Tesis Maestría: Genética Vegetal. Facultad de Cs. Agrarias. Universidad de Rosario (FCA, UNR), Santa Fe, Argentina / EEA INTA Pergamino, Buenos Aires, Argentina.

Ignjatović Micić, D., Ristić, D. Babić, V., Andjelković, V., Marković, K., Vančetović, J. (2013) Genetic assessment of maize landraces from former Yugoslavia. Genetika 45, Issue 2 (405-417). https://doi.org/10.2298/GENSR1302405I

Instituto Nacional de Tecnología Agropecuaria (INTA) (1972). Carta de suelos de la República Argentina.

Javed, R. M., Iqbal, S., Ullah, M.R., Khan, A., Iqbal, M.S., Ullah, M.U., Rehman, F. U., Khan, Saqib, M.S., Ali, S. (2021). Phenotypic and molecular divergence in maize (Zea mays L.) ecotypes. Pak. J. Agri. Sci. 58:1777-1787. https://doi.org/10.21162/PAKJAS/21.1469

Joshi, B., Rawat, J., Adhikari, B., & Pokhrel, R. (2020). SSR Markers Based Genetic Diversity in Nepalese Maize Landraces. SAARC Journal of Agriculture, 18(1), 23–37. https://doi.org/10.3329/sja.v18i1.48379

Kleinhofs, A., Kilian, A., Saghai Maroof, M. A., Biyashev, R. M., Hayes, P., Chen, F. Q., Lapitan, N., Fenwich, A., Blake, T. K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J. D., Hoffman, D., Skadsen, R., Steffenson, B. J. (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theoretical and Applied Genetics 86:705-712. https://doi.org/10.1007/BF00222660

Labate, J. A., Lamkey, K. R., Mitchell, S. E., Kresovich, S., Sullivan, H., Smith, J. S. C. (2003) Molecular and historical aspects of Corn Belt Dent diversity. Crop Science 43:80.91. https://doi.org/10.2135/cropsci2003.8000

Lopes, C. A., Rodríguez, M. E., Querol, A., Bramardi, S. J., Caballero, A. C. (2006) Relationship between molecular and enological features of Patagonian wine yeasts: relevans in selection protocols. World Journal of Microbiology and Biotechnology 22:827-833. https://doi.org/10.1007/s11274-005-9110-4

López, C. G., Eyhérabide, G. H., Lorea, R. D., Delucchi, C., Percibaldi, N. M., Castelarin, J., Pedrol, H., Borrás, F. (2005) Selección de poblaciones locales de maíz como fuente de alelos favorables para el mejoramiento en un híbrido flint x dentado. Actas VIII Congreso Nacional de Maíz. Rosario. Pp: 358-360.

Lucchin, M., Barcaccia, G., Parrini, P. (2003) Characterization of a flint maize (*Zea mays L. convar. Mays*) Italian landrace: I. Morpho-phenological and agronomic traits. Genetic Resource and Crop Evolution 50:315-327. https://doi.org/10.1023/A:1023578207258

Luna, J. T., Safont Lis, J. (1978) El maíz en la Argentina: Vulnerabilidad y Recursos Genéticos.Ciencia e investigación. Tomo 3 – 4 – 5 y 6: 83 – 90.

Maize Data Bank www.agron.missouri.edu/Coop/SSR-Probes/SSR1.html.

Dias, P. M. B., Julier, B., Sampoux, J. P., Barre, P., Dall'Angol, M. (2008) Genetic diversity in red clover (*Trifolium pretense* L.) revealed by morphological and microsatellite (SSR) markers. Euphytica 160:189-205. http://dx.doi.org/10.1007/s10681-007-9534-z

Paz, N. M., Schlatter, A. R., Letis, G., Ferrer, M. (2005) Caracterización molecular de poblaciones locales de maíz. Actas VIII Congreso Nacional de Maíz. Rosario. Pp: 361-363.

Paz, N. M. (2009) Caracterización de la variabilidad genética en poblaciones locales de maíz mediante marcadores microsatélites. Tesis Maestría en Genética Vegetal; Área Mejoramiento Genético. UNR/INTA. Magister Scientiae, Mayo, 2009. 93p.

Pilling, D., Bélanger, J., Diulgheroff , S., Koskela, J., Leroy, G., Mair, G. and Hoffmann, I. (2020) Global status of genetic resources for food and agriculture: challenges and research needs : Global status of genetic resources for food and agriculture, Genetic Resources, 1(1), pp. 4–16. https://doi.org/10.46265/genresj.2020.1.4-16

Prevosti, A. (1974) La distancia genética entre poblaciones. Miscellanea Alcobé. Universidad de Barcelona, 109-118.

Presello, D. A., Ferrer, M., Solari, L., Céliz, A. (1996) Resistencia al virus del Mal de Río Cuarto en variedades locales argentinas de maíz. RIA, 27 (1): 19 a 26.

Presello, D.A., Iglesias, J., Botta, G., Reid, L. M., Lori, G. A., Eyhérabide, G. H. (2006) Stability of maize resistance to the ear rots caused by Fusarium graminearum and F. verticilloides in Argentinian and Canadian environments. Euphytica. 147: 403-407. https://doi.org/10.1007/s10681-005-9037-8

Reif, J. C., Melchinger, A. E., Xia, X. C., Warburton, M. L., Hoisington, D. A., Vasal, S. K., Srinivasan, G., Bohn, M., Frisch, M. (2003) Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. Crop Science 43:1275- 1282. https://doi.org/10.2135/cropsci2003.1275

Rivas, J. G., Gutierrez, A. V., Defacio, R. A., Schimpf, J., Vicario, A. L., Hopp, H. E., Paniego, N. B. and Lía, V. V. (2022) Morphological and genetic diversity of maize landraces along an altitudinal gradient in the Southern Andes. PLoS ONE 17(12): e0271424. https://doi.org/10.1371/journal.pone.0271424.

Rohlf, F. J. (2002) NTSYSpc: Numerical Taxonomy System, ver. 2.1. Exeter Publishing, Ltd. Setauket. NY.

Salhuana, W., Pollak, L. M., Ferrer, M. E., Paratori, O., Vivo, G. (1998) Breeding Potential of Maize Accessions from Argentina, Chile, USA, and Uruguay. Crop Sci. 38:866-872. https://doi.org/10.2135/cropsci1998.0011183X003800030040x

Secretaría de Agricultura, Ganadería, Pesca y Alimentación (1997). Resolución 757/9: MAIZ: establécese un Reglamento Técnico de Identidad de Maíz Flint o Plata. 13 de octubre de 1997 https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-757-1997-46664/texto

Solari, L. R. (2007) IV Catálogo de Germoplasma de Maíz. Buenos Aires: INTA. 78p.:il. + CD-ROM. ISBN 978-987-521-293-0.

Sparks, A. 2018. "nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R." The Journal of Open Source Software, 3(30), 1035. https://doi.org/10.21105/joss.01035.

Tao, K., Li, Y., Hu, Y., Li, Y., Zhang, D., Li, C., He, G., Song, Y., Shi, Y., Li, Y., Wang, T., Lu, Y., Liu, X. (2023) Overexpression of ZmEXPA5 reduces anthesis-silking interval and increases grain yield under drought and well-watered conditions in maize. Mol Breeding 43, 84. https://doi.org/10.1007/s11032-023-01432-x

Torres-Morales, B., Rocandio-Rodríguez, M., Santacruz-

Varela, A., Córdova-Téllez, L., Coutiño-Estrada, B., López Sánches, H. (2023) Genetic diversity characterization of maize populations using molecular markers. Italian Journal of Agronomy: 18 (3):2206. https://doi:10.4081/ija.2023.2206

Troyer, A. F., Openshaw, S. J., Knittle, K. H. (1988) Measurement of Genetic Diversity Among Popular Commercial Corn Hybrid Crop Sci. 28: 481-485. https://doi.org/10.2135/cropsci1988.0011183X002800030010x

Vigouroux, Y., Glaubitz, J. C., Matsuoka, Y., Goodman, M. M., Sánchez, J., Doebley, J. (2008) Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. Am J Bot 95:1240–1253. doi: https://doi.org/10.3732/ajb.0800097

Warburton, M. L., Xianchun, X., Corssa, J., Franco, J., Melchinger, A. E., Frisch, M., Bohn, M., Hoisington, D. (2002) Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Science 42:1832-1840. https://doi.org/10.2135/cropsci2002.1832

Wright, S. (1978) Evolution and the genetics populations. Vol. 4, Variability within and among natural populations. Univ. Chicago Press, Chicago.

Xiang, K., Yang, K. C., Pan, G. T., Reid, L. M., Li, W. T., Zhu, X., Zhang, Z. M. (2010) Genetic diversity and classification of maize landraces from China's Sichuan Basin based on agronomic traits, quality traits, combining ability and SSR markers. Maydica 55(1):85-93.

Zuliani, P., Defacio, R., Lavalle, A., Bramardi, S. (2018). Comparación de técnicas de Análisis Multivariado mediante simulación para caracterización de recursos fitogenéticos en función de caracteres susceptibles a interacción genotipo-ambiente. Revista FAVE Sección Ciencias Agrarias, Universidad Nacional del Litoral. 17(1): 75-86. https://doi.org/10.14409/fa.v17i1

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 26–38
DOI: 10.46265/genresj.ZDEW3901
https://www.genresj.org
ISSN: 2708-3764

Phenotypic variability of *Smallanthus sonchifolius* germplasm of Peru

Angel Esteban Santa Cruz-Padilla*[, a], Jorge Luis Vásquez-Orrillo*[, a], Silvia Yanina Rodríguez López[a], Araceli Eugenio Leiva[a], Ricardo Manuel Bardales-Lozano[b], Juan F. Seminario[c] and Hipolito Murga-Orrillo[d]

[a] *Instituto Nacional de Innovación Agraria (INIA), Estación Experimental Agraria Baños del Inca, Subdirección de Recursos Genéticos, Jirón. Wiracocha s/n, 06004, Los Baños del Inca, Perú*
[b] *Universidad Nacional de la Amazonía Peruana (UNAP), Departamento Académico de Suelos y Cultivos, Facultad de Agronomía, Jirón Nauta, 16002, Iquitos, Perú*
[c] *Universidad Nacional de Cajamarca. Facultad de Ciencias Agrarias, Programa de Raíces y Tubérculos Andinos. Av. Atahualpa 1050. C.P. 06003. Cajamarca, Cajamarca, Perú*
[d] *Universidad Nacional Autónoma del Alto Amazonas (UNAAA). Escuela Profesional de Ingeniería Agrónoma, Prol. Libertad 1220 - 1228 - Yurimaguas, Alto Amazonas, Loreto, Perú*

Abstract: *Smallanthus sonchifolius* (yacon) is a functional food native to the South American Andes. Its tuberous root and leaves are the main parts consumed; however, few studies have been carried out on its phenotypic variability. This study aimed to characterize 214 yacon accessions from the Germplasm Bank of the Instituto Nacional de Innovación Agraria (INIA), Peru. Twelve qualitative and seven quantitative variables were used. Accession Y-74 showed the largest leaf dimensions, while Y-28 showed the highest productivity per plant. Multiple correspondence analysis and principal component analysis revealed that the variables propagule color, leaf shape, root pulp color, leaf length and width, root weight per plant, and yield contributed significantly to the discrimination and identification of promising accessions. The geographical grouping of the accessions showed differences between accessions from the north and south of Peru. The qualitative phylogenetic tree showed 12 morphological groups discriminated mainly by leaf morphology and root characteristics, while the dendrogram analysis identified four clusters, with Cluster II standing out with an average yield of 73.5t/ha of tuberous roots. These results are important, as they allowed the identification of promising accessions and useful traits that can contribute to improving productivity and promoting the expansion of yacon cultivation at national and international levels.

Keywords: Germplasm, phenotypic, functional food, yacon, Andes

Citation: Santa Cruz-Padilla, A. E., Vásquez-Orrillo, J. L., Rodríguez López, S. Y., Eugenio Leiva, A., Bardales-Lozano, R. M., Seminario, J. F. and Murga-Orrillo, H. (2025) "Phenotypic variability of *Smallanthus sonchifolius* germplasm of Peru", *Genetic Resources*, 6(12), pp. 26–38. doi: 10.46265/genresj.ZDEW3901

Introduction

Smallanthus sonchifolius, known as yacon, is a perennial species native to the Andes of South America (Caballero and Colonia, 2018; De Sales *et al*, 2021). It is cultivated from Venezuela to northern Argentina, between 900 and 3,500masl (Huaycho *et al*, 2016). However, its remarkable plasticity has facilitated its adaptation to climates outside the Andes (Seminario *et al*, 2003; Mansilla *et al*, 2006; Wagner *et al*, 2019), in countries such as the Czech Republic, the United States, Brazil (Quaresma *et al*, 2020), New Zealand and Germany (Lachman *et al*, 2007). In the Peruvian Andes, tuberous root yields vary from 8 to 96t/ha depending on the genotype (Santa Cruz and Vásquez-Orrillo, 2023).

Yacon maintains a historical and cultural value, as it has been an important functional food for Andean populations since pre-Columbian times (Huaycho *et al*, 2016; Lopera-Marín, 2020). The roots and leaves have benefits for human health. The roots are usually eaten fresh, and the leaves in infusions (Lebeda *et al*, 2004; Moreira *et al*, 2020). The health benefits of yacon are due to its antioxidant, antimicrobial, hypolipidemic, antidiabetic and even anticancer properties

* Corresponding authors: Angel Esteban Santa Cruz-Padilla (asantacruz@inia.gob.pe), Jorge Luis Vásquez-Orrillo (jorge.vasquez.orrillo@gmail.com)

Received: 14.03.2025 | Accepted: 11.06.2025 | Published online: 18.08.2025

(Baek *et al*, 2018; Myint *et al*, 2019; Adriano *et al*, 2019; Minchola-Castañeda *et al*, 2022). The consumption of extracts from the leaves and tuberous roots of yacon regulates the glycemic state and increases the concentration of insulin in the blood (Santos *et al*, 2017; Ferraz *et al*, 2020); it also contributes to the reduction of body weight (Honoré *et al*, 2018). In this context, the therapeutic benefits of yacon highlight the need to conserve and document the genetic heritage of this species (García *et al*, 2022; Wagner *et al*, 2019).

Yacon germplasm exhibits phenotypic variability among different accessions. Variations in shape, weight and oligofructan content have been revealed in tuberous roots (Valentová *et al*, 2006), and differences in isoenzymes and phenolic content in leaves (Valentová *et al*, 2006; Mansilla *et al*, 2006). Morphological characterization has identified multiple morphotypes and ecotypes of both cultivated and wild yacon (Polanco and García, 2013; Ignacio *et al*, 2017). Genetic diversity analysis using molecular markers has shown distinct groupings among accessions, with the highest diversity observed in central Peru (Mansilla *et al*, 2006). Variations in reproductive biology, including flowering time and pollen viability, have also been reported among accessions (Mansilla *et al*, 2010). Furthermore, studies have found differences in total phenolic content, antioxidant activity and chemical composition among local yacon phenotypes (Lachman *et al*, 2007; Russo *et al*, 2015). This phenotypic variability of yacon makes it a valuable resource for breeding programmes and agroindustrial applications.

In Peru, the Instituto Nacional de Innovación Agraria (INIA) conserves yacon germplasm from 11 regions distributed throughout the Andes, currently counting 214 accessions. This diversity has highlighted the need to update phenotypic characterization studies. In this context, it is hypothesized that this germplasm has significant phenotypic variability, which will allow the identification of accessions with superior agronomic characteristics, suitable for use in future genetic improvement programmes. This study aimed to characterize the phenotypic variability of 214 yacon accessions from the INIA Germplasm Bank, conserved at the Estación Experimental Agraria Baños del Inca, Cajamarca, Peru.

Materials and methods

Plant material

The study was carried out between June 2021 and March 2022. A total of 214 yacon accessions were used, originating from the regions of Piura, Cajamarca, Amazonas, La Libertad, Ancash, Pasco, Junín, Cusco, Ayacucho, Apurímac and Puno (Figure 1 and Supplemental Table 1) and conserved *ex situ* since 1986. These accessions are part of the Andean Roots Germplasm Bank of INIA, Estación Experimental Agraria Baños del Inca, Cajamarca, Peru.

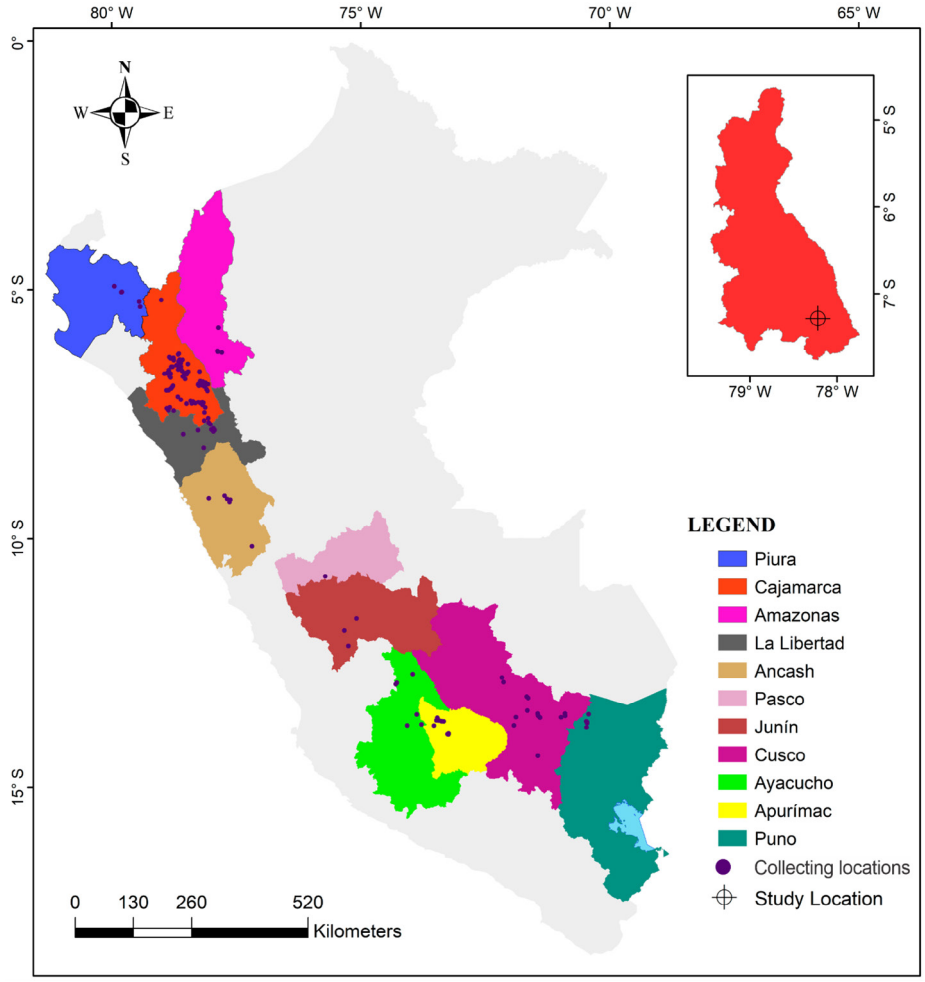


Figure 1. Map of collecting locations of 214 accessions of *Smallanthus sonchifolius* collected in 11 regions of Peru and conserved in the INIA Germplasm Bank.

Experimental area

In the INIA Cochamarca Experimental Annex, located at 7.28137 S, 78.21987 W and 2,835masl, in the province of San Marcos, Cajamarca region (Figure 1). The experiment was installed in an area of 1,796m^2, prepared to a depth of 30cm, where island guano (1,200kg/ha), diammonium phosphate (164kg/ha) and potassium chloride (73kg/ha) were applied before sowing. Each experimental unit had an area of 6.3m^2 (0.9m x 7m). Planting was done in a systematic way (each accession in a row), at distances of 0.90m between rows and 0.50m between plants (22,220 plants per hectare). Each experimental unit consisted of 15 plants, 10 plants being evaluated and 5 plants as borders.

Soil analysis of the experimental plot indicated the following: pH of 5.7, organic matter 2.62%, 33.39ppm of phosphorus and 165ppm of potassium, and sandy loam texture (Laboratorio de suelos, aguas, abonos y foliares de la Estación Experimental Agraria Baños del Inca - INIA). During this period of study execution, a mean temperature of 14.3ºC, a minimum of 7.7ºC and a maximum of 20.8ºC were recorded; and the mean monthly rainfall was 117.9mm (SENAMHI, 2024).

Qualitative and quantitative descriptors used in characterization

Twelve qualitative morphological descriptors and seven quantitative ones were used (Table 1), which were assessed over several agricultural campaigns. Colors were defined using the Royal Horticultural Society colour chart (RHS, 2001).

Table 1. Qualitative and quantitative descriptors used in the characterization of yacon accessions

Descriptor type	Descriptor	Acronym	Period of assessment
Qualitative	Secondary stem color and its distribution	SSCD	Preflowering
	Stem branching	SBRA	Preflowering
	Pigmentation of the vein on the underside of the leaf	PVUL	Preflowering
	Leaf blade shape	LBSH	At 50% flowering
	Shape of the leaf base	STLB	At 50% flowering
	Leaf blade edge	LBED	At 50% flowering
	Ray flower shape	RFSH	At 50% flowering
	Petal tooth slit depth	PTSD	At 50% flowering
	Storage root surface color	SRSC	At harvest
	Flesh color of the storage root	FCSR	At harvest
	Clefts in the storage roots	CSRO	At harvest
	Color of the propagules	COPR	At harvest
Quantitative	Number of stems per plant	NSPL	At the end of flowering
	Plant height (cm)	PLHE	At the end of flowering
	Leaf length (cm)	LELE	At 50% flowering
	Leaf width (cm)	LEWI	At the end of flowering
	Weight of storage roots per plant (kg)	WSRP	At harvest
	Number of storage roots per plant	NSRP	At harvest
	Yield of storage roots (t/ha)	YOSR	At harvest

Data analysis

The characterization data were subjected to multivariate statistical analyses. Initially, the overall structure of the dataset was explored using a factor analysis of mixed data (FAMD), which simultaneously integrated both qualitative and quantitative traits. To further investigate specific patterns, a multiple correspondence analysis (MCA) was applied to the qualitative variables, and a principal component analysis (PCA) to the quantitative ones. In addition, associations among the quantitative traits were assessed using Pearson's correlation coefficient.

To define phenotypically differentiated groups among the accessions, a hierarchical cluster analysis was performed, employing Euclidean distance as the dissimilarity measure and the complete linkage method for clustering. The optimal number of clusters was determined by inspecting the resulting dendrograms, selecting the cut-off point that maximized within-group homogeneity. The biological validity of the clusters was confirmed by evaluating their phenotypic coherence based on the descriptors analyzed.

To assess quantitative differences among the defined clusters, mean comparisons of the traits were conducted using Tukey's HSD test ($p < 0.05$). The analyses were conducted using the Factoextra (Kassambara & Mundt, 2020) and FactoMineR (Lê *et al*, 2008) packages for MCA and PCA, respectively. Dendrograms were generated using the cluster (Maechler *et al*, 2021) and circlize (Gu *et al*, 2014) packages, while visualization of results was carried out with ggplot2

(Wickham, 2016) and iTOL: Interactive Tree Of Life (Letunic & Bork, 2024). Mean comparisons were performed using the AgroR package (Shimizu *et al*, 2023). All analyses were conducted in the RStudio statistical software (R Core Team, 2024).

Results

The phenotypic characterization data recorded for the 214 yacon accessions conserved in the INIA Germplasm Bank are presented in Supplemental Table 2. These include 12 qualitative and 7 quantitative traits, which were used to assess phenotypic variability.

Factor analysis of qualitative and quantitative traits

Figure 2 shows a joint analysis using the 'scree plot' and 'variable contribution' on the first two principal components, based on a mixed data set including 19 traits (qualitative and quantitative) analyzed for the 214 accessions. The scree plot (Figure 2A) shows the values of the first ten components, with the first and second components having the highest values adding up to 8.55, out of a total value of 22.53. In the contribution of variables (Figure 2B), those above the mean contribution line stand out: COPR, FCSR, LELE, LEWI, PTSD, SRSC, LBED and SSCD (for explanation of acronyms see Table 1). Of these, COPR and FCSR are the most relevant qualitative variables, while LELE and LEWI are the most prominent among the quantitative variables. Due to these findings, and considering that the traits WSRP and YOSR are relevant for breeding programmes, further analysis was conducted to explore in detail the specific contribution of each type of variable to obtain a more complete understanding of their influence on the dataset.

Qualitative characteristics of yacon

The characters showed a cumulative variability between Dim1 and Dim2 of 19.9%. The COPR and FCSR traits showed the highest similarity in the MCA (Figure 3A), emphasizing their high contribution to the observed differentiation. In addition, a pattern of grouping of the accessions according to their geographical origin was evident, showing that accessions from Piura and La Libertad were associated with those from Cajamarca; similarly, accessions from Puno and Cusco showed proximity. Individual associations were also identified, such as those from Junín and Pasco (Figure 3B).

Hierarchical analysis of qualitative yacon characters

The hierarchical analysis presented in Figure 4 shows the formation of 12 groups based on qualitative morphological traits, which are structured according to vegetative and reproductive characteristics of the accessions, as well as geographical and possibly environmental factors. The groups are mainly distinguished by leaf morphology (shape, base and margin), branching pattern, as well as surface and pulp coloration of the reservoir root and propagules.

The analysis of Table 2 and the information in Supplemental Table 1 shows that the regions with the lowest representation are Junín (9 accessions), Puno (14 accessions), La Libertad (9 accessions) and Pasco (1 accession), with only two clusters for Junín and Puno, three for La Libertad, and one for Pasco. This low frequency is possibly due to the fact that the accessions from these regions exhibited lower morphological diversity, were restricted to specific characteristics, or there was limited availability of accessions for analysis. The presence of accessions from Junín and Puno in only a few clusters suggests reduced variability, which could indicate local adaptations that have occurred over a shorter evolutionary period or under more homogeneous environmental conditions.

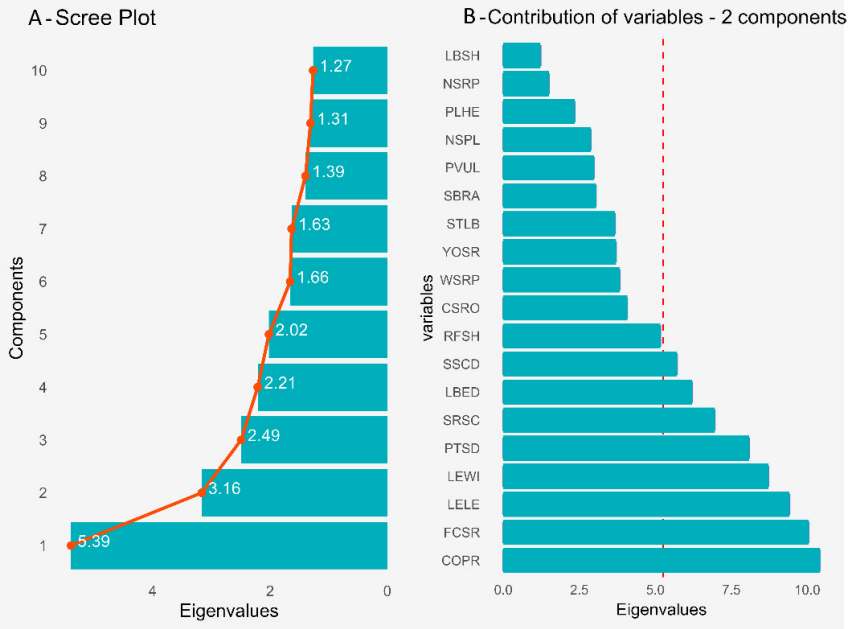


Figure 2. Factor analysis of a mixed data set in 214 yacon (*Smallanthus sonchifolius*) accessions from INIA, Peru. A, Scree plot of ten principal components; B, Contribution of qualitative and quantitative variables. Acronyms are the same as those in Table 1.

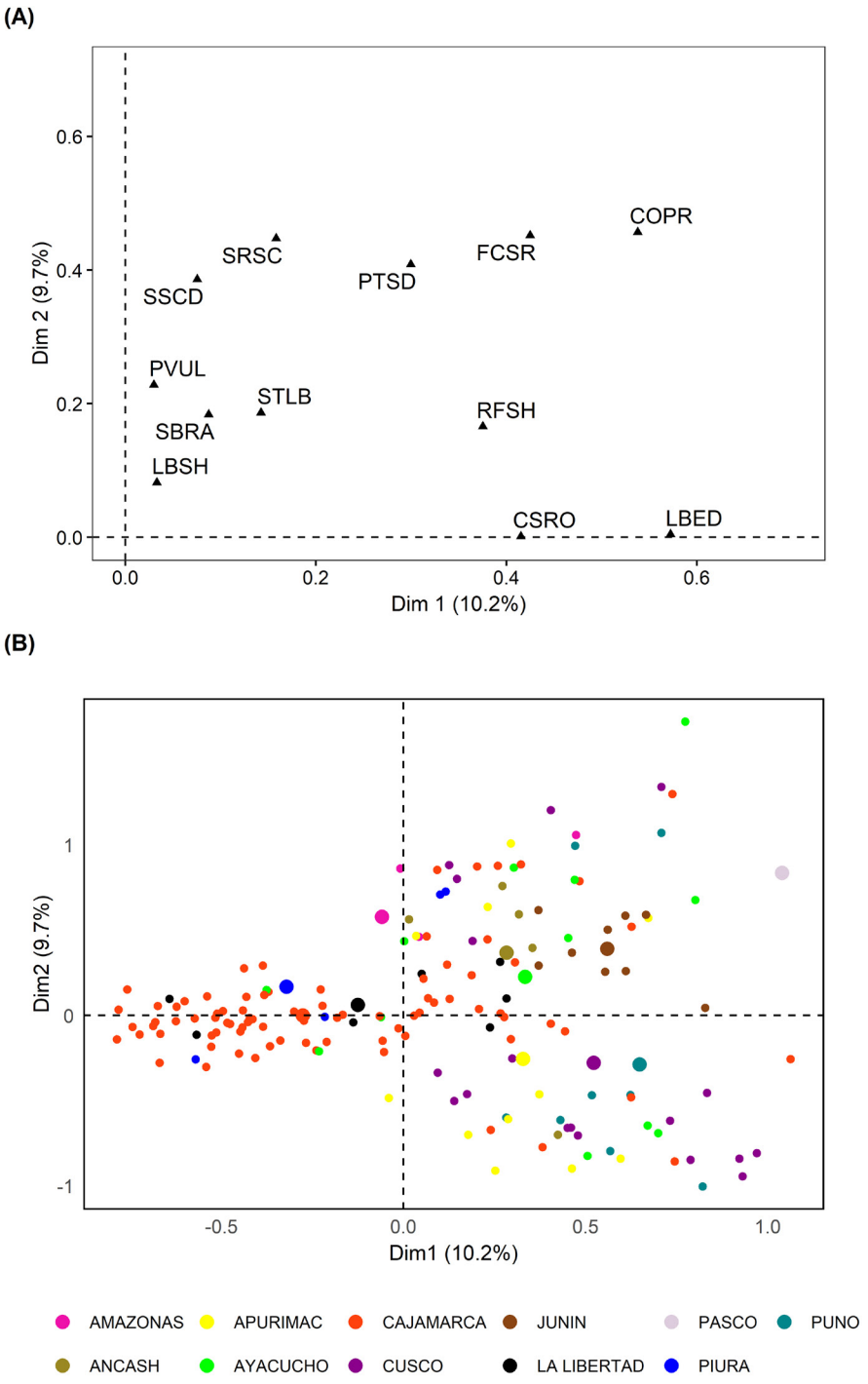


Figure 3. Multiple correspondence analysis (MCA) of twelve qualitative traits in 214 yacon (*Smallanthus sonchifolius*) accessions from INIA, Peru. A, Contribution of qualitative variables to the MCA; B, Geographical clustering based on collection locations. Acronyms correspond to those listed in Table 1. Small dots represent individual accessions; larger dots indicate the centroid of each geographic group.

On the other hand, regions such as Cajamarca and Apurímac, with a high frequency of accessions (presence in seven clusters), showed greater morphological diversity. This is reflected in a wide range of characteristics such as leaf shape, flesh color and stem structure, indicating a wide genetic variability and adaptive capacity.

Quantitative characteristics

Plants had between 1 and 5 stems (NSPL), with a mean of 2.71 (Supplemental Table 2). Fourteen accessions reached the minimum value, while three accessions reached the maximum. In the case of plant height (PLHE), a minimum value of 41.6cm and a maximum value of 164cm were observed, corresponding to accessions Y-206 and Y-154, respectively. On the other hand, accession Y-74 had the longest leaves (LELE) with 25.4cm, while Y-174 recorded the lowest value with 9.6cm. The latter also had the lowest leaf width (LEWI) value with 9.2cm, while the highest value was measured in accession Y-65 with 29cm.

Accessions Y-111, Y-114, Y-137, and Y-182 coincided in the lowest storage root weight (WSRP) with 0.5kg, while accession Y-28 reached the maximum value with 4.7kg, positioning itself as a promising accession due to its high

Table 2. Qualitative morphological group distribution in relation to the main traits of the yacon accessions, as identified through hierarchical cluster analysis.

Cluster	N° of accessions	Regions of provenance	Characteristics
I	74	Cajamarca, Amazonas, La Libertad, Piura	Stem with purple secondary color at nodes and internodes; triangular leaves, with predominant basal branching; surface of storage roots light yellow; yellow-orange pulp; dark greyish purple propagules.
II	40	Apurímac, Ayacucho, Cajamarca, Cusco, La Libertad, Piura	Triangular leaves; elliptic to oblong ligulate flower; variability in flesh color: yellowish white, yellow-orange and light orange.
III	9	Junín	Stem with purple secondary color in internodes; triangular leaves with sagittate base; absence of pigmentation on the veins on the underside of the leaf; light greyish purple storage root surface ; light orange and orange-yellow flesh.
IV	7	Amazonas, Apurímac, Cajamarca, Cusco, Puno, Junín	Secondary color absent on stem; triangular leaves with subhastate base; dark greyish purple storage root surface; white flesh; purplish red and white propagules.
V	15	Amazonas, Ancash, Apurímac, Ayacucho, Cajamarca, Piura	Secondary color absent on stem; triangular leaves; light greyish purple storage root surface; yellowish white or orange-yellow flesh color of the storage roots; purplish red and dark greyish purple propagules.
VI	40	Ancash, Apurímac, Ayacucho, Cajamarca, Cusco, Puno	Stem with purple secondary color at nodes and internodes; triangular leaves with hastate base; light yellow storage root surface, with white flesh mottled with greyish purple or reddish purple tones; white and purplish red propagules.
VII	7	Apurímac, Cajamarca, Cusco	No secondary color on stem; triangular leaves with subhastate base; light greyish purple storage root surface; yellowish white or orange-yellow flesh color of the storage roots; purplish red and dark greyish purple propagules.
VIII	10	Amazonas, Ancash, Apurímac, Ayacucho, Cajamarca, La Libertad	Green stem with purple tinges at nodes and internodes; triangular leaves with subhastate base; light yellow storage root surface; purplish red to dark greyish purple propagules.
IX	6	Cajamarca, La Libertad	Branching along the entire length of the stem; triangular leaves with truncate or subhastate base; light yellow storage root surface; yellow-orange flesh color of the storage roots; white and dark greyish purple propagules.
X	3	Ayacucho, Cusco	Branching along the entire length of the stem; deltoid leaves with truncate base; oblong ligulate flower; light yellow storage root surface; white flesh color of the storage roots; violet-blue propagules.
XI	1	Cajamarca	Stem with purple secondary color at nodes and internodes; basal branching; leaves cordate with lobed base; light yellow storage root surface; light orange flesh color of the storage roots; purplish red with white propagules .
XII	2	Pasco, Ayacucho	Stem with purple secondary color at nodes and internodes; branching along stem; triangular leaves with sagittate base; light greyish purple storage root surface; light orange or yellowish white with reddish purple pits flesh color of the storage roots; purplish red propagules.

productivity. Regarding NSRP, accessions Y-42, Y-103, Y-126, Y-181 presented the lowest number of storage roots (NSRP) with six units, while accession Y-144 reached the maximum production with 32 units. The accessions ranged from 10t/ha to 94t/ha of storage root yield (YOSR). Finally, the coefficients of variation ranged between 18.9% and 45.7%; NSPL, WSRP, NSRP and YOSR were the characters with a variability higher than 30%.

Principal component analysis and correlation

The PCA results (Figure 5) show that the first two principal components together account for 63.80% of the total variance (PC1 = 41.60%; PC2 = 22.20%). Using this analysis, the accessions were grouped into six categories based on the flesh color of the storage root. Accessions with white and

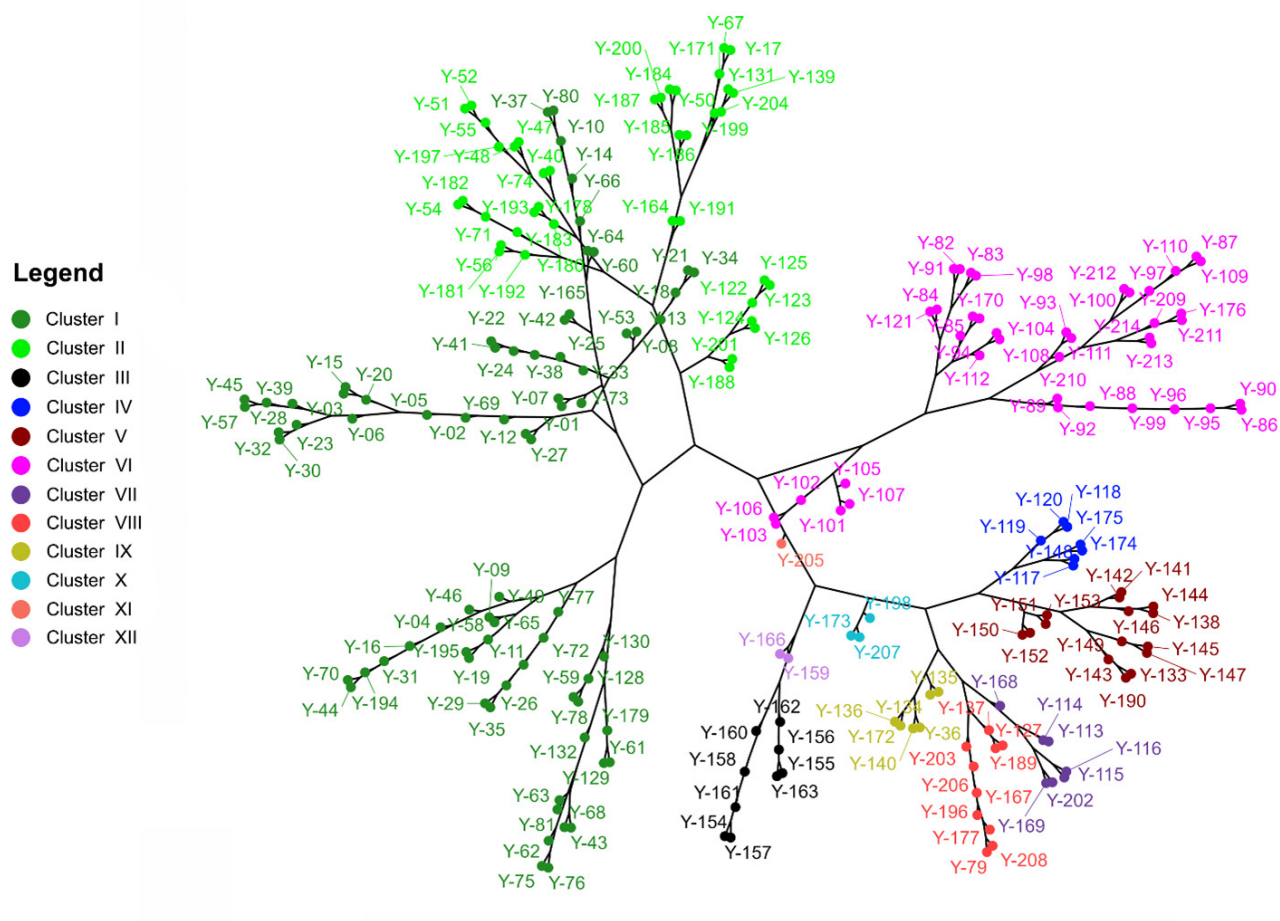


Figure 4. Phylogenetic hierarchical tree of 214 accessions of yacon (*Smallanthus sonchifolius*) based on 12 qualitative characters of the germplasm of INIA, Estación Experimental Agraria Baños del Inca, Cajamarca. Clusters are colour-coded and numbered as in Table 2.

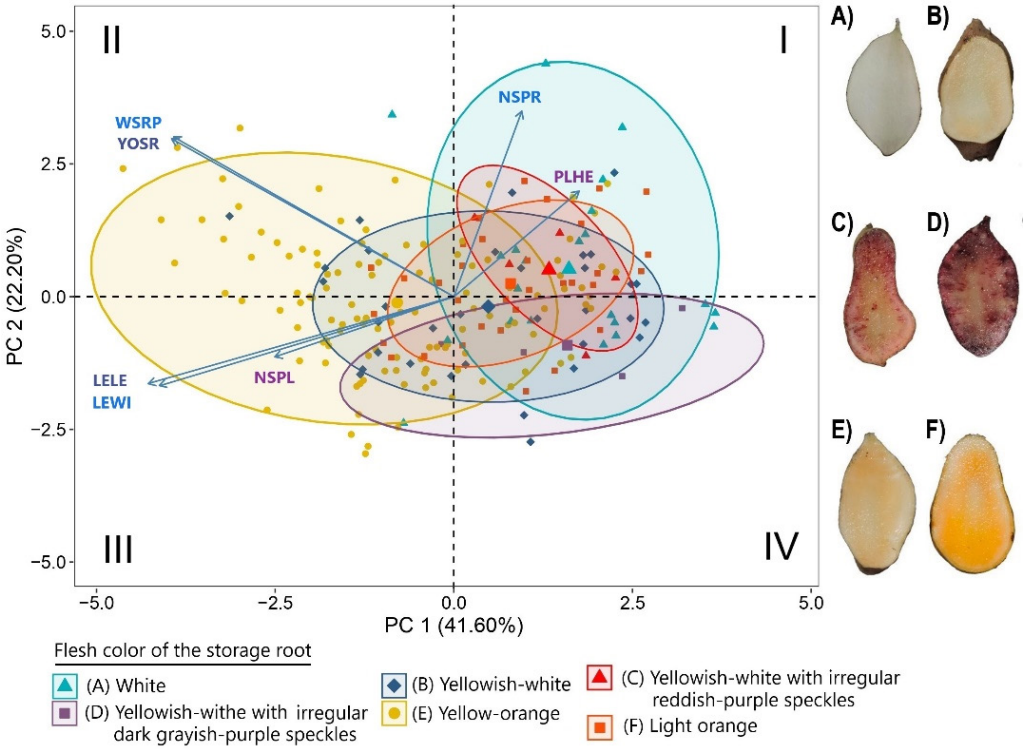


Figure 5. Principal component analysis on 214 accessions of yacon (*Smallanthus sonchifolius*) from INIA, Peru, grouped according to the flesh color of the storage root (with 95% confidence ellipses): A, white; B, yellowish-white; C, yellowish-white with irregular reddish-purple speckles; D, yellowish-white with irregular dark greyish-purple speckles; E, yellow-orange; F, light orange. Acronyms are the same as those in Table 1.

light orange flesh are mainly distributed in quadrants I and IV; accessions with yellowish-white flesh are distributed in the four quadrants; accessions with yellowish-white flesh with irregular reddish-purple speckles are mainly located in quadrant I; accessions with yellowish-white flesh with irregular dark greyish-purple speckles are found in quadrant IV; and accessions with yellowish-orange flesh are distributed in quadrants II and III.

In this case, WSRP and YOSR vectors point towards the upper left quadrant suggesting a strong association of accessions with negative PC1 and positive PC2 values, mainly related to yellow-orange pulp color. On the other hand, PLHE and NSRP show a higher correlation with positive PC1 and PC2 values, being associated with pulp color accessions A, B, C and F.

The variables LELE, LEWI and NSPL correlate with negative values in both PC1 and PC2, suggesting their relationship with accessions in the lower left quadrant.

Correlation between quantitative traits

The correlation matrix for quantitative traits is presented in Table 3. A strong positive correlation was observed between leaf length and leaf width. Both traits also showed moderate positive correlations with the weight of storage roots per plant and with total root yield. The weight of storage roots per plant exhibited a high correlation with total yield. In contrast, the number of storage roots per plant did not show a significant correlation with either root weight or yield.

Plant height was negatively correlated with leaf length, leaf width, and the number of stems per plant. Additionally, the number of storage roots per plant showed negative correlations with leaf dimensions.

Hierarchical analysis of quantitative yacon characters

The circular dendrogram in Figure 6, along with the corresponding information in Table 4, shows a grouping of the accessions in four clusters. Cluster I, with 120 accessions, has a mean of three stems per plant and 116.3cm plant height. This group exhibits a mean leaf length of 20.1cm and leaf width of 20.9cm. The weight of storage roots per plant is 1.8kg, with 13.1 storage roots per plant and a mean yield of 36.91t/ha. This cluster is characterized by mean values for vegetative development and root production compared to the other clusters.

Cluster II is composed of 19 accessions that together have the most outstanding agronomic characteristics of all the groups. With a plant height of 104.3cm and a mean of three stems per plant, the accessions in this group have the largest leaf dimensions, with 21.1 and 21.9cm length and width, respectively. This cluster is particularly distinguished by its high root productivity, with 3.6kg of storage root weight per plant, 15.2 storage roots per plant, and a mean yield of 73.5t/ha.

Cluster III groups a total of 55 accessions with a plant height of 115.4cm and a mean of 2.3 stems per plant. The leaf dimensions of this group are 16.0 cm leaf length and 15.9 cm leaf width, indicating moderate leaf development compared to Clusters I and II. Furthermore, in comparison with Clusters I, II and IV, it presents the lowest values in storage root weight per plant (1.3kg), number of storage roots per plant (11.6) and yield (26.5t/ha).

Cluster IV includes 20 accessions, which represent 9.3% of the total. These accessions are distinguished by a plant height of 131.5cm, the highest among the clusters. In addition, it has a mean of 1.7 stems per plant; together with a reduced leaf development in length and leaf width, with 14.9 and 14.7cm, respectively. In terms of production, the storage root weight per plant is 1.46kg, with 21 storage roots per plant and a yield of 37.37t/ha. This cluster stands out for its high plant size and a higher number of roots per plant.

Descriptive analysis and comparison of means of quantitative traits of yacon

The analysis of the mean values of the quantitative traits (Table 4) revealed significant differences ($p < 0.05$) among the clusters, indicating a clear structuring of the yacon accessions into four groups with distinct agronomic profiles. These groups enable the identification of materials with potential for different objectives: selection aimed at high yields (Cluster II), balance between growth and productivity (Cluster I), or evaluation of accessions with agronomic limitations that may require specific improvements (Clusters III and IV).

Table 3. Correlation matrix among quantitative characters. Significant correlations at *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns: not significant. The character acronyms are the same as those used in Table 1.

Character	NSPL	PLHE	LELE	LEWI	WSRP	NSRP	YOSR
NSPL	1.00						
PLHE	-0.23 ***	1.00					
LELE	0.34 ***	-0.23 ***	1.00				
LEWI	0.30 ***	-0.18 **	0.90 ***	1.00			
WSRP	0.21 **	-0.10 ns	0.39 ***	0.35 ***	1.00		
NSRP	-0.09 ns	0.22 **	-0.23 ***	-0.24 ***	0.10 ns	1.00	
YOSR	0.20 **	-0.11 ns	0.37 ***	0.35 ***	0.98 ***	0.11 ns	1.00

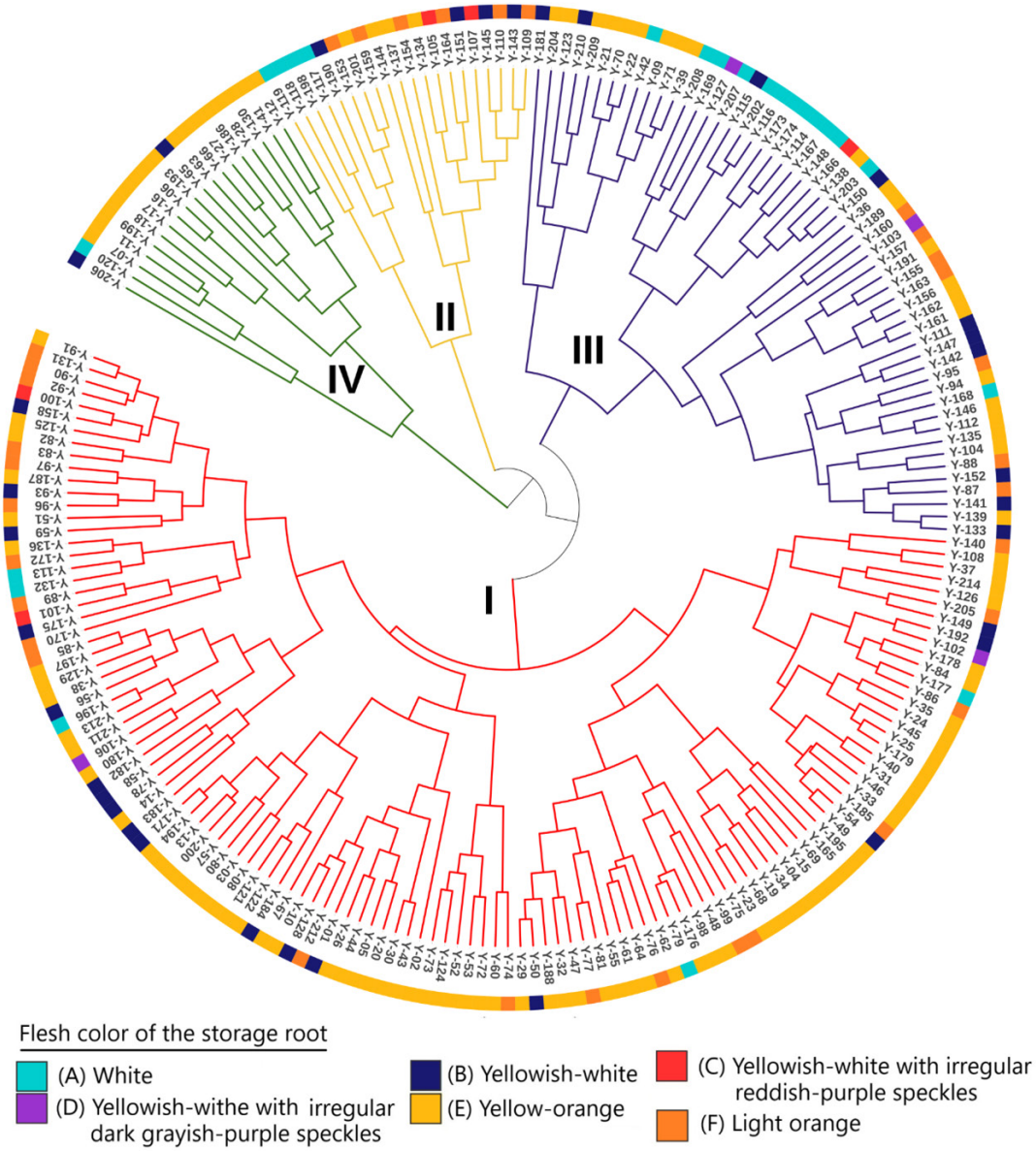


Figure 6. Circular dendrogram of yacon (*Smallanthus sonchifolius*) germplasm of INIA, Peru

Table 4. Descriptive analysis and comparison of means between quantitative traits of the clusters. The character acronyms are the same as those used in Table 1. SD, standard deviation; CV, coefficient of variation. *, means followed by the same letter in the rows do not differ statistically from each other, by Tukey's test ($p < 0.05$).

Character	Mean character values* and SD				
	Cluster I	Cluster II	Cluster III	Cluster IV	CV(%)
NSPL	3.00±0.77 a	3.05±0.52 a	2.29±0.71 b	1.75±0.55 c	26.51
PLHE (cm)	116.33±20.86 b	104.31±25.5 b	115.45±28.05 b	131.5±16.97 a	19.78
LELE (cm)	20.15±2.27 a	21.42±3.21 a	16.03±3.65 b	14.93±2.09 b	14.77
LEWI (cm)	20.96±3.34 a	21.99±4.01 a	15.92±4.25 b	14.76±2.72 b	18.82
WSRP (kg)	1.87±0.59 b	3.68±0.56 a	1.34±0.50 c	1.46±0.66 c	30.94
NSRP	13.18±3.52 bc	15.26±3.33 b	11.64±3.27 c	21.05±3.95 a	25.43
YOSR (t)	36.91±11.93 b	73.58±11.11 a	26.58±9.98 c	29.2±13.18 c	31.29

Discussion

Qualitative characteristics

The qualitative traits contributed heterogeneously to the phenotypic variability of 214 yacon accessions (Figure 3A), explaining 19.9 % of the total variability in the first two dimensions of the analysis. Although the contribution was moderate, it was observed that the color of the propagules and flesh color of the storage root traits stand out for their high discriminatory capacity between accessions, indicating their relevance in group differentiation. In particular, the relevance of FCSR is supported by previous studies that identified it as one of the three most important traits for evaluating yacon hybrids (Vegas *et al*, 2015).

The geographical analysis in Figure 3B revealed that phenotypic variability is influenced by the adaptation of the accessions to specific environmental conditions, observing clustering patterns according to their geographical proximity. Accessions from Piura, La Libertad and Cajamarca formed particular and related groups, probably due to ecological, anthropogenic and genetic conditions, suggesting a strong relationship of the accessions with the environment where they thrive (Da Silva *et al*, 2019). This finding is consistent with those obtained by Polanco and García (2013) who noted that yacon genotypes are adapted and specialized to specific agro-ecological conditions.

Complementarily, the hierarchical analysis presented in Figure 4 and detailed in Table 2 provides a more detailed view of the morphological diversity of the accessions, classifying them into 12 qualitative morphological groups according to their vegetative and reproductive characteristics. This grouping reflects the genetic complexity and adaptation of the plants to different ecological conditions. The differences observed in branching patterns, leaf morphology (including shape, base and margin), as well as the coloration of the storage root and propagules, support a grouping based on their phenotypic characteristics.

The Junín and Pasco accessions showed phenotypic characteristics differentiated from the rest, probably due to the influence of unique microenvironments and genetic factors, which would have driven the evolution and differentiation of these accessions. This finding is supported by molecular studies conducted by Mansilla *et al* (2006) and Soto (2012), who identified accessions specific to central and southern Peru, while in the north, they showed greater homogeneity. The results suggest the existence of important centres of diversity for the conservation, genetic improvement and sustainable use of yacon.

Quantitative characteristics

The descriptive analysis of quantitative traits revealed variability among the accessions with evident differences in plant height and yield traits. PCA (Figure 5) showed the greatest contribution of leaf size and yield traits in the phenotypic differentiation of the yacon accessions. A significant proportion of accessions with yellow-orange flesh color were associated with the WSRP and YOSR vectors, showing a phenotypic differentiation centred on the storage root, suggesting that these accessions were adapted to optimize the accumulation of reserves. This result is congruent with Polanco and García (2013) who determined that yacon has been subjected to anthropogenic selection aimed at obtaining highly productive storage root genotypes.

Accessions with white and light orange flesh were grouped with those exhibiting yellowish-white flesh marked by irregular reddish-purple speckles or irregular dark greyish-purple speckles. This grouping, as observed in Figure 5, shared morphological traits related to PLHE and, to a lesser extent, to NSRP. PLHE was inversely correlated with LELE and LEWI (Table 3). This indicates that accessions with higher plant height had smaller leaf dimensions, while those with lower plant height had larger leaf dimensions (Table 4). In contrast, accessions with yellowish-white flesh exhibited greater dispersion in the four quadrants, indicating greater variability, probably associated with their phenotypic plasticity. Given the relationship of the traits assessed in the PCA, we can select LELE, LEWI, WSRP and YOSR as valuable traits to discriminate accessions within the species.

According to the quantitative traits, the accessions were distributed into four clusters (Figure 6). The analysis of the distribution of accessions suggested that accessions with larger leaf dimensions were associated with higher yields, since a larger leaf area implies greater light uptake, a larger surface area for gas exchange, and greater accumulation of water and nutrients. Consequently, photosynthate production increased, leading to a greater biomass in the storage roots. Leaf dimensions and their relationship with yield have been correlated in other crops such as potato (*Solanum tuberosum* L.) and tomato (*Solanum lycopersicum* L.) (León-Burgos *et al*, 2021), as well as in common bean (*Phaseolus vulgaris* L.) (Warnock *et al*, 2006). These studies suggest that increased photosynthate accumulation in sink organs is related to optimal development of the source organs. However, further studies are required to determine the direct impact of leaf size on yield.

On the other hand, clusters grouping lower-yielding accessions showed limitations in biomass mobilization to storage roots, possibly attributable to vegetative or environmental factors influencing the phenotype. This finding is consistent with Douglas *et al* (2007) who established a significant positive relationship between yield and both planting time and climatic conditions. This observation suggests the need to investigate genotype–environment interaction to identify accessions that maximize the translocation of assimilates to storage roots under different environmental conditions.

Descriptive analysis and comparison of means of yacon quantitative traits (Table 4) provided valuable information on variability and performance of the accessions. These results showed significant relationships between WSRP, YOSR and leaf dimensions. This indicates that selection of promising individuals should focus on clusters with significant and outstanding traits in yield and associated traits (foliage) to maximize productivity in future breeding programmes. This finding coincides with the study by Rodríguez López *et al* (2022), who identified promising genotypes based on their yield and morphological characteristics such as leaf area and number of stems, among others.

Conclusions

The qualitative traits COPR and FCSR, together with the quantitative traits LELE and LEWI, were key determinants in the phenotypic differentiation of the 214 yacon accessions.

Morphological variability exhibited a clear geographical structuring. Accessions from northern Peru (Piura, La Libertad and Cajamarca), the south (Cusco and Puno) and the central region (Junín and Pasco) formed well-defined groupings based on phenotypic similarity.

The hierarchical analysis based on quantitative traits identified Cluster II, comprising 19 accessions, as having the greatest agronomic potential, with an average yield of 73.5t/ha, a storage root weight of 3.6kg per plant, and an average of 15.2 storage roots per plant.

Positive correlations were observed between YOSR and both WSRP, LELE and LEWI, suggesting that foliar development may serve as a reliable predictor of yield performance.

Supplemental data

Supplemental Table 1. Geographical origin and coding of 214 accessions of yacon from the INIA Germplasm Bank, Cajamarca, Peru.

Supplemental Table 2. Agromorphological characterization data of 214 yacon accessions from the INIA Germplasm Bank, Peru.

Acknowledgements

This work was supported by the Subdirección de Recursos Genéticos de la Estación Experimental Agraria Baños del Inca of the Instituto Nacional de Innovación Agraria (INIA). The authors would like to thank Armando Linares Estrada and Sebastián Llico Sánchez for their invaluable support in the field.

Author contributions

Angel Esteban Santa Cruz-Padilla and Jorge Luis Vásquez-Orrillo: conceptualization, formal analysis, writing – original, research, data curation, resources, methodology, proofreading and editing; Silvia Yanina Rodríguez López and Araceli Eugenio Leiva: research, data curation, resources, methodology, proofreading and editing; Ricardo Manuel Bardales-Lozano and Hipolito Murga-Orrillo: formal analysis, writing – original, research, methodology, proofreading and editing; Juan F. Seminario: research, resources, methodology, proofreading and editing.

Conflict of interest statement

The authors have declared that no competing interests exist.

References

Adriano, L., Dionísio, A., De Abreu, F. Carioca, A., Zocolo, G., Wurlitzer, N., De Oliveira Pinto, C., de Oliveira, A., and De Carvalho Sampaio, H. (2019). Yacon syrup reduces postprandial glycemic response to breakfast: A randomized, crossover, double-blind clinical trial. *Food Research International*, 126, 108682. https://doi.org/10.1016/j.foodres.2019.108682

Baek, S., Choi, N., Lee, K.-P., Jhun, H., and Kim, J. (2018). *Smallanthus sonchifolius* leaf attenuates neuroinflammation. *Journal of Exercise Nutrition & Biochemistry*, 22(2), 31-35. https://doi.org/10.20463/jenb.2018.0014

Caballero, L. and Colonia, A. (2018). Yacón como planta promisoria en el manejo de enfermedades. *Investigaciones Andina*, 20(36), 145-157. https://dialnet.unirioja.es/servlet/articulo?codigo=9359322

Da Silva, D., De Oliveira, F., Quaresma, M., Erlacher, W., and Mendes, T. (2019). Yacon production at different planting seasons and growing environments. *Bioscience Journal*, 35(4), 992-1001. https://doi.org/10.14393/BJ-v35n4a2019-42091

De Sales, R., De Oliveira, E., Xavier, A., De Oliveira, F., Pezzopane, J., Da Silva, D., and Da Silva Berilli, S. (2021). Base temperature, cycle duration, and thermal constant for yacon culture. *Acta Scientiarum. Agronomy*, 44, e52623. https://doi.org/10.4025/actasciagron.v44i1.52623

Douglas, J., Follett, J., Douglas, M., Deo, B., Scheffer, J., Littler, R. and Manley-Harris, M. (2007). Effect of environment and time of planting on the production and quality of yacon (*Smallanthus sonchifolius)* storage roots. *New Zealand Journal of Crop and Horticultural Science*, 35(1), 107-116. https://doi.org/10.1080/01140670709510174

Ferraz, A., Garcia, J., Costa, M., De Almeida, C., Gregolin, C., Alves, P., Hasimoto, F., Berchieri-Ronchi, C., Dos Santos, K., and Corrêa, C. (2020). Yacon (*Smallanthus sonchifolius*) use as an antioxidant in diabetes. En V. R. Preedy (Ed.), Pathology (pp. 379-386). Academic Press. https://doi.org/10.1016/B978-0-12-815972-9.00036-6

García, D., Sotelo, A., Malpica, E., Álvarez, H., Norabuena, E., Gonzáles, T., and Sumarriva, L. (2022). Impacto del helado dietético con yacón (*Smallanthus sonchifolius*) en la hipoglicemia y aceptabilidad. *Nutrición Clínica y Dietética Hospitalaria*, 42(2), 142-149. https://doi.org/10.12873/422garcia

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* 30(19), 2811–2812. doi: https://doi.org/10.1093/bioinformatics/btu393

Honoré, S., Grande, M., Gomez, J., and Sánchez, S. (2018). Smallanthus sonchifolius (Yacon) Flour Improves Visceral Adiposity and Metabolic Parameters in High-Fat-Diet-Fed Rats. *Journal of Obesity*, 2018(1), 5341384. https://doi.org/10.1155/2018/5341384

Huaycho, H., Aruquipa, R., Mercado, G., Trigo, R., Bosque, H. y Condori, J. (2016). Conocimientos tradicionales en yacón o aricoma (*Smallanthus sonchifolius*) en comunidades de Mocomoco, Coroico e Irupana de La Paz. *Revista de Investigación e Innovación Agropecuaria y de Recursos Naturales*, 3(2), 152-165. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2409-

16182016000200005&lng=es&tlng=es

Ignacio, S., Camarena, F., Baudoin, J. y Blas, R. (2017). Ethno-Botany and in-situ conservation of the genetic diversity of arracacha (*Arracacia xanthorrhiza* Bancroft), yacon (*Smallanthus sonchifolious* H. Robinson), and wild relatives. *Peruvian Journal of Agronomy*, 1(1), 21-31. http://dx.doi.org/10.21704/pja.v1i1.1064

Kassambara, A. and Mundt, F. (2020). Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package Version 1.0.7. url: https://CRAN. R-project.org/package=factoextra

Lachman, J., Fernández, E., Viehmannová, I., Šulc, M. and Èepková, P. (2007). Total phenolic content of yacon (*Smallanthus sonchifolius*) rhizomes, leaves, and roots affected by genotype. *New Zealand Journal of Crop and Horticultural Science*, 35(1), 117-123. https://doi.org/10.1080/01140670709510175

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* 25(1), 1–18. doi: http://dx.doi.org/10.18637/jss.v025.i01

Lebeda, A., Dolezalová, I. and Dolezal, K. (2004). Variation in morphological and biochemical characters in genotypes of maca and yacon. *Acta Horticulturae*, 629, 483-490. https://doi.org/10.17660/ActaHortic.2004.629.62

León-Burgos, A., Beltrán-Cortes, G., Barragán-Pérez, A., and Balaguera-López, H. (2021). Distribution of photoassimilates in sink organs of plants of Solanaceas, tomato and potato. A review. *Ciencia y Agricultura*, 18(3), 79-97. https://doi.org/10.19053/01228420.v18.n3.2021.13566

Letunic, I., and Bork P. (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool, *Nucleic Acids Research*, Volume 52, Issue W1, 5 July 2024, Pages W78–W82, https://doi.org/10.1093/nar/gkae268

Lopera-Marín, J., Angulo-Arizala, J., Murgueitio-Restrepo, E. y Mahecha-Ledesma, L. (2020). Producción de tubérculos y biomasa aérea del yacón, *Smallanthus sonchifolius* (Poepp.) H. Rob. (Asteraceae), para alimentación animal en el trópico alto colombiano. *Livestock Research for Rural Development*. 32 (135). http://www.lrrd.org/lrrd32/8/jjlop32135.html

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2021). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.2. url: https://CRAN.R-project.org/package=cluster

Mansilla, R., López, C., Blas, R., Chia, J. y Baudoin, J. (2006). Análisis de la variabilidad molecular de una colección peruana de *Smallanthus sonchifolius* (Poepp & Endl) H. Robinson "YACÓN". *Ecología Aplicada*, 5(1-2), 75-80. https://doi.org/10.21704/rea.v5i1-2.320

Mansilla, R., López, C., Flores, M. y Espejo, R. (2010). Estudios de la biología reproductiva en cinco accesiones de *Smallanthus sonchifolius* (Poepp. & Endl.) Robinson. *Ecología Aplicada*, 9(2), 167-175. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-22162010000200012&lng=es&tlng=es.

Minchola-Castañeda, K., Luzuriaga-Tirado, E., Montalvo-Rodríguez, A., Moncada-Carrera, J., Morales-Ibañez, F. y Gil-Reyes, W. (2022). Propiedades beneficiosas del yacón (*Smallanthus sonchifolius*) en la salud. *Más Vita*, 4(3), 87-98. https://doi.org/10.47606/ACVEN/MV0135

Moreira, R., Redko, F., Ulloa, J., Flor, S., Tulino, M., Muschietti, L., and Carballo, M. (2020). Toxicogenetic evaluation of *Smallanthus sonchifolius* (yacon) as a herbal medicine. *Journal of Ethnopharmacology*, 257. 112854. https://doi.org/10.1016/j.jep.2020.112854

Myint, P., Dao, T., and Kim, Y. (2019). Anticancer Activity of *Smallanthus sonchifolius* Methanol Extract against Human Hepatocellular Carcinoma Cells. *Molecules*, 24(17), 3054. https://doi.org/10.3390/molecules24173054

Polanco, M. and García, M. (2013). Caracterización morfológica y molecular de materiales de yacón (*Smallanthus sonchifolius* Poep. & Endl) H. Robinson colectados en la ecorregión Eje Cafetero de Colombia. *Revista de Investigación Agraria y Ambiental*, 4(2), 97-116. https://doi.org/10.22490/21456453.981

Quaresma, M., De Oliveira, F., Amaral, J., Parajara, M., Dalvi, L., and Teixeira, A. (2020). Planting methods and depths for yacon (*Smallanthus sonchifolius*) crops. *Revista Colombiana de Ciencias Hortícolas*, 14(2), 249-256. https://doi.org/10.17584/rcch.2020v14i2.9562

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/

Reis, F., Marques, C., Sales de Moraes, A., and Masson, M. (2021). Effect of processing methods on yacon roots health-promoting compounds and related properties. *Trends in Food Science & Technology*, 113, 346-354. https://doi.org/10.1016/j.tifs.2021.05.010

RHS (2001). The Royal Horticultural Society Colour Chart.

Rodríguez, S., Seminario, A., Vásquez, V. y Seminario, J. (2022). Rendimiento agronómico de ocho cultivares de yacón [*Smallanthus sonchifolius* (Poepp. & Endl.) H. Rob.] del norte peruano. *Siembra*, 9(1). e3630 https://doi.org/10.29166/siembra.v9i1.3630

Russo, D., Malafronte, N., Frescura, D., Imbrenda, G., Faraone, I., Milella, L., Fernandez, E., and De Tommasi, N. (2015). Antioxidant activities and quali-quantitative analysis of different *Smallanthus sonchifolius* [(Poepp. and Endl.) H. Robinson] landrace extracts. *Natural Product Research*, 29(17), 1673-1677. DOI: https://doi.org/10.1080/14786419.2014.990906

Santa Cruz Padilla, A. E. y Vásquez Orrillo, J. L. (2023). Catálogo de yacón del Banco de Germoplasma del INIA. En Instituto Nacional de Innovación Agraria. Instituto Nacional de Innovación Agraria. https://repositorio.inia.gob.pe/handle/20.500.12955/2237

Santos, K. dos, Bueno, B., Pereira, L., Francisqueti, F., Braz, M., Bincoleto, L., Da Silva, L., Ferreira, A., Nakamune, A. C. de M. S., Chen, C.-Y. O., Blumberg, J., and Corrêa, C. (2017). Yacon (*Smallanthus sonchifolius*) Leaf Extract Attenuates Hyperglycemia and Skeletal Muscle Oxidative Stress and Inflammation in Diabetic Rats. *Evidence-Based Complementary and Alternative Medicine*, 2017(1), 6418048. https://doi.org/10.1155/2017/6418048

Seminario, J., Valderrama, M. y Manrique, I. (2003). El Yacon Fundamentos para el Aprovechamiento de un Recurso Promisorio. https://cipotato.org/wp-content/uploads/2014/07/Yacon_Fundamentos_password.pdf

SENHAMI (2020). Servicio Nacional de Meteorología e

Hidrología del Perú-Datos hidrometeorológicos. url: https://www.senamhi.gob.pe/servicios/?p=estaciones

Shimizu, G., Marubayashi, R., and Goncalves, L. (2023). AgroR: Experimental Statistics and Graphics for Agricultural Sciences. url: https://agronomiar.github.io/AgroRpackage/index.html

Soto Torres J. V. (2012). Evaluación de la diversidad genética de colecciones de *Smallanthus sonchifolius* (Poepp. & Endl.) "Yacón" en el Perú. Tesis de Maestría. La Molina – Perú. url: https://hdl.handle.net/10568/126005

Valentová, K., Lebeda, A., Doležalová, I., Jirovský, D., Simonovska, B., Vovk, I., Kosina, P., Gasmanová, N., Dziechciarková, M., and Ulrichová, J. (2006). The biological and chemical variability of yacon. *Journal of agricultural and food chemistry*, 54(4), 1347-1352. DOI: https://doi.org/10.1021/jf052645u

Vegas, D., Bracamonte, O. y Valladolid, A. (2015). Caracterización morfológica de seis variedades parentales de yacón (*Smallanthus sonchifolius*) y trece cruzas obtenidas de un plan de hibridación. *Revista Peruana de Biología*, 22(2), 175-192. https://doi.org/10.15381/rpb.v22i2.11352

Wagner, M., Kamp, L., Graeff-Hönninger, S., and Lewandowski, I. (2019). Environmental and Economic Performance of Yacon (*Smallanthus sonchifolius*) Cultivated for Fructooligosaccharide Production. *Sustainability*, 11(17), 4581. https://doi.org/10.3390/su11174581

Warnock, R., Valenzuela, J., Trujillo, A., Madriz, P. y Gutiérrez, M. (2006). Área foliar, componentes del área foliar y rendimiento de seis genotipos de caraota1. *Agronomía Tropical*, 56(1), 21-42. url: https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0002-192X2006000100002&lng=es&tlng=es

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. url: https://ggplot2.tidyverse.org

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 39–56
DOI: 10.46265/genresj.APNR6909
https://www.genresj.org
ISSN: 2708-3764

The promise of access and benefit-sharing is met through holistic policy reform: Insights from Colombia's genetic diversity and innovation landscape during COP16

Bob Kreiken[a,b], **Lotte Asveld**[a,c]

[a] *Section Ethics and Philosophy of Technology, Delft University of Technology, Jaffalaan 5, 2628 BX, Delft, the Netherlands*
[b] *Centre for Genetic Resources, Wageningen University & Research, Droevendaalsesteeg 4, 6708 PB, Wageningen, the Netherlands*
[c] *Section Biotechnology and Society, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, the Netherlands*

Abstract: To tackle the global biotechnological innovation divide, Parties to the Convention on Biological Diversity (CBD) are negotiating policies to fairly share the benefits from the use of digital sequence information (DSI) on genetic resources. The policies aim to transfer money, knowledge and technologies from technology-rich developed to biodiversity-rich developing countries in order to bolster the latter's capacities to achieve the CBD's objectives. However, by focusing predominantly on scientific capacities, these policies overlook the complex interactions between various actors, conditions and infrastructures that collectively constitute a country's innovation capacity. In the first-time application of the National Innovation System approach in this policy context, we identify many factors contributing to an innovation gap in Colombia, the host country of COP16, resulting in barriers to study and valorize biodiversity and in lost opportunities for the country to benefit from new technologies. This analysis calls for consideration of broader policy reforms in access and benefit-sharing (ABS) negotiations, and illustrates how holistic policy interventions are needed in countries that benefit from ABS instruments to effectively use financial, scientific and technological resources. Without such an approach, efforts to enhance benefit-sharing from genetic resources and DSI risk reinforcing inequalities in innovation capacity. Finally, we discuss actions countries could take to use their current resources better, as well as how scientists and companies as users of genetic resources and DSI can pursue mutual interests by tackling innovation bottlenecks.

[Para una versión en español del resumen, por favor consulte los Datos suplementarios – For a Spanish version of the abstract, please see Supplemental data*]*

Keywords: Bioprospecting, access and benefit-sharing, digital sequence information, capacity building, distributive justice, Cali Fund, innovation divide, national innovation system

Citation: Kreiken, B. and Asveld, L. (2025) "The promise of access and benefit-sharing is met through holistic policy reform: Insights from Colombia's genetic diversity and innovation landscape during COP16", *Genetic Resources*, 6(12), pp. 39–56. doi: 10.46265/genresj.APNR6909.

Introduction

The completion of the human genome kick-started the 21st century for biotechnology and bioinformatics. With the rapid decrease in sequencing costs, large swathes of genetic sequence data from wild and domesticated species are being generated. These data help researchers and companies understand the threats species face and identify valuable genetic traits in them, such as drought resistance or the ability to break down plastic. Globally, however, there is a growing divide between countries with and without this capacity to reap scientific and economic benefits.

For many years, most benefits from the use of digital sequence information (DSI) on genetic resources have been accrued by high-income countries (HIC), while most biodiversity, and therefore potential DSI, is found in low- and middle-income countries (LMIC). This inequality has been the subject of access and benefit-sharing (ABS) negotiations under the UN Convention on Biological Diversity (CBD) (Rohden & Scholz, 2022). At COP16, held in Cali, Colombia, in 2024, the CBD negotiated the functioning of a multilateral mechanism for benefit-sharing from the use of DSI, including the Cali Fund for the disbursement of monetary benefits, and

* Corresponding author: Bob Kreiken (b.e.kreiken@tudelft.nl)

Received: 28.02.2025 | Accepted: 07.07.2025 | Published online: 19.08.2025

called upon large and medium-scale businesses that use DSI to contribute 1% of their profit or 0.1% of their revenue (CBD, 2024). This mechanism, which, according to some, could potentially generate USD billions per year (LSE Roundtable Team, 2024), is expected to be used by recipient governments to fund conservation projects, meeting the self-identified needs of Indigenous peoples and local communities (IPLC), technology transfer and capacity-building. Scientists and companies are expected to contribute to the latter activities under the banner of non-monetary benefit-sharing.

Literature on a multilateral benefit-sharing mechanism has focused on alignment with research needs, its underlying ethical principles and directions for allocation of the funds (Bagley, 2021; Deplazes-Zemp, 2019; Scholz *et al*, 2022). While the Cali Fund details are being further negotiated, it remains unclear why certain countries fare better in developing their innovation capacity to access, generate and utilize genetic resources and DSI than others. Without that knowledge, benefit-sharing from the Cali Fund risks being ineffective and even unjust. COP host Colombia, an upper-middle-income megadiverse country with advanced science but a small biotech sector, and a likely beneficiary, is an excellent case study to investigate this research question.

Initially, innovation scholarship assumed a linear relationship between government-funded basic and applied research, the development of products and their diffusion in society (Godin, 2006). ABS frameworks arguably mirror this view by regarding research on genetic resources as a stepping stone for commercial bioprospecting and benefit redistribution by governments (Secretariat of the Convention on Biological Diversity, 2011). Biotechnological trajectories, however, are embedded in and formed by institutions and their interactions (Chaturvedi, 2005; Hall, 2005). From early on, the capacity to create and share benefits as an incentive to promote conservation has been part of the rationale behind ABS policies (Sirakaya, 2022). However, assumptions that benefit-sharing automatically translates to enhanced innovation capacities are far too simple. Bilateral ABS agreements have long been criticized for oversimplifying how genetic resources are used in research and development (R&D) (Sherman *et al*, 2025). The factors that make R&D in a country possible in the first place are, however, still overlooked in the ABS literature.

It is important that ABS policies also recognize this institutional complexity so that to-be-shared benefits strengthen these interactions. That gap in understanding is evident in the recurring tendency to attribute scientific capacity development challenges to resource deficiencies. For example, making more data, information and communication technologies, and training resources available may increase individual scientists' capacity but obscures "insidious" patterns of inequality (Bezuidenhout *et al*, 2017). Precisely because knowledge production is sustained by institutional, economic, organizational and political factors (Mormina, 2019), taking into account and strengthening the knowledge structures wherein monetary and non-monetary benefits are created and received is as important as facilitating benefit-sharing itself. In a nutshell, these insights call for holistic, country- and issue-specific capacity-building and investments by the Cali Fund and by users of genetic resources and DSI.

The moral value at stake here is the fair and equitable sharing of benefits from genetic diversity. We reiterate two distributive justice claims here. Distributive justice requires a fair distribution of both benefits and scientific capabilities to create benefits (Mormina, 2019). That means that aside from an equal distribution of resources and opportunities in science, structural biases and barriers in the use of genetic resources and DSI are dismantled. Furthermore, the (non-) monetary benefit transfers do not change the overall direction of R&D, but distributive justice demands from users of genetic resources and DSI an integration of the needs and priorities of beneficiaries upstream at the onset of the R&D cycle (Kreiken & McCarthy, 2025; De Jonge & Korthals, 2006). So, instead of maintaining the status quo, HIC as dominant valorizers of DSI and LMIC as beneficiaries, we argue that the CBD and its stakeholders should target countries' innovation capacity gaps to create and retain benefits.

The reason for undertaking this study is to assess how a country's creation and retention of benefits from genetic diversity is influenced by institutional, economic, historical, organizational and political factors. We now turn to the policy rationale behind this study in relation to ABS policymaking and the ongoing development of the Cali Fund.

While most capacity-building programmes of the ABS Initiative and Global Environment Facility focus on legislative capacities to implement ABS policies, there are so far fewer programmes focused on scientific and innovation capacity deficits to use DSI and genetic resources. Recipients of the Cali Fund are expected to primarily direct funding towards activities that contribute to conservation and sustainable use of biodiversity, which can include scientific research and capacity-building to "generate, access, use, analyse and store [DSI]" (article 18 of Annex Decision 16/2 (CBD, 2024)).

Innovation and institutional capacities to valorize scientific research on DSI and genetic resources are overlooked, however. Currently, there are indications that the vast majority of their economic value ('the pie') is captured at the end of the bioprospecting value chain in patents acquired in HIC (Dunshirn & Zhivkoplias, 2024). While historically LMIC have benefited greatly from conserving and developing their biodiversity, for example, by having a rich crop and animal breed variety with nutritional and medicinal value, not all of these benefits have translated to financial gains or technological development. This valorization gap should be considered significant in the context of the premise of ABS to transfer money back to LMIC for conservation and capacity-building purposes. Charting out a path to economic self-sustenance in ABS policy is important because the contributions to the Cali Fund so far remain voluntary, making expectations about it being a sustainable source of finance for LMIC perhaps unrealistic. Additionally, the COVID-19 pandemic laid bare the vaccine dependency of LMIC, leading to calls for greater biotechnological sovereignty (Guzman *et al*, 2024).

Simply put, if the innovation divide is left unchanged, the monetary benefits that LMIC will receive through ABS ('crumbs') are marginal relative to the economic gains realized in HIC in the long term. For context, in unequal exchanges in raw materials and labour with the Global North, the losses the South incurs exceed the aid it receives thirtyfold (Hickel *et al*, 2022). Disregarding the innovation divide is a missed opportunity because countries with genetic diversity-based industries may direct innovations and tax revenue to nationally relevant goals, including conservation and scientific research, also because we assume that R&D activities in LMIC are more easily matched to the needs of the country and its vulnerable groups than downstream R&D activities in HIC. In addition, we expect that companies will be more willing to contribute to the

Fund if beneficiary countries have clearer ideas of issues that can be addressed through the Fund and have long-term plans for greater economic self-sustenance. These assumptions do not disregard the need for fair and equitable benefit-sharing.

Therefore, this article's insights into the factors that hamper or boost scientific research and innovation for conserving and sustainably using genetic diversity, contribute to informing investment priorities for Cali Fund recipients and broader business engagement. In the next sections, we first elaborate on the National Innovation System model that guides our analysis and data collection. After an overview of Colombia's relevant laws, state of biodiversity and bioeconomy, the findings are categorized per aspect of the value chain and linked back to components of the analytic model. Finally, we make a call to action to rethink domestic and ABS policymaking and the Cali Fund's investment priorities.

Materials and methods

Framework: National Innovation Model

In this section, we explain how the recognition of institutions and interactions enhances our ability to answer the research question. Figure 1 represents a simplified value chain of genetic resources and DSI in Colombia, according to a linear innovation view.

To integrate institutional complexity, science policy analysts have used the National Innovation System (NIS) since the 1980s to analyze individual systems of innovation and their interactions, like the alignment of education with business priorities (Godin, 2009). The commonly used framework for NIS is shown in Figure 2 and includes actors and processes that enable knowledge- and innovation-based economic development (Kuhlmann & Arnold, 2001). As a whole, the NIS model reflects the underlying mechanics of a society's innovation capacity, which is "the context-specific range of skills, actors, practices, routines, institutions and policies needed to put knowledge into productive use in response to an evolving set of challenges, opportunities, and technical and institutional contexts" (Hall, 2005).

Following this definition, we include various users and providers of genetic resources, contextualize the work in relation to ABS and science policy in Colombia, and consider biodiversity loss as the main challenge and the bioeconomy as the main opportunity. For the purpose of the article, and not uncommonly, we enlarge the basic NIS model with three additions to form a nature-based biotechnological innovation system (see Figure 3). In line with the potential of NISs for positive environmental impact (Brás & Robaina, 2024; Fernandes *et al*, 2022), we hypothesize that increased benefit-sharing contributes to the conservation and sustainable use of biodiversity. Because genetic resources and DSI can be considered inputs to the innovation system (Bruynseels, 2020), we include a 'natural system' and a detailed

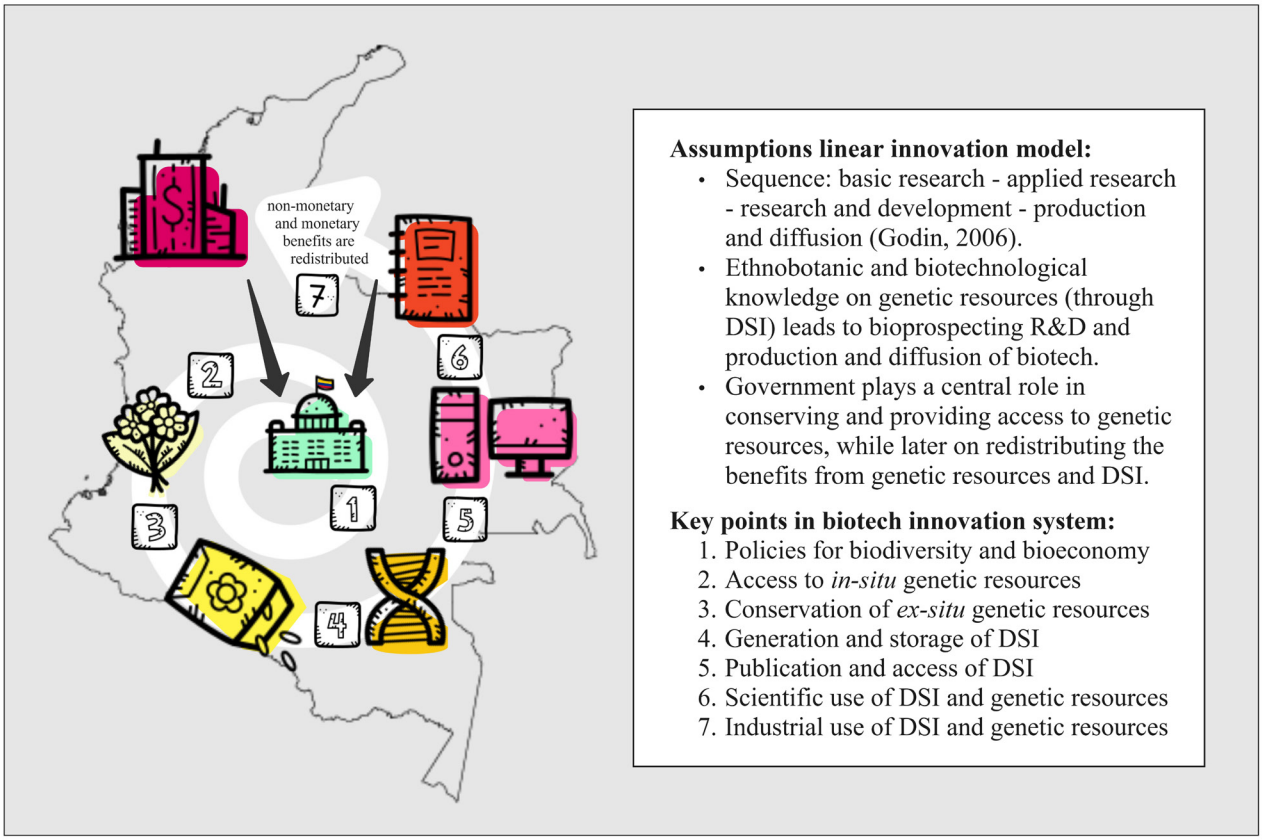


Figure 1. Linear representation of innovation in Colombia. Adapted imagery from Icons8.

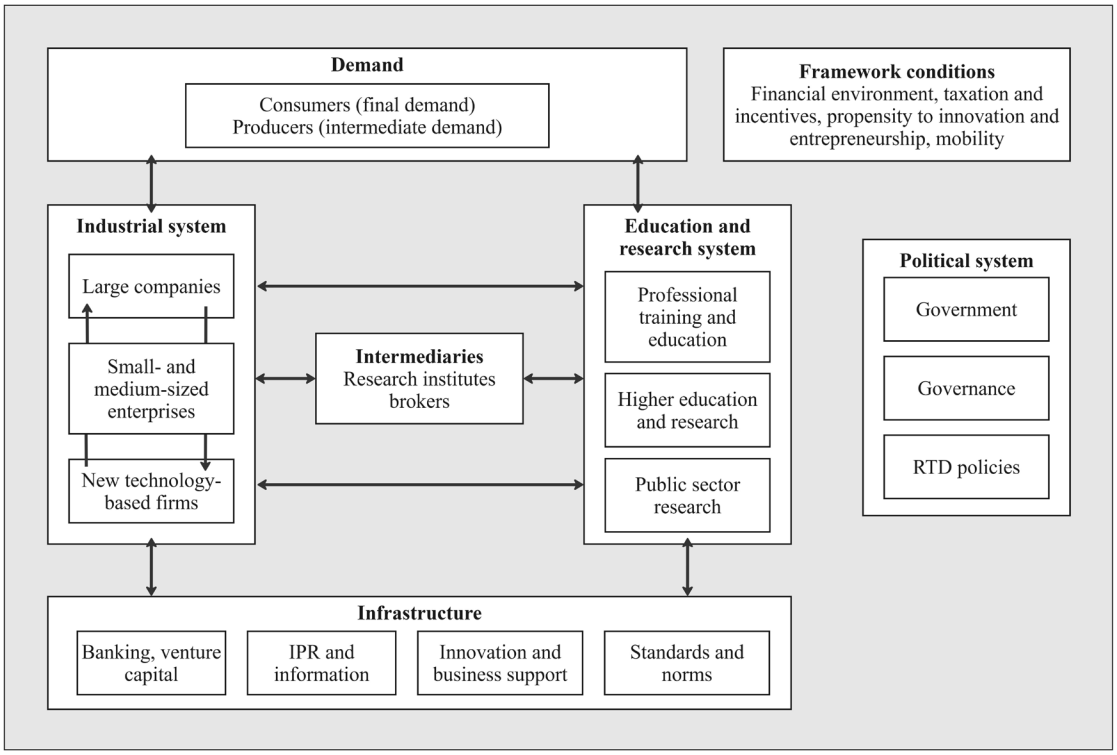


Figure 2. The National Innovation System model (adapted from Kuhlmann & Arnold (2001)). The NIS shows all institutions and actors in various systems that play a role in driving a country's innovation. Well-performing linkages between systems, represented by the arrows, are equally important. For example, actors in the industrial system react to consumer demand and government demand for R&D by commercializing innovations that were developed or co-developed with actors in the education and research systems. Activities in the industrial, education and research systems are influenced by a country's infrastructure, which can range from (un)available venture capital to code of conduct and strong/weak protection of intellectual property rights.

subcategory for DSI-related research infrastructure. Secondly, traditional knowledge associated with genetic resources and DSI is included in the education and research system. Thirdly, because supranational science, technology and innovation policies are gaining more influence on NISs (Weerasinghe *et al*, 2024), and because we want to know the (potential) impact of international ABS policies, we include the 'international policy and political system'. Altogether, Figure 3 shows that the value chain of genetic resources and DSI, as represented in Figure 1, is sustained and influenced by various systems and interactions.

Although this is the first application of the NIS model in this policy context, we are cognizant of the empirical gap and challenges with its application to developing countries. In a general sense, the developing context is characterized by weaker intellectual property rights (IPRs), incremental technological development, unstructured business interactions, and low levels of knowledge, demand and investment (Egbetokun *et al*, 2017). Many developing countries also lack adequate data to allow for international comparison (Weerasinghe *et al*, 2024).

Data collection

At the start, we conducted a short scoping review of literature and policy documents on biodiversity research, biotechnology and the bioeconomy. With approval from a human research ethics committee, online and in-person semi-structured interviews with professionals throughout the enlarged NIS were conducted during one month of fieldwork during and after COP16 in the fall of 2024 (Table 1). COPs are a good field site because host countries position themselves strongly with regard to the CBD's objectives (Lee *et al*, 2021), and because they have an unprecedented concentration of stakeholders. The research questions and conceptual framework were revised cyclically during and after the fieldwork (Lew, 2010). Beforehand, interviewees were identified and contacted via LinkedIn, based on their contributions to relevant research articles and webinars. Further interviews were secured at COP16, which was separated into a Blue Zone for negotiations and associated events, and a Green Zone for more Colombia-specific events. Furthermore, two business conferences were attended, the Expo Bioingredientes in Cali and the Open Innovation and Investor Summit in Bogotá. Visits to the biochemical laboratory of Icesi University in Cali, and the bioprospecting laboratory of INVEMAR in Santa Marta, complemented findings on research infrastructure. Finally, three ecotours to Farallones, Chingaza and Tayrona Park helped to understand the conservation context. The English and Spanish transcripts from the 53 interviews were open-coded and thereafter clustered under one or more of the NIS model's components. Preliminary findings were presented to bystanders in the hall of the COP16 Blue Zone's venue, and later in a seminar on the COP outcome and DSI at Universidad de Los Andes for several key stakeholders. Interviewees were requested to validate the results section and give written permission to be cited anonymously or with their full name.

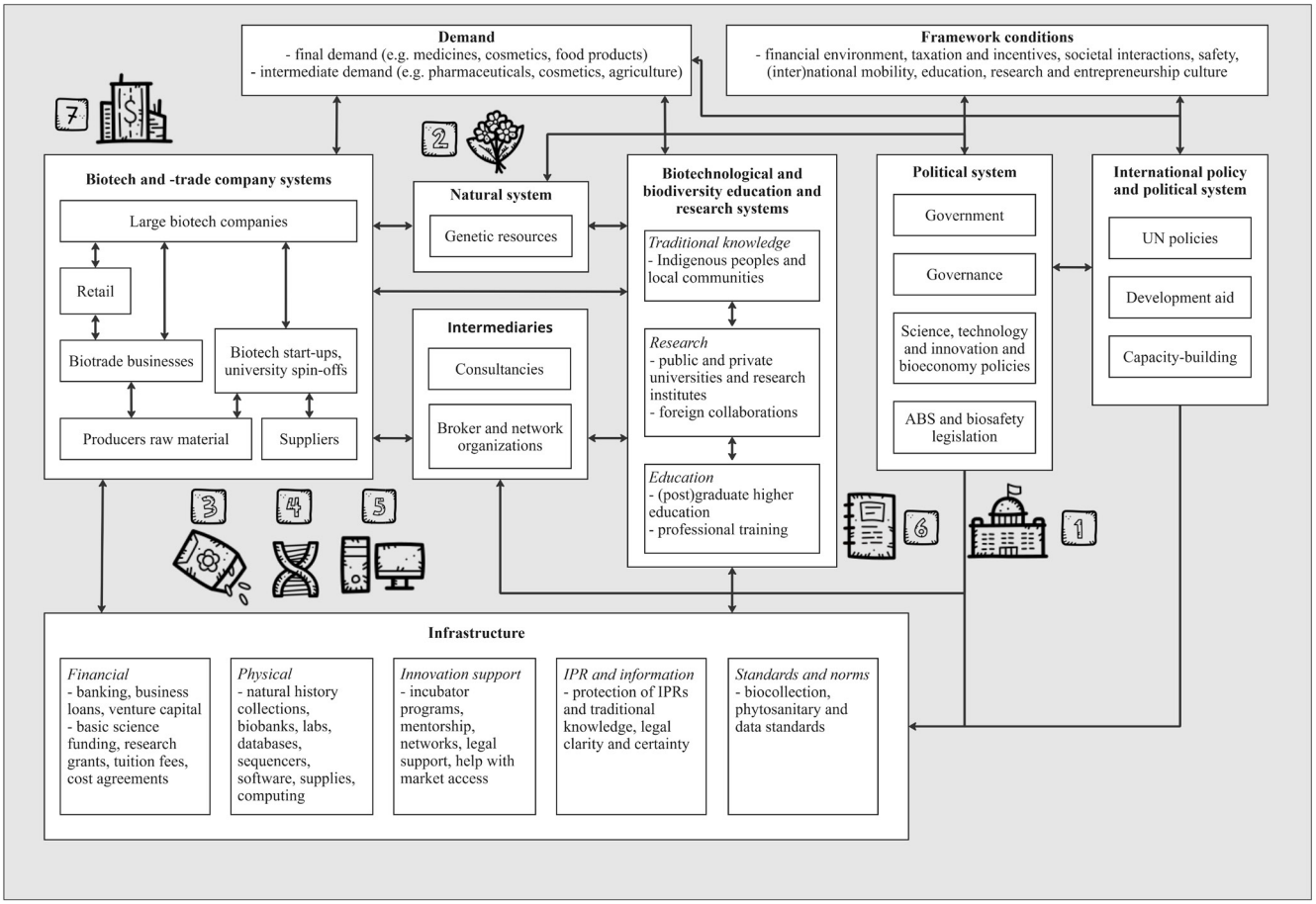


Figure 3. Overview of a nature-based biotechnological innovation system. The different aspects of the value chain of genetic resources and DSI are positioned near the related systems, and new relationships are included compared to Figure 2. The international policy and political system interacts with the political system (e.g. domestic implementation of UN policies, tax system as framework condition), demand (e.g. accessibility to new markets), and infrastructure (e.g. technology transfer and monetary benefit-sharing). The political system influences access to genetic resources and the conservation of the natural system, and potentially the availability of intermediaries (e.g. funding of incubator programmes).

Results

The results are represented in order of the key aspects of the value chain (Figure 1) and with the relevant NIS components with which they interact (Figure 3).

Policies for biodiversity and bioeconomy

NIS components: framework conditions, natural, research and political systems

Colombia has clear policies to boost its nature-based biotechnological capacity that are grounded in its natural, political, historical and socio-economic context. In between the Pacific and Atlantic oceans and divided by the Andes and Amazon, Colombia is the world's second most biodiverse country. Because of its proximity to key markets, biodiversity is described as an international "competitive advantage" (Melgarejo, 2013) and regarded as an opportunity for nationwide cultural and economic transformation (Aparicio, 2022). Minister of the Environment and COP16 Chair Susana Muhamad (resigned in February 2025) aspired to increase the share of GDP from the bioeconomy from 0.8% to 3% by 2030 (The City Paper Bogota, 2024a). Internationally, the government repeatedly associates the country with its natural wealth. In his COP16 opening speech, President Gustavo Petro referenced the words of Indigenous peoples about creating the "idea [of Colombia] as a world power of life as a national mission" (Presidencia de la República de Colombia, 2024). Petro envisions development without neoliberalism, the use of fossil fuels, or the extraction from nature, and "harmony with nature" is included in the National Development Plan 2022–2026 (Vallejo Zamudio, 2023).

The minority Afro-Colombian and Indigenous populations have endured violence, land-grabbing and subjugation to Western scientific ontologies and Christianity, first under Spanish colonization, later under Colombian governments, and recently by corporations and armed groups (Goyes & South, 2016; Chaves-Agudelo *et al*, 2015). In a "vicious cycle of biopiracy", genetic resources and traditional knowledge from these marginalized and impoverished groups are at risk of misappropriation (Goyes & South, 2016). Biopiracy, particularly DSI-enabled 'digital biopiracy', also causes national concerns, leading Susana Muhamad to call for measures to ensure "sovereignty over genetic information" (The City Paper Bogota, 2024b).

Colombia has a broad legal basis for IPR to protect innovations abroad (ProColombia, 2024), which is similar to that of other countries. Biopiracy related to genetic resources

Table 1. Overview of interviewees in Colombia (details available in Supplemental Material 1)

Type of system and work	No. of interviews
Education and research	
Animal biology	4
Botany and crop research	8
Omics and bioinformatics specialists	4
(Industrial) biotechnology and -chemistry	5
Students biochemistry (group interview)	7
Bioprospecting specialists	6
Law, ethics and human rights	6
Subtotal	40
Biotrade, -tech and -economy	
Biotrade companies	3
Biotech start-ups and spin-offs	4
Bioeconomy experts	1
Innovation broker	1
Subtotal	9
Politics and policy	
Policymaker and diplomat	1
Politician	1
Embassy worker	1
Subtotal	3
Other	
Non-governmental organization	1
Total	53

through patent acquisition is inhibited by Colombia's IPR system. Particularly important to genetic resources is Andean Decision 486 (Andean Community, 2000) on the Common Provisions on Industrial Property which contains various provisions to prevent biopiracy (Salas, 2020), most importantly: (1) the requirement in Article 3 that biological and genetic heritage and traditional knowledge underlying inventions was acquired in accordance with the law, so as not to breach provisions of Decision 391 on ABS, (2) the exclusion in Article 15 sub b of patents on biological processes and material and genomes or germplasm, and (3) a requirement in article 26 sub h to disclose an access contract in the patent filing if traditional knowledge of IPLCs was obtained.

The state of biodiversity in Colombia today is heavily influenced by the aftermath of the 2016 peace agreement between the government and the FARC guerrilla group, which has enabled increased deforestation and illicit coca production, but also biodiversity exploration (Huddart *et al*, 2022; Irwin, 2023). Without measures to curb the expansion of agriculture, a major employment sector in Colombia, biodiversity loss could accelerate by 38 to 52% by 2033 (Guerrero-Pineda *et al*, 2022). Faced with the need to conserve and simultaneously sustainably use biodiversity and the need to integrate thousands of people from previous conflict zones back into society, then President Manuel Santos reinvigorated the 2015 Colombia Bio programme, a nationwide policy agenda focusing on biodiversity research, bioprospecting, product valorization, institutional strengthening of value chains and public awareness of biodiversity (Irwin, 2023). According to both interviewees and expert institutes, Colombia Bio is internationally recognized as an exemplary bioeconomy programme. Colombia defines the bioeconomy as an "economy that efficiently and sustainably manages biodiversity and biomass to generate new products and processes with added value, based on knowledge and innovation" (Consejo Nacional de Política y Economía Social, Citation, 2018), p. 26, as translated in Johnson *et al* (2022)). The Bioeconomy Mission, a national policy launched in 2020, has five focus areas: biodiversity and ecosystem services, sustainable agricultural production, biomass and green chemistry, biointelligent Colombia, and health and well-being. Central to achieving each of these goals is boosting the use of biotechnology, omics and bioinformatics.

Access to *in situ* genetic resources

NIS components: framework conditions, research, industry, natural, political and international policy system, physical and IPR and information infrastructure

Considerable barriers to researchers' access to and collection of Colombian genetic resources are administrative, legislative burdens, customs and safety challenges. For fair access to genetic resources, a balance must be struck between user burdens and user rights (Collins *et al*, 2020). But among biodiversity researchers, Colombia's access regulation is notoriously burdensome, sometimes leading to researchers giving up a study (Fernández, 2011; Wight, 2019).

To date, Colombia has not ratified the Nagoya Protocol . The legal bases for ABS are Article 81 of the Constitution (Senado de la República de Colombia, 1991), Andean Decision 391 (Andean Community, 1996), and various subsequent decrees (Reep, 2025). Access permits are evaluated and granted by the genetic resources team in the Ministry of Environment and Sustainable Development. Users usually have to report yearly to this team and negotiate another contract in case of commercial interest. In addition, users have to comply with the National Parks Service's and other regulations for responsible and sustainable sampling. This patchwork of regulation means that to access just one sample, a foreign scientist may have to acquire seven different documents (Collins, 2019).

While most interviewees did not express concern with the objective and content of the ABS legislation, they experienced high red tape and delays related to its implementation, which affected graduate and short research projects the most. This, in turn, affects international collaborations, as exemplified by the experience of a Colombian university biologist:

> *"One time, I had my application filed already three months before a research visit of five months in Germany. After personally returning to Colombia, it took another three months before the sample could be exported to Germany, long after the visit ended."*

Interviewees report stories of researchers secretly shipping samples in their luggage to avoid delays. ABS's

adverse impacts on research are not well-received. Among interviewees, law enforcement gaps in extractive industries created the strong feeling that regulation "harms honest people while bad people continue to destroy biodiversity." Fortunately, scientific access has become easier over the past years with regulatory changes. Another promising development is that Colombian institutions are increasingly signing a Memorandum of Understanding with international collaborators to facilitate standardized access to samples.

However, red tape still looms large for commercial research, which is key to kickstarting the bioeconomy (Silvestri, 2016). Back in 2013, less than a third (27%) of all bioprospecting permits were accepted and three-quarters of applications took longer than eight months to be processed (Güiza & Bernal Camargo, 2013). The red tape and delays caused a high degree of informality, estimated at three-quarters (77%) of bioprospecting activities (Güiza & Bernal Camargo, 2013). Apparently, some companies manage the business risks posed by red tape by delaying permit applications for genetic resources with unknown or prior obvious commercial potential until after completing R&D and reaching the final investment decision stage. According to an interviewed policymaker, the major cause of the permit delays is insufficient human resources in the team to handle the requests, which are only increasing. Legal unclarity and "coordination failure between institutions" are to blame for the delays (Güiza & Bernal Camargo, 2013). But delays are also caused frequently by users who submit insufficient and inaccessible documentation. Interviewees shared that in-house legal counsel for scientists is essential to gaining permits fast and avoiding legal repercussions. This highlights how administrative burdens disproportionately affect small research institutions and companies.

Until recently, there were no specific procedures for access to genetic resources on Indigenous and Afro-Colombian lands, complicating bilateral ABS negotiations (Silvestri, 2016). The state pursued an extractivist policy for genetic resources tailored to industrial interests, while Indigenous peoples were hardly consulted (Nemogá, 2014). The Interior Ministry has to verify whether a consultation with IPLCs is necessary before the genetic resources team can grant an access permit. But history-related distrust and the self-protective attitude of IPLCs, that an interviewee describes as "¿Gano yo?" ("What do I win from this?"), in combination with legal unclarity, have made such negotiations very complicated. Regulatory changes alone will not rebuild that trust.

Sampling also involves costs for already constrained research budgets. The government Instituto Agropecuario de Colombia charges between 500 and 3,000USD for risk analysis services before seed of a species can be imported (AgriBrasilis, 2022). All subsequent importers receive a waiver, disincentivizing first-users to pay the fees for species without direct economic benefit.

The prohibitive cost or lack of cargo services is another challenge. Interviewees have experienced degradation or destruction of samples due to delays in customs, caused by personnel's distrust of equipment like nitrogen containers, and because samples were not stored in the right conditions during transit at some airports in Colombia. This has considerable negative effects because fieldwork in remote regions is often too costly to undertake twice.

A last factor of limitation in sampling is violence. While the safety situation has improved considerably since 2016, narco-trafficking is rampant in remote regions. One research team was limited to conducting research directly in the surroundings of an army base and later had to abort the project due to a deteriorating security situation. We assume that such security precautions increase fieldwork costs and limit participatory processes with IPLCs.

Conservation of *ex situ* genetic resources

NIS components: framework conditions, research, political and international policy system, financial, physical, and standards and norms infrastructure

Colombia's biocollections mostly face financial and organizational challenges, while there are opportunities to be found in increased research, education and benefit-sharing activities. *Ex situ* conservation is organized at various scales, including in botanical gardens, the Future Seeds genebank of the Centro Internacional de Agricultura Tropical (CIAT), and in four public research institutes, namely AGROSAVIA for crop research, Sinchi for Amazonian research, INVEMAR for marine and coastal research and the Humboldt Institute for nationwide biodiversity research. Universities also maintain their own, sometimes outsourced biocollections, and some IPLCs store seeds of nutritious and culturally relevant plants in community seedbanks.

Generally, interviewees reported insufficient funding for biobanking, although agricultural research receives more support than biodiversity research. Economic challenges are energy price swings and the salaries of permanent contract staff. Meanwhile, research funding is decreasing and the government's re-valuation of grants at the end of each year creates a lot of job insecurity, making it hard for institutes to retain their staff. Interviewees indicated that the decoupling of biocollection funding from research project funding would be desirable.

University biocollections experience unclear assignments of responsibilities and degrading infrastructure. Some scientists have a curator's responsibility on top of their day-to-day research tasks, leading to decreased vigilance for incidents. At one time, a researcher lost almost a complete tissue collection when a freezer thawed without raising an alarm. The lack of dedicated curators also means that access to others' samples becomes more dependent on personal favours by the researcher who facilitates access, thereby slowing down overall research. An opportunity for cost reduction lies in the centralization of biocollections and service provision.

An interviewee noted that here, again, violence is a risk. During the major unrest of the National Strike in Cali in 2021, protestors blocked off entire roads, including towards Palmira, where the Future Seeds bank is located. Only at the last moment, a truck carrying liquid nitrogen was exempted by the protestors, showing the external fragility of even the most secure collections.

Altogether, these threats to biocollections must be seen in the light of sequencing efforts, since when funding for sequencing finally becomes available, sample quality must be maintained, especially for long reads. Additionally, the informational value of DSI that becomes available through sequencing builds on the characterization work and advanced regeneration practices at biocollections. Enthusiasm for the use of next-generation technologies could bias capacity-building efforts toward sequencing and ignore current capacity deficits in *ex situ* conservation. However, both

capacities need to be strengthened to realize mutual benefits.

Interviewees also considered it important that samples in biocollections "do not just sit there" and additional measures are taken for value creation. The Humboldt Institute's seed bank in Boyacá, which has species from the whole country, for example, hired an ethnobotanist to add more value to the collection for society. The marine research institute, INVEMAR, is exploring the expansion of its natural history museum, which is currently limited to a small exposition in the wet collection, to teach the public about marine biodiversity. Apart from creating value for the public, biocollections can address concerns specific to IPLCs by helping them conserve seeds that are vulnerable to weathering in glass and mason jars with training, freezers, and seed repatriation. These ideas highlight that capacity-building activities aimed at biocollections should not only focus on conserving genetic diversity *ex situ* but also support biocollections' aspirations to maximize value for the public and stakeholders.

Generation and storage of DSI

NIS components: research, industry and political system, physical infrastructure

Contrary to 'each its own sequencer' thinking, the capacity to generate DSI in Colombia is primarily constrained by import costs of reagents, infrequent maintenance of sequencing equipment, and customs issues to import sequencing, laboratory equipment and reagents. This, among other factors, has caused significant biodiversity data gaps in Colombia. DSI is only available for one in twenty species, with the vast majority of the available data describing bacteria or being related to just a few projects (Noreña et al, 2018). To increase the availability of genome sequence data, in 2019, a new node of the Earth Biogenome Project network was founded, EBP-Colombia, which was also embedded in bioeconomy programmes like Colombia Bio (Huddart *et al*, 2022). However, there have been no updates for some years now, raising the impression that the project has been discontinued due to dried-up funding.

Despite decreasing costs, sequencing is still a costly exercise in Latin America (Noreña *et al*, 2018; Vilaça *et al*, 2024). In Colombia, it is much cheaper to ship samples abroad for sequencing. Although some institutions have sequencing capacity, others face limited or no access to these machines. Additionally, to make a purchase of such equipment cost-effective it is necessary to process a high volume of samples. Maintenance costs and delivery, as well as reagent costs, however, form the major bottleneck. There are long waiting times for maintenance workers to repair machines. Furthermore, whereas researchers in the USA can order reagents and get them delivered almost instantly, Colombian researchers have to wait for extremely lengthy periods, frequently more than a couple of months. The first cause for this delay is bureaucracy in academic institutions, which restricts purchase authorization to a small number of people. Secondly, obligated by national import regulations, researchers have to submit orders to licensed intermediaries that can import the reagents. But because there are few such intermediaries in Colombia, companies can charge higher prices, which is the main reason why reagents are two to five times as expensive as the original price in the exporting country. When the order is finally shipped, delays, damage and loss in customs are possible:

> *"A reagent for a RT-LAMP test took three months to arrive. The kit has a pH indicator, which usually is cherry red, which turns yellow with a positive test. But the kit arrived orange, meaning it can't be used anymore. We discovered that the cold chain broke during the shipping process because customs did not put it in a fridge for four days." [University biologist]*

With significant delays, technology software and service support by the sequencer vendor can become obsolete. One research team that faced more than two years of delay, therefore, renegotiated with the technology provider for a newer sequencer machine.

Conversations with interviewees suggest that to improve this part of the value chain, waste of research budget and time can be avoided by having the government shake up the intermediary market to enable researchers to access reagents more cheaply. Existing sequencing capacity can be used more efficiently if institutions advertise and rent or centralize their sequencers to achieve cost reduction. Greater sequencing capacity could come from investments in new businesses that produce reagents or provide maintenance in the Latin American region.

A recommendation to increase the availability of DSI from remote regions of Colombia is for research institutions to set up collaborations with businesses that collect biodiversity data through environmental DNA (eDNA) and other techniques for conducting environmental impact assessments. Although large-scale storage of DSI is organized by genetic databases in other countries, institutions may need local data servers and portals for digital genebanking and pre-analyses. Here, negotiations with software and cloud providers form an opportunity for cost reduction.

A mentality shift in biological research is probably also needed. According to an interviewee, instead of the "catch them all" mindset in biodiversity sequencing, researchers could perhaps better focus on collaborations on the generation of fewer high-quality DSI. Likewise, countries with overlapping biodiversity may avoid duplicate work and achieve scale benefits by forming regional collaborations wherein sampling, *ex situ* conservation, and the generation and storage of DSI are coordinated across borders. In Biodiversity Genomics Europe, for instance, tasks, resources, lessons learned and capacities gained along the genomic pipeline are distributed over institutions in different countries. Because diversity in ABS regulations may pose an issue, Colombia could best collaborate inside the Andean Community with Bolivia, Ecuador and Peru, as each country's ABS legislation builds on Decision 391 (Ljungqvist *et al*, 2025). Lastly, through international collaborations with producers of sequencing equipment, which are under the scope of the Cali Fund (CBD, 2024), researchers in LMIC can gain access to grants for generating DSI (PacBio, 2020).

Publication of and access to DSI

NIS components: education, research and international policy system, financial infrastructure

While challenges for LMIC scientists are reported with regard to data access and compliance with FAIR data standards in the literature (Shanahan & Bezuidenhout, 2022; Bezuidenhout *et al*, 2017), no particular issues were reported by the Colombian interviewees. When new biodiversity data standards are adopted, scientists require both capacity

building and efforts to demonstrate their benefits – an experience that is common across countries. There are, however, significant challenges with regard to publishing DSI that originated from IPLC territories. Their unfamiliarity with DSI and distrust make initial conversations between them and scientists difficult. There are no standards yet for respecting traditional knowledge and maintaining best practices for intellectual property. IPLCs also reject unrestricted open data:

> *"When we have spoken with them about data sharing, they usually say: "We need some specific rules and safeguards about publishing the DNA data and how we will be reflected in these publications" They have their own needs that we, as scientists and data managers, need to meet." [Biodiversity data specialist]*

The alignment of data-driven research with community needs is recognized as an ethical priority by the C3Biodiversidad consortium (C3Biodiversidad, 2018a). Lessons to navigate this engagement are found in a collaboration of the Humboldt Institute and the organization Wise Ancestors with a Paisa community in Antioquia to produce two reference genomes for two critically endangered birds (Wise Ancestors, 2024). Five local collaborators received one year of salary and a broad training in sampling, bioinformatics and biomonitoring while Wise Ancestors guided them in genomics-based conservation management actions. Because the generation of a genome sequence as a research outcome is not directly relevant to the community, the project also worked on developing ecotourism as an alternative livelihood, and helped the collaborators to cultivate the edible mortiño berries (*Vaccinium myrtillus*), which benefit both the Antioquia Brushfinch (*Atlapetes blancae*), also called the 'Montañerito Paisa', and the community. Increasing the social value of genome sequencing required a new mindset:

> *"I think in the long run, with these social benefits, resources are better spent because the project is filling a community need that helps them to conserve nature. There is a growing detachment between what people learn in universities and the needs of Colombia. Studying the gene for a bird to be blue or yellow has very important scientific value, but that kind of information may not be in Colombia's list of highest priorities." [Gustavo A. Bravo, Curator of Ornithology at Instituto Alexander von Humboldt]*

For this change of mindset, university curricula have to include more ethics and responsible business conduct. It also necessitates a reorganization of research funding because, in a project with a strong participatory component like this, a sequence can be five times as expensive to obtain. Those costs illustrate that while scientists in the DSI discussions managed to be exempted from monetary benefit-sharing, the organization of research funding is indeed related to financial benefits for communities of interest. Yet, it is so far not clear to the project leaders how these benefits can be sustained over time because a genome publication is a one-time event and scientific project funding will dry up at some point. That means that sequencing projects by themselves are inadequate sources of funding compared to standard conservation funding.

DSI and genetic resources use in science

NIS components: framework conditions, education, research, industry, political and international policy systems, financial and physical infrastructure

In Colombia, advanced scientific capacity is held back by structural underfunding, brain drain, a lack of cyberinfrastructure, and inconsistent and increasingly short research grants. Opportunities arise mainly in public–private research partnerships and international scientific collaboration. Interviewees indicated that there already is advanced scientific knowledge, lab quality and technological development in Colombia. However, relative to comparable countries in the region, it has notably fewer biotechnological and bioinformatics publications (Benitez-Paez, 2010; Martinez *et al*, 2014). This may be explained by the following factors.

The first reason is related to education. In higher education, less than 10% of students follow bioeconomy-related disciplines (Alviar *et al*, 2021). For university graduates, there are few research positions, to such an extent that fewer are enrolling in PhD programmes, and many emigrate. Some researchers with a strong motivation to give back to the country eventually return, but their patience gets exhausted too. This poor economic perspective is also experienced by employed researchers, exemplified by a story from a public university scientist:

> *"In just one year, all three of my PhD students left because they were not being paid by the university as contractually promised. No research was done. Many professors tell me that the only students who end up graduating are the ones who are independently wealthy or those with scholarships." [University biotechnologist]*

This outflow of trained personnel, in combination with short-term research grants, places a burden on principal investigators who have to repeatedly train new researchers.

Another limiting factor is the cyberinfrastructure. Colombia has sufficient developers, administrators and bioinformatics expertise to run such infrastructure, but the availability of high-performance computing, computational training resources and data storage is limited (De Vega *et al*, 2020; C3Biodiversidad, 2018b). In other words, the strong human resource component is suppressed by a weak physical resource component (Figure 3).

Both interviewees and C3Biodiversidad (2018b) indicated that insufficient and unstable investment in R&D is the major challenge to knowledge production. Between 2000 and 2020, total R&D expenditure as a share of GDP in Colombia was, on average, 0.22% (World Bank, 2024). Although not uncommon in Latin America, that is very low when compared to regional leader Brazil, with 1.11%, and the OECD average of 2.40% over that same period. Meanwhile, under the current government of Gustavo Petro, the budget for higher education is declining by one quarter in real terms, even though the ambition was to increase R&D expenditure to 0.5% of GDP by 2026 (Fernández, 2023). As research grants decrease in size, competition among researchers increases, and the scope of projects becomes more limited. Structural underfunding and swings in government budgets particularly affect public universities, which receive less funding through tuition fees and private investment than private universities.

Interestingly, the tendency of new governments to tie science funding to political themes, thereby limiting research groups' consistency, has had a knock-on effect on public universities' collaborations with companies that prefer long-term and stable research relationships. Additionally, funding often does not arrive in time. Ten percent of the Sistema General de Regalías, a system for the distribution of royalties from industry, is invested in science. Yet only half of the budget was delivered during its first year of operation, raising suspicions of corruption with interviewees (Organisation for Economic Co-operation and Development (2014), p. 118). Irregularities and lost resources are recurrent to this day.

Opportunities mainly appear in public–private collaborations. The share of business expenditure in R&D (BERD) in Colombia in 2020 stood at 0.15% compared with the OECD median of 1.15%. Four-fifths of businesses in Colombia are reported to not invest in R&D (DANE (2023), as cited in Organisation for Economic Co-operation and Development (2024a)). BERD can complement university funding when basic research is co-financed by applied research. For example, in the Icesi Sustainable Industries and Applied Science Labs in Cali, one of the most advanced in Latin America, companies hire university researchers to design, validate or prototype bioprocess tests. EAFIT, a private university in Medellín, secured high-performance computing, known as the Apolo platform, in partnership with Purdue University, which now enables them to conduct paid services for companies. Hybrid positions wherein a researcher or student works part-time in a company are also smart from both a science funding and valorization standpoint. Despite these synergies, public–private research collaborations can have trade-offs, including the skewing of the research agenda toward commercial applications over basic research that can benefit the conservation and sustainable use of biodiversity, and a possible misalignment between the private sector's need to maintain a quick pace and patience-requiring participatory processes with stakeholders.

There are also improvements to be made in international science policies. Slightly less than half (43.5%) of Colombian scientists experienced feelings of language-based discrimination in article revisions and rejections (Ramírez-Castañeda, 2020). The country has low English literacy, and translating takes up a lot of research time and budget. High fees for open-access publishing can also be a barrier to publishing in high-impact and high-visibility journals.

Bilateral science diplomacy can help to enhance international scientific collaborations without the need for extra funding. The United Kingdom embassy in Bogotá contributed to the creation of BRIDGE Colombia, a network of Colombian and British scientists. Academics in BRIDGE Colombia then secured funding from the Global Challenges Research Fund for a flagship project called GROW Colombia, which aims to boost the country's bioeconomic innovation capacity. Bilateral science diplomacy can also help to flag science issues directly to governments or agencies and businesses. Feedback from scientists on the prohibitive cost of sequencing in Colombia, for example, was communicated through the embassy to British technology providers.

> *"Science diplomacy helps to connect UK science with Colombian priorities. There was a time when the UK came to 'teach' Colombian scientists. That is over now as a result of their joint research work. The bilateral collaboration operates now under a logic of equitable research partnerships. That is possible because Colombian science has been advancing to a point at which scientists from both countries work on a peer-to-peer basis." [Luis Calzadilla, Head of Science and Innovation at UK Embassy in Bogota]*

This facilitative role in accessing funding and enhancing equity in research is very important. Interviewees remarked that they increasingly have to find non-Colombian collaborators to be able to apply for grants abroad, which feels awkward and exploitative. In the opposite direction, it was felt by an interviewee that some foreigners just seek collaboration with Colombian scientists to access samples, without leaving meaningful work on the publication for their counterparts. This type of sample-focused "helicopter research" in collaborations with Colombian institutions is already reported in human genomics (Cock-Rada & Gomez, 2018).

Industrial use of genetic resources and DSI

NIS components: framework conditions, demand, education, research, industry, political and international policy systems, intermediaries, financial and IPR and information infrastructure

There is a major valorization gap in the country that is caused by a lack of entrepreneurship support and culture, investment unavailability, administrative delays, export challenges and tax burdens. Colombia has fewer large bioprospecting centres and biotechnology-based companies than similar countries in the region (Bueno & Ritoré, 2019). The country seems to suffer from "an absence of scientific governance" to direct R&D toward commercialization and toward knowledge gaps in genetic diversity conservation (Bueno & Ritoré, 2019).

Once university scientists find bioprospecting value, they struggle to valorize the findings. Marine botanist Enrique Peña at Universidad del Valle in Cali initiated studies for the application of two species of invasive red algae (*Sargassum fluitans* and *Sargassum natans*) that pollute the tourist beaches of San Andrés Island in the Caribbean as a feedstock for fertilizer production. Building a pilot plant on campus would cost approximately 1 million USD. Although Peña has already attracted interest from companies, domestically and abroad, funding or investment to initiate building the plant is still insufficient. Peña noted that an entrepreneurial mindset is still uncommon in public universities.

Private universities, once founded by business leaders, offer more support for entrepreneurship and valorization, such as business and finance courses in life sciences programmes, incubators, and mentorship programmes. Sciphage is a biotechnological start-up with its own R&D and some patents from research at Universidad de los Andes. It develops phage therapy for the livestock sector, an alternative method to antibiotics that has various health and environmental benefits (Mishra *et al*, 2024). Sciphage has already built a production plant outside Bogotá and is seeking investment to

scale up production. The Pontificia Universidad Javeriana has another model where it creates spin-offs with its own patents. Dreembio is one spin-off that develops phytomedicines for cancer treatment based on ethnobotanical plant knowledge. The company is working with campesino or farmer communities to develop the raw material value chain.

These companies clearly show that innovation can have socio-environmental benefits. According to Arturo Luna, the former Minister of Science, Technology and Innovation, while the bioeconomy does require advanced biotechnology, it can include low-tech businesses too:

> *"There are already low-tech businesses in Colombia that add value to biodiversity and fit perfectly with the bioeconomy model. We have to invest in these businesses because this virtuous cycle will generate employment and financial resources and opportunities for biodiversity conservation. However, in order to add value to our biodiversity, more and sustained investments in biotechnology are required." [Arturo Luna, freelance]*

Take, for example, Kahai S.A. It managed to produce the jungle cacay tree (*Caryodendron orinocense*) on plantations. Because the nut's oil outperforms argan oil in some cosmetic applications, there is an enormous biotrade potential. To collaborate with international development agencies and have a better investment case, the company built a strong corporate social responsibility (CSR) component, including a reforestation programme and inclusion of IPLCs.

These three home-grown companies represent textbook examples of successful bioprospecting and biotrade. But they face numerous challenges, which are illustrative of the kinds of issues that limit the development of genetic resources and DSI in the natural products industry.

First, there is an investment gap in Colombia. Regular investors, like banks, lack advanced knowledge of biotrade and biotech. Throughout the country, there is an over-demanding investment culture with sometimes "aggressive pushing for unrealistic targets" and requests for proof of traction and several letters of intent from buyers before pre-seed or seed funding can be acquired. In other countries, interviewees experienced more eager investors who understand that break-even is far away and that not all investees will succeed.

The second major challenge relates to the issuing of sales permits by Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA), the government agency that approves products for human and animal health. Biotech companies experience the licensing process as a severe regulatory obstacle, which is attributed to the agency's perceived unfamiliarity with novel technologies, insufficient capacity and personal hesitancy to take risks to avoid liability. Biotrade companies also struggle, primarily with delays. Both Kahai S.A. and Pangea Natural Products, a company that sells medicinal herbal supplements, waited roughly five years for approval of their product, though the hefty permit costs of over 3,000USD were already paid. For starting companies, these costs are high, and delays can mean 'make it or break it.'

Another hurdle is the lack of domestic demand for natural products based on scientific research:

> *"Demand creation is a cultural process where you teach about the environment and the values of the product, but also [the positive] impacts for communities." [Legal expert]*

Colombia has the character of a 'follower market' and its consumer culture requires that companies first demonstrate product success abroad. While adapting to customer needs, companies struggle with strict import regulations such as those imposed by the EU. Network brokers, like local chambers of commerce, have an important role in helping companies access markets. Reducing global conference fees would help companies gain exposure to investors and clients.

The fourth challenge is Colombia's tax and fiscal policy, which, according to interviewees, stifles the growth of their companies and makes them strongly consider commercializing abroad. Compared with an average of 23.7%, Colombia's 35% corporate income tax rate is the highest of all OECD countries (Organisation for Economic Co-operation and Development (2024b), p. 20-26). Although the 19% value-added tax is close to the OECD average, it is still among the highest in Latin America. A policy proposal for a phased decrease in tax rates for small and large businesses is being debated in parliament.

Because the results show that universities strategically use patents for founding startups and spin-offs, it is also important to discuss the challenges and opportunities for the IPR system (described earlier in section 'Policies for biodiversity and bioeconomy'). In 2024, the Colombian government issued a compulsory license to start producing generic and more affordable versions of the drug Dolutegravir for tackling an emerging HIV crisis. This type of action, however, could have negative effects such as diplomatic repercussions, arbitration or decreased market entry by foreign drug companies (Landis, 2024). Such risks may be lower if Colombia can source more products from domestic companies. But here again, the country's funding landscape will need drastic adjustment to guide startups with acquired patents through the financially challenging Valley of Death phase between R&D and revenue generation.

On the other hand, there are opportunities at a national and international level. To begin with, universities should continue investing in their technology transfer and innovation offices that support researchers during R&D and connect them to industry. On top of that, universities can cover the costs of researchers' patent applications and create room in work schedules for entrepreneurship (Calza *et al*, 2020). At a national level, the government can implement patent pilot programmes to review the efficacy of patent dispute resolution by courts (Salas, 2020). Bilateral diplomacy related to IP can also facilitate the protection and export of Colombian products and technology. One tool is the Patent Prosecution Highway, which countries use to fast-track patent applications by companies that acquired patents in another country. To stimulate bilateral technology transfer between Colombia and countries with divergent IPR systems, it can develop mutual transfer agreements that contain Colombia's ABS provisions but are flexible enough to incorporate the other country's priorities (Fajardo *et al*, 2025).

There are also unresolved tensions in the IPR system with regard to the protection and empowerment of IPLCs,

caused by CBD decisions that reinforced national sovereignty over genetic resources to the detriment of Indigenous self-governance (Fredriksson, 2019). Martha Gomez Lee, an ABS scholar, argues that the self-governance of IPLCs can be promoted by embedding article 31 of the UN Declaration on the Rights of Indigenous Peoples in the ABS system, which states: "… They also have the right to maintain, control, protect and develop their intellectual property over such cultural heritage, traditional knowledge, and traditional cultural expressions." Empowerment of IPLCs during the development of intellectual property may carry implications for the governance of DSI:

> *"In my opinion, the crux of the matter is that even before intellectual property rights, any initial digitization of a genetic resource and its deposit in any database must have explicit Prior Informed Consent." [Martha Gomez Lee, teacher-researcher in ABS and traditional knowledge at Universidad Externado de Colombia]*

Finally, it is worth noting here that companies think that working with campesinos is much easier than with IPLCs because the former are generally better organized (so engagement has a faster pace) and the latter have complex cosmovisions to engage with. Also, companies can use publicly available scientific knowledge on the function of species that may have been discovered by IPLCs in the pre-ABS era. This perhaps explains why there are, to date (and as confirmed by Martha Gomez Lee), zero signed annexes by providers of traditional knowledge as an intangible component to genetic resources, although the latest figures show that around 20 commercial ABS contracts and 400 ABS contracts overall have been concluded with the Ministry (Ministerio de Ambiente y Desarrollo Sostenible, 2021).

Discussion

Despite Colombia's large bioeconomic potential, many structural challenges remain. In Figure 4, we present the current issues (in text boxes) and provide recommendations (in italics). The majority of issues relate to underfunding, investment and regulatory capacity, and red tape. Existing research capacity is not used efficiently because of siloed R&D, high equipment and maintenance costs, and outflow of trained personnel. Moreover, the companies with bioeconomic potential that do emerge struggle to grow inside Colombia. Without considering these domestic policy issues, monetary payments, scientific capacity building and technology transfer as typical ABS tools, will have little effect on achieving that potential. Here, we reflect on what could be accomplished by Colombia itself, how specific needs could be addressed by ABS policies, the Cali Fund and non-monetary benefit-sharing from DSI, and lastly, the value of innovation system models for both holistic and targeted policy interventions.

Although subsequent governments have expressed strong bioeconomic intent, policies are being discontinued, and there is a coordination failure between government bodies. This aligns with Aparicio's (2022) finding that Colombia's narrative is focused on future high-tech-driven biodiversity exploration as a precondition for the bioeconomy, while questions over the current performance of value chains are pushed aside. On the other hand, the findings showed that the fostering of cultural pride in home-grown science and companies may be the next step for growing national interest in biodiversity and bioprospecting.

Colombia's combination of underinvestment and high taxation further lowers the creation of and erodes benefits from genetic diversity. This shows that bioeconomy policy should expand its scope from technology and biodiversity to bureaucracy, government capacity, science funding, factors for business growth and retention of companies. Immediate priorities to tackle are the delays in various permits and the degradation of biocollections. In our opinion, the bureaucracy, particularly, is a 'talent grinder' for people who intentionally do research or business for the benefit of the country. Greater engagement of scientists and entrepreneurs in policy-making that concerns them is recommended in order to help address their needs and decrease red tape. Mobilizing intermediaries and network organizations would be useful for this purpose and for enabling public–private partnerships that help scientists access new funding sources (Figure 4). Also, the export focus of companies due to a lack of domestic demand leads to missed opportunities to address local needs through valuable products, retain jobs and tax revenue. The government should therefore stimulate more demand for biotech and biotrade products and services (Weerasinghe *et al*, 2024), and, although very complex, incentivize private sector alignment with green rather than fossil-fuel biotech pathways:

> *"With some companies, there is a mentality that when you provide jobs, you already do enough. But business conduct is not per se sustainable and not per se benefit-sharing." [Ana Maria Castillo, Director of Competitiveness and Internationalization at the Cali Chamber of Commerce]*

The analysis has various implications that (Non-)Parties and observers to the CBD must consider. First, they should engage potential beneficiaries of the Cali Fund, like Colombia, to critically evaluate the effects of various policies on the full value chain of genetic resources and DSI. To enhance synergies between the political system and international policy and political system (Figure 4), (Non-)Parties to the CBD can link the biodiversity-focused negotiations to bilateral and multilateral diplomatic processes over scientific, technological and industrial cooperation, IP and technology transfer, taxation, development aid, and customs. Furthermore, the NIS model and similar approaches can help to diagnose issues and identify bottlenecks in countries, and steer a more effective and efficient use of monetary benefits from the Fund. For instance, it is smarter to invest in low-hanging fruits and missing cogs, such as public–private partnerships and reagent cost reduction for sequencing than in new sequencers. Additionally, investing in biodiversity-positive and IPLC-inclusive companies may have a higher return on investment in the long term than grants for short-term conservation projects. Ultimately, (Non-)Parties and observers to the CBD have to look further than physical infrastructural (Figure 4) capacity needs related to "generate, access, use, analyse and store" DSI (CBD, 2024). To embed this holistic perspective in ABS policy, Parties to the CBD are recommended to incorporate standard language related to innovation systems in the negotiation documents on DSI and

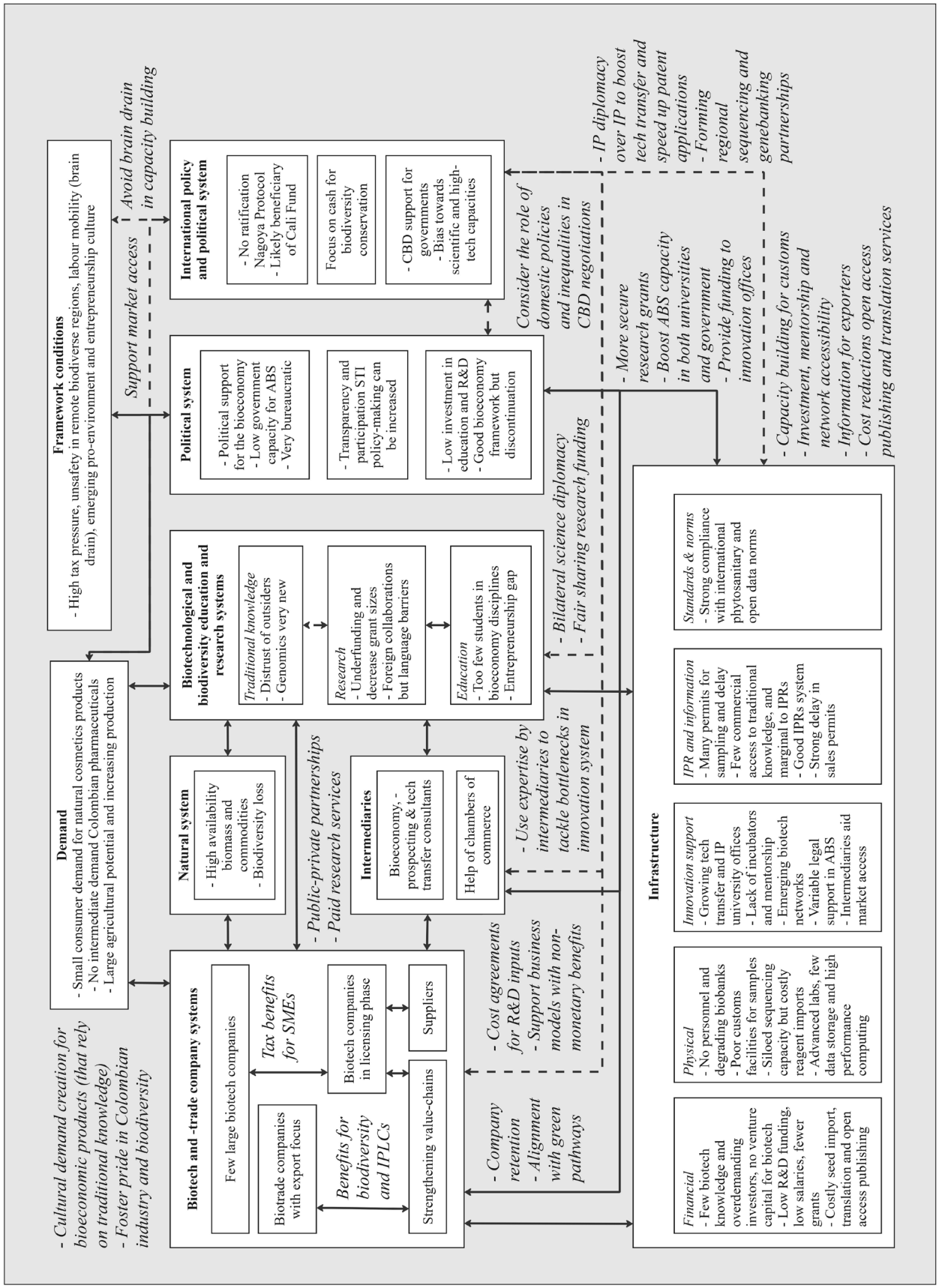


Figure 4. Identified challenges in Colombia's nature-based biotechnological innovation system with areas for improvement presented in italics. Acronyms: small and medium enterprises (SMEs); research and development (R&D); science, technology and innovation (STI); intellectual property (IP); intellectual property rights (IPRs).

ABS, for example, 'take into account and strengthen countries' capacity to innovate with genetic resources and DSI'.

The decision to distribute the Fund's money via national biodiversity funds instead of project-based applications deserves attention, too. Some companies' patent royalties that have been paid to the Ministry of the Environment and Sustainable Development through ABS mechanisms have not been distributed yet. Thus, governments' capacity to distribute funds and IPLCs' capacity to receive them deserve immediate attention to maintain company engagement and minimize overhead costs. Also, because many beneficiary countries are unequal (Colombia's Gini coefficient is among the highest in the world (Organisation for Economic Co-operation and Development, 2024a), care is warranted so that benefits are directed to the most disadvantaged actors in beneficiary countries. Apart from IPLCs, that would be female researchers (Paz & Pardo-Díaz, 2024), particularly in public universities in the underdeveloped regions of Colombia.

Seeing these issues, new opportunities for benefit-sharing emerge. Even though research and academia remain, rightly so, 'off the hook' from monetary benefit-sharing, their handling of resources certainly involves distributive questions. ABS policy should coordinate more with international efforts to map and decrease costs for LMIC scientists, such as open-access publishing and conference fees. And, as the Wise Ancestors and BRIDGE projects illustrate, North-South scientific collaborations and capacity-building projects may consider the fair allocation of research funding and joint application for research grants in HIC.

Although unconventional, there is also a big potential for non-monetary benefits from companies. To understand this, we draw on the concept of political corporate social responsibility (CSR). It holds that in the era of globalization, countries, especially developing ones, cannot fully regulate business conduct and that CSR activities have an increasingly political nature. This not only manifests itself in voluntary actions and initiatives for self-regulation but also in collaborations with governments to fill governance gaps or to provide public goods like science (Azizi, 2020; Frynas & Stephens, 2015). The COP16 decision reflects a narrow sense of justice in exchange by stating that users who make a payment "are considered to have fairly and equitably shared monetary benefits" (CBD (2024), para 15). But do benefits have to necessarily take the form of tick-the-box payments? Our findings show a clear need for companies in LMIC to access investment and loans, and mentorship in R&D, company leadership, market access and regulatory affairs. Eligible companies for payment to the Fund could play a huge role here and, in some cases, pursue mutual interests. For this, the CBD could also undertake a networking function to match companies between HIC and LMIC as a form of development aid (an example is PUM in the Netherlands), thereby allowing for better alignment of R&D than general capacity-building projects. Parties could consider legitimizing existing or new activities as monetary benefit-sharing. The risk is that the functioning of the Fund for biodiversity finance is undermined, and that the cost figures of the activities could potentially be skewed to meet the 1% of profit or 0.1% of revenue mark. In any case, initiatives should build and not erode trust in the UN and recipient governments. Analysts using the NIS approach may include an international industry system to guide these interactions, especially for analyzing cross-border industries.

On the nature-based biotechnological NIS (Figure 3), we note that the wealth of findings and the wide scope of the study limit the discussion of specific interactions between systems or actors in detail. While the article adopts a nationwide lens, many interactions happen on a local or regional level, especially in innovation hubs such as Medellín. A closer look may reveal regionally different interactions and systems, and possibly inequalities. Scholars using the NIS approach may reconsider the marginal position of traditional knowledge by including it under a separate system. It is recommended that their and civil society's perspectives be included more strongly in further research, although an effort was made to secure a balanced selection of interviewees (Table 1). We emphasize that because of the approach, we highlighted issues while the situation is not black and white. Many issues are also found in other countries, including HIC, and Colombia, as a case study, is likely not unique in that respect. More comprehensive studies and texts would allow for more comparison and insights into strengths and weaknesses. It is important to tailor the NIS to the unique context of the country of analysis, since the promotion of 'one-size-fits-all' ideas for development can disregard other ways of learning (Casadella & Tahi, 2023).

Conclusion

Facing accelerating technological progress and growing innovation divides, the CBD stands at a crossroads. Will it finally consider the fair distribution of scientific and innovation capacities and thereby put shared benefits to use effectively? Clearly, a business-as-usual continuation of capacity-building and benefit-sharing activities is unfit for current and emerging biotechnological innovation trajectories. The global political push to share benefits from DSI, therefore, has to be coupled with concerted efforts for policy reform in beneficiary countries, in ABS, and unconventionally in other policy domains. Holistic policy interventions, by principle, require governments to consult those who create benefits, educators, scientists, CEOs, and, of course, IPLC representatives. But perhaps not all can or should be solved by governments. Governance gaps invite country-specific contributions by users of genetic diversity that match technological niches and satisfy mutual interests. We eagerly await new research on innovation systems in the context of bioprospecting, so that policymakers can learn from countries' best practices and take advantage of regional opportunities.

Supplemental data

Supplemental Material 1. List of interviewees
Supplemental Material 2. Resumen en español (Spanish abstract)

Author contributions

Bob Kreiken: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing

Lotte Asveld: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – review & editing

Conflict of interest statement

The authors report no conflict of interest.

Ethics statement

The methods in this research were approved by the Human Research Ethics Committee of Delft University of Technology, the Netherlands.

Acknowledgements

We would like to thank all interviewees very much for their time and hospitality, and in particular, Alexandra Reep, Andrew Crawford, Catalina Arteaga Florez, Javier Gomez, Juliana Castro Londoño and Martha Cepeda for their outstanding contributions during the fieldwork. We thank Andrea Gammon and colleagues of the Biotech and Society section at Delft University of Technology for their helpful comments on the article.

References

AgriBrasilis (2022). Colombia Depends on Seed Imports, but Faces Regulatory Difficulties. https://agribrasilis.com/2022/09/21/colombia-depends-on-seed-imports-and-faces-regulatory-difficulties/

Alviar, M., García-Suaza, A., Ramírez-Gómez, L., Villegas-Velásquez, S. (2021). Measuring the Contribution of the Bioeconomy: The Case of Colombia and Antioquia. *Sustainability* 13(4), 2353. https://doi.org/10.3390/su13042353

Andean Community. (1996). Decision No. 391 of the Commission of the Andean Community Establishing the Common Regime on Access to Genetic Resources. https://www.wipo.int/wipolex/en/legislation/details/9446

Andean Community. (2000). Decision No. 486: Common Industrial Property Regime. World Intellectual Property Organization. https://www.wipo.int/wipolex/en/legislation/details/9451

Aparicio, A. (2022). The road ahead: Narratives and imaginaries of the value of biodiversity in shaping bioeconomy policy in Colombia. *Tapuya: Latin American Science, Technology and Society* 5(1), 2059137. https://doi.org/10.1080/25729861.2022.2059137

Azizi, S. (2020). Political Corporate Social Responsibility (CSR), Development, and Business Legitimacy. In Handbook of Business Legitimacy, ed. J. D. Rendtorff (Springer International Publishing), 1295–1308. https://doi.org/10.1007/978-3-030-14622-1_69

Bagley, M. A. (2021). "Just" Sharing: The Virtues of Digital Sequence Information Benefit-Sharing for the Common Good. *Harvard International Law Journal*, 63(1). https://journals.law.harvard.edu/ilj/wp-content/uploads/sites/84/HLI101_crop-3.pdf

Benítez-Páez, A., Cárdenas-Brito, S. (2010). [Bioinformatics in Colombia: State of the art and perspectives]. *Biomedica: Revista Del Instituto Nacional De Salud* 30(2), 170–177.

Bezuidenhout, L. M., Leonelli, S., Kelly, A. H., Rappert, B. (2017). Beyond the digital divide: Towards a situated approach to open data. *Science and Public Policy* 44(4), 464–475. https://doi.org/10.1093/scipol/scw036

Brás, G. R., Robaina, M. (2024). National innovation systems and sustainable environmental performance: A cross country analysis. *Environmental Challenges* 16, 100978. https://doi.org/10.1016/j.envc.2024.100978

Bruynseels, K. (2020). When nature goes digital: Routes for responsible innovation. *Journal of Responsible Innovation* 7(3), Article 3. https://doi.org/10.1080/23299460.2020.1771144

Bueno, J., Ritoré, S. (2019). Bioprospecting Model for a New Colombia Drug Discovery Initiative in the Pharmaceutical Industry. In Analysis of Science, Technology, and Innovation in Emerging Economies, ed. C. I. Pardo Martínez, A. Cotte Poveda, S. P. Fletscher Moreno (Springer International Publishing), 37–63. https://doi.org/10.1007/978-3-030-13578-2_3

Calza, F., Ferretti, M., Panetti, E., & Parmentola, A. (2021). Moving drug discoveries beyond the valley of death: The role of innovation ecosystems. *European Journal of Innovation Management* 24(4), 1184–1209. https://doi.org/10.1108/EJIM-11-2019-0342

Casadella, V., Tahi, S. (2023). National Innovation Systems in Low-Income and Middle-Income Countries: Re-evaluation of Indicators and Lessons for a Learning Economy in Senegal. *Journal of the Knowledge Economy* 14(3), 2107–2137. https://doi.org/10.1007/s13132-022-00945-8

CBD. (2024). Article 18 of Annex Decision 16/2. Secretariat of the Convention on Biological Diversity. https://www.cbd.int/doc/decisions/cop-16/cop-16-dec-02-en.pdf

Chaturvedi, S. (2005). Evolving a National System of Biotechnology Innovation: Some Evidence from Singapore. *Science, Technology and Society* 10(1), 105–127. https://doi.org/10.1177/097172180401000106

Chaves-Agudelo, J. M., Batterbury, S. P. J., Beilin, R. (2015). "We Live From Mother Nature": Neoliberal Globalization, Commodification, the "War on Drugs," and Biodiversity in Colombia Since the 1990s. *Sage Open* 5(3), 2158244015596792. https://doi.org/10.1177/2158244015596792

Cock-Rada, A. M., & Ossa Gomez, C. A. (2018). Leveraging International Collaborations to Advance Genomic Medicine in Colombia. In Genomic Medicine in Emerging Economies (pp. 49–69). Elsevier. https://doi.org/10.1016/B978-0-12-811531-2.00009-6

Collins, E. (2019, November 15). Hello from Colombia! Experiment. https://experiment.com/u/tuef5A (Accessed: 28 February, 2025).

Collins, J. E., Sirakaya, A., Vanagt, T., Huys, I. (2020). Developing a Methodology to Balance Benefit-Sharing: Application in the Context of Biodiversity Beyond National Jurisdiction. *Genetic Resources* 1(1), 24–39. https://doi.org/10.46265/genresj.2020.1.24-39

Consejo Nacional de Política Económica y Social. (2018). Política de Crecimiento Verde (CONPES 3934). Departamento Nacional de Planeación. https://colaboracion.dnp.gov.co/cdt/conpes/econ%C3%B3micos/3934.pdf

C3Biodiversidad. (2018a). Policy Briefing: Developing a research cyberinfrastructure in Colombia [Policy Brief]. Bridge Colombia.

C3Biodiversidad. (2018b). Developing a research cyberinfrastructure in Colombia [Policy Brief]. Bridge Colombia.

De Jonge, B., Korthals, M. (2006). Vicissitudes of benefit sharing of crop genetic resources: downstream and upstream. *Developing World Bioethics* 0(0), 061004045703001. https://doi.org/10.1111/j.1471-8847.2006.00167.x

De Vega, J. J., Davey, R. P., Duitama, J., Escobar, D., Cristancho-Ardila, M. A., Etherington, G. J., Minotto, A., Arenas-Suarez, N. E., Pineda-Cardenas, J. D., Correa-

Alvarez, J., Camargo Rodriguez, A. V., Haerty, W., Mallarino-Robayo, J. P., Barreto-Hernandez, E., Muñoz-Torres, M., Fernandez-Fuentes, N., Di Palma, F., the Colombian Cyberinfrastructure Consortium for Biodiversity. (2020). Colombia's cyberinfrastructure for biodiversity: Building data infrastructure in emerging countries to foster socioeconomic growth. *Plants, People, Planet* 2(3), 229–236. https://doi.org/10.1002/ppp3.10086

DANE. (2023). Boletín Técnico: Encuesta de Inversión en I+D 2021.

Deplazes-Zemp, A. (2019). A global biodiversity fund to implement distributive justice for genetic resources. *Developing World Bioethics* 19(4), 235–244. https://doi.org/10.1111/dewb.12230

Dunshirn, P., & Zhivkoplias, E. (2024). Conducting marine genetic research for whom? Mapping knowledge flows from science to patents. Npj Ocean Sustainability, 3(1), 50. https://doi.org/10.1038/s44183-024-00088-0

Egbetokun, A., Oluwadare, A. J., Ajao, B. F., Jegede, O. O. (2017). Innovation systems research: An agenda for developing countries. *Journal of Open Innovation: Technology, Market, and Complexity* 3(4), 1–16. https://doi.org/10.1186/s40852-017-0076-x

Fajardo, O., Dorado, F., & Lora, A. (2025). The Potential for Bioeconomy and Biotechnology Transfer and Collaboration Between Colombia and China. *Sustainability* 17(11), 5083. https://doi.org/10.3390/su17115083

Fernandes, A. J. C., Rodrigues, R. G., Ferreira, J. J. (2022). National innovation systems and sustainability: What is the role of the environmental dimension? *Journal of Cleaner Production* 347, 131164. https://doi.org/10.1016/j.jclepro.2022.131164

Fernández, F. (2011). The Greatest Impediment to the Study of Biodiversity in Colombia. *Caldasia* 33(2), iii–v.

Fernández, J.E. (2023, October 18). Presupuesto General de la Nación de 2024 dejó pocos recursos a la ciencia, tecnología e innovación. Infobae. https://www.infobae.com/colombia/2023/10/19/presupuesto-general-de-la-nacion-de-2024-dejo-pocos-recursos-a-la-ciencia-tecnologia-e-innovacion/

Fredriksson, M. (2021). Dilemmas of protection: Decolonising the regulation of genetic resources as cultural heritage. *International Journal of Heritage Studies* 27(7), 720–733. https://doi.org/10.1080/13527258.2020.1852295

Frynas, J. G., Stephens, S. (2015). Political Corporate Social Responsibility: Reviewing Theories and Setting New Agendas. *International Journal of Management Reviews* 17(4), 483–509. https://doi.org/10.1111/ijmr.12049

Godin, B. (2006). The Linear Model of Innovation: The Historical Construction of an Analytical Framework. *Science, Technology, & Human Values* 31(6), 639–667. https://doi.org/10.1177/0162243906291865

Godin, B. (2009). National Innovation System: The System Approach in Historical Perspective. *Science, Technology, & Human Values* 34(4), 476–501. https://doi.org/10.1177/0162243908329187

Goyes, D. R., South, N. (2016). Land-grabs, Biopiracy and the Inversion of Justice in Colombia. *British Journal of Criminology* 56(3), 558–577. https://doi.org/10.1093/bjc/azv082

Guerrero-Pineda, C., Iacona, G. D., Mair, L., Hawkins, F., Siikamäki, J., Miller, D., Gerber, L. R. (2022). An investment strategy to address biodiversity loss from agricultural expansion. *Nature Sustainability* 5(7), 610–618. https://doi.org/10.1038/s41893-022-00871-2

Güiza, L., Bernal Camargo, D. R. (2013). Bioprospecting in Colombia. *Universitas Scientiarum* 18(2), 153. https://doi.org/10.11144/Javeriana.SC18-2.bc

Guzman, C., Mattar, S., Alvis-Guzman, N., Hoz, F. D. la, & Arias, E. (2024). Biotechnological sovereignty is not a mere nationalist concept, it is a necessity for Colombia and Latin America. *Cadernos de Saúde Pública* 40(9), e00202323. https://doi.org/10.1590/0102-311xen202323

Hall, A. (2002). Innovation systems and capacity development: An agenda for North-South research collaboration? *International Journal of Technology Management & Sustainable Development* 1(3), 146–152. https://doi.org/10.1386/ijtm.1.3.146

Hall, A. (2005). Capacity development for agricultural biotechnology in developing countries: An innovation systems view of what it is and how to develop it. *Journal of International Development* 17(5), 611–630. https://doi.org/10.1002/jid.1227

Hickel, J., Dorninger, C., Wieland, H., Suwandi, I. (2022). Imperialist appropriation in the world economy: Drain from the global South through unequal exchange, 1990–2015. *Global Environmental Change* 73, 102467. https://doi.org/10.1016/j.gloenvcha.2022.102467

Huddart, J. E. A., Crawford, A. J., Luna-Tapia, A. L., Restrepo, S., Di Palma, F. (2022). EBP-Colombia and the bioeconomy: Genomics in the service of biodiversity conservation and sustainable development. *Proceedings of the National Academy of Sciences* 119(4), e2115641119. https://doi.org/10.1073/pnas.2115641119

Irwin, A. (2023). Expeditions in post-war Colombia have found hundreds of new species. But rich ecosystems are now under threat. *Nature* 619(7970), 450–453. https://doi.org/10.1038/d41586-023-02300-6

Johnson, F. X., Canales, N., Fielding, M., Gladkykh, G., Aung, M. T., Bailis, R., Ogeya, M., Olsson, O. (2022). A comparative analysis of bioeconomy visions and pathways based on stakeholder dialogues in Colombia, Rwanda, Sweden, and Thailand. *Journal of Environmental Policy & Planning* 24(6), 680–700. https://doi.org/10.1080/1523908X.2022.2037412

Kreiken, B., & McCarthy, A. (2025). Scientists and the Sovereigns: The Distributive Justice Implications of Digital Sequence Information Governance Under the Convention on Biological Diversity. In C. Didier, A. Béranger, A. Bouzin, H. Paris, & J. Supiot (Eds.), Engineering and Value Change (Vol. 48, pp. 125–143). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-83549-0_8

Kuhlmann, S., Arnold, E. (2001). RCN in the Norwegian National Innovation System (Background report no. 12). Technopolis Group. https://ris.utwente.nl/ws/portalfiles/portal/15070352/RCN_in_the_Norwegian_Research_and_Innovation_Syste_1_.pdf

Landis, A. (2024, September 13). In a historic move, Colombia bypasses patent to access HIV drug. Al Jazeera. Available at: https://www.aljazeera.com/features/longform/2024/9/13/in-a-historic-move-colombia-bypasses-a-patent-to-access-a-key-hiv-drug (Accessed: 10 June 2025).

Lee, S. H., Kang, Y. H., Dai, R. (2021). Toward a More Expansive Discourse in a Changing World: An Analysis of Political Leaders' Speeches on Biodiversity. *Sustainability* 13(5), 2899. https://doi.org/10.3390/su13052899

Lew, A.A. (2011). Defining and redefining conceptual frameworks for social science field research. In M.C. Hall (Ed.), *Fieldwork in Tourism: Methods, Issues and Reflections* (pp. 19-34). Routledge.

Ljungqvist, G. V., Weets, C. M., Stevens, T., Robertson, H.,

Zimmerman, R., Graeden, E., & Katz, R. (2025). Global patterns in access and benefit-sharing: A comprehensive review of national policies. *BMJ Public Health* 3(1), e001800. https://doi.org/10.1136/bmjph-2024-001800

LSE Roundtable Team. (2024). Identifying Ways Forward: LSE Roundtable on Biodiversity Finance and Digital Sequence Information. Zenodo. https://doi.org/10.5281/ZENODO.13935826

Martinez, H., Jaime, A., Camacho, J. (2014). Biotechnology profile analysis in Colombia. *Scientometrics* 101(3), 1789–1804. https://doi.org/10.1007/s11192-014-1408-2

Melgarejo, L. M. (2013). Bioprospecting as a possible development mechanism for Colombia. *Acta Biologica Colombiana* 18. 19-30.

Ministerio de Ambiente y Desarollo Sostenible. (2021). Régimen de acceso a recursos genéticos y sus productos derivados en Colombia [presentation]. Available at: https://www.minambiente.gov.co/wp-content/uploads/2021/12/Presentacion-Regimen-de-acceso-a-recursos-geneticos-y-sus-productos-derivados-en-Colombia.pdf

Mishra, V., Bankar, N., Tiwade, Y., Ugemuge, S. (2024). How Phage Therapy Works, Its Advantages and Disadvantages: Mini Review. *Journal of Pure and Applied Microbiology* 18(1), 177–184. https://doi.org/10.22207/JPAM.18.1.49

Mormina, M. (2019). Science, Technology and Innovation as Social Goods for Development: Rethinking Research Capacity Building from Sen's Capabilities Approach. *Science and Engineering Ethics* 25(3), 671–692. https://doi.org/10.1007/s11948-018-0037-1

Nemogá, G. R. (2014). Biodiversity research and conservation in Colombia (1990–2010): The marginalization of indigenous peoples' rights. *Canadian Journal of Latin American and Caribbean Studies / Revue Canadienne Des Études Latino-Américaines et Caraïbes* 39(1), 93–111. https://doi.org/10.1080/08263663.2014.978166

Noreña – P, A., González Muñoz, A., Mosquera-Rendón, J., Botero, K., Cristancho, M. A. (2018). Colombia, an unknown genetic diversity in the era of Big Data. BMC Genomics 19(S8), 859. https://doi.org/10.1186/s12864-018-5194-8

Organisation for Economic Co-operation and Development. (2014). OECD Territorial Reviews: Colombia 2014 (Paris: OECD Publishing). http://dx.doi.org/10.1787/9789264224551-en

Organisation for Economic Co-operation and Development. (2024a). OECD Economic Surveys: Colombia 2024 (Paris: OECD Publishing). https://doi.org/10.1787/a1a22cd6-en

Organisation for Economic Co-operation and Development. (2024b). Corporate Tax Statistics 2024 (Paris: OECD Publishing). https://doi.org/10.1787/9c27d6e8-en.

PacBio. (2020). Colombian SMRT Grant winners hope HiFi sequencing will help save critically endangered toads. Available at: https://www.pacb.com/blog/colombian-smrt-grant-winners-toads/ (Accessed: 10 June 2025)

Paz, A., Pardo-Díaz, C. (2024). Female researchers are underrepresented in the Colombian science infrastructure. *PLOS ONE* 19(3), e0298964. https://doi.org/10.1371/journal.pone.0298964

Presidencia de la República - Colombia. (2024, October 21). Palabras del Presidente de Colombia Gustavo Petro durante la Ceremonia de Apertura de la COP16 [Video]. YouTube. https://www.youtube.com/watch?v=-BHA5QcPEr8&ab_channel=PresidenciadelaRep%C3%BAblica-Colombia

ProColombia. (2024). Legal guide to doing business in Colombia. Available at: https://investincolombia.com.co/en/resources/legal-guide-2024 (Accessed: 10 June 2025).

Ramírez-Castañeda, V. (2020). Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological sciences. *PLOS ONE* 15(9), e0238372. https://doi.org/10.1371/journal.pone.0238372

Reep, A. (2025). A Human Rights Approach to Benefit-Sharing from the Use of Digital Sequence Information (DSI) [Policy brief]. Dejusticia. https://publicaciones.dejusticia.org/items/7a9d1b95-e629-4eac-b094-0d5b30a248ba

Rohden, F., & Scholz, A. H. (2022). The international political process around Digital Sequence Information under the Convention on Biological Diversity and the 2018–2020 intersessional period. *Plants, People, Planet* 4(1), Article 1. https://doi.org/10.1002/ppp3.10198

Salas, P. (2020). Legal Protection of Sustainable Design in Colombia. Rev. Prop. Inmaterial, 30, 73.

Scholz, A. H., Freitag, J., Lyal, C. H. C., Sara, R., Cepeda, M. L., Cancio, I., Sett, S., Hufton, A. L., Abebaw, Y., Bansal, K., Benbouza, H., Boga, H. I., Brisse, S., Bruford, M. W., Clissold, H., Cochrane, G., Coddington, J. A., Deletoille, A.-C., García-Cardona, F., … Overmann, J. (2022). Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation. *Nature Communications* 13(1), Article 1. https://doi.org/10.1038/s41467-022-28594-0

Scholz, A. H., Lange, M., Habekost, P., Oldham, P., Cancio, I., Cochrane, G., Freitag, J. (2021). Myth-busting the provider-user relationship for digital sequence information. *GigaScience* 10(12), Article 12. https://doi.org/10.1093/gigascience/giab085

Secretariat of the Convention on Biological Diversity. (2011). Introduction to access and benefit-sharing [Fact sheet]. https://www.cbd.int/abs/infokit/revised/web/all-files-en.pdf

Senado de la República de Colombia. (1991). Constitución de Colombia. https://www.senado.gov.co/images/Archivospdf/elsenado/Normatividad/constitucion_politica.pdf

Shanahan, H., Bezuidenhout, L. (2022). Rethinking the A in FAIR Data: Issues of Data Access and Accessibility in Research. Frontiers in Research Metrics and Analytics 7, 912456. https://doi.org/10.3389/frma.2022.912456

Sherman, B., Adhikari, K., & Balaji, S. (2025). The impact of access and benefit-sharing measures on the use and exchange of genetic resources for food and agriculture and associated traditional knowledge [Draft Report]. Food and Agricultural Organization, CGRFA-20/25/3.2/Inf.1. Available at: https://openknowledge.fao.org/items/cd0b96c7-8524-4ecb-bc75-d8c6fc97cf20

Silvestri, L. C. (2016). Access to genetic resources and benefit-sharing in Colombia: challenges of the legal framework. Investigación & Desarrollo 24, Article 1. https://doi.org/10.14482/indes.24.1.8682

Sirakaya, A. (2022). Where access and benefit-sharing comes from: A historical overview, *Genetic Resources*, 3(6), pp. 74–88. doi: https://doi.org/10.46265/genresj.PPUF5169.

The City Paper Bogota. (2024a, October 25). COP16: Colombia presents ambitious Biodiversity Action Plan. Available at: https://thecitypaperbogota.com/news/cop16-colombia-presents-ambitious-biodiversity-action-plan/ (Accessed: 28 February 2025).

The City Paper Bogota. (2024b, October 23). COP16: Colombia calls for Genetic "Sovereignty" and Global Fund to share Data. Available at: https://thecitypaperbogota.

com/news/cop16-colombia-calls-for-genetic-sovereignty-and-global-fund-to-share-data/ (Accessed:10 June 2025).

Vallejo Zamudio, L. E. (2023). El Plan Nacional de Desarrollo 2022-2026: Colombia, potencia mundial de la vida. *Apuntes Del Cenes* 42(76), 7–13. https://doi.org/10.19053/01203053.v42.n76.2023.16467

Vilaça, S. T., Vidal, A. F., Pavan, A. C. D., Silva, B. M., Carvalho, C. S., Povill, C., Luna-Lucena, D., Nunes, G. L., Figueiró, H. V., Mendes, I. S., Bittencourt, J. A. P., Côrtes, L. G., Costa Canesin, L. E., Oliveira, R. R. M., Damasceno, R. P., Vasconcelos, S., Barreto, S. B., Tavares, V., Oliveira, G., … Aleixo, A. (2024). Leveraging genomes to support conservation and bioeconomy policies in a megadiverse country. *Cell Genomics* 4(11), 100678. https://doi.org/10.1016/j.xgen.2024.100678

Villegas, F. (2024, June 16). Bioeconomía, una gran oportunidad para Colombia. Cambio. https://cambiocolombia.com/medio-ambiente/bioeconomia-una-gran-oportunidad-para-colombia

Weerasinghe, R. N., Jayawardane, A. K. W., Huang, Q. (2024). Critical inquiry on National Innovation System: Does NIS fit with developing countries? *Sustainable Technology and Entrepreneurship* 3(1), 100052. https://doi.org/10.1016/j.stae.2023.100052

Wight, A. J. (2019). In Colombia, biodiversity researchers seek relief from regulatory red tape. *Science*. https://doi.org/10.1126/science.aax1404

Wise Ancestors. (2024). New Conservation Challenges to Strengthen Conservation for Two Bird Species Endemic to the Antioquia Department of Colombia. Available at: https://www.wiseancestors.org/blog/new-conservation-challenges-to-conserve-two-bird-species-endemic-to-the-antioquia-department-of-colombia (Accessed: 10 June 2025).

World Bank (2024, September 30). Research and development expenditure (% of GDP) [Dataset]. World Bank Indicators. https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 57–72
DOI: 10.46265/genresj.FQLF1923
https://www.genresj.org
ISSN: 2708-3764

Ethnography of traditional healers and their indigenous medicinal plants in southern Philippines: Implications for conservation and sustainable use

Bhegie May G. Buay[a,d], Catherine Hazel M. Aguilar[b], Maria Celeste N. Banaticla-Hilario[c], Chenny Rose M. Rodriguez[d] and Florence L. Zapico[d,*]

[a] *Department of Biology, Institute of Mathematics, Arts, and Sciences, Davao del Sur State College, Digos City, The Philippines*
[b] *Global Crop Diversity Trust, Platz der Vereinten Nationen, 7 53113 Bonn, Germany*
[c] *Plant Biology Division, Institute of Biological Sciences, University of the Philippines-Los Banos, College, Laguna, The Philippines*
[d] *Department of Biology, College of Natural Sciences and Mathematics, Mindanao State University - General Santos, The Philippines*

Abstract: This paper investigates Indigenous medicinal plants, the threats they face and the healing knowledge and profiles of traditional healers in the Sarangani uplands, Southern Philippines. During field and community floral inventories, 39 medicinal plant species were documented, belonging to 18 orders, 20 families, and 31 genera. While this study unveiled diverse utilization of medicinal plants, interviewed healers unfortunately revealed local losses which they attributed to (1) climate change, (2) overharvesting, (3) forest denudation, and (4) the shift to over-the-counter medicines. Additionally, the gradual erosion of healing knowledge was ascribed to (1) Christianization suppressing traditional healing practice, (2) local losses of medicinal plants, (3) shift in culture and lifestyle brought by increasing market integration, (4) reluctance of tribal healers to share healing knowledge, (5) devaluation of Indigenous knowledge by the younger generation, (6) advanced ages of knowledge keepers, and (7) the oral nature of mentoring. Moreover, this paper reports that conservation was accomplished mainly through continuous utilization/cultivation and the judicious collection of medicinal plants. These efforts are, however, grossly insufficient and without complementary *in situ* and *ex situ* conservation initiatives, these invaluable genetic treasures will face local extinction. In addition to Sarangani's medicinal plant losses, the valuable cache of associated traditional knowledge will likewise be lost, resulting in a culturally impoverished and less resilient community. Finally, to foster inclusivity, promote knowledge pluralism, and aid in the preservation of traditional healing knowledge, the involvement of healers in the crafting of a comprehensive healthcare strategy for Southern Philippines is recommended.

Keywords: Traditional healers, Indigenous medicinal plants, conservation, utilization, Sarangani Province

Citation: Buay, B. M. G., Aguilar, C. H. M., Banaticla-Hilario, M. C. N., Rodriguez, C. R. M. and Zapico, F. L. (2025) "Ethnography of traditional healers and their indigenous medicinal plants in southern Philippines: Implications for conservation and sustainable use", *Genetic Resources*, 6(12), pp. 57–72. doi: 10.46265/genresj.FQLF1923.

Introduction

In biodiversity-rich, low to middle-income countries like the Philippines, traditional healing knowledge and medicines remain the primary refuge of ethnic peoples for their healing and therapeutic needs. In rural communities, medicinal plants are abundant, freely available and reputed to be efficacious based on millennia of traditional use (Mahmoud and Gairola, 2013; Bankole *et al*, 2015; Barata *et al*, 2016). Indigenous groups have been utilizing Indigenous medicinal plants (IMPs) since time immemorial, leading to a wealth of accumulated traditional healing knowledge (THK). Even up to this time, developed and developing countries alike rely on IMPs as direct sources of medicine or as raw materials for the processing of therapeutic drugs (Miano *et al*, 2011; Ambasta *et al*, 2016), thereby resulting in an increasing demand for these priceless natural resources.

Studies from Nepal (Aryal *et al*, 2016) and Nigeria (Stoffersen *et al*, 2011; Borokini *et al*, 2013) report on age-

* Corresponding author: Florence L. Zapico (florence.zapico@gmail.com)

Received: 02.05.2025 | Accepted: 01.08.2025 | Published online: 17.09.2025

old THK and utilization of IMPs by local communities as well as their economic potentials (Batugal *et al*, 2004). Known as ethnomedicine, this field of anthropology deals with medicinal plants and the wealth of healing knowledge of traditional healers (Cotton, 1996; Cheikhyoussef *et al*, 2011). THK is, therefore, an indispensable source of socioculturally coherent information about IMPs, their Indigenous uses, and their natural habitats. Nonetheless, THK is vulnerable to sociocultural and ecological transformations that define traditional communities in contemporary times.

Considered as one of the 18 mega-biodiverse countries in the world, the Philippines ranks fifth in terms of plant species richness (CBD, 2025). Recently, Meniza *et al* (2024) documented 1,500 species of medicinal plants in the Philippines, with over one-third (or 530 species) found in Mindanao. Recent studies (Dapar *et al*, 2020; Paraguison *et al*, 2020; Alinsug *et al*, 2022; Cabugatan *et al*, 2022; Ilagan *et al*, 2022) revealed the prevalent use of medicinal plants in rural Mindanao communities. However, these priceless genetic resources and associated Indigenous knowledge are threatened by emerging social-ecological realities in these areas. Among the identified major pressures to medicinal plant species in the Philippines are agricultural expansion, deforestation, mining, environmental degradation, unregulated resource extraction and climate change (Mendoza *et al*, 2016; Cordero *et al*, 2022; Agduma *et al*, 2023; Belgica *et al*, 2024). Moreover, Alinsug *et al* (2022), Dapar *et al* (2020), Fiscal (2017), and Ong and Kim (2014) reported about the propensity of the younger generation to embrace modernization, resulting in knowledge erosion. Other identified causes of knowledge erosion are acculturation, outmigration and increasing access to over-the-counter medicine (Dapar *et al*, 2020; Cordero *et al*, 2022).

At the national level, the Philippine government has been promoting the shift to IMPs, given the exorbitant prices of modern medicines. After thorough evaluation, the Department of Health (DOH) endorsed ten medicinal plants, viz. *Senna alata* (L.) Roxb., *Momordica charantia* L., *Allium sativum* L., *Psidium guajava* L., *Vitex negundo* L., *Combretum indicum* (L.) De Filipps, *Blumea balsamifera* (L.) DC., *Ehretia microphylla* Lam., *Peperomia pellucida* (L.) Kunth and *Clinopodium douglasii* (Benth.) Kuntze for widescale use (Dapar *et al*, 2020). Another initiative of DOH was the promotion of herbal medicine gardens in rural communities and the integration of traditional healing into mainstream healthcare (Maramba-Lazarte, 2020). Unfortunately, there is a paucity of information about traditional healthcare practitioners, especially those residing in far-flung areas.

Sarangani Province in Southern Mindanao is home to Lumad (tribes of non-Muslim ethnicity) groups such as B'laans, Tagakaulos, and T'bolis. These groups inhabit remote and inaccessible upland areas and consequently receive no (or very little) basic social services from the local government. Moreover, the lack of government presence in these areas and an official census of the tribal population have resulted in a dearth of information about their exact numbers. These Lumads are, therefore, the most disadvantaged sector of Philippine society, whose dependence on local resources was highlighted when the COVID-19 pandemic isolated their remote communities. This study was thus designed to (1) profile traditional healers, (2), document Sarangani medicinal plants and how they are used in healing rituals, (3) investigate threats to Sarangani medicinal plants, (4) carry out initiatives for evidence-based *ex situ* and *in situ* conservation of Sarangani medicinal plants and (5) recommend for the inclusion of traditional healers in the crafting of a comprehensive health strategy. It is envisioned that meticulous documentation of IMP utilization and conservation status will contribute significantly to the preservation of Sarangani traditional healing knowledge for future generations.

Materials and methods

Study locations

Figure 1 shows a location map of Sarangani Province with all study sites indicated. This study was conducted in seven upland villages (Sitios) in four towns (viz. Kinam and Banlas in Malapatan, Datal Anggas and Ihan in Alabel, Miasong and Ligaya in Glan and Gasie in Maasim) in Sarangani Province. In Malapatan and Glan, B'laans are the predominant tribal group. T'bolis inhabit Maasim while B'laans and Tagakaulos reside in the upland villages of Alabel. The study sites and their geophysical coordinates are shown in Table 1.

Preliminary preparations and ethical considerations

Letters seeking permission for the study were sent to the provincial governor, the mayors, heads of agriculture offices, and barangay (village) captains before field visits. Subsequently, both oral and written consents were secured from healers before the study commenced. Elderly healers who were illiterate appended their thumbprints to the consent form. The participants were informed about the study, its objectives and the information to be solicited from them. They were assured that their identities would be maintained with the utmost confidentiality. Additionally, they were apprised of their right to decline to respond to inquiries or to terminate the key informant interview (KII) at any moment for any personal reason.

Research design, respondent selection, inclusion and exclusion criteria

This study is a qualitative research that employed ethnographic methods, such as direct participant observation, interviews with key informants and a semi-structured questionnaire to chronicle the traditional knowledge of local healers based on their lived experiences. Purposive sampling based on information from sitio officials was utilized to identify potential study respondents. Subsequently, the researchers scoured the remote upland areas of Alabel, Glan, Malapatan and Maasim to locate identified healers. While the municipalities of Maitum, Malungon and Kiamba were initially considered for inclusion in the study, several factors rendered this difficult, if not outright impossible. In Maitum, traditional healers visited by the team declined requests for an interview, believing that divulging secret information would diminish their healing powers. As for the municipalities of Malungon and Kiamba, locals reported shifting to modern medicines and that no more healers were found in the sitios. Moreover, the remoteness and isolation of some villages and safety concerns prevented the researchers from penetrating these far-flung areas.

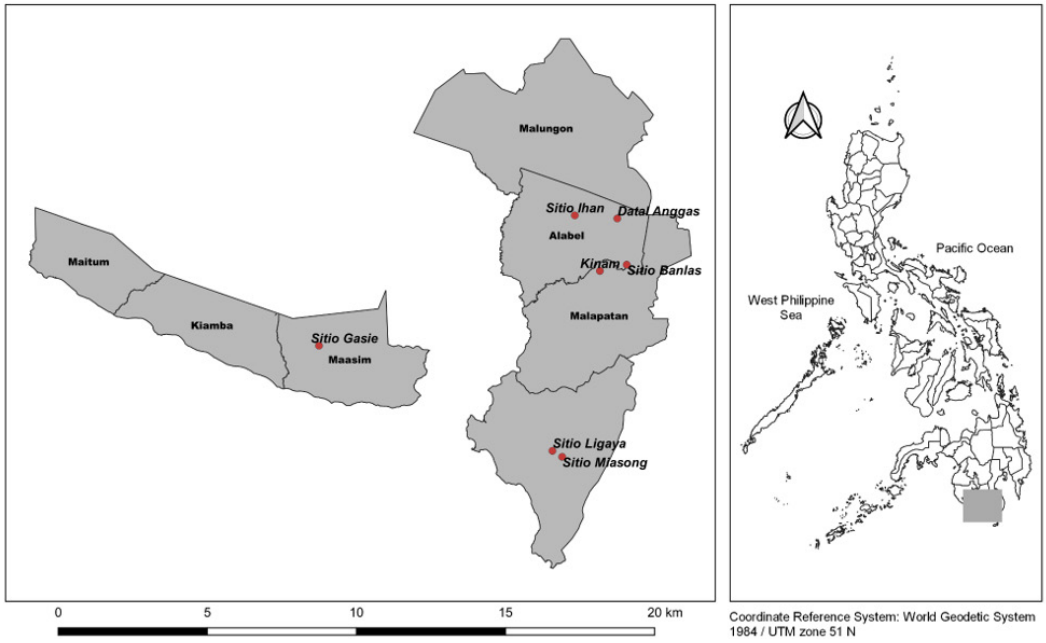


Figure 1. Map of Sarangani Province showing the municipalities where the study was conducted

Table 1. Study sites, their geophysical coordinates and dominant ethnic group/s. IP, Indigenous People; masl, metres above sea level.

Sitio	Municipality	GPS Coordinates	Altitude (masl)	Dominant IP Group
Datal Anggas	Alabel	N 06°11.995' E 125°27.158'	900	B'laan, Tagakaulo
Sitio Ihan	Alabel	N 06°12'21.7" E 125°22'21.7"	905	B'laan, Tagakaulo
Sitio Ligaya	Glan	N 05°46'17.5" E 125°19'50.5"	426	B'laan
Sitio Miasong	Glan	N 05°45.638' E 125°20.954'	428	B'laan
Sitio Banlas	Malapatan	N 06°06.508' E 125°28.065'	342	B'laan
Kinam	Malapatan	N 06°06.125' E 125°25.0158'	406	B'laan
Sitio Gasie	Maasim	N 05°57.495' E 124°54.083'	1,016	T'boli

The inclusion criteria encompassed traditional healers of any gender, aged 25 years or older, who utilized IMPs for healing and provided explicit consent to participate in the study. A prospective study participant had to be a recognized traditional healer within the community. Conversely, individuals who shifted to modern medicine and those who declined invitations for participation were excluded from the study.

Data collection methodologies and analyses

Ethnobotanical assessment

The semi-structured questionnaire used consisted of open-ended questions, which enabled respondents to articulate comprehensive views regarding traditional healing knowledge in their own words. This questionnaire contained four main sections: (1) demographic profiles of healers and knowledge acquisition, (2) medicinal plants and their modes of utilization, (3) medicinal plant sources, processing and storage, and (4) threats to medicinal plants and conservation awareness. After pre-testing, the validated questionnaire (see Supplemental Material 1) was used as the main data-gathering tool during face-to-face interviews. The questionnaire, written in English, was administered using the predominant dialect of Sarangani (Cebuano) or the tribal dialect through a local intermediary for respondents who could only converse in their native language.

To validate the results of the questionnaire and delve deeper into the healing knowledge of the Sarangani ethnic groups, key informant interviews were carried out. As a qualitative in-

depth interview method, a key informant interview is flexible and extracts knowledge from key respondents who have particularly informed perspectives and first-hand knowledge about the topic at hand. Being a descriptive study, answers to the questionnaire and key informant interviews were analyzed thematically, classified a posteriori, and discussed narratively to minimize bias in the conclusions. Subsequently, relevant information was presented in tables for easier reference and interpretation.

Inventory, collection, identification and conservation of Indigenous medicinal plants

Field observation and community walks were conducted to triangulate information and corroborate data from the questionnaires and key informant interviews, thereby lending credence to the results. With the assistance of the healers, medicinal plants were identified and gathered (with consent) from backyards, home gardens, roadsides and adjacent woodlands. Two distinct collection methods were utilized, contingent on the biological characteristics of the plants. For sexually reproducing plants, viable seeds were collected and placed in labelled coin envelopes. On the other hand, vegetatively propagated herbs and seedlings were immediately planted in black polypropylene seedling bags filled with garden soil. Both coin envelopes and seedling bags were labelled with passport information such as the plant's local name, collection date and site, the name of the household head (if collected in home gardens), and its geophysical coordinates. During community walks, the healers initially recognized medicinal plants by their ethnic names, which were subsequently verified through appropriate references. Collected medicinal plants were classified according to their growth habits and their habitats, among other information. Taxonomic identification was done up to the species level whenever possible using suitable references and consultation with a botanist. No scientific identification was conducted for some IMPs growing deep in the forests, and the researchers relied solely on the healers' memory and knowledge. Subsequently, collected seeds from Sarangani medicinal plants were sent to the National Plant Genetic Resources Laboratory of the University of the Philippines in Los Banos, Laguna for conservation purposes.

Results

Profile of the traditional healers

The demographic profiles of the traditional healers are presented in Table 2. Of the 12 traditional healers who consented to participate in the study, 5 were from Malapatan, 3 were from Alabel, and Glan and Maasim had 2 healers each. Female healers and those with B'laan ancestry predominated at 11 and 8, respectively, with 8 healers having no formal education and the rest having 1–6 years of primary education because of the distance of schools. Moreover, healers' ages varied across a broad spectrum (27 to 110 years old), with a mean age of 63.9 years. In terms of healing experience, traditional practitioners ranged from being neophytes (1 year) to veteran healers (90 years), with most of them falling within the range of less than 20 years of experience. In terms of religion, the healers professed to be Christians, a fact borne out by small chapels established by Protestant missionaries from the lowlands. When asked about their sources of healing knowledge, all respondents received information through oral transmission, with ten healers learning from their parents and grandparents. The remaining healers availed of instruction from government-sponsored seminars or the radio. Unfortunately, a traditional healer refused to participate in an interview and disclose healing knowledge, as doing so could jeopardize her healing abilities.

Table 2. Demographic profiles of the traditional healers.

Demographic Profiles	Details	Frequency
Gender	Male	8%
	Female	92%
Ethnicity	B'laan	67%
	T'boli	17%
	Aklanon	8%
	Lowland Tribes	8%
Education	Primary	36%
	Secondary	0%
	Tertiary	0%
	None	64%
Mean age		63.9
Number of years practising healing	≤ 20 years	75%
	≥ 20 years	25%
Sources of healing knowledge	Oral transmission	83%
	Government-sponsored seminars/radio	17%
Perceived health status	Healthy	100%
	Sickly	0%
Religion	Christian	100%

Moreover, the interviewed tribal healers disclosed that the ability to heal is a gift, and that it stays in the family. While most healers use plants alone to treat sick individuals, a few invoke the spirits through incantations.

Biodiversity assessment of Indigenous medicinal plants

A total of 39 IMP species were documented based on the accounts of the traditional healers (Table 3). However, only 32 species, belonging to 31 genera, 18 orders and 20 families, were collected during community walks. Seven taxa were identified up to the genus level only as they lack reproductive structures needed for species level identification. The comparable number of families and orders reflect the phylogenetic diversity of the medicinal plant species in the Sarangani upland communities. Families Asteraceae and Lamiaceae were the most represented with five species each. These were followed by families Euphorbiaceae, Moraceae, Poaceae and Zingiberaceae with two species each.

Of the collected plants, 26 were foraged within the community (backyards, roadsides and transition zones between forests and villages), while 6 species were collected in the forests by the healers. Moreover, of the 39 plant species identified by healers as having medicinal properties, 21% were trees, 56% were herbs, 5% were climbers or vines, and the remaining 18% were classified as shrubs or small trees.

Indigenous medicinal plants and traditional healing knowledge

From the semi-structured questionnaires and key informant interviews, information about the preparation and utilization of IMPs was collated. Most herbal remedies were prepared using primarily leaves (82%), while others used stems (or tree bark), roots, and plant sap. Moreover, traditional healers favoured decoction (85%) over fresh material use in their herbal concoctions. For some medicinal preparations, tree bark (or other tougher parts of the plant) was mixed with coconut oil and used as a liniment by spreading liberally on the skin. As for fresh leaves, some were directly applied to the skin as a poultice. Moreover, except for T'kulu (*Tibouchina* sp.), T'kas (*Elytropappus* sp.), Kataas (*Colocasia esculenta* (L.) Schott), K'lol (*Tinospora crispa* (L.) Hook. f. & Thomson), Langka (*Artocarpus heterophyllus* Lam.) and Luy-a (*Zingiber officinale* Roscoe), which were combined with other plants in coconut oil during the preparation of the herbal remedies, the remainder of the medicinal plants were given as single preparations.

Incidentally, the traditional healers did not have any concept of dosage and were not mindful of the proportions of plant parts to water or oil that they used in their preparations. Traditional healers likewise disclosed that most of the IMPs (56%) were readily available in the neighbourhood, did not need to be planted, and could be foraged anytime. Conversely, IMPs (predominantly trees and shrubs) found in the forests were less accessible, and their harvesting posed significant risks. Shown in Table 3 are the IMPs of the ethnic groups along with their scientific and local names, their habit, utilization and their availability in the upland areas of Sarangani Province.

Traditional healers in Sarangani Province classified human ailments into nine categories: (1) gastro-intestinal diseases, (2) head, ear, nose, and throat diseases, (3) cuts or wounds, (4) diseases of the bone and muscles, (5) urinary ailments, (6) cough/colds, (7) fever, (8) mouth/tongue/tooth problems and (9) other ailments (Table 3). In the upland areas, the more prevalent disorders were cough, loose bowel movement (LBM), kabuhi, stomachache and headache. A diagnosis of kabuhi is made when the patient experiences dizziness, nausea, chills and severe abdominal pain. This disorder is akin to heartburn, hyperacidity, or gastroesophageal reflux disease, depending on the nature and severity of symptoms. Furthermore, the healers revealed their capability to treat buyag, a culturally recognized ailment that can be loosely defined as a malady resulting from a curse imposed by someone with malevolent intent, an 'evil eye', or nature spirits. This supposed ailment necessitates incantations in addition to the IMPs.

Figure 2 shows some indigenous medicinal plants used by the traditional healers. As for their healing practices, traditional healers disclosed that decoctions of combined langka leaves/ ginger and Kataas/K'lol are effective against kabuhi. Among the T'bolis in Maasim, stems and leaves of T'kulu and T'kas are chopped into small pieces and put in a receptacle containing coconut oil. The resulting liniment or salve (haplas) is used to relieve shoulder pain and stomachache. For cough and colds, leaf decoctions of gabon (*Blumea balsamifera*), lagundi (*Vitex negundo*), mayana (*Coleus scutellarioides*), tawa-tawa (*Euphorbia hirta*), kalabo/ balbas pusa (*Orthosiphon aristatus*) and mertaan (*Ficus septica*) are used by healers. For persons with LBM, root decoction of banlo-banlo (*Galinsoga* sp.) and leaf decoctions of star apple (*Chrysophyllum cainito*) and white flower (*Andrographis paniculata*) are given. In contrast, a tawa-tawa leaf decoction is given to a person suffering from headaches. On the other hand, patients with dengue are given papaya (*Carica papaya*) and tawa-tawa leaf decoctions, while fresh samples of mertaan, root decoctions of bulong baltang (*Heliotropium indicum*) and leaf decoctions of blibid (*Eleusine indica*) are given for bughat. Bughat, which loosely translates to relapse, is used to describe a condition when a person who is recovering from illness or who has shortly recovered from it becomes sick again. In addition, sili (*Capsicum* sp.) and tawal leaf decoctions are used to treat snake bites while gabon and galong are used for urinary tract infections by the traditional healers. To treat fever, healers use alingatong or *Dendrocnide* sp. (despite its itchiness), angelika/skaan bulan (*Kalanchoe pinnata*) and native sibuyas (*Allium* cepa). Regarding their utilization, tawa-tawa, kyama (*Stevia rebaudiana*) and mertaan have a wide range of medicinal uses. For instance, traditional healers use tawa-tawa for the treatment of malaria, cough, swelling, stomachache and headache, whereas kyama is utilized to address stomachache, toothache, dysmenorrhea, flatulence, and to facilitate wound cleansing. In addition, mertaan serves as a remedy for toothache, cough and bughat, while *Coleus amboinicus*, commonly referred to as oregano, is utilized for cough treatment and is widely utilized by the tribes. Moreover, another remarkable observation was related to plants having redundant utilities. For instance, tribal healers utilized ten, six, and five IMPs for common disorders like cough, LBM, and stomachache, respectively.

Table 3. Medicinal plants used in Sarangani Upland Communities. LBM, loose bowel movement; UTI, urinary tract infection; IUCN Categories: LC, least concern; EN, endangered; DD, data deficient.

Local Name	English/ Common name	Scientific name	Family	Order	Used for	Plant habit	Plant part used	Mode of preparation	Mode of administration	Where harvested	Source	Method of use	Availability	IUCN
Alingatong	Stinging nettle	*Dendrocnide* sp.	Urticaceae	Rosales	Fever	Tree	Leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	Singly	Less available	
Ampalaya	Bitter gourd	*Mimordica charantia*	Cucurbitaceae	Cucurbitales	Cough	Climber	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Angelika / Skaan Bulan	Katakataka / Air plant / Cathedral bells	*Kalanchoe pinnata* (Lam.) Pers.	Crassulaceae	Saxifragales	Fever; Cough	Herb	Leaf	Decoction; Fresh	Drunk; External/ rubbing	Home garden	Cultivated	Singly	Less available	
Avocado	Avocado	*Persea americana* Mill.	Lauraceae	Laurales	LBM	Tree	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	LC
Banlo-banlo	Gallant soldier	*Galinsoga* sp.	Asteraceae	Asterales	LBM	Herb	Root	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Blilid	Goose Grass	*Eleusine indica* (L.) Gaertn.	Poaceae	Poales	Bughat (Relapse)	Herb	Leaf	Decoction	Drunk	Ubiquitous	Wild-harvested	Singly	Less available	LC
Bulong baltang	Indian heliotrope	*Heliotropium indicum L.*	Boraginaceae	Boraginales	Bughat (Relapse)	Herb	Root	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Gabon	Sambong	*Blumea balsamifera* (L.) DC.	Asteraceae	Asterales	Cough; UTI	Shrub	Leaf	Decoction	Drunk	Collected from the Wild; Home garden	Wild-harvested; Cultivated	Singly	Readily available	LC
Galong						Tree	Leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	Singly	Less available	
Kalabo / Balbas pusa	Cat's whiskers	*Orthosiphon aristatus* (Blume) Miq.	Lamiaceae	Lamiales	Cough; Cold	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Less available	
Kape	Coffee	*Coffea arabica* L.	Rubiaceae	Rubiales	LBM	Shrub	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	EN
Kataas	Taro	*Colocasia esculenta* (L.) Schott	Araceae	Alismatales	Kabuhi	Herb	Leaf	Fresh	External/ rubbing	Collected from the wild	Wild-harvested	With other plant	Less available	LC

Table 3 continued

Local Name	English/ Common name	Scientific name	Family	Order	Used for	Plant habit	Plant part used	Mode of preparation	Mode of administration	Where harvested	Source	Method of use	Availability	IUCN
K'lol		*Tinospora crispa* (L.) Hook.f. & Thomson		Ranunculales	Kabuhi	Climber	Root	Fresh	External/ rubbing	Collected from the wild	Wild-harvested	With other plant	Less available	
Kusol	Aromatic ginger	*Kaempferia galanga* (L.)	Zingiberaceae	Zingiberales	Earache	Herb	Leaf	Fresh	Ear Drops	Home garden	Cultivated	Singly	Less available	DD
Kyama	Candy leaf	*Stevia rebaudiana* (Berdoni) Berdoni	Asteraceae	Asterales		Herb	leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	Singly	Less available	
Lagano	Oregano	*Coleus amboinicus* Lour.	Lamiaceae	Lamiales	Cough	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Lagundi	Lagundi	*Vitex negundo* L.	Lamiaceae	Lamiales	Cough	Shrub	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	LC
Langka	Jackfruit	*Artocarpus heterophyllus* Lam.	Moraceae	Rosales	Kabuhi	Tree	Leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	With other plant	Readily available	
Layet langas					Headache; LBM	Herb	Leaf	Decoction	Drunk	Ubiquitous	Wild-harvested; Cultivated	Singly	Less available	
Luy-a	Ginger	*Zingiber officinale* Roscoe	Zingiberaceae	Zingiberales	Kabuhi	Herb	Stem	Decoction	Drunk	Collected from the wild; Home garden	Wild-harvested; Cultivated	With other plant	Readily available	DD
Mayana	Coleus	*Coleus scutellarioides* (L.) Benth.	Lamiaceae	Lamiales	Cough	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Mertaan	Fig tree	*Ficus septica* Burm. f.	Moraceae	Rosales	Toothache; Cough; Relapse	Shrub	Leaf	Fresh	Poultice	Collected from the wild	Wild-harvested	Singly	Less available	LC
Nabol					Tongue Sore	Tree	Sap	Fresh		Home garden	Cultivated	Singly	Readily available	
Native sibuyas	Onion	*Allium cepa* L.	Amaryllidaceae	Asparagales	Fever	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Papaya	Papaya	*Carica papaya* L.	Caricaceae	Brassicales	Dengue	Tree	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	DD

Table 3 continued

Local Name	English/ Common name	Scientific name	Family	Order	Used for	Plant habit	Plant part used	Mode of preparation	Mode of administration	Where harvested	Source	Method of use	Availability	IUCN
Sili	Sili	*Capsicum* sp.	Solanaceae	Solanales	Snake Bite	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Sislok					Wound Cleaning	Herb	Leaf	Fresh	Poultice	Ubiquitous	Wild-harvested; Cultivated	Singly	Less available	
S'lot	Carabao grass	*Paspalum conjugatum* P.J. Bergius	Poaceae	Poales	Vomiting	Herb	Leaf	Decoction	Drunk	Ubiquitous	Wild-harvested; Cultivated	Singly	Less available	LC
Star apple	Star apple	*Chrysophyllum cainito* L.	Sapotaceae	Ericales	LBM	Tree	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	LC
Talil					Vomiting	Herb	Stem	Decoction	Drunk	Collected from the wild	Wild-harvested	Singly	Less available	
Tambisan	Mugwort	*Artemisia* sp.	Asteraceae	Asterales	Cough	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Tawal					Snake Bite	Tree	Leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	Singly	Readily available	
Tawa-tawa	Asthma plant	*Euphorbia hirta* L.	Euphorbiaceae	Malpighiales	Malaria; Cough; Swelling;	Herb	Leaf; root	Decoction	Drunk	Collected from the wild; Home garden	Wild-harvested; Cultivated	Singly	Less Available	
T'kas		*Elytropappus* sp.	Asteraceae	Asterales	Shoulder pain	Shrub	Stem	Decoction	Drunk	Collected from the wild	Wild-harvested	With other plant	Less Available	
T'kulu		*Tibouchina* sp.		Myrtales	Shoulder pain	Shrub	Leaf	Decoction	Drunk	Collected from the wild	Wild-harvested	With other plant	Less Available	
Tuba-tuba	Tubang bakod	*Jatropha curcas* L.	Euphorbiaceae	Malpighiales	Mouth sore; Headache	Shrub	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	LC
Vicks Sibu	Mint	*Mentha* sp.	Lamiaceae	Lamiales	Fever; Cough	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
White flower	Green Chiretta	*Andrographis paniculata* (Burm. f.) Wall. ex Nees	Acanthaceae	Lamiales	LBM	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	
Wuh					Toothache	Herb	Leaf	Decoction	Drunk	Home garden	Cultivated	Singly	Readily available	

Figure 2. Images of medicinal plants collected in Sarangani upland communities. (a) Blilid (*Eleusine indica*); (b) Mertaan (*Ficus septica*); (c) white flower (*Andrographis paniculata*); (d) Alingatong (*Dendrocnide* sp.); and (e) Balbas pusa (*Orthosiphon aristatus*).

Indigenous medicinal plants: threats and conservation initiatives

Questionnaire administration and in-depth conversations with key informants were carried out to ascertain threats to their medicinal plant resource base. Figure 3 illustrates the challenges traditional healers encounter in the Sarangani uplands about their THK and the IMPs they utilize. The cultural practice of swiddening or alnigo (Figure 3a) before upland rice cultivation led to significant deforestation and considerable biodiversity losses (Figure 3c). Another legitimate concern raised by the locals during field observations was the considerable deforestation caused by unregulated resource extraction in the mountainous regions. Figure 3b, conversely, depicts an elderly tribal healer who has surpassed the remarkable age of 100 years but who was still active on the farm. Regrettably, these elderly healers expressed their sorrow over their failure to pass on THK to the younger generation, attributing this to a notable disinterest from the latter.

Upon the demise of these knowledge keepers, the invaluable information they possess will be irretrievably lost to humanity. Finally, the waning interest of the younger generation in traditional healing and their preference for modern medicines (Figure 3d) were also identified as significant threats to the perpetuation of traditional healing practices and the conservation of IMPs. Conversations with tribal healers indicated that although certain medicinal plants could be harvested year-round from home gardens and forest fringes, other IPMs were diminishing due to various stresses.

From 2016 to 2017, the prolonged effects of El Niño resulted in wide-scale losses of flora in the Sarangani uplands. Other threats identified by the healers were climate change-related devastation (such as heavy rains and flash floods), pest infestation, unregulated harvesting, forest denudation, and the shift to modern medicines. The healers also disclosed that they had to walk long distances to harvest rarely occurring medicinal plants previously abundant in the communities. Moreover, residents from relatively accessible settlements such as Lamlifew (Malungon) and Sitio Malaya (Kiamba) disclosed that they forsook traditional remedies in favour of modern medicines several years back, owing to the perceived superiority of the latter. Furthermore, the healers acknowledged the potential extinction threats of medicinal plants if conservation efforts are not implemented. Through continuous cultivation of IMPs in their backyards and home gardens, the healers inadvertently conserved these invaluable genetic treasures. Furthermore, the healers disclosed that to avert the extinction of IMPs in their natural habitats, they exclusively harvest what is necessary.

Discussion

Sarangani traditional healers: A dying breed?

One important feature of the Sarangani traditional healthcare system is the preponderance of female healers – a fact that emphasizes their vital roles in the traditional healing system. This association of women with traditional healing practices transcends cultures due to their nurturing nature and profound knowledge of plants that thrive in their communities. Struthers (2003) reported that Aboriginal women healers in the United States and Canada generally practised holistic healing using age-old methods passed down by their ancestors. However, the idea that women are the predominant healers is not a universal truth, as there are cultures where male healers serve as the primary practitioners (Gessler *et al*, 1995; Semenya and Potgieter, 2014). In Morocco, Bakker (1992) documented the notable rise of Berber women as traditional healers when their male counterparts lost influence due to political upheavals in the region. French colonization and the consequent succession of the Moroccan Makhzen in the political sphere stripped male healers of prestige, leading to the remarkable emergence of women healers to fill this healthcare vacuum. This is one case

Figure 3. Threats to medicinal plants and traditional healing knowledge in the Sarangani uplands: a, ongoing Alnigo (or swiddening); b, advanced ages of knowledge keepers and non-transmission of healing knowledge to succeeding generations; c, denudation of mountains; d, proliferation of modern medicines.

wherein politics, and not modernization, was the identified cause of shifting gender roles in the traditional healthcare system. Another study about Busoga healthcare practitioners in Uganda (Isiko, 2018) revealed that while men and women are equally accepted in traditional medicine, the power and influence that they wield depend on societal expectations relating to gender roles.

It must likewise be emphasized that traditional healers possess a critical understanding of medicinal plants, owing to their long years of experimentation and use. Acquiring and mastering utilization of such material requires significant time and effort. This knowledge, transmitted to them through generations, enables them to provide healing and therapy for poverty-stricken communities beyond the reach of basic social services. However, one factor that does not bode well for the perpetuation of traditional knowledge was the advanced ages of the Sarangani knowledge keepers. Similarly, Teves *et al* (2023) reported that Eskaya healers in Bohol, Philippines, are dwindling in numbers because of old age-related deaths. These are elderly people who become weak and decrepit with advancing years, lack of proper nutrition and inadequate health care due to the remoteness of their villages. Moreover, traditional healing knowledge, being passed on to chosen members of the family, limits its transmission, especially when the appointed successor does not show any interest in learning these skills. Consequently, when these elderly healers pass on, the priceless healing knowledge that they possess will forever be lost to humanity.

Sarangani Indigenous medicinal plants: Current state and major pressures

This study revealed a significant number of IMPs and a diverse repertoire of healing rituals and knowledge among the visited ethnic communities in the Sarangani uplands. This prevalent use of medicinal plants for healing in other areas of Mindanao has been reported in the Mount Matutum Protected Landscape (Alinsug *et al*, 2022), Agusan del Sur (Dapar *et al*, 2020), Davao Occidental (Paraguison *et al*, 2020; Cabugatan *et al*, 2022) and Surigao del Sur (Ilagan *et al*, 2022). Community inventory also revealed that Sarangani medicinal plants with high utilization were more frequently found closer to home. This fact underscores the direct correlation between the availability and cultural importance of medicinal plants to the Sarangani tribes. With greater accessibility, healers have more opportunities to test by trial and error the efficacy of potentially medicinal plants (Vandebroek *et al*, 2008). Furthermore, in Sarangani upland tribal communities, a direct relationship between the redundancy of medicinal plant use and the more prevalent physical ailments was established. This relationship was likewise borne out by studies done by Kunwar *et al* (2015),

Kumar and Bussman (2011), and Vandebroek *et al* (2008) in Nepal, India, and Bolivian Andes, respectively.

However, during recent years, the multi-fold effects of rapidly evolving social-ecological scenarios have severely impacted the Sarangani traditional agroecosystem, resulting in wide-scale natural resource losses. Moreover, species introductions from the lowlands resulted in a mosaic of Indigenous and non-indigenous medicinal plants, giving the healers a more expansive repertoire on which to experiment. Study results revealed, however, that IMPs (especially those thriving in the wild) are declining in terms of numbers in their natural habitats. Among the identified pressures leading to these losses are climate change-related devastation (such as heavy rains and flash floods), pest infestation, unregulated harvesting and forest denudation. In addition, IMPs thriving in transition areas and roadsides are more often subjected to anthropogenic pressures, resulting in their diminished numbers. While some IMPs are still available, several could no longer be found in the vicinity, and tribal healers reported travelling long distances to harvest them. In Limpopo Province, South Africa, locals identified over-harvesting and indiscriminate collection of IMPs as factors leading to large-scale losses (Mathibela *et al*, 2015). Moreover, several studies in Sri Lanka and Northern Ethiopia ascribe losses of IMPs to ecological devastation, climate change-associated devastation, agricultural expansion and human habitation, among other factors (De Silva and Wettasinghe, 2004; Mesfin *et al*, 2013). Owing to the burgeoning global population, the increased demand for modern medicines as well as the international trade for medicines by pharmaceutical companies, IMPs are disappearing in the wild because of overharvesting (Soetan and Aiyelaagbe, 2009; Oladele *et al*. 2011; Otieno and Analo, 2012; Ganie *et al*. 2015). Another study by Birhanu *et al* (2015) identified the shift to modern medicines as a factor leading to the extinction of IMPs. In the Sarangani uplands, conversations with locals revealed the use of modern medicines, such as paracetamol (for headaches), loperamide (for LBM), phenylephrine HCL or Neozep (for colds), Bioflu (colds and flu) and mefenamic acid (toothache). To prevent infection, tribal people open an amoxicillin capsule and sprinkle its contents on an open wound.

Accessibility of the tribal communities was also identified as a factor contributing to medicinal plant losses in the Sarangani uplands. Villages like Lamlifew (Malungon) and Sitio Malaya (Kiamba) are relatively accessible and can be reached by 4x4 vehicles and single motorcycles. The former community is dominated by B'laans and is frequently frequented by tourists and researchers for its School of Living Tradition for abaca weaving. Local inhabitants of Lamlifew enjoy the creature comforts of modernization (i.e. running water, electricity, cellphone service, internet service and cable television, among others). Sitio Malaya, on the other hand, is inhabited by T'bolis who manage a thriving abaca industry for local and international markets. All household needs (food and medicines) are purchased from the lowland markets. In both these places, traditional healing had been wholly abandoned for modern medicine. In a study done in the Himalayas, Tali *et al* (2014) identified agricultural expansion, road improvement, overgrazing and deforestation as threats leading to the extinction of medicinal plants. Furthermore, increased accessibility and foreign introductions resulted in a mosaic of Indigenous and non-indigenous medicinal plants in the Sarangani ethnic communities, resulting in a more diversified pharmacopoeia. Needless to say, medicinal plants remain the sole refuge of remote tribal communities in Sarangani Province, where the provision of basic social services is scanty at best and non-existent at worst. The need for self-sufficiency was highlighted during the COVID-19 pandemic (2019–2021) when these communities were isolated, travel was prohibited, food/resource supply chains were disrupted, and prices of commodities soared beyond the reach of many.

Furthermore, the practice of traditional healing was more prevalent in rural communities with very little/no access to modern medicines owing to geographical isolation and economic reasons (Maramba-Lazarte, 2020). In another study done in six middle-income countries (China, Ghana, India, Mexico, Russia and South Africa), Oyebode *et al* (2016) identified factors leading to the decline in the use of medicinal plants in these countries. These were the significant shifts in social trends and cultural beliefs, as well as the political support and provision of resources for training, practising and increasing public awareness of modern medicines.

Traditional healing knowledge: Headed for oblivion

In the remote Sarangani uplands, conversion to Christianity primarily resulted in the abandonment of belief in nature spirits. For some healers, however, some vestiges of spiritism persisted. In fact, this belief that human activities intersect with the goings-on in the spiritual world persists even with the incursion of modernization and Christianization. Similarly, local inhabitants in Namibia, South Africa and Bangladesh believe that nature spirits influence their health and that the efficacy of IMPs hinges on a complete understanding of their physical and spiritual purposes (Motaleb *et al*, 2010; Cheikhyoussef *et al*, 2011; Stofferson *et al*, 2011). This persistence of spiritism in healing practices, which has been labelled as pagan and backwards, has caused a significant number of Sarangani locals to forsake traditional medicine for modern medicine.

In the Sarangani uplands, losses of THK can be ascribed to additional pressures such as acculturation, education, increasing accessibility of the communities, the lure of modernization/technology, aging knowledge keepers (healers) and the oral nature of knowledge transmission that predisposes it to loss (Posey, 1996). Mahwasane *et al* (2013) concurred that the THK is gradually becoming extinct in the absence of a writing system and because healers do not keep written records. In South Africa, THK has been largely ignored because, being verbally transmitted, it could not be included in school curricula or policy documents (Mathibela *et al*, 2015). Consequently, huge volumes of THK worldwide remain undocumented, thereby underscoring now more than ever the need to record THK and conserve IMPs, before they completely disappear. In addition, the unwillingness of some healers (particularly in Kiamba and Maitum) to divulge healing knowledge that they consider secret information likewise exacerbates THK losses. This unwillingness to share THK was also reported by Kala *et al* (2006) and Giday *et al* (2003). Others are wary about researchers coming to their communities and extracting traditional knowledge that had been passed on to them by their ancestors. The lack of written

records due to the oral nature of knowledge transmission compounds these losses, while the secretive nature of some healers dooms knowledge transmission to permanent loss.

Further exacerbating these losses is the diminished intergenerational knowledge transfer of THK. Consequently, there is no mentoring for the next generation and, if the elderly healers pass on, their THK will be lost forever. In the Sarangani uplands, THK is losing its appeal among the younger generation, who consider these as backward, primitive, unchristian and inferior to modern medicine. In rural Malaysia, the oral nature of mentoring has also caused failure in THK transmission to the next generation (Batugal *et al*, 2004). Consequently, the older generation of healers will have no pool of recruits to inherit their knowledge and practices. Further contributing to the decline in the practice of THK is the proliferation of private and state-sponsored health facilities all over the province, especially in more accessible areas.

Future directions

Sarangani traditional healthcare system: Ways forward

In Sarangani Province, healers' knowledge results from learning and social interactions between knowledge keepers and those who seek it. However, the majority of THK practitioners still lack scientific justifications for their methodologies. Although they believe in the outcomes of their practices, they lack understanding of the mechanisms that produce these effects. Interviewed healers were unable to elucidate the mechanisms and efficacy of the remedies.

Bannerman (1977) emphasized the necessity of assessing THK through the lens of contemporary science, thereby enhancing beneficial practices and safeguarding against detrimental ones. Rather than depending on trial and error resulting from arbitrary screening methods, well-documented THK could assist scientists in identifying plants with potential medical characteristics. Scientific validation of IMPs through phytochemical, toxicological and pharmacological studies is thus warranted to eliminate quackery associated with THK, warn against inappropriate usage, and identify potential sources of medicines. However, one crucial matter that needs to be considered is the lack of coherence among healers about modes of preparation and proper dosages of IMPs (Wilcox and Bodeker, 2004). Once the effectiveness of THK is properly assessed, the next logical step is its integration into a nation's healthcare delivery system.

Integrating THK into national healthcare will preserve the Indigenous peoples' cultural heritage for future generations (Mahwasane *et al*, 2013). One way to do this is by including traditional healers in the crafting of a comprehensive healthcare strategy that is grounded in inclusivity, knowledge pluralism and sharing of information. In 1997, the Philippine government passed the Traditional and Alternative Medicine Act, which affirmed government support for THK (Maramba-Lazarte, 2020). Kaido (1997) proposed that traditional healers could be significantly used if they were organized and trained. Empowering traditional healers to participate in the crafting of local healthcare strategies for disadvantaged people whose views are oftentimes ignored during deliberations is also suggested. Moreover, policy positions relating to THK promotion should be more nuanced, and the use of IMPs must not be encouraged when there are reasons to doubt their effectiveness in comparison to modern medicine. Ultimately, collaborative initiatives by the scientific community and governments are essential to generate the impetus for such actions. Dissemination of pertinent research findings and facilitating further education and training programmes for traditional healers will empower them to document and share their significant Indigenous knowledge independently. The use of audiovisual media, espoused by Bidwell *et al* (2011), can serve as an effective means for knowledge transfer.

Biocultural conservation

A comprehensive inventory of IMPs must be undertaken before any conservation initiative commences. Soetan and Aiyelaagbe (2009) stated that IMPs, when not recorded or classified, can hamper conservation efforts. The importance of documenting IMPs as a prelude to their conservation was also espoused by Hamilton (2003). In the Sarangani upland communities, emerging sociocultural scenarios underscore the need for complementary *ex situ* and *in situ* conservation of IMPs. Through continuous utilization and cultivation of IMPs in their backyards and home gardens, the healers are unknowingly conserving these priceless genetic resources. The healers also practice judicious collection of IMPs and only harvest when these are needed. In a rural Nigerian community, Oladele *et al* (2011) reported that the cultivation of IMPs contributed to their conservation while Batugal *et al* (2014) espoused continuous cultivation as an efficient conservation strategy. In Africa, Okigbo *et al* (2008) maintained that only through sustainable conservation practices can a constant supply of IMPs be ensured for prolonged periods.

In Southern Italy, regional and national parks (considered protected areas) that feature medicinal plants conserved *ex situ* attract ethnobotanists from all over the country because of the joint preservation of IMPs and associated THK of the rural people (Menale *et al*, 2016). Along these lines, community-initiated *ex situ* conservation of Sarangani IMPs and THK must be undertaken in collaboration with academia, the local government and other stakeholders to preclude further losses of IMPs. In doing this, priority should be given to IMPs having redundant utilities since this will help offset pressure on IMPs that are considerably threatened in the wild.

As for the IMPs collected in the Sarangani uplands, these were sent to the Philippine national genebank for *ex situ* conservation. This backup collection will serve as an 'insurance policy' against the extinction of medicinal plant species amidst widespread environmental devastation in their natural habitats in the Sarangani uplands. Additionally, these IMPs are potential sources of new drugs for present and future health needs, not just of the locals but also of the communities at large. The need to preserve THK via knowledge transfer is crucial for the perpetuation of tribal culture and the promotion of local communities' rights over their plant genetic resources. Moreover, THK, when harmonized with frontier knowledge, may prove useful in addressing problems that may arise in the future. Recognizing the significance of the knowledge held by these healers and meticulously recording essential elements with sensitivity and respect for intellectual property rights are paramount in assisting the healers (Mathibela *et al*, 2015). Therefore, it is important to have a governance framework to effectively protect the THK of the tribes for posterity.

Limitations of the study

This study covered a limited number of study sites and gleaned valuable healing knowledge from a handful of traditional healers. Some of the sitios were inaccessible due to geographic isolation and difficult terrain, while others, unfortunately, still experience peace and order issues brought by insurgent groups, specifically in the towns of Banlas, Malapatan, Datal Anggas and Alabel. In addition, the danger posed by armed jihadist Muslim groups also prevented the research team from exploring areas in Maasim and Maitum. As for the healers themselves, some were unwilling to share healing knowledge in the belief that doing so would diminish their powers, while others were wary about sharing traditional knowledge with outsiders who they believe are merely out to extract information from them. Moreover, in some sitios, the locals had completely shifted to modern medicine and traditional healers were nowhere to be found. Furthermore, on the part of the researchers, the expensive and lengthy process of securing government permits for every community visited often discourages them from visiting more communities. Consequently, as some of the healers did not impart their THK, it is possible that some of the IMPs were not documented in this study.

Conclusion

This study unveiled the healing knowledge of local healers and the diversity of medicinal plants in the Sarangani uplands. Results, however, revealed that these IMPs are declining in numbers because of environmental degradation and the shift to modern medicines, among other pressures. Also revealed by this study are THK losses due to the advanced ages of knowledge keepers and the shifting preferences of the Sarangani locals towards modern medicine. To preclude further losses of IMPs, the following are recommended: (1) complementary *in situ* and *ex situ* conservation, (2) rehabilitation of the natural habitats of IMPs, (3) documentation of the THK of elderly healers before they pass on, (4) further scientific studies to validate THK, (5) establishment of a school of living traditions for the preservation of THK and its transmission to the younger generation, (6) protection of IMPs from exploitation and biopiracy through appropriate policy and legislation and (7) promotion of value-adding, utilization and widescale cultivation of IMPs. Having a continuous supply of IMPs (along with their associated traditional knowledge) to promote and support the health of the locals will bring the community (and eventually the Philippines) closer to the attainment of its Sustainable Development Goals (SDGs), such as improved health (SDG3), empowerment of the tribal women (SDG5), and self-sustainability (SDG11) in the face of disasters and isolation. Furthermore, when validated by scientific knowledge and promoted, these IMPs can be a potential livelihood source for tribal households, resulting in improved household income. Through these interventions, it is envisaged that the Sarangani IMPs and their associated THK will be preserved for future generations of tribes. Finally, the conservation of THK will likewise make tribal households more resilient against the ravages of climate change, pandemics, and other future problems and perturbations.

Traditional healers' organization, empowerment and participation are essential for developing suitable and culturally sensitive healthcare interventions for the local community.

Supplemental data

Supplemental Material 1. Questionnaire about Sarangani Indigenous medicinal plants

Acknowledgments

The authors extend their most profound thanks to the tribal healers who took part in the study.

Authors' contributions

FLZ conceptualized the study, supervised its implementation, revised the first draft, and co-wrote the final paper with MCBH, who also carried out confirmation of taxonomic identification of medicinal plants and finalization of the article. BMGB carried out the majority of field work, wrote the first draft and did data analysis. CRMR provided good quality pictures for the paper, helped in field collection and in the processing of the paper for publication while CHMA assisted in field preparations, collection and *ex situ* conservation of collected medicinal plants at the Philippine National Genebank.

Conflict of interest statement

The authors have declared that no competing interests exist.

Ethics statement

A free and prior informed consent for this study was incorporated in the application sent to the National Commission on Indigenous Peoples (NCIP) by Sarangani Province as part of its 'Ridge to Reef' project funded by the Peoples' Survival Fund.

References

Agduma, A.R., Garcia, F.G., Cabasan, M.T., *et al*, 2023. Overview of priorities, threats, and challenges to biodiversity conservation in the southern Philippines. Reg. Sustain 4, 203e213. https://doi.org/10.1016/j.regsus.2023.05.003

Alinsug, M.V., Estandarte, M.H.G., Somodio, E.M.N., Sabarity, M.J.J., Deocaris, C.C., 2022. Biodiversity of ethnomedicinal plants from the B'laan tribe in Mount Matutum Protected Landscape, Southern Mindanao, Philippines. Biodiversitas, 23:1, 554-563. https://doi.org/10.13057/biodiv/d230160

Ambasta, S.K., Kumari, S., Yadav, A.K., Trivedi, I., Prasad, B., Sinha, U.K., 2016. Medicinal plants of Bihar and its neighboring region which needs attention for their conservation. European Journal of Biomedical and Pharmaceutical Sciences, 3(4), 554-550. https://www.researchgate.net/publication/327210058_medicinal_plants_of_Bihar_and_its_neighboring_region_which_needs_attention_for_their_conservation

Aryal, K.K., Dhimal, M., Pandey, A., Pandey, A.R., Dhungana, R., Khaniya, B.N., Mehta, R.K., Karki, K.B., 2016. Knowledge

Diversity and Healing Practices of Traditional Medicine in Nepal. Kathmandu, Nepal: Nepal Health Research Council. https://www.researchgate.net/publication/313159390_Knowledge_Diversity_and_Healing_Practices_of_Traditional_Medicine_in_Nepal

Bakker, J., 1992. The Rise of female healers in the Middle Atlas, Morocco. Social Science and Medicine, 35(6); 819-829. https://doi.org/10.1016/0277-9536(92)90082-2

Bankole, A. E., Adekunle, A. A., Sowemimo, A. A., Umebese, C. E., Abiodun, O., Gbotosho, G. O., 2015. Phytochemical screening and in vivo antimalarial activity of extracts from three medicinal plants used in malaria treatment in Nigeria. DOI https://doi.org/10.1007/s00436-015-4747-x

Bannerman, R.H., 1977. WHO's programme: the approach will focus on the psychosocial and anthropological aspects of traditional medicine, on acupuncture and other healing methods, and on the claims made for herbs and medicinal plants. World Health, (November, 16-17). World Health Organization. https://iris.who.int/handle/10665/326066

Barata, A.M., Rocha, F., Lopes, V., Carvalho, A. M., 2016. Conservation and sustainable uses of medicinal plants and aromatic plants genetic resources on the worldwide for human welfare. Industrial Crops and Products. https://doi.org/10.1016/j.indcrop.2016.02.035

Batugal, P.A., Kanniah, J., Sy, L., Oliver, J.T., 2004. Medicinal plants research in Asia, Volume 1: The framework and project workplans. International Plant Genetic Resources Institute – Regional Office for Asia, the Pacific and Oceania (IPGRI-APO), Serdang, Selangor DE, Malaysia. ISBN 929043-615-8. https://cgspace.cgiar.org/server/api/core/bitstreams/edad56b4-e6ce-4c1e-9328-8eb4f0fe1d33/content

Belgica, T.H.R., Suba, M.D., Alejandro, G.J.G., 2024. Botanical assessment and conservation status of medicinal plants in mountain range of Malinao Albay, Philippines. Biodiversitas, 25:4, 1413-1419, DOI: https://doi.org/10.13057/biodiv/d250409

Bidwell, N.J., Winschiers-Theophilus, H., Koch Kapuire, G., & Chivuno-Kuria, S., 2011. Situated interactions between audiovisual media and African herbal lore. In Proceedings of the 2nd ACM SIGCHI Workshop on Walking Together to Design (pp. 1-4). ACM. https://doi.org/10.1007/s00779-010-0337-1

Birhanu, T., Abera, D., Ejeta, E., 2015. Ethnobotanical study of medicinal plants in selected Horro Gudurru Woderas, Western Ethiopia. Journal of Biology, Agriculture and Healthcare. ISSN 2224-3208. https://www.researchgate.net/publication/279512308_Ethnobotanical_Study_of_Medicinal_Plants_in_Selected_Horro_Gudurru_Woredas_Western_Ethiopia

Borokini, T.I., Ighere, D.A., Clement, M., Ajiboye, T.O., Alowonle, A.A., 2013. Ethnobiological Survey of Traditional Medicine Practice for Women's Health in Oyo State. Journal of Medicinal Plants Studies, 1(5): 17-29. https://www.plantsjournal.com/archives/2013/vol1issue5/PartA/2.1.pdf

Cabugatan, M.A.D., Ong, R.L.J.T., Mancao, L.S., *et al*, 2022. Ethnobotanical survey on medicinal plants used by the Manobo tribe of Don Marcelino, Davao Occidental, Philippines. Asian Journal of Biological and Life Sciences, 11, 492e504. https://doi.org/10.5530/ajbls.2022.11.67

Cheikhyoussef, A., Shapi, M., Matengu, K., Ashekele, H. M., 2011. Ethnobotanical study of indigenous knowledge on medicinal plant use by traditional healers in Oshikoto region, Namibia. Journal of Ethnobiology and Ethnomedicine. https://ethnobiomed.biomedcentral.com/articles/10.1186/1746-4269-7-10

Convention on Biological Diversity (CBD). Philippines - Country Profile. Biodiversity Facts: Status and trends of biodiversity, including benefits from biodiversity and ecosystem services. https://www.cbd.int/countries/profile?country=ph

Cordero, C.S., Meve, U., Alejandro, G.J.D., 2022. Ethnobotanical documentation of medicinal plants used by the indigenous Panay-Bukidnon in Lambunao, Iloilo, Philippines. Frontiers in Pharmacology, 12, 1-20. doi: https://doi.org/10.3389/fphar.2021.790567

Cotton, C. M., 1996. Ethnobotany: Principles and applications. Chichester, England: John Wiley and Sons. https://doi.org/10.1021/JM9701841

Dapar, M.L.G., Alejandro, G.J.D., Meve U., Liede-Schumann, S., 2020. Quantitative ethnopharmacological documentation and molecular confirmation of medicinal plants used by the Manobo tribe of Agusan del Sur, Philippines. Journal of Ethnobiology and Ethnomedicine, 16:14. https://doi.org/10.1186/s13002-020-00363-7.

De Silva, M. A., Wettasinghe D. T., 2004. Sri Lanka conservation and sustainable use of medicinal plants. IUCN – The World Conservation Union Sri Lanka. https://iucn.org/sites/default/files/2022-05/sri_lanka_conserv_sustain_use_med_plants.pdf

Fiscal, R.R., 2017. Ethnomedicinal plants used by the traditional healers in Laguna, Philippines. Asia Pacific Journal of Multidisciplinary Research, 5(4), 132-137. https://www.researchgate.net/publication/322929882_Ethnomedicinal_Plants_Used_by_Traditional_Healers_in_Laguna_Philippines

Ganie, S. H., Upadhyay, P., Das, S., Sharma, M.P., 2015. Authentication of medicinal plants by DNA markers. Plant Gene, 4, 83-99. https://doi.org/10.1016/j.plgene.2015.10.002

Gessler, M.C., Msuya, D.E., Nkunya, M.H.H., Schir, H.M., Tanner, M., 1995. Traditional healers in Tanzania: Socio-cultural profile and three short portraits. Journal of Ethnopharmacology, 48, 145-160. https://pubmed.ncbi.nlm.nih.gov/8719975/

Giday, M., Asfaw, Z., Elmqvist, T., Woldu, Z., 2003. An ethnobotanical study of medicinal plants used by the Zay people in Ethiopia. Journal of Ethnopharmacology, 85(1):43–52. PMID: 12576201. https://pubmed.ncbi.nlm.nih.gov/12576201/

Hamilton, A., 2003. Medicinal plants, conservation and livelihoods. Biodiversity and Conservation, 13: 1477-1577. https://www.researchgate.net/publication/227147775_Medicinal_Plants_Conservation_and_Livelihoods

Ilagan, V.A.D., Alejandro, G.J.D., Paraguison, D.J.B., *et al*., 2022. Ethno-pharmacological documentation and molecular authentication of medicinal plants used by the Manobo and Mamanwa tribes of Surigao del Sur, Philippines. Biodiversitas, 23, 3185e3202. https://doi.org/10.13057/biodiv/d230646

Isiko, A.P., 2018. Gender roles in traditional healing practices in Busoga. Retrieved from https://hdl.handle.net/1887/63215

Kaido, T.L., Veale, J.H., Havlika, I., Rama, B.K., 1997. Preliminary screening of plants used in South Africa as traditional herbal remedies during pregnancy and labour. Journal of Ethnopharmacology, 55, 145-160. https://pubmed.ncbi.nlm.nih.gov/9080339/

Kala, C.P., Dhyani, P.P., Sajwan, B.S., 2006. Developing the

medicinal plants sector in northern India: challenges and opportunities. Journal of Ethnobiology and Ethnomedicine, 2:32. https://ethnobiomed.biomedcentral.com/articles/10.1186/1746-4269-2-32

Kumar, M., Bussmann, R.W., 2011. Ethnomedicinal and ecological status of plants in Garhwal Himalaya, India. Journal of Ethnobiology and Ethnomedicine, BioMed Central Ltd; 7(1):32. Available from: http://www.ethnobiomed.com/content/7/1/32

Kunwar, R., Acharya, R.P., Chaodhary, C.L., 2015. Medicinal plant dynamics in indigenous medicines in farwest Nepal. Journal of Ethnopharmacology, 163:210–9. PMID: 25655999. https://doi.org/10.1016/j.jep.2015.01.035

Mahmoud, T., Gairola, S., 2013. Traditional Knowledge and Use of Medicinal Plants in the Eastern Desert of Egypt: A Case Study from Wadi El-Gemal National Park. Journal of Medicinal Plants Studies, 1, 10-17. https://www.scirp.org/reference/referencespapers?referenceid=3143895

Mathibela, M.K., Egan, B.A., Du Plessis, H.J., Potgieter, M.H., 2015. Socio-cultural profile of Bapedi traditional healers as indigenous knowledge custodians and conservation partners in the Blouberg area, Limpopo Province, South Africa. Journal of Ethnobiology and Ethnomedicine, 11:49. DOI https://doi.org/10.1186/s13002-015-0025-3

Mahwasane, S.T.L., Middleton, L.N., Boaduo, N., 2013. An ethnobotanical survey of indigenous knowledge on medicinal plants used by the traditional healers of the Lwamondo area, Limpopo province, South Africa. South African Journal of Botany, 88, 69–75. https://www.ethnopharmacologia.org/prelude2020/pdf/biblio-hm-55-mahwasane.pdf

Maramba-Lazarte, C. 2020. Benefits of Mainstreaming Herbal Medicine in the Philippine Healthcare System. Acta Medica Philippina, 54(1). https://doi.org/10.47895/amp.v54i1.1078

Menale, B., De Castro, O., Cascone, C., Muoio, R., 2016. Ethnobotanical investigation on medicinal plants in the Vesuvio National Park (Campania, Southern Italy). Journal of Ethnopharmacology. http://dx.doi.org/10.1016/j.jep.2016.07.049

Mendoza, L.A., Lagbas, A.J., Buot Jr., I.E., 2016. Conservation status of the plant species in selected areas with frequent human activities in Roosevelt Protected Landscape, Bataan, Luzon Island, Philippines. Thailand Nat His Mus J, 10(2), 79-115. https://www.researchgate.net/publication/312551606_Conservation_Status_of_the_Plant_Species_in_Selected_Areas_with_Frequent_Human_Activities_in_Roosevelt_Protected_Landscape_Bataan_Luzon_Island_Philippines

Meniza, J.F., Pasco, M.M., Alimbon, J.A., 2024. A review of ethnobotanical studies reveals over 500 medicinal plants in Mindanao, Philippines. Plant Diversity, 46, 551-564. https://pubmed.ncbi.nlm.nih.gov/39290882/

Mesfin, K., Tekle, G., Tesfay, T., 2013. Assessment of threatening factors of medicinal plants species in Samre District, South-eastern Tigray, Northern Ethiopia. Journal of Medicinal Plant Studies, 1, 38-42. ISSN 2320-3862. https://www.plantsjournal.com/vol1Issue1/Issue_july_2013/14.1.pdf

Miano, R.S., Picardal, J.P., Alonso, C.A., Reuyan, D., 2011. Ethnobotanical inventory and assessment of medically-important plant roots in Cebu Island, Philippines. Asian Journal of Biodiversity. CHED Accredited Research Journal, Category A, 2, 81-102. ISSN 2094-1519. https://www.researchgate.net/publication/314913900_Ethnobotanical_Inventory_and_Assessment_of_Medically-Important_Plant_Roots_in_Cebu_Island_Philippines

Motaleb, M.A., Firoz, R., Adrika, A., Khan, N.A., 2010. Approaches to conservation of medicinal plants and traditional knowledge: A focus on the Chittagong Hill Tracts. International Union for Conservation of Nature, Keidanren Nature Conservation Fund, Bangladesh Country Office, Dhaka, Bangladesh, pp viii+30. https://portals.iucn.org/library/sites/library/files/documents/2010-030.pdf

Okigbo, R.N., Eme, U.E., Ogbogu, S., 2008. Biodiversity and conservation of medicinal and aromatic plants in Africa. Biotechnology and Molecular Biology Reviews, 3(6), 127-134. ISSN 1538-2273. https://academicjournals.org/journal/BMBR/article-full-text-pdf/2EFEB8C40238

Oladele, A. T., Alade, G. O., Omobuwajo, O.R., 2011. Medicinal plants conservation and cultivation by traditional medicine practitioners (TMPs) in Aiyedaade Local Government Area of Osun State, Nigeria. Agriculture and Biology Journal of North America. ISSN 2151-7517. https://doi.org/10.5251/abjna.2011.2.3.476.487

Ong, H.G., Kim, Y.D., 2014. Quantitative ethnobotanical study of the medicinal plants used by the Ati Negrito indigenous group in Guimaras Island, Philippines. Journal of Ethnopharmacology, 157, 228-242. https://doi.org/10.1016/j.jep.2014.09.015

Otieno, N. E., Analo, C., 2012. Local indigenous knowledge about some medicinal plants in and around Kakamega forest in western Kenya. F1000Research. https://doi.org/10.12688/f1000research.1-40.v2

Oyebode, O., Kandala, N.B., Chilton, P.J,, Lilford, R.J., 2016. Use of traditional medicine in middle-income countries: a WHO-SAGE study. Health Policy and Planning, 31, 984–99, doi: https://doi.org/10.1093/heapol/czw022

Paraguison, L.D.R., Tandang, D.N., Alejandro, G.J.D., 2020. Medicinal plants used by the Manobo tribe of Prosperidad, Agusan del Sur, Philippines: An ethnobotanical survey. Asian Journal of Biological and Life Sciences, 9, 326e333. https://doi.org/10.5530/ajbls.2020.9.49.

Posey, D.A., 1996. Protecting indigenous peoples' rights to biodiversity. Environment: Science and Policy for Sustainable Development, 38(8), 6-45. DOI: https://doi.org/10.1080/00139157.1996.9930990.

Semenya, S.S., Potgieter, M.J., 2014. Bapedi traditional healers in the Limpopo Province, South Africa: Their socio-cultural profile and traditional healing practice. Journal of Ethnobiology and Medicine, 10(4). https://ethnobiomed.biomedcentral.com/articles/10.1186/1746-4269-10-4

Soetan, K. O., Aiyelaagbe, O. O., 2009. The need for bioactivity-safety evaluation and conservation of medicinal plants – A review. Journal of Medicinal Plants Research, 3(5), 324-328. ISSN 1996-0875. https://www.researchgate.net/publication/242035192_The_need_for_bioactivity-safety_evaluation_and_conservation_of_medicinal_plants_-_A_review

Stofferson, A., Winstrup, M., Nieminen, R., Allerton, T., 2011. The sustainability of medicinal plant use in the local culture in Ongeluksnek, Eastern Cape, South Africa. Interdisciplinary Land Use and Natural Resource Management. Faculty of Life Sciences, University of Copenhagen. https://sluse.dk/project/South-Africa_medicinal_plants_and_traditional_healing_in_contemporary_rural_south_africa.pdf

Struthers, R., 2003. The Artistry and Ability of Traditional Women Healers, Health Care for Women International, 24:4, 340-354. https://pubmed.ncbi.nlm.nih.gov/12746005/

Tali, B.A., Ganie, A. H., Nawchoo, I. A., Wani, A. A., Reshi, Z. A., 2014. Assessment of threat status of selected endemic medicinal plants using IUCN regional guidelines:

A case study from Kashmir Himalaya. Journal for Nature Conservation. https://doi.org/10.1016/j.jnc.2014.06.004

Teves, R.M., Tangtengco, O.A.G., Sumatra, R.U., Carag, H.M., Isidro-Lapena, J.S., 2023. Ethnomedicinal survey of valuable plants used by Eskaya traditional healers in Bohol Island, Philippines. Acta Med Philipp., 57(3), 17-27. https://doi.org/10.47895/amp.vi0.3883.

Vandebroek, T.E., Goetghebeur, P., Sanca, S., Arrazola, S., Van Damme, P., 2008. The relationship between plant use and plant diversity in the Bolivian Andes, with special reference to medicinal plant use. Human Ecology, 36(6):861–79. https://www.researchgate.net/publication/226977713_The_Relationship_Between_Plant_Use_and_Plant_Diversity_in_the_Bolivian_Andes_with_Special_Reference_to_Medicinal_Plant_Use

Wilcox, M.L., Bodeker, G., 2004. Traditional Herbal Medicines for Malaria. British Medical Journal, 329, 1156-1159. https://doi.org/10.1136/bmj.329.7475.1156

World Health Organization (WHO), International Union for Conservation of Nature and Natural Resources (IUCN), World Wide Fund for Nature (WWF), 1993. Guidelines on conservation of medicinal plants. ISBN 2-8317-0136-8. https://iris.who.int/bitstream/handle/10665/41651/2831701368_en.pdf

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 73–82
DOI: 10.46265/genresj.BPGX5961
https://www.genresj.org
ISSN: 2708-3764

Exploring the situation of transboundary breeds in Europe for their effective management and conservation

Eléonore Charvolin[a,†,*], Dimitrios Tsiokos[b,†,*], Grégoire Leroy[c] and Christina Ligda[d]

[a] *Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France*
[b] *Hellenic Agricultural Organization–DIMITRA, RIAS, 58100 Paralimni Giannitsa, Greece*
[c] *Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla, 00153 Roma, Italy*
[d] *Hellenic Agricultural Organization–DIMITRA, VRI, 57001 Thermi-Thessaloniki, Greece*

Abstract: Geographical distribution plays a crucial role in the effectiveness of breeding and conservation programmes, especially for livestock breeds with a small population size. Among these, transboundary breeds present unique challenges and opportunities for conservation efforts. This study specifically examines the case of transboundary breeds in Europe and the associated challenges. Population and descriptive data were sourced from the Domestic Animal Diversity Information System (DAD-IS) to assess their current state of monitoring and management. The analysis revealed that 42% of the 6,460 National Breed Populations reported in Europe are transboundary, with 25% occurring exclusively within the region (Europe). Alarmingly, 85% of European transboundary breeds are classified as 'at risk' or have an 'unknown' conservation status, a fact that further accentuates the urgent need for improved sustainable management. This paper identifies key data gaps, for instance related to common understanding of concepts used by managers, and proposes improvements to enhance the monitoring, conservation and management of transboundary breeds in Europe.

Keywords: Transboundary breeds, DAD-IS, conservation, animal genetic resources, livestock diversity

Citation: Charvolin, E., Tsiokos, D., Leroy, G. and Ligda, C. (2025) "Exploring the situation of transboundary breeds in Europe for their effective management and conservation", *Genetic Resources*, 6(12), pp. 73–82. doi: 10.46265/genresj.BPGX5961.

Introduction

Livestock diversity in Europe has been shaped by the collective management and selection practices of farmers, evolving within the region's political, economic and social environment. In Europe, systematic performance recording, animal identification and pedigree recording have contributed to the development of numerous livestock breeds (FAO, 2015). However, changes in market demand, environmental factors and political instability continue to impact farmers' selection decisions, directly influencing farmers' preferences and have a direct impact on a breed's population size (Cao et al, 2021; Verrier et al, 2015). A variety of migration, introgression and isolation events over time have influenced the domestic diversity within and among countries (Leroy et al, 2015). While some breeds remain confined to specific geographical areas, others are distributed across multiple countries or even globally. It is estimated that, on a global scale, half of all breeds are transboundary, i.e. shared between at least two different countries (FAO, 2024a; see Box 1 for definitions). Transboundary breeds (TBs) may be native to their current locations or introduced as exotic breeds, recently or a long time ago, sufficient to be considered adapted to the local environment (Box 1).

The conservation of animal genetic resources (AnGR) for food and agriculture became a priority in the late 20th century, driven by concerns over the loss of genetic diversity. This was motivated by observant breeders in some countries, who realized the risk early on. The next steps were taken by scientists, providing solid scientific evidence and contributing to enhanced awareness concerning the threats to global biodiversity. The understanding that the disruption of

* Corresponding authors: Eléonore Charvolin (eleonore.charvolin-lemaire@inrae.fr), Dimitrios Tsiokos (tsiokosd@gmail.com)
† Equal contribution, listed in alphabetical order

Received: 24.01.2025 | Accepted: 07.07.2025 | Published online: 26.09.2025

Box 1. Definitions of breed and breed classification concepts

Breed (FAO, 2000): is a homogenous, subspecific group of individuals with definable and identifiable external characters that enable it to be separated by visual appraisal from other similarly defined groups within the same species. Alternatively, it is a homogeneous group where geographical separation from phenotypically similar groups has led to general acceptance of its separate identity.

National Breed Population (NBP) (FAO, 2021): The NBP refers to the animals belonging to one specific breed in one country. In case it is a local breed, occurring only in one country, it represents the breed. In case it is part of a transboundary breed, occurring in more than one country, it represents only a part of a breed.

Adaptedness classification (FAO, 2015): The DAD-IS adaptedness classification differentiates exotic breeds from native/indigenous breeds and from locally adapted breeds. The latter correspond, in the classification, to non-native breeds that have become adapted to the local environment and/or production system.

Locally adapted breeds have been initially defined as breeds of exotic origin imported into the national territory from another country that show limited genetic relationship with the original population. Later the definition has been broadened as breeds which have been in the country for a sufficient time to be genetically adapted to one or more of the traditional production systems or environments in the country.

Native Breed (FAO, 2001): (also termed Indigenous Breeds, autochthonous) originating from, adapted to and utilized in a particular geographical region, form a subset of the **Locally Adapted Breed** (broader definition). It refers to a breed in its country of origin (i.e. the country where the breed was created originally from genetic material that was available when the initial breed development commenced). It is important to note that a breed may be a native breed in more than one country depending on the history of the breed.

historical links between breeders and the breeds developed by their ancestors had led to a decline in those breeds' population was the driving force for taking actions to interrupt this trend. Threats such as indiscriminate crossbreeding, the introduction of exotic breeds, weak institutional policies and economic challenges continue to endanger genetic diversity (FAO, 2015). Conservation efforts encompass a wide range of activities, including awareness raising, strengthening governance, adding value to natural resources and implementing conservation schemes (Ligda et al, 2013; Sponenberg et al, 2019). Genetic diversity losses and conservation issues are not limited to local breeds, as international breeds can also be subject to genetic erosion phenomena (Ablondi et al, 2022). Relevant strategies for AnGR management also concern international breeds due to their importance in ensuring food security and public health globally. Sustainable management of both local and international breeds is essential for global food security and agricultural resilience (Lefevre et al, 2024). Europe initiated conservation efforts in the late 1970s; however, global consensus on preservation of livestock diversity only emerged after the publication of the first report on the *State of the World's Animal Genetic Resources for Food and Agriculture* (FAO, 2007a). This led to the development of the *Global Plan of Action for Animal Genetic Resources* (GPA) (FAO, 2007b).

This framework outlines four strategic priorities: (1) characterization, inventory and monitoring of the breeds, (2) sustainable use and development, (3) conservation, and (4) policies, institutions and capacity building. The Domestic Animal Diversity Information System (DAD-IS), developed by the Food and Agriculture Organization of the UN (FAO), serves as a comprehensive database supporting these conservation and management efforts worldwide. DAD-IS contains information on breed characteristics, uses, geographic distribution and demographics, and images and tools for generating reports. The recorded information, based on agreed definitions on terms and classification criteria and provided by nominated National Coordinators, enables the monitoring of breeds at global, regional and national levels (Polack et al, 2022). Moreover, DAD-IS is the data source for two UN Sustainable Development Goals (SDG) indicators (2.5.1b and 2.5.2), as approved by the FAO Commission on Genetic Resources for Food and Agriculture (CGRFA) (FAO, 2011; FAO, 2013). In particular, SDG indicator 2.5.2 takes into account the proportion of breeds considered at risk of extinction, on the basis of their population size estimation. Initially focusing on local breeds, the focus of the indicator is currently extended to transboundary ones (FAO, 2024b).

Efficient breeding and conservation programmes for breeds with small population sizes are particularly challenging due to several reasons (Biscarini et al, 2015). These breeds often face geographic constraints, policy limitations and economic disadvantages, creating a cycle that restricts their development. However, the presence of such breeds across multiple countries presents opportunities for collaborative management. Understanding breed distribution across Europe, along with different environments and breeding conditions, can improve resilience assessments and utilization of genetic resources. Advances in molecular genetics enable systematic genomic studies that can further support the classification and management of TBs.

Definitions of specific terms used in this article are given to help the reader understand the situation of TBs (Box 1).

This study aims to explore the current status of TBs in Europe using DAD-IS data. It provides information on the distribution of TBs around Europe, their Risk Status at local and regional level and adaptedness classifications. The outcomes reveal the complexity of factors that impact the decisions concerning the management of TBs and focusses on certain areas for further analysis, using a case study approach.

Material and methods

Data for this analysis were extracted from FAO's DAD-IS using the 'Data export' tool. The data used include the descriptive and population data files, as well as the transboundary list file. The study included all 31 registered species in the 45 countries classified as 'Europe and Caucasus' (FAO, 2007b). All data was provided by the National Coordinators (NCs) appointed by the countries. The analysis was based on the population and historical data available as of 8 August 2023.

The analysis was built on the **National Breed Populations (NBPs)**, as defined in Box 1. The descriptive data included information on the Risk Status, Geographic and Adaptedness classification provided by DAD-IS. The population data file contained the information on the population size of the breeds for all years with recorded data. Based on this information, an NBP was classified into the Risk Status categories. An NBP was reported as TB when this NBP was linked (by the NC) to a breed in the TB list provided in DAD-IS. This is done automatically by DAD-IS, without taking into consideration the number of NBPs linked to the transboundary name. DAD-IS classified NBPs as local (reported in only one country), regional transboundary (reported as TB, only by countries from a given region, e.g. Europe), or **international transboundary** (reported as TB by countries from multiple regions). Breeds were further classified by DAD-IS as **native, locally adapted, or exotic**, based on reported data (Box 1).

Statistical comparisons were made to assess patterns in breed distribution, risk status, and adaptedness. Associations between categorical variables were analyzed using the Chi-square test via the Crosstabs procedure in SPSS version 27.0 (IBM Corp., 2020). Statistical significance was set at $p < 0.001$. Data gaps and inconsistencies were analyzed to identify possible reporting issues and areas for improvement.

Results

Assessment of the current situation

Among the 6,460 NBPs analyzed, 3,761 (58.2%) are local breeds, while 2,699 (41.8%) are transboundary. Of the transboundary NBPs (Figure 1), 674 (25%) are European TBs (found only within Europe), and 2,025 (31.3%) are international TBs (present in at least one non-European region).

These findings highlight the importance of coordinated conservation strategies within Europe. The presence of TBs across multiple countries suggests opportunities for shared conservation programmes.

TBs are reported by all countries in Europe, as shown in Figure 2. The TB share among NBPs exhibits considerable variation, ranging from 11,6% (Spain) to 77,1% (Ireland). No significant correlation is found between the share of TBs among NBPs and the number of NBPs reported by the countries.

The number of countries that declare a specific TB linked to their NBP may reflect the dynamics of cooperation among

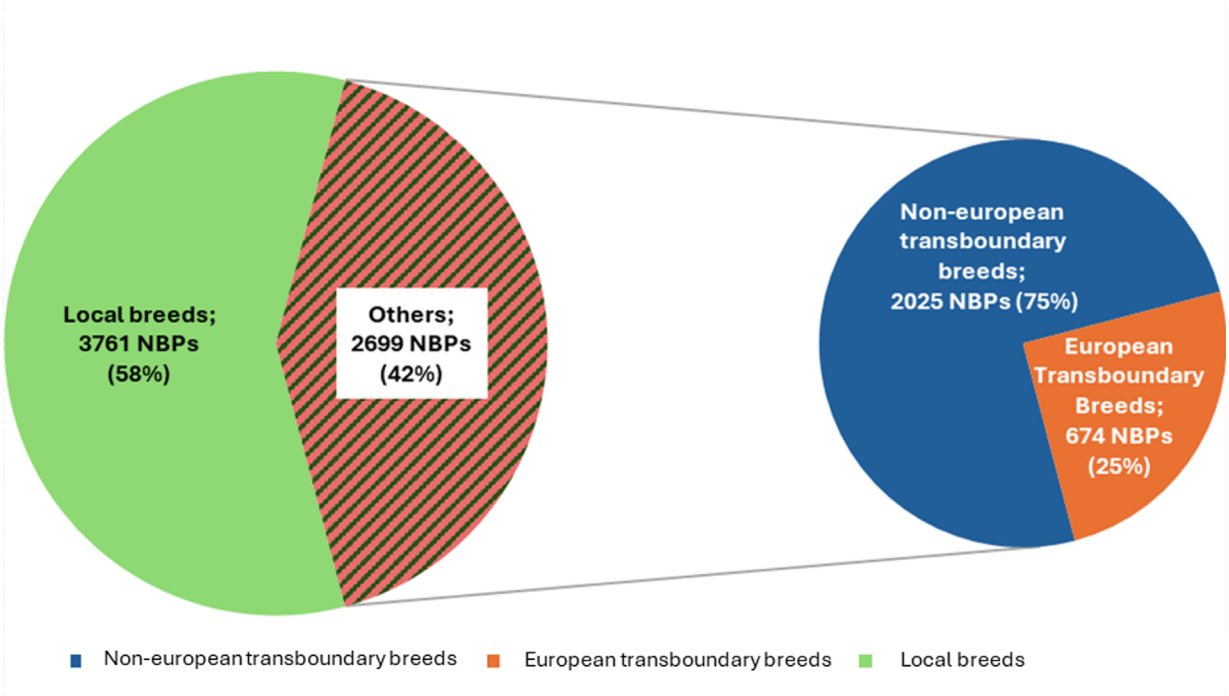


Figure 1. Repartition of the European National Breed Populations (NBPs) following their geographical classification, their corresponding number of NBPs (in brackets, the corresponding proportions). Data source DAD-IS.

the countries. However, further information on the historical evolution of the NBPs, the ongoing breeding or conservation programmes, organizational and social aspects is required.

Figures 3a and 3b provide information on the distribution of regional TBs between countries, in mammalian and avian species. In Figure 3a, the results are presented separately for the five mammalian species (cattle, goat, horse, pig and sheep), for which breed-related information is reported by almost all countries in the European region (cattle and sheep in 100% of countries, goats in 93%, and horses and pigs in 91%) (FAO, 2007a). Similarly, the results in Figure 3b concern avian species, for which breed-related information is reported by 50% or more European countries.

These figures present a general overview of the distribution of TBs at species level and do not infer a species impact on the number of countries linked to a TB.

Figure 3a (mammalian species) shows that a significant number of TBs is linked to only one NBP. This percentage varies between 25.9% (21 NBP) for sheep and 50% (7 NBP) for swine. In cattle, 13 (41.9%) of the 31 European TBPs are reported by only one country. Similar trends are observed in avian species (Figure 3b); however, this is referred to an overall smaller number of TBs and lower number of countries reporting breed data on these species. Those results underline a substantial gap between breeds classified as TBs according to DAD-IS rules and the fact that those breeds are actually reported by more than one country.

The fact that an NBP has not been linked to an existing TB name could be attributed to a different perception of the concept of TB; however, several other reasons may also exist. Therefore, a case study approach would be required to better understand the situation and propose a step forward.

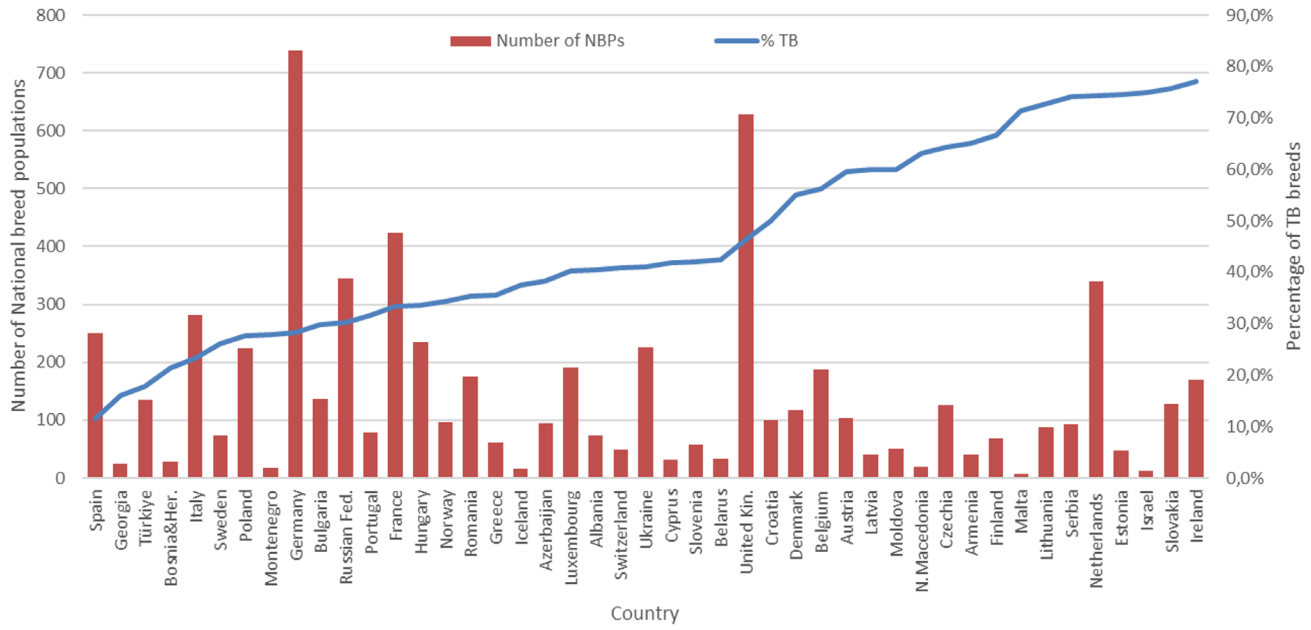


Figure 2. Number of National Breed Populations (NBPs) and percentage of Transboundary Breeds (TBs) declared per European country, ordered by increasing share of TBs.

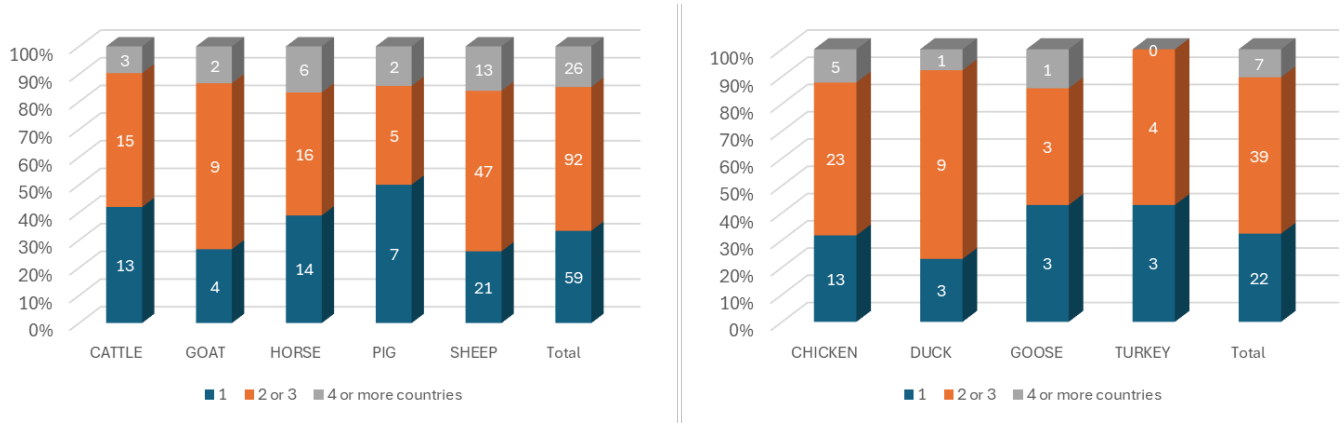


Figure 3. Regional transboundary breeds reported by one, two or three, and four or more countries across mammalian (a) and avian species (b).

Adaptedness classification

Considering the classification by adaptedness, from the total of 2,699 registered (regional and international) NBPs, 15% were classified as native, 10% as locally adapted and 36% as exotic (Table 1). The remaining 38% NBPs were not classified in any category, as no relevant data has been reported by any country. The substantial number of NBPs with no information on adaptedness classification in Europe (260 NBPs only reported in Europe) indicate that further progress is still needed in this domain. The high percentage of unclassified breeds suggests the need for clearer reporting standards across countries.

Nevertheless, compared to the corresponding figures at global level (45% of non-available classification, FAO 2023), the results of the European region are more complete. The high percentage of non-classified NBPs is attributed to several capacity or policy-related factors, reflecting the organization of AnGR management in each reporting country and the specific considerations on the definition of terms, at country level. Nevertheless, this gap may hinder the efficient management of AnGR, as the missing information could result in misleading reports and biased decisions. From the remaining, 26% (177 NBPs) are classified as exotic, thus they are considered as introduced from another country and have not been developed for sufficient time in the country.

Compared to international ones, European TBs display a larger proportion of locally adapted and native breeds ($P < 0{,}0001$), which could be partially explained by the fact that locally adapted and native breeds are more regionally distributed than the exotic ones, which have a wider global expansion.

The distribution of NBPs in adaptedness classes in European countries can be considered indicative of the current picture of breeds; however, an in-depth analysis of relevant case studies is needed. The data show that in 14 countries no information is available concerning the adaptedness classification of their recorded NBPs (linked to TB), while in 21 countries, more than 80% of their NBPs have recorded information on the adaptedness classification (Figure 4).

Table 1. Number and percentage of international and regional national breed populations linked to transboundary breeds in Europe in DAD-IS (classified as native/locally adapted/exotic)

	Breed classification (geographic)		
Breed classification (adaptedness)	**International**	**Regional (Europe)**	**Total**
	No (%)	**No (%)**	**No (%)**
Native	225 (11)	184 (27)	409 (15)
Locally adapted	218 (10)	53 (7)	271 (10)
Exotic	809 (39)	177 (26)	986 (36)
Non-available	773 (38)	260 (38)	1,033 (38)
Total	2,025 (100)	674 (100)	2,699 (100)

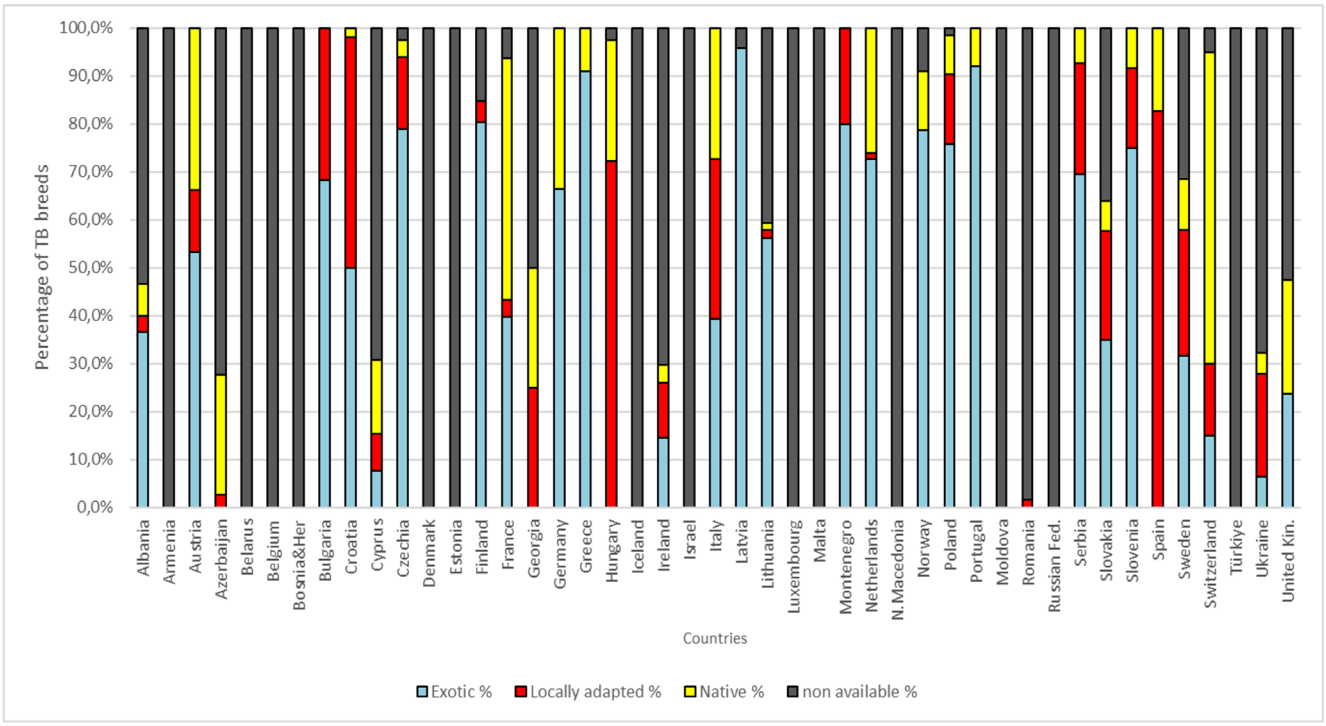


Figure 4. Adaptedness classification of national breed populations linked to transboundary breeds (TB) in Europe

Risk status

Specific thresholds differentiated by species are defined in various countries through national and international regulations that provide the framework for AnGR management. The current analysis is based on the local risk status, following FAO's classification system, which is recognized as the global reference according to CGRFA (FAO, 2013). In the case of TBs, in DAD-IS the risk status is estimated at local, regional and global levels, provided that the country has been reporting population data for the past ten years. The estimation of risk status at different levels, in the case of TBs, could support their efficient conservation management, as different approaches are needed when a TB is at risk in one country, but not at risk in other(s), or when the TB is at risk everywhere.

Among the 674 NBPs found in more than one European country, 44.2% were considered at risk in at least one country (Table 2). Moreover, considering that the existence of sufficient recorded data and knowledge of population trends is a key point for the efficient management of breeds, 45% of NBPs in the 'unknown' category, i.e. without population data provided over the last ten years, could be interpreted as a first sign of being at risk. Therefore, this could be considered an early indicator of the current or potential risk at national level of the nearly 90% of NBPs that are transboundary (reported only in Europe). However, a more in-depth analysis of these data is needed, as NBPs (from TBs) could be at risk in one country and not in others, as previously mentioned.

Therefore, further examining the data in Table 3, the NBP's risk status was calculated under different combinations for the five species considered in this study. In Table 3, it is shown that 42% of goat and 30% of sheep NBPs linked to a TB, were reported at risk in all countries. When the 'unknown' category is added, this percentage increased to 75% and 70%, respectively. The higher percentage of sheep and goat NBPs that were at risk in all countries could be explained, at least to some extent, by the different evolution of sheep and goat populations compared to cattle and pigs, for which intensification has been more widespread.

Table 2. Local risk status of all national breed populations related to TBs reported by at least two countries in Europe.

Breed classification (geographic)		At Risk	Cryoconserved only	Extinct	Not at Risk	Unknown	Total
International	Count	844	2	108	230	841	2,025
	Row N %	41.7	0.1	5.3	11.4	41.5	100.0
Regional	Count	298	1	33	39	303	674
	Row N %	44.2	0.1	4.9	5.8	45.0	100.0
Total	Count	1,142	3	141	269	1,144	2,699
	Row N %	42.3	0.1	5.2	10.0	42.4	100.0

Table 3. Local risk status of transboundary breeds present in more than one country, (%) by species

	Not at risk in at least one country (%)	At risk in all countries (%)	At risk or unknown or extinct in all countries (%)
Cattle	11.76	5.88	88.24
Goat	8.33	41.67	91.67
Horse	22.73	9.09	77.27
Pig	14.29	14.29	85.71
Sheep	26.67	30.00	73.33
Total	21.19	22.88	78.81

The diversity of cases is presented with selected examples in Table 4. The Hutsul horse reported as TB by eight countries, is at risk in six countries and unknown in the remaining two, due to lack of data. The Hutsul horse is categorized as native in Hungary, Austria and Poland, while it is defined as locally adapted in Slovakia and Czechia, and exotic in Germany. The field was not filled in the remaining two countries. The TB is at risk at regional level. The case of Precoce sheep, reported by France (native), Spain (locally adapted) and Portugal (exotic) is a similar case, as the breed is at risk at national and regional levels. However, the case of the Ouessant sheep – present in six countries, native in France and exotic in the Netherlands, Germany and Czechia (no classification in Belgium and Denmark) – differs as it is at risk in all countries, but the breed is not at risk at regional level. The Podolian cattle represents a case where the breed is developed in the country of origin (Italy), bred in high numbers, while a small population (at risk) is bred in another country (Serbia). In the DAD-IS data analyzed, there was no information on the links and exchanges between breeders' associations or on the genetic differences between NBPs.

Discussion

This study, based on the data reported in DAD-IS, provides a general overview of the status of TBs in Europe and reflects the quality of data and level of recorded information. In Europe, NBPs linked to TBs account for a large part (42%) of NBPs reported. This number could reach up to 80% of NBPs in some countries, with varied percentages across countries. The interpretation of this variation is not straightforward, as various factors contribute to this picture. AnGR have evolved under specific conditions (physical, technical, social, political and organizational) in each country, forming the current breeds and links among countries. Furthermore, AnGR management is under a country's sovereignty, and national decisions and measures are in accordance with global (FAO) and European guidelines.

Approaches to identify the link between NBPs and TBs may differ among countries according to national needs and perceptions of terms. Our results highlight that, depending on the species, a percentage between 20% and 50% of European TBs are reported by only a single country. This raises concerns about data accuracy and consistency. Although this issue appears less significant at the global level – where only 17% of TBs are reported by a single country – it remains noteworthy in the European context. In Europe, several such breeds are not linked to the TB list by their country of origin. Examples include the Turopolje pig from Croatia, the Asino Sardo donkey from Italy and the Olkusz sheep from Poland. In some cases, this cannot be considered a reporting gap, as it is a justified decision by the NC. Such decisions could be based on specific national policies and priorities. Certain countries could have chosen to classify their native breeds as local, given that global indicators mainly focus on local breeds.

Table 4. Transboundary breed cases examples

Species	Transboundary breed name	Local breed name	Country	SDG local risk status	Adaptedness	Regional risk status
Horse	Hutsul	Huculł	Poland	At risk	Native	At risk
		Hutsul	Romania	Unknown	No info	
		Huculł	Slovakia	At risk	Locally adapted	
		Huculł	Hungary	At risk	Native	
		Huzule	Germany	At risk	Exotic	
		Gutsul	Ukraine	At risk	No info	
		Huzule	Austria	At risk	Native	
		Huculsky kun	Czechia	At risk	Locally adapted	
Cattle	Podolian	Podolica	Italy	Not at risk	Native	Not at risk
		Podolian	Serbia	At risk	Locally adapted	
Sheep	Precoce	Merino Precoz	Spain	At risk	Locally adapted	At risk
		Merina Precoce	Portugal	At risk	Exotic	
		Mérinos précoce	France	Unknown	Native	
Sheep	Ouessant	Ouessant	France	At risk	Native	Not at risk
		Ouessant	Belgium	At risk	No info	
		Ouessant	Netherlands	At risk	Exotic	
		Ouessant Schaf	Germany	At risk	Exotic	
		Kesantská ovce	Czechia	At risk	Exotic	
		Ouessant	Denmark	At risk	No info	

The global indicators employed to assess goal attainment may introduce biases in data entry into global databases. These biases can also arise from the implementation of national indicators or adherence to national legislation. An illustrative example of a national rules-based decision is the case of the Icelandic Horse, which is reported by several countries, even though Iceland itself does not link it to the TB list. This is due to national legislation prohibiting the re-importation of Icelandic horses, which results in the national population being considered genetically isolated (Campana et al, 2011). In addition to these cases, certain widely used industrial pig and poultry lines – such as the Topigs Lignée E (TN Tempo) pig line, currently reported only by the Netherlands – may appear underreported due to limited data exchange between private breeders and NCs. It cannot be excluded that in a number of cases, a breed in the TB list is actually present in only one country; however, it is likely that the greater part of this gap is caused by underreporting in countries. To address this issue, NCs could be encouraged to review the breeds listed in Supplemental Table 1 – particularly those listed in the first column (TB reported by only one country) – and assess whether NBPs from their own country should also be linked to this list. Complementarily, FAO, in collaboration with NCs and relevant stakeholders, should undertake efforts to clarify the definitions and utilization of breed classifications. This includes revisiting the criteria used to associate NBPs with TB names, to ensure greater consistency and completeness in reporting.

Several actions initiated within the framework of the European Regional Focal Points for Animal Genetic Resources (ERFP) aim to facilitate exchange among countries and promote agreement on common perspectives. At the same time, however, it is important to recognize that the existing variation among countries could also be beneficial, as it reflects diverse needs and priorities. Genomics could play a key role in the management of TBs and provide strong scientific evidence to justify further steps. The developments in genetics and the wide availability of genomic tools could reveal genetic diversity between NBPs that have evolved in different environments and have been bred under separate breeding programmes for certain time periods. Whole genome sequencing (WGS) could offer valuable information on the demographic history of populations to support decisions relevant to the management of TBs.

Thus, the management of TBs can be approached considering the existing definitions of breeds. According to the definition used by FAO (FAO, 1999), a breed is not a simple genetic concept, but it also has a social dimension. Breeds are developed by the farmers' breeding practices, which are not exclusively technically driven. Several exogenous factors have an impact on them. In certain cases, such factors could radically change the environment in which the farmers work. Generalized guidelines and rules cannot be easily applied in the case of TBs. A common understanding of the term 'transboundary breed', based on both genetic and social parameters, could help in developing appropriate approaches. However, one should be very sceptical in defining specific criteria that could be used in all cases, as the decision is not clearly technical, but also political, to a large extent.

Besides the aspect of demographic data quality, which results in an unknown risk status, the outcomes of our analysis highlight certain inconsistencies related to the approach followed in setting the TB groups, as shown by the number of TB groups, to which only one NBP is linked. Taking this into account, the results can be considered as a first step in choosing breed cases that could be further analyzed, incorporating additional information.

By examining TBs through case studies, either in neighbouring countries or across more widely spread breeds, useful conclusions could be drawn that may serve as guiding principles on a broader level. The knowledge of whether breeders cooperate already, i.e. through the exchange of breeding animals, is important for grouping TB cases and could be essential for making decisions in terms of *ex situ* conservation. Mainstream breeds that are exported to other countries and are raised under separate breeding programmes, or with continuous import of breeding animals or semen, fall into this category; however, further elements are needed to decide on implementing common breeding programmes. Which criteria could be used to consider these NBPs as the same breed? Furthermore, changes in European borders over the past 25 years, which have resulted in some TBs being native to multiple countries with different local risk status, present an interesting case study, linking genetics with economic, social and political changes.

The level (local, regional or global) at which the risk status is estimated is a crucial point for TB management. The relevance of risk status at regional level is questionable when two (or more) NBPs have been developed separately for years. Besides that, the risk status classification is one of the critical points leading to reluctance in linking a NBP to an existing transboundary name. This is demonstrated in the current analysis by the number of NBPs that are declared as transboundary by only one country (listed in Supplemental Table 1).

The various combinations of risk status at local and regional levels could be the starting point of the case study approach, following the general approach presented in Figure 5. Supplemental Table 1 provides the list of European TBs that could be examined following this general approach.

Conclusions

This article provides new evidence on the status of TBs in Europe, assessing the quality of data and the frequency of updates to population data and relevant fields, as reported in DAD-IS. These outcomes could be useful in improving data quality and population management by enhancing data exchange and communication among countries.

Data analysis revealed several inconsistencies, as is the case with TBs linked to only one NBP, which opens the discussion on the definitions of TBs and the criteria applied to link a NBP to a TB. Under which conditions would it be feasible to establish unified criteria, including historical data and genetic information? The lack of standardized definitions and consistent risk classifications across countries poses significant challenges for TB management. Collaborative breeding programmes, genetic studies and enhanced data sharing could improve conservation outcomes. Advances in genomics can further clarify genetic relationships among NBPs, supporting more effective conservation planning.

As discussed above, TBs cannot be examined exclusively through demographic data and genetic information, as several technical, social and political aspects shape future management opportunities. Thus, a case study approach is recommended for the analysis of breeds, either *in situ* or *ex situ*, when they

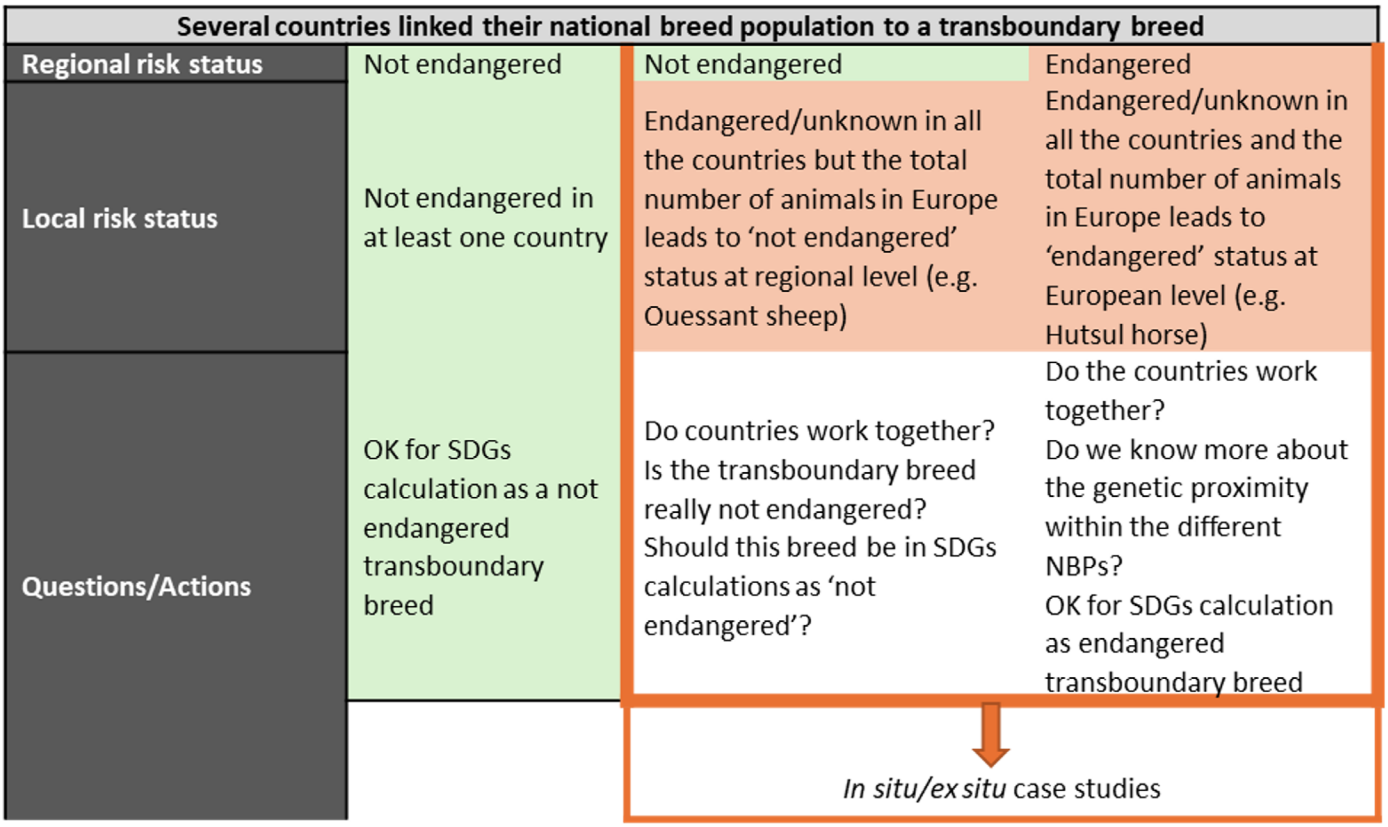


Figure 5. Questions and actions depending on transboundary breeds' risk status at different levels, depending on the national breed population's status (NBPs) and the impact on the calculation of the indicators for the Sustainable Development Goals (SDGs).

participate in common breeding programmes, frequently exchange breeding animals, or share a common history and/or environment.

This discussion is in accordance with the recommendations of the Animal Genetic Resources Strategy for Europe, that promotes *in situ* and *ex situ* strategies for TBs (ERFP, 2021). Specific actions are foreseen to improve the knowledge on TBs, support the exchange between actors involved in the conservation and breeding programmes of these breeds and promote cooperation in this field (ERFP, 2021). In this complex context, the advances in genomics and the progress of relevant research could further enrich existing knowledge on TBs and support their sustainable management.

Supplemental data

Supplemental Table 1. Distribution of European transboundary breeds (TBs) and their risk status

Acknowledgements

This work was supported by the European Regional Focal Point for Animal Genetic Resources (ERFP) and realized in the frame of the Task Force on Transboundary Breeds (https://www.animalgeneticresources.net/wp-content/uploads/2023/09/6d_ERFP_GA_2024_TF-TB.pdf).

The authors acknowledge the financial contributions of European Countries to ERFP to continue its activities and support the *in situ* and *ex situ* conservation and sustainable use of animal genetic resources (AnGR) and facilitate the implementation of FAO's Global Plan of Action for AnGR in Europe.

Author contributions

All the authors have conceptualized the analyses and drafted the manuscript. Eleonore Charvolin and Dimitrios Tsiokos performed the data analyses. All the authors have substantially contributed to the interpretation, drafting, revision and final version of the manuscript.

Data availability

Data used in this study were extracted from FAO DAD-IS and are publicly available (https://www.fao.org/dad-is/data/en/).

Conflict of interest statement

The authors declare that they have no conflicts of interest.

References

Ablondi, M., Sabbioni, A., Stocco, G., Cipolat-Gotet, C., Dadousis, C., Kaam, J. T. V., Finocchiaro R., Summer A. (2022). Genetic diversity in the Italian Holstein dairy cattle based on pedigree and SNP data prior and after genomic

selection. *Frontiers in veterinary science* 8:773985. doi: https://doi.org/10.3389/fvets.2021.773985

Biscarini, F., Nicolazzi, E., Stella, A., Boettcher, P., Gandini, G. (2015). Challenges and opportunities in genetic improvement of local livestock breeds. *Front Genet.* 6: 33. doi: https://doi.org/10.3389/fgene.2015.00033

Campana, M. G., Stock, F., Barrett, E., Benecke, N., Barker, G. W. W., Seetah, K., Bower, M. A. (2012). Genetic stability in the Icelandic horse breed. *Animal genetics*, 43(4), 447-449.

Cao, J., Baumung, R., Boettcher, P., Scherf, B., Besbes, B. and Leroy G. (2021). Monitoring and Progress in the Implementation of the Global Plan of Action on Animal Genetic Resources. *Sustainability* 13(2):775. https://doi.org/10.3390/su13020775

ERFP (2021). Animal Genetic Resources Strategy for Europe. https://www.animalgeneticresources.net/wp-content/uploads/2022/03/Final_AnGR-Strategy-_022022.pdf

FAO (1999). The global strategy for the management of farm animal genetic resources. Executive Brief. Rome.

FAO (2000). The World Watch List for Domestic Animal Diversity. Rome.

FAO (2001). Animal Genetic Resources Information; Preparation of the first report on the State of the World's animal genetic resources, Annex2. Rome.

FAO (2007a). The State of the World's Animal Genetic Resources for Food and Agriculture, edited by Barbara Rischkowsky & Dafydd Pilling. Rome.

FAO (2007b). Global Plan of Action for Animal Genetic Resources and the Interlaken Declaration, Rome. https://openknowledge.fao.org/handle/20.500.14283/a1404e

FAO (2011). Report of the Thirteenth Regular Session of the Commission on Genetic Resources for Food and Agriculture https://www.fao.org/4/mc192e/mc192e.pdf

FAO (2013). Guidelines In vivo conservation of animal genetic resources. http://www.fao.org/3/a-i3327e.pdf

FAO (2015). The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture, Scherf, B.D., Pilling, D., (Eds), CGRFA, Rome.

FAO (2021). Eleventh session of the Intergouvernmental Technical Working Group on animal genetic resources for food and agriculture - Status of the development of the Domestic Animal Diversity Information System - CGRFA/WG-AnGR-11/21/5, https://openknowledge.fao.org/items/0ff821b5-1043-4471-a083-61948eb9fbf6

FAO (2023). Status and trends of Animal Genetic Resources, CGRFA Intergovernmental Technical Working Group on Animal Genetic Resources for Food and Agriculture, FAO, Rome https://openknowledge.fao.org/server/api/core/bitstreams/c8121641-5385-404f-bf98-6a584bec2e7b/content

FAO (2024a). Status and trends of Animal Genetic Resources, CGRFA Intergovernmental Technical Working Group on Animal Genetic Resources for Food and Agriculture, FAO, Rome, https://openknowledge.fao.org/server/api/core/bitstreams/b8c17b9b-f931-4b65-bfbb-0a83a89db912/content

FAO (2024b). Thirteenth Session of the Intergovernmental Technical Working Group on Animal Genetic Resources for Food and Agriculture - CGRFA/WG-AnGR-13/24/Report, https://openknowledge.fao.org/items/9fdbfef3-85e6-4141-8448-803c995254cc/full

IBM Corp. (2020). IBM SPSS Statistics for Windows (Version 27.0) [Computer software]. IBM Corp.

Lefevre, F., Bojkovski, D., BouDagher Kharrat, M., Bozzano, M., Charvolin-Lemaire, E., Hiemstra, S.J., Kraigher,H., Laloë, D., Restoux, G., Sharrock, S.,Sturaro, E., vanHintum, T., Westergren, M., Maxted, N., GenResBridge Expert Panel (2024). European genetic resources conservation in a rapidly changing world: three existential challenges for the crop, forest and animal domains in the 21st century. *Genetic Resources* 5(9),13–28. https://doi.org/10.46265/genresj.REJR6896

Leroy, G., Danchin-Burge, C., Palhière, I., SanCristobal, M., Nédélec, Y., Verrier, E., and Rognon, X. (2015). How do introgression events shape the partitioning of diversity among breeds: a case study in sheep. *Genetics Selection Evolution* 47, 1-14. DOI https://doi.org/10.1186/s12711-015-0131-7

Ligda, C., & Casabianca, F. (2013). Adding value to local breeds: challenges, strategies and key factors. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales*, 53, 107-116.

Polak, G., Sosin, E., Martyniuk. E. (2022). FAO Commission on Genetic Resources for Food and Agriculture: what it does and how it supports the livestock sector. *Animal Science and Genetics* vol. 18, no 4 DOI: https://doi.org/10.5604/01.3001.0016.2197

Sponenberg, D. P., Martin, A., Couch, C., & Beranger, J. (2019). Conservation strategies for local breed biodiversity. *Diversity*, 11(10), 177.

Verrier, E., Audiot, A., Bertrand, C., Chapuis, H., Charvolin, E., Danchin-Burge, C., Danvy, S., Gourdine, J.L., Gaultier, P., Guémené, D., Laloë, D., Lenoir, H., Leroy, G., Naves, M., Patin, S. and Sabbagh, M. (2015). Assessing the risk status of livestock breeds: a multi-indicator method applied to 178 French local breeds belonging to ten species. *Animal Genetic Resources* 57, 105–118. FAO doi: https://doi.org/10.1017/S2078633615000260

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 83–94
DOI: 10.46265/genresj.JDUT8893
https://www.genresj.org
ISSN: 2708-3764

A fast and effective method to distinguish cultivated fonio species: conservation and evaluation perspectives

Sandrine Causse*[,a,b], Thomas Kaczmarek[a,b,c], Cécile Dubois[a,b], Enoch G. Achigan-Dako[d], Joseph Adjebeng-Danquah[e], Richard Y. Agyare[e], Louise Akanvou[f], Yacoubou Bakasso[g,n], Mamadou B. Barry[h], Baye M. Diop[i], Mame C. Gueye[i], Abdou R. Ibrahim Bio Yerima[d,j], Happiness O. Oselebe[k], Sani Saidou Idi[l,n], Edak A. Uyoh[m], Sylvie Vancoppenolle[a,b], Adeline Barnaud[c], Claire Billot[a,b], Jean-François Rami[a,b], Christian Leclerc[a,b]

[a] *CIRAD, UMR AGAP Institut, Montpellier, France*
[b] *AGAP Institut, University of Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France*
[c] *DIADE, University of Montpellier, IRD, CIRAD, Montpellier, France*
[d] *Unit of Genetics, Biotechnology and Seed Sciences, GBioS, Faculty of Agronomic Sciences, University of Abomey-Calavi, Cotonou, Republic of Benin*
[e] *Council for Scientific and Industrial Research—Savanna Agricultural Research Institute (CSIR-SARI), Ghana*
[f] *Département Ressources génétiques, CNRA, Abidjan, Côte d'Ivoire*
[g] *Department of Biology, Faculty of Science and Technic, Abdou Moumouni University, Niamey, Niger*
[h] *IRAG, Conakry, Guinea*
[i] *ISRA, CERAAS, Thiès, Senegal*
[j] *Department of Rainfed Crop Production (DCP), National Institute of Agronomic Research of Niger (INRAN), Niamey, Niger*
[k] *Center for Crop Improvement, Nutrition & Climate Change (CCINCC), Ebonyi State University, Abakaliki, Nigeria*
[l] *Department of Plant Production and Biodiversity, Faculty of Agronomic and Ecologic Sciences, University of Diffa, Diffa, Niger*
[m] *Department of Genetics and Biotechnology, University of Calabar, Calabar, Nigeria*
[n] *Laboratory for the Management and Valorization of Biodiversity in the Sahel (GeVaBioS), Abdou Moumouni University, Niamey, Niger*

Abstract: Plant genetic resources characterization is essential for their conservation and their use in both breeding strategies and adaptation to global change. This is all the more important for species often neglected by research such as fonio. Fonio refers to two indigenous small millets grown in West Africa, white and black fonio (*Digitaria exilis* and *Digitaria iburua,* respectively). This research was carried out to develop a simple and reliable method to identify the two cultivated species of fonio in the context of genebank collection. A morphometric analysis was performed on seeds of 98 accessions of *D. exilis* and 20 accessions of *D. iburua*. Morphometric characters measured were seed dimensions, shape and colour. We showed that the major delimiting criterion was the seed width and that the seeds of black fonio were wider than those of white fonio. The proposed method, based on seed morphometrics, could be applied systematically in conservation routine to guarantee the accuracy of the passport data in fonio collections, as well as to identify fonio remains for archaeological studies.

Keywords: plant genetic resources, fonio, morphometrics, seed width, genebanks

Citation: Causse, S., Kaczmarek, T., Dubois, C., Achigan-Dako, E. G., Adjebeng-Danquah, J., Agyare, R. Y., Akanvou, L., Bakasso, Y., Barry, M. B., Diop, B. M., Gueye, M. C., Ibrahim Bio Yerima, A. R., Oselebe, H. O., Saidou Idi, S., Uyoh, E. A., Vancoppenolle, S., Barnaud, A., Billot, C., Rami, J.-F. and Leclerc, C. (2025) "A fast and effective method to distinguish cultivated fonio species: conservation and evaluation perspectives", *Genetic Resources*, 6(12), pp. 83–94. doi: 10.46265/genresj.JDUT8893.

* Corresponding author: Sandrine Causse (sandrine.causse@cirad.fr)

Received: 22.04.2025 | Accepted: 01.08.2025 | Published online: 29.09.2025

Introduction

Crop genetic resources refer to the diversity of traditional landraces and modern cultivated varieties, including their crop wild relatives. The number of cultivated crops has been drastically reduced by the intensification of agriculture since the 20th century. Today, around 30 species are used to satisfy 90% of humanity's needs, whereas 100 species were used at the beginning of the 20th century (Gepts, 2006). In this context of crop genetic erosion, *in situ* and *ex situ* conservation of a wide range of crop diversity is essential to ensure food security and to face global changes (FAO, 2010; Khoury *et al*, 2014; FAO, 2020).

The *ex situ* approach involves safeguarding crop diversity outside of its native environment, typically within conservatories, or specialized infrastructures such as seedbanks. The primary goal is to conserve and propagate crop genetic resources and make them available for research, breeding and cultivation. This is particularly crucial for neglected and underutilized species (NUS) that received limited scientific attention (Stamp *et al*, 2012; Hunter *et al*, 2019; Ulian *et al*, 2020), whereas they could be used to face global changes and improve the quality and sustainability of food production (Ulian *et al*, 2020).

Describing and characterizing NUS accessions that are preserved in *ex situ* collections is essential for their management and sustainable use. Accurate documentation of accessions enables informed decisions to be made on conservation, research, breeding and potential use (Weise *et al*, 2020). However, this conservation approach requires the availability of accurate passport data, notably to avoid any taxonomic misidentification (Guzzon *et al*, 2018) or geographic location errors, which can introduce spatial bias into databases and distort large-scale biodiversity analyses (Beck *et al*, 2014).

Among NUS, fonio is a key cereal, native to West Africa, with valuable nutritional and agronomic qualities. Fonio is highly adapted to harsh environmental conditions and plays a crucial role in food security within developing economies. The potential of fonio has earned it recognition by the Value Addition in Cereal Systems (VACS) initiative as a top cereal for West Africa (Karl *et al*, 2024). The accuracy of passport data is particularly critical in the case of fonio. Fonio comprises two similar species with tiny seeds, both grown in West Africa, sometimes in the same localities; identifying the two species is not obvious. The most common, *Digitaria exilis* Stapf, is known as white fonio and is cultivated in an area stretching from Senegal to Nigeria. The second, *Digitaria iburua* Stapf, is named black fonio and its distribution is limited to northern Nigeria, Togo and Benin (Animasaun *et al*, 2018). The black and white denomination for fonio refers to the colour of the seed husks, which are likely to be more intensely dark brown for *D. iburua* than *D. exilis* (Adoukonou-Sagbadja A-H, 2010) (Figure 1). However, this criterion can vary within the two species, leading to confusion. The variability of this trait, plus the fact that fonio species are sometimes not distinguished by their common name (Blench, 2016), can lead to misidentification in collections.

Figure 1. Pictures of: a, black fonio **(***Digitaria iburua)* and b, white fonio seeds **(***Digitaria exilis)*.

Improving the accuracy of genebank passport data concerning the identification of *D. exilis* and *D. iburua* is a key issue that needs to be overcome to preserve fonio genetic resources, and make their adaptive potential available to farmers. Fonio identification could be based on vegetative, floral or spikelet characteristics, or molecular markers. For example, the growth habits of white and black fonio differ (Figure 2), but using this trait as an identification criterion requires seed growing, as well as access to large-scale cultivation areas or costly infrastructures. In addition, genebank collections often contain only limited seed samples for some fonio accessions. On the other hand, genetic identification methods, based on microsatellite genotyping or genome sequencing, require small samples (Mondini *et al*, 2009). All these methods are expensive, destructive and time-consuming.

The aim of this work was to develop a low-cost, non-destructive and rapid method, based on seed morphology, in order to assign fonio accessions to either white or black fonio species. To date, the seed morphometrics approach has never been applied to fonio crops. Such an approach is highly relevant for fonio genebank collections to improve the quality of associated passport data and enhance the value of these collections.

Materials and methods

Plant material and sampling

Fonio accessions are maintained in the seed collection in Montpellier, France, at the GAMéT Resource Centre (ARCAD) and the French National Research Institute for Sustainable Development (IRD) and in national genebank collections across West Africa. These collections of seeds (paddy grains), sometimes in small amounts (less than one gram), have been built up since 1977 thanks to collection missions to farmers in the areas of origin and thanks to partnerships between French and African research institutions involved in various research projects.

A sample of 118 accessions of *D. exilis* and *D. iburua* (98 and 20 respectively, Supplemental Table 1) previously sequenced in Abrouk *et al* (2020) and Kaczmarek *et al* (2025) was selected to maximize geographical coverage (Figure 3) and thus climatic diversity. Only accessions whose species had been genetically validated were selected.

For each accession, one seed sample was prepared in a 2ml microtube, weighing between 110mg and 151.4mg. The variability in sample mass was linked to the quantity of seeds available and to the difficulty of handling very small seeds (in the millimetre range, Figure 1).

Seed image analysis

For each accession, the seed sample was poured and carefully laid on a 13.5 × 10.5cm surface of a flatbed scanner (Epson Expression 10000XL) to be scanned on a green background (Canson-C200040066). An average of 232 seeds per accession was scanned. In total, 27,345 seeds were analyzed. The images were saved in .tif format (800dpi resolution).

The images were analyzed with the Rigatoni v.0.9.3 R package (Rami, 2022), based on the EBImage R package. The Rigatoni package was designed to analyze seed images acquired by scanning. In contrast to the colour of background pixels, the algorithm detects objects in an image and characterizes their size, shape and colour, with a total of 27 descriptors. Each seed was individually cropped in the image using the krnel function.

Preliminary tests were carried out to calibrate the seed

Figure 2. Pictures of several plants of: a, black fonio *(Digitaria iburua)* and b, white fonio *(Digitaria exilis)*.

detection algorithm and define size and colour thresholds to avoid the detection of artefacts. It was observed that fonio seeds cannot be smaller than 500px (minimum size threshold for an object), nor larger than 2,000px (maximum size threshold). To find the seeds in the image, the hue range of the green background was set between 68° and 92° and the brightness threshold was set at 0.001.

Dimensions in pixels were converted to tenths-of-a-millimeter (tmm), and the area was converted from pixels to square tenths-of-a-millimetre (tmm^2). The hexadecimal colour code was also determined from each seed's cropped image. For data analysis, this colour code in RGB components was then converted into H, S and V components with the colorspace v2.0-3 R package (Zeileis *et al*, 2020). The advantage of the HSV colour system is that it is based on components perceived by humans to describe colours: hue (tint or predominant colour), saturation (colour intensity) and value (brightness), allowing intuitive interpretation of colour variations (Hema *et al*, 2019).

Preliminary exploratory data analysis

Data validation

A graphical exploratory analysis was performed on the values measured per seed, and revealed distributions that were sometimes highly asymmetrical with very extreme and therefore suspicious values. An automatic extreme values filter was applied with the boxplot.stats() function, using the interquartile range (IQR) and a whiskers coefficient of 3 times this length to eliminate only highly improbable values. For each descriptor and accession, any value below Q1 - 3IQR, or above Q3 + 3IQR was considered an extreme outlier, with Q1 as the lower quartile and Q3 as the upper quartile. Seeds with at least one descriptor presenting an extreme outlier value were excluded from the data. For further data analysis, the morphometric values measured per seed were summarized, for each accession, by their median value.

Morphometric and colour descriptors

An exploratory analysis of the 27 descriptors was carried out to assess their variation and reduce their number in the event of strong correlations. A graphical method was used to explore the relationship between descriptors in pairs (matrix of graphs, not shown), initially considering separately each category: size, shape, pixel intensity, colour, contour (see Table 1 for descriptor details). Thousand-grain weight was added as a usual descriptor for cereals. Descriptors were selected in such a way as to retain only those descriptors that made sense, i.e. those that provided specific and easily understandable information on the sample variability. In the case of highly correlated variables, the selected variable was the one that made more sense. For example, s.area (seed area) and s.radius.mean (mean seed radius, Table 1) were highly correlated (Pearson correlation value 0.99), hence seed area, which made more sense than s.radius.mean, was selected. Variables representing statistical dispersion parameters such as standard deviation, median absolute deviation, or quantile, were not selected. All variables that were complementary to each other, such as the H, S and V colour parameters, were selected.

Statistical analysis of morphometric data

Descriptive statistics were carried out using the stat.desc() function of the pastecs v.1.3.21 R package (Grosjean *et al*,

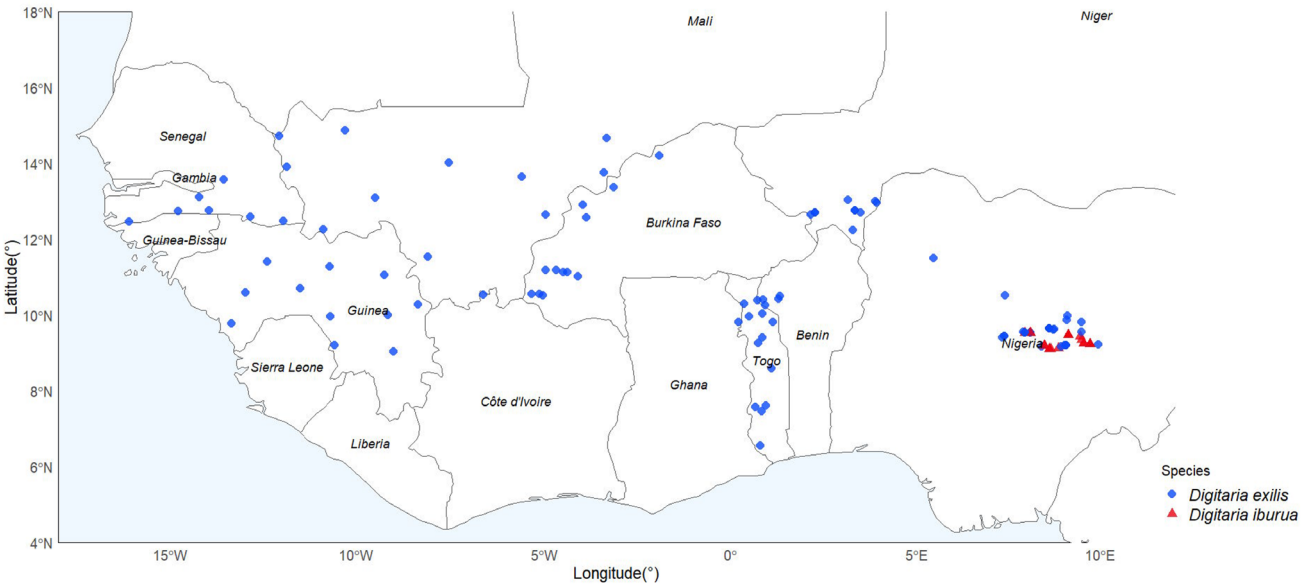


Figure 3. Geographical map of West Africa showing the spatial distribution of the 118 accessions of *D. exilis* and *D. iburua* used for seed measurements.

2018) in order to characterize the two species. Histograms of the median values per accession for each species (Supplemental Figure 1) showed non-normal (skewed and/or over-spread) distributions for most morphometric variables. These data were not suitable for the application of a t-test, and the two fonio species were compared by the Mann-Whitney-Wilcoxon rank test (wilcox.test() function). Morphometric diversity of the 118 accessions was explored using principal component analysis (PCA). PCA was performed on the seven seed descriptors (Table 1), using the PCA() function in the FactoMineR v.2.6 R package (Lê *et al*, 2008).

Table 1. Characteristics of the morphometric and colour descriptors. *, descriptors selected for the study; tmm, tenths-of-a-millimeter; tmm^2, square tenths-of-a-millimeter; sqrt, square root.

Descriptor	Characteristics
	Seed size (provided by Rigatoni package)
bbox.width*	Object bounding box width (in pixels, converted to tmm) measuring seed width
bbox.height*	Object bounding box height (in pixels, converted to tmm) measuring seed length
s.area*	Area, number of pixels in the shape (converted to tmm^2)
s.perimeter	Perimeter, number of pixels in the boundary of the object (converted to tmm)
s.radius.mean	Mean radius (in pixels), average radius value from the centre of shape to boundary (converted to tmm)
s.radius.sd	Standard deviation of the radius values (in pixels)
s.radius.max	Max radius (in pixels), largest radius value from the centre of shape to boundary (converted to tmm)
s.radius.min	Min radius (in pixels), shortest radius value from the centre of shape to boundary (converted to tmm)
	Seed shape (provided by Rigatoni package)
m.eccentricity*	Elliptical eccentricity, values ranging from 0 (perfect circle) to 1 (straight-line). Calculated with the longest axis (majoraxis) and the shortest axis (minoraxis) of the best-fitting ellipse: sqrt(1-minoraxis2/majoraxis2).
m.majoraxis	Largest axis of the best-fitting ellipse (in pixels, converted to tmm)
m.cx, m.cy	Centre of the best-fitting ellipse coordinates (in pixels)
m.theta	Object angle (in radians)
	Pixels intensity (provided by Rigatoni package)
b.mean	Average of pixel intensity in the shape
b.sd	Standard deviation of pixel intensity in the shape
b.mad	Median absolute deviation of pixel intensity in the shape
b.q (b.q001, b.q005, b.q05, b.q095, b.q099)	Quantile intensity of pixel intensity in the shape
	Seed contour (provided by Rigatoni package)
poi.x, poi.y	Pole of inaccessibility coordinates, coordinates of the point farthest away from the boundary of the object
poi.dist	Longest distance to the boundary of the object (in pixels, converted to tmm)
	Seed colour (obtained by converting the RGB code provided by Rigatoni package)
H*	Hue (in degrees), predominant colour or tint, (values ranging from 0 to 360°)
S*	Saturation, intensity of colour pigmentation (values ranging from 0 to 1)
V*	Value, brightness of the colour (values ranging from 0 to 1)
	Thousand-grain weight (calculated by accession)
TGW*	Thousand seed weight (in grams)

Results

The objective of this study was to develop an affordable, non-destructive and rapid method based on seed morphology to categorize fonio accessions into white or black fonio species.

Outliers detection

The 3IQR method discarded 1.1% of the 27,345 seeds analyzed. Nine accessions had no outliers. For the remaining 109 accessions, the percentage of outliers varied from 0.4% to 8.8% (Supplemental Table 2). No link could be established between the percentage of outliers and the parameters structuring the sampling design (species, country of origin and collection date).

Figure 4 shows an example of how outliers were detected for seeds with attached pedicels and seeds with open glumes.

Morphometrics description of black and white fonio

D. iburua and *D. exilis* differed significantly for all variables (Wilcoxon test at the p-value threshold < 0.05). It can be noted that seed area (s.area) and seed width (bbox.width) showed no overlap at all between the two species for our sample of accessions (Supplemental Figure 1). The results presented below are based on the descriptors' median values.

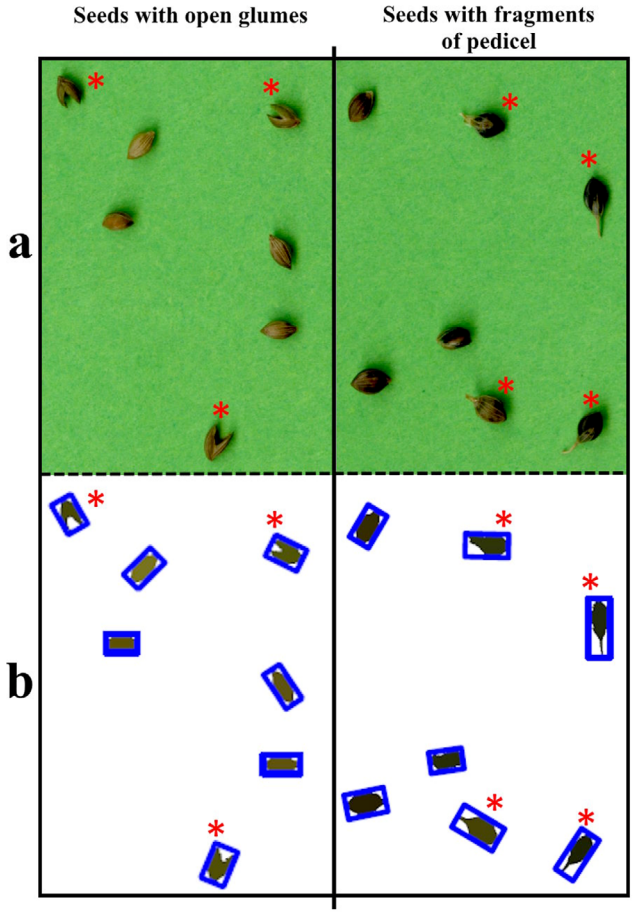


Figure 4. Picture showing the two main types of outliers (*). The two columns on the left show seeds with open glumes and, on the right, seeds with attached pedicels. The two rows represent: a, original seeds picture; b, bounding boxes plotted by the image analysis.

Size and shape analysis

D. iburua seeds were significantly wider (+19%), longer (+13%) and heavier (31%) than *D. exilis* seeds (Wilcoxon test, $p < 0.001$, Supplemental Figure 1). *D. iburua* seeds were 19.1tmm long (bbox.height) and 11.1tmm wide (bbox. width), while *D. exilis* seeds were 16.9tmm long and 9.4tmm wide. The thousand-grain weight was equal to 0.71g for *D. iburua* and 0.54g for *D. exilis* (Table 2).

For seed width (bbox.width), the range of variation between the minimum and the maximum values showed a clear demarcation between the two species (Supplemental Figure 1). Indeed, seed width varied for *D. exilis* from 8.2 tmm to 10.1tmm as compared to 10.5tmm to 12.0tmm for *D. iburua* accessions (Table 2). Seed area (s.area) values also highly differed (Wilcoxon test, $p < 0.001$), revealing a distinct separation between the two species. For *D. exilis*, seed area varied from 90.6tmm^2 to 133.7tmm^2 whereas for *D. iburua*, seed area varied from 139.8tmm^2 to 162.7tmm^2 (Table 2). These results confirmed that *D. iburua* seeds were significantly bigger than those of *D. exilis*.

Furthermore, m.eccentricity values were significantly (Wilcoxon test, $p < 0.05$) higher for *D. exilis* (0.82) than for *D. iburua* (0.80). The distribution of values for each species (Supplemental Figure 1) showed a second peak at high values of m.eccentricity (around m.eccentricity = 0.83), indicating that some *D. exilis* accessions had, on average, more elongated seeds.

Colour analysis

Hue values were significantly different (Wilcoxon test, $p < 0.001$) between *D. iburua* (H = 51.9°) and *D. exilis* (H= 56.4°). Both values were in the yellow range, but *D. iburua* seeds had a warmer yellow-red hue than *D. exilis* seeds whose hue was closer to pure yellow (Supplemental Figure 2). Moreover, colour brightness (V) values were significantly (Wilcoxon test, $p < 0.001$) lower for *D. iburua* (V = 0.21) than for *D. exilis* (V = 0.32); reversely, colour saturation (S) values were significantly (Wilcoxon test, $p < 0.001$) higher for *D. iburua* (S = 0.89) than for *D. exilis* (S= 0.77) (Table 2). Those contrasted V and S values, associated with H values correspond to the contrasted colour of the seed husks, which are dark brown for *D. iburua* and light brown for *D. exilis* (Figure 1).

Morphometric diversity

The PCA analysis revealed a clear distinction between white and black fonio (Figure 5). The first four principal components covered 99% of the total variance (Supplemental Figure 3).

Table 2. Descriptive statistics for size and shape analysis of each fonio species. N, number of accessions.

Species	Variables	Minimum value (Min)	Maximum value (Max)	Median	Mean	Standard deviation (SD)	Coefficient of variation (CV, %)
Digitaria exilis (N = 98)	s.area (tmm²)	90.63	133.72	113.28	112.80	71.91	7.5
	bbox.width (tmm)	8.24	10.09	9.37	9.34	0.10	3.4
	bbox.height (tmm)	14.34	19.05	16.86	16.85	0.95	5.8
	m.eccentricity	0.74	0.86	0.82	0.82	0.00	2.9
	H (°)	53.81	62.56	56.40	56.41	1.89	2.4
	S	0.65	0.87	0.77	0.76	0.00	5.5
	V	0.22	0.48	0.32	0.32	0.00	11.6
	TGW (g)	0.31	0.70	0.54	0.53	0.00	12.2
Digitaria iburua (N = 20)	s.area (tmm²)	139.82	162.70	153.23	152.86	46.02	4.4
	bbox.width (tmm)	10.54	11.96	11.11	11.14	0.14	3.3
	bbox.height (tmm)	17.89	20.32	19.11	19.17	0.61	4.1
	m.eccentricity	0.74	0.84	0.80	0.80	0.00	3.4
	H (°)	48.24	71.71	51.86	53.63	35.54	11.1
	S	0.80	0.94	0.89	0.89	0.00	3.7
	V	0.11	0.37	0.21	0.23	0.01	35.3
	TGW (g)	0.59	0.88	0.71	0.73	0.01	9.8

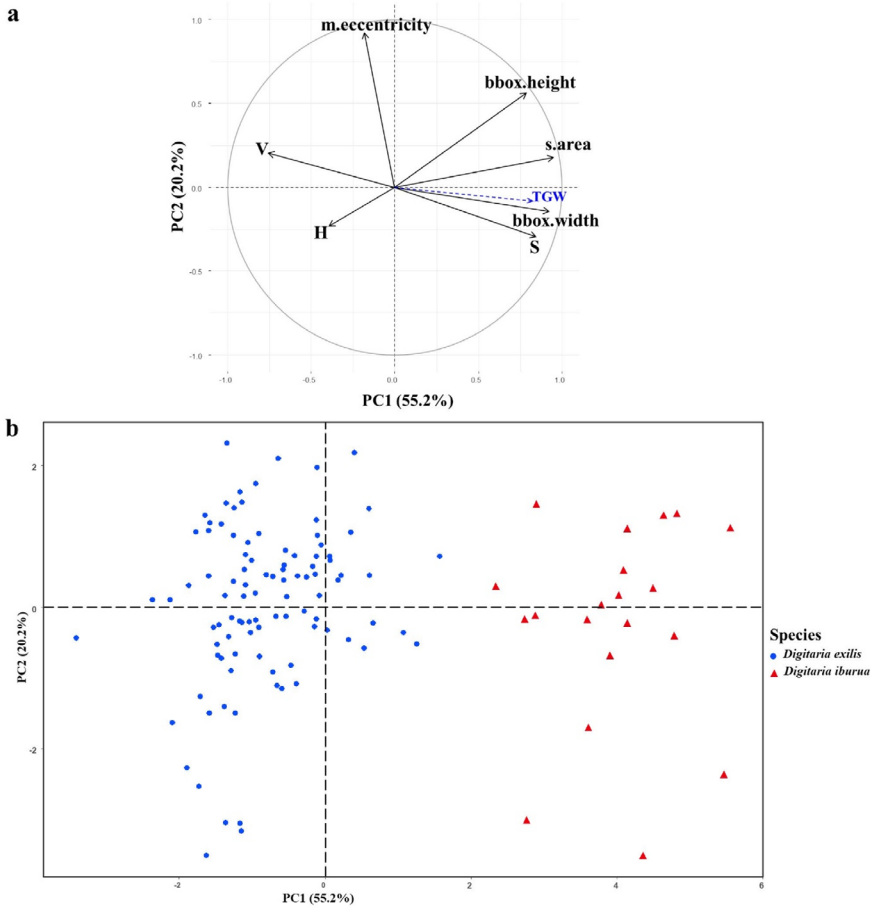


Figure 5. Principal component analysis carried out on the seed morphometric measurements of the 118 fonio accessions (n = 98 for *Digitaria exilis* and n = 20 for *Digitaria iburua*). a, Correlation circle for the first two principal components, where bbox.width correspond to width, bbox.height to length, S to saturation of colour, H to hue of colour, V to brightness of colour, m.eccentricity to eccentricity, s.area to seed area, and TGW to thousand-grain weight. The supplementary variable TGW was coloured in blue. b, Scatterplot of the 118 fonio accessions projected on the first two principal components plane.

The first principal component (PC1, 55.2% of the total variance) opposed seed size parameters and saturation (S) of the seed colour (positive values), with brightness (V) of the seed colour (negative values, Figure 5a). Seed size refers to seed area (s.area), seed width (bbox.width) and seed length (bbox.height), which were among the most influential characters on PC1 (Supplemental Table 3). Variability of s.area and bbox.width variables was almost entirely represented by PC1 (cos2: resp. 90% and 85%, Supplemental Table 3). The PC2 (20.2% of the total variance) was mainly related to the m.eccentricity variable (contribution: 60.7%), which is a component of the seed shape, and to a lesser extent to seed length (bbox.height, contribution: 22.5%). The variability of the m.eccentricity variable is almost entirely represented by PC2 (cos2: 86%). PC3 (13.9% of the total variance) was mainly related to the tint (H, contribution: 80%). For PC4, which accounted for less than 10% of total variability, the brightness (V) of the seed colour was the most influential character (contribution: 37%) (Supplemental Table 3).

The first axis (Figure 5b) completely differentiated the two fonio species, with black fonio (right) seeds characterized by higher values of area, width, colour saturation (S) and height, and lower values of brightness (V), compared to white fonio (left) with the opposite characteristics. Independent of this clear structure separating the species on the first axis, the second axis showed, within each species, a gradient of variation in seed shape (m.eccentricity), from the least to the most elongated. Illustration of accessions by their country of origin (Figure 6) revealed that the roundest seeds (bottom of axis 2) originated from Nigeria only. Moreover, the Nigerian accessions were projected over the same range along this axis for both species. On the opposite side and on the top of axis 2, the white fonio accessions with the most tapered seeds mainly originated from Mali and Senegal.

The first factorial plane concentrated 75% of the total variability and made it possible to characterize the morphometric differences between the seeds of the two species. The next PCs, calculated on the remaining variability, did not appear to be informative. They showed either particularities for some accessions (PC3), or more continuous variations which could not be linked to explanatory factors and therefore could not be interpreted (PC4).

Differentiation of the two species with only one morphometric descriptor

Figure 7 clearly showed that the seed width alone perfectly separated the two fonio species. Descriptive statistics (Table 2) specified the limiting values observed for these data: up to a value of seed width equal to 10.09tmm for *D. exilis* (n = 98) and from 10.54tmm for *D. iburua* (n = 20). In addition, the seed area parameter, which is strongly dependent on seed width and height, was also a differentiating trait between the two species (Supplemental Figure 4).

Discussion

This work sought to distinguish the two species of cultivated fonio based on the characteristics of their seeds in a context of genebank conservation. Seed morphometrics was used as a rapid, low-cost and non-destructive method to identify the two fonio species. Morphometrics has largely

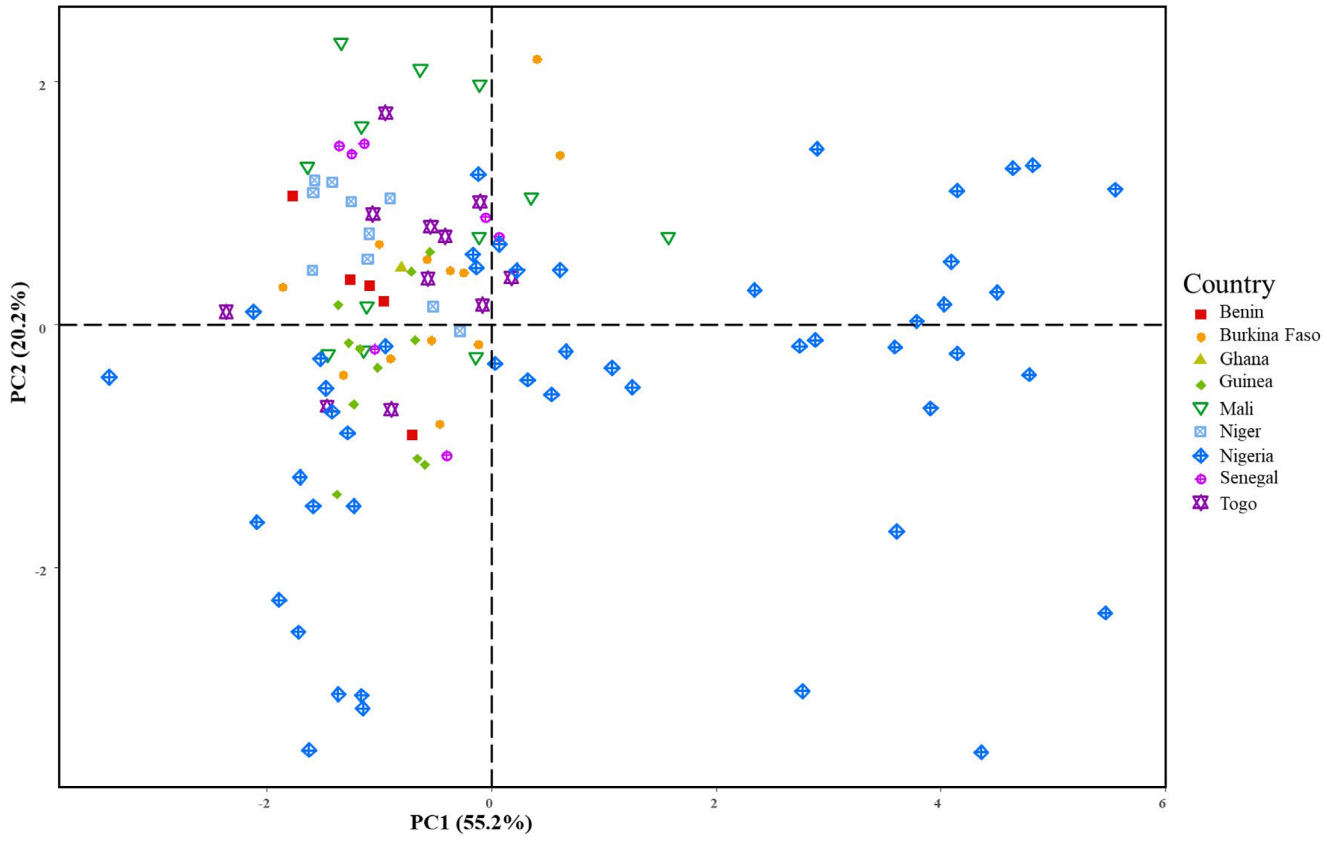


Figure 6. Principal component analysis carried out on the seven selected morphometric variables and the 118 fonio accessions. The geographic origin of individuals is represented by symbols of different shapes and colours.

been used to describe and compare organism shapes, allowing reliable species identification. This was confirmed on organisms as varied as orchids (Chemisquy *et al*, 2009), mosquitoes (Chaiphongpachara *et al*, 2022), indigo plants (Soladoye *et al*, 2010), wheat species (Goriewa-Duba *et al*, 2018) and olive (Terral *et al*, 2004; Newton *et al*, 2014).

The tiny size of the seeds of both species is a source of practical difficulties, particularly for handling and visual identification. We showed that the Rigatoni R package (Rami, 2022) could be adapted to detect very small objects in images. While this package has been initially developed for the analysis of irregular shapes like peanut pods, we used here a smaller number of Rigatoni variables since fonio seeds have an oblong ellipsoid shape (Idu *et al*, 2008).

One consequence of small seed size is that seed lots can be heterogeneous due to the residual presence of undesirable biological material (pedicels, open glumes, foreign seeds), sand or stones in seed samples despite careful cleaning of the sample before measurement, as suggested by Koreissi-Dembélé *et al* (2013). A large number of seeds (over 200) per accession were scanned to ensure reliable results by limiting the influence of outliers and being able to detect them using robust quantitative methods. We implemented a fast method to remove extreme outliers from the analysis to make the identification more robust. Depending on the accession, between 0.4% and 8.8% of seeds were identified as extreme outliers.

White and black fonio are very similar as they share many agromorphological characteristics. According to Adoukonou-Sagbadja *et al* (2007), a clear-cut separation of both species was not possible using agromorphological traits such as plant height, number of tillers, leaf length, fresh and dry biomass weight, panicle length, and yield. This study focused on seed morphology and seed weight. We showed that seeds of *D. iburua* were significantly larger, heavier and had a more intense and darker brown colour than those of *D. exilis*. Seed width clearly distinguished *D. exilis* from *D. iburua* in our sample of accessions. Our results thus confirmed a difference in seed size between the two species, as previously noted by Echendu *et al* (2009) and Jideani (2012). The distinctiveness of *D. exilis* and *D. iburua* was also confirmed with respect to seed weight (Aliero *et al*, 2002; Nyam *et al*, 2017). We showed that seed area was a descriptor also separating the two species; however, we focused on seed width, a one-dimensional parameter whose variations are straightforward to interpret. It is worth noting that the differences between species, in seed size, weight and colour, were revealed despite the different conditions under which the accessions were collected and conserved. The species effect on these morphometric characteristics, therefore, appears stable, as it is greater than eventual environmental effects. Inversely, the more or less tapered shape of the seeds did not show a clear difference between species, but varied in an apparently structured way according to country of harvest (PCA, axis 2, Figure 6). The projection on a map (Supplemental Figure 5) of seed shape values (m.eccentricity descriptor cut into classes), at harvesting sites, visually confirmed that the most tapered *D. exilis* seeds came mainly from the northern edge of the sampling zone (Senegal, Mali, northern Burkina Faso, Niger), in drier climatic regions. This trend requires more in-depth studies in relation to climatic data.

The number of accessions used in our study differed

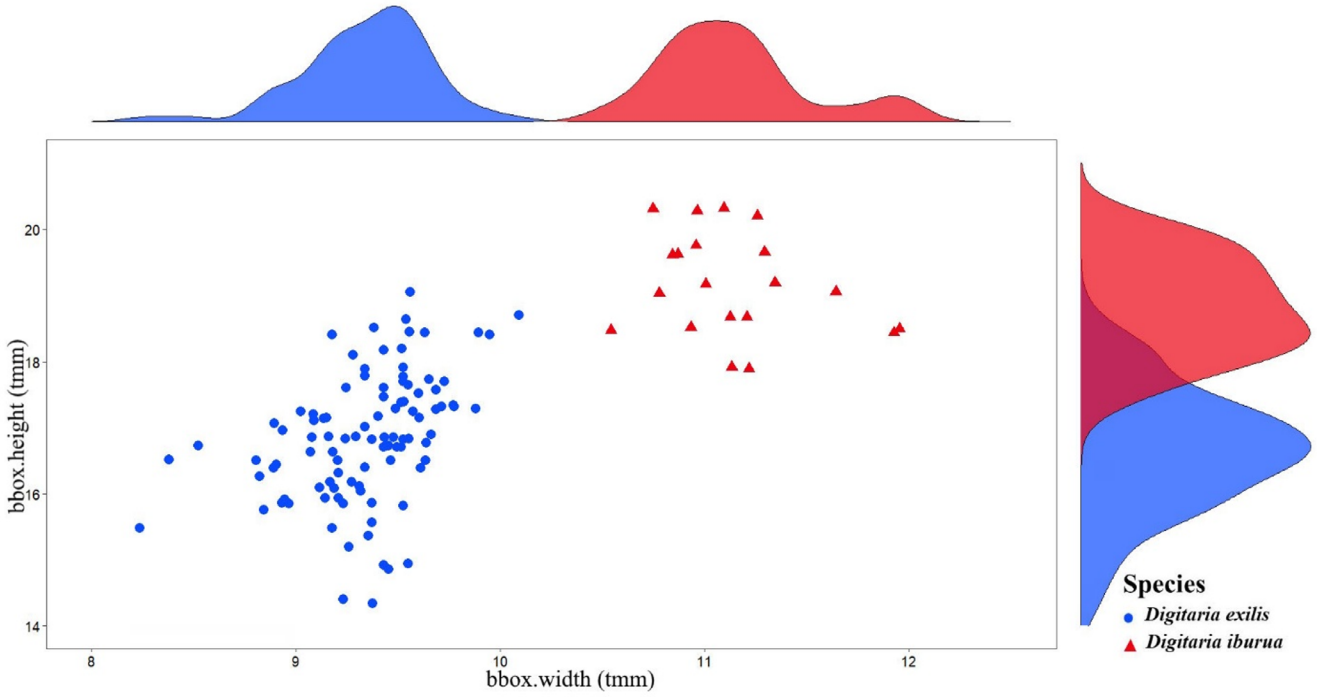


Figure 7. Scatterplot with marginal distribution of seed width (x) and seed length (y) of the 118 fonio accessions.

markedly between the two species. This is partly because black fonio is predominantly grown in central Nigeria today, from where samples are available in genebanks. Moreover, some geographical origins could not be used for *D. exilis* because of incomplete passport data. We can't rule out the possibility that a more balanced sampling design, both in terms of species frequency and geographic distribution, might nuance our results, and that the species classification might be less categorical in a context of greater morphometric variability. However, our sampling design was based on genetic studies that already maximized the diversity and species geographical range in their sampling. Despite sampling constraints, PCA provided some confidence in the results. The first principal component (Figure 6), which differentiates the two species by seed size and colour saturation, did not appear to show any structure related to geographical origin. Further studies focusing on the effect of geographic origin on seed morphology should investigate this issue more precisely, both within and between species.

Our morphometric results also open up new avenues for research in archaeobotany. The morphometric approach proposed in this paper could be used to identify fonio in archaeological records, contributing to the reconstruction of the evolutionary history of both cultivated fonio and its wild relatives. Morphometrics of ancient seeds would provide a better understanding of the domestication and diffusion histories of *D. exilis* and *D. iburua* in West Africa. Indeed, morphometrics on crops allowed to confidently distinguish between domesticated versus wild forms and trace the evolution of cultivated forms through space and time, as in the case of grapevine (Bouby *et al*, 2013; Rôs *et al*, 2014; Bonhomme *et al*, 2021; Ucchesu *et al*, 2024) or date palm (Terral *et al*, 2012; Gros-Balthazard *et al*, 2016). Additional research with carbonization experiments is however needed to see the impact of charring on seed morphology in both species (Ivorra *et al*, 2024).

In order to preserve the existing fonio diversity from genetic erosion, germplasm collection and *ex situ* conservation is a necessity (Dansi *et al*, 2010). Correct taxonomic identification is all the more crucial for the plant genetic resources that are currently underrepresented in *ex situ* conservation facilities worldwide, as is the case for fonio, and therefore have a high priority for future collecting missions and urgent conservation measures (Guzzon *et al*, 2018). Labelling of accessions is essential to enhance the value of these collections and to make accessions usable. By enhancing passport data and combining it with additional information from other fields, new knowledge about plant genetic resources can be generated, which is crucial for the sustainable management of genebank collections.

In the case of fonio species, the seed width could be used as the sole criterion for a simple and inexpensive method to make their taxonomic identification and genebank passport data more reliable. This approach is particularly useful when genetic data are not available. As a new genebank conservation procedure, we suggest that the method proposed in this paper be applied to fonio accessions already conserved in genebank collections to ensure the reliability of fonio identification in passport data. It could also be systematically applied to any new fonio accession before its integration in genebank collections, especially for fonio originating from regions where both species are grown.

Supplemental data

Supplemental Table 1. Sampling details: number of accessions by countries and species.

Supplemental Table 2. Information on the 118 accessions analyzed.

Supplemental Table 3. Results for the Principal Component Analysis carried out on the seven selected morphometric variables and the 118 fonio accessions.

Supplemental Figure 1. Histogram of median morphometric values by accession and TGW (thousand grain weight) for the two species.

Supplemental Figure 2. Position of the two fonio species, according to their median hue values, on the diagram representing a part of the sequence of Hue.

Supplemental Figure 3. Principal component analysis carried out on the seven selected morphometric variables and the 118 fonio accessions. Scree Plot: percentage of total variation explained by each principal component.

Supplemental Figure 4. Scatterplot with marginal distribution of seed area (x) and seed length (y) of the 118 fonio accessions.

Supplemental Figure 5. Projection on a map of seed shape values (m.eccentricity descriptor cut into classes).

Data and code availability statement

Dataset supporting the results of this article are available via Dataverse: https://dataverse.cirad.fr/dataset.xhtml?persistentId=doi:10.18167/DVN1/ZFZTWP

R script for the different analyses carried out throughout the paper are available on the CIRAD Gitlab platform: https://gitlab.cirad.fr/agap/fonio/seedfonio

Author contributions

SC, TK, AB, CB, and CL designed the research. EGAD, ARIBY, SS, YB, BMD, MCG, RYA, JAD, LA, MMB, EAU, HOO, SSI, SV, TK, AB and CB contributed to the sampling of the biological material or the curation of collections. SC generated the data. SC analyzed the data with inputs from CD, JFR, TK, AB, CB, and CL. SC and CL wrote the paper with substantial inputs from TK, AB, CD, CB.

Conflict of interest statement

The authors have no conflicts of interest to report.

Acknowledgments

The exploration of fonio diversity over the past decade has been facilitated by various research initiatives. These projects received funding from WAAPP/PPAAO 2A (CERA58ID06 SE), Agropolis Fondation (specifically the Agropolis Resource Center for Crop Adaptation and Diversity

[ARCAD] project and the Cultivar project—ID 1504-007) under the Investissements d'Avenir programme (Labex Agro: ANR-10-LABX-0001-01) within the I-SITE MUSE framework (ANR-16-IDEX-0006). Additionally, support came from the French government ANR project (AfriCrop project, ANR-13-BSV7-0017) and the European Union Horizon 2020 research and innovation programme (EWA-BELT, 862848, 'Linking East and West African farming systems experience into a BELT of sustainable intensification').

We would also like to thank CRB GAMéT for providing us with seeds for our study (Centre de Ressource Biologique GAMéT, ARCAD – UMR AGAP Institut de Montpellier. https://doi.org/10.18167/infrastructure/00007)

References

Abrouk M., Ahmed H.I., Cubry P., Šimoníková D., Cauet S., Pailles Y., Bettgenhaeuser J., Gapa L., Scarcelli N., Couderc M., Zekraoui L., Kathiresan N., Čížková J., Hřibová E., Doležel J., Arribat S., Bergès H., Wieringa J.J., Gueye M., Kane N.A., Leclerc C., Causse S., Vancoppenolle S., Billot C., Wicker T., Vigouroux Y., Barnaud A., Krattinger S.G. (2020). Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate. *Nat Commun.*, 11, 4488. doi: https://doi.org/10.1038/s41467-020-18329-4

Adoukonou-Sagbadja, H., Wagner, C., Dansi, A., Ahlemeyer, J., Daïnou, O., Akpagana, K., Ordon F., Friedt W. (2007). Genetic diversity and population differentiation of traditional fonio millet (*Digitaria* spp.) landraces from different agro-ecological zones of West Africa. *Theor. Appl. Genet.*, 115, 917-931. doi: https://doi.org/10.1007/s00122-007-0618-x

Adoukonou-Sagbadja, A.H. (2010). Genetic Characterization of Traditional Fonio Millets (*Digitaria exilis, D. iburua* STAPF) Landraces from West-Africa: Implications for Conservation and Breeding. PhD thesis, Justus-Liebig University, Giessen, Germany.

Aliero, A. A. & Morakinyo, J.A. (2002). Characterization of *Digitaria exilis* (Kipp) StapF and *D. iburua* Stapf Accessions. Nigerian Journal of Genetics, 16, 10-21. doi: https://doi.org/10.4314/njg.v16i1.42277

Animasaun, D. A., Awujoola, K. F., Oyedeji, S., Morakinyo, J. A., and Krishnamurthy, R. (2018). Diversity level of genomic microsatellite among cultivated genotypes of *Digitaria* species in Nigeria. *African Crop Science Journal*, 26(2), 305-313. doi: https://doi.org/10.4314/acsj.v26i2.11

Beck, J., Böller, M., Erhardt, A., Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10-15. doi: https://doi.org/10.1016/j.ecoinf.2013.11.002

Blench, R. (2016). Finger millet: the contribution of vernacular names towards its prehistory. *Archaeol Anthropol Sci* 8, 79–88. doi: https://doi.org/10.1007/s12520-012-0103-6

Bonhomme, V., Terral, J.F., Zech-Matterne, V., Ivorra, S., Lacombe, T., Deborde, G., Kuchler, P., Limier, B., Pastor, T., Rollet, P., Bouby, L. (2021) Seed morphology uncovers 1500 years of vine agrobiodiversity before the advent of the Champagne wine. *Sci Rep.*, 11(1):2305. doi: https://doi.org/10.1038/s41598-021-81787-3

Bouby L., Figueiral I., Bouchette A., Rovira N., Ivorra S., Lacombe T., Pastor T., Picq S., Marinval P., Terral J.F. (2013) Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* L.) during Roman times in Southern France. *PLoS One*, 8(5): e63195. doi: https://doi.org/10.1371/journal.pone.0063195

Chaiphongpachara, T., Changbunjong, T., Sumruayphol, S., Laojun, S., Suwandittakul, N., Kuntawong K. (2022). Geometric morphometrics versus DNA barcoding for the identification of malaria vectors *Anopheles dirus* and *An. baimaii* in the Thai-Cambodia border. *Sci Rep*, 12, 13236. doi: https://doi.org/10.1038/s41598-022-17646-6

Chemisquy, M. A., Prevosti, F. J., and Morrone, O. (2009). Seed morphology in the tribe Chloraeeae (Orchidaceae): combining traditional and geometric morphometrics. *Botanical journal of the Linnean Society,* 160(2), 171-183. doi: https://doi.org/10.1111/j.1095-8339.2009.00968.x

Dansi, A., Adoukonou-Sagbadja, H., and Vodouhè, R. (2010). Diversity, conservation and related wild species of Fonio millet (*Digitaria* spp.) in the northwest of Benin. *Genetic Resources and Crop Evolution*, 57, 827-839. doi: https://doi.org/10.1007/s10722-009-9522-3

Echendu, C. A., Obizoba, I. C., Anyika, J. U., & Ojimelukwe, P. C. (2009). Changes in chemical composition of treated and untreated hungry rice "Acha" (*Digitaria exilis*). *Pakistan journal of nutrition*, 8(11), 1779-1785. doi: https://doi.org/10.3923/pjn.2009.1779.1785.

FAO. (2010). The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture. Rome. url: https://www.fao.org/3/i1500f/I1500F.pdf

FAO. (2020). How the world's food security depends on biodiversity. Rome. url: https://www.fao.org/3/cb0416en/CB0416EN.pdf

Gepts, P. (2006). Plant Genetic Resources Conservation and Utilization: The Accomplishments and Future of a Societal Insurance Policy. *Crop Science*, 46, 2278-2292. doi : https://doi.org/10.2135/cropsci2006.03.0169gas

Goriewa-Duba, K., Duba, A., Wachowska, U., and Wiwart, M. (2018). An evaluation of the variation in the morphometric parameters of grain of six *Triticum* species with the use of digital image analysis. *Agronomy*, 8(12), 296. doi: https://doi.org/10.3390/agronomy8120296

Gros-Balthazard, M., Newton, C., Ivorra, S., Pierre, M. H., Pintaud, J. C., and Terral, J. F. (2016). The domestication syndrome in *Phoenix dactylifera* seeds: toward the identification of wild date palm populations. *PloS ONE*, 11(3), e0152394. doi: https://doi.org/10.1371/journal.pone.0152394

Grosjean, P., Ibanez, F., Etienne, M. (2018). pastecs: Package for Analysis of Space-Time Ecological Series. url: https://github.com/SciViews/pastecs

Guzzon, F., and Ardenghi, N.M.G. (2018). Could taxonomic misnaming threaten the *ex situ* conservation and the usage of plant genetic resources? *Biodiversity and Conservation*, 27, 1157-1172. doi: https://doi.org/10.1007/s10531-017-1485-7

Hema, D., and Kannan, D. S. (2019). Interactive color image segmentation using HSV color space. *Sci. Technol. J*, 7(1), 37-41. doi: https://doi.org/10.22232/stj.2019.07.01.05

Hunter D., Borelli T., Beltrame D.M.O., Oliveira C.N.S., Coradin L., Wasike V.W., Wasilwa L., Mwai J., Manjella A., Samarasinghe G.W.L., Madhujith T., Nadeeshani H.V.H., Tan A., Ay S.T., Güzelsoy N., Lauridsen N., Gee E., Tartanac F. (2019). The potential of neglected and underutilized species for improving diets and nutrition. *Planta*, 250, 709-729. doi: https://doi.org/10.1007/s00425-019-03169-4

Idu, M., J. U. Chokor, and O. Timothy. (2008). Effect of Various Hormones on the Germination of Fonio-*Digitaria exilis* L. *International Journal of Botany*, doi: https://doi.org/10.3923/ijb.2008.456.460

Ivorra, S., Tengberg, M., Bonhomme, V., Kaczmarek, T., Pastor, T., Terral, J.F., Gros-Balthazard, M. (2024). Leveraging the potential of charred archaeological seeds for reconstructing the history of date palm. *Journal of Archaeological Science*, 170, 106052. doi: https://doi.org/10.1016/j.jas.2024.106052.

Jideani, I.A. (2012). *Digitaria exilis* (acha/fonio), *Digitaria iburua* (iburu/fonio) and *Eleusine coracana* (tamba/finger millet) Non-conventional cereal grains with potentials. *Scientific Research and Essays*, 7, 3834-3843. doi: https://doi.org/10.5897/SRE12.416

Kaczmarek, T., Cubry, P., Champion, L., Causse, S., Couderc, M., Orjuela, J., ... & Leclerc, C. (2025). Independent domestication and cultivation histories of two West African indigenous fonio millet crops. Nature Communications, 16(1), 4067. doi: https://doi.org/10.1038/s41467-025-59454-2

Karl, K., MacCarthy, D., Porciello, J., Chimwaza,G., Fredenberg, E., Freduah, B.S., Guarin, J., Mendez Leal, E., Kozlowski, N., Narh, S., Sheikh, H., Valdivia, R., Wesley, G., Van Deynze, A., van Zonneveld, M., Yang, M.(2024). Opportunity Crop Profiles for the Vision for Adapted Crops and Soils (VACS) in Africa. doi: https://doi.org/10.7916/7msa-yy32

Koreissi-Dembélé, Y., Fanou-Fogny, N., Hulshof, P. J., and Brouwer, I. D. (2013). Fonio (*Digitaria exilis*) landraces in Mali: Nutrient and phytate content, genetic diversity and effect of processing. *Journal of Food Composition and Analysis*, 29(2), 134-143. doi: https://doi.org/10.1016/j.jfca.2012.07.010

Khoury, C.K., Bjorkman, A.D., Dempewolf H., Ramirez-Villegas J., Guarino L., Jarvis A., Rieseberg L.H., Struik P.C. (2014). Increasing homogeneity in global food supplies and the implications for food security. *Proc Natl Acad Sci U S A,* 111 (11), 4001-4006. https://doi.org/10.1073/pnas.1313490111

Lê, S., Josse, J. and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1). pp. 1-18. doi: https://doi.org/10.18637/jss.v025.i01

Mondini, L., Noorani, A., and Pagnotta, M. A. (2009). Assessing Plant Genetic Diversity by Molecular Tools. *Diversity*, 1(1), 19-35. doi: https://doi.org/10.3390/d1010019

Newton, C., Lorre, C., Sauvage, C., Ivorra, S., and Terral, J.-F. (2014). On the origins and spread of *Olea europaea* L. (olive) domestication: evidence for shape variation of olive stones at Ugarit, Late Bronze Age, Syria - a window on the Mediterranean Basin and on the westward diffusion of olive varieties. *Vegetation History and Archaeobotany,* 23(5), 567-575. doi: https://doi.org/s00334-013-0412-4

Nyam, D., Kwon-Ndung, E. and Ap, W. (2017). Genetic Affinity and Breeding Potential of Phenologic Traits of Acha (fonio) in Nigeria. *Journal of Scientific and Engineering Research*, 4(10), 91-101. url: https://irepos.unijos.edu.ng/jspui/handle/123456789/1966

Rami, J.F. (2022). Rigatoni: Object detection (typically grains) in images. R package version 0.9. url: https://github.com/jframi/rigatoni

Rôs, J., Evin, A., Bouby, L., and Ruas, M. P. (2014). Geometric morphometric analysis of grain shape and the identification of two-rowed barley (*Hordeum vulgare* subsp. *distichum* L.) in southern France. *Journal of Archaeological Science*, 41, 568-575. doi: https://doi.org/10.1016/j.jas.2013.09.015

Schloerke B., Cook D., Larmarange J., Briatte F., Marbach M., Thoen E., Elberg A., Crowley J. (2024). GGally: Extension to 'ggplot2'. R package version 2.2.1. url: https://ggobi.github.io/ggally/

Soladoye, M. O., Sonibare, M. A., and Chukwuma, E. C. (2010). Morphometric Study of the Genus *Indigofera* Linn. (Leguminosae-Papilionoideae) in South-Western Nigeria. *International Journal of Botany*, 6: 343-350. doi: https://doi.org/10.3923/ijb.2010.343.350

Stamp, P., Messmer, R. and Walter, A. (2012), Competitive underutilized crops will depend on the state funding of breeding programmes: an opinion on the example of Europe. *Plant Breeding*, 131: 461-464. doi: https://doi.org/10.1111/j.1439-0523.2012.01990.x

Terral, J. F., Alonso, N., Capdevila, R. B. I., Chatti, N., Fabre, L., Fiorentino, G., Marinval, P., Pérez-Jordà, G., Pradat, B., Rovira, N., Paul, A. (2004). Historical biogeography of olive domestication (*Olea europaea* L.) as revealed by geometrical morphometry applied to biological and archaeological material. Journal of Biogeography, 31. 63-77. doi: https://doi.org/10.1046/j.0305-0270.2003.01019.x

Terral, J.F., Newton, C., Ivorra, S., Gros-Balthazard, M., Morais, C., Picq, S., Tengberg, M., Pintaud, J.C. (2012). Insights into the historical biogeography of the date palm (*Phoenix dactylifera* L.) using geometric morphometry of modern and ancient seeds. *Journal of Biogeography*, 39. 929-941. doi: https://doi.org/10.1111/j.1365-2699.2011.02649.x

Ucchesu, M., Depalmas, A., Sarigu, M., Gardiman, M., Lallai, A., Meggio, F., Usai, A., Bacchetta, G. (2024). Unearthing Grape Heritage: Morphological Relationships between Late Bronze–Iron Age Grape Pips and Modern Cultivars. *Plants*, 13(13): 1836. doi: https://doi.org/10.3390/plants13131836

Ulian, T., Diazgranados M., Pironon S. *et al* (2020). Unlocking plant resources to support food security and promote sustainable agriculture. *Plants, People, Planet*, 2: 421-445. doi: https://doi.org/10.1002/ppp3.10145

Weise, S., Lohwasser, U., Oppermann, M. (2020). Document or Lose It - On the Importance of Information Management for Genetic Resources Conservation in Genebanks. *Plants* 2020, 9, 1050. doi: https://doi.org/10.3390/plants9081050

Zeileis, A., Fisher, J. C., Hornik, K., Ihaka, R., McWhite, C. D., Murrell, P., Stauffer, R., and Wilke, C. O. (2020). colorspace: A Toolbox for Manipulating and Assessing Colors and Palettes. *Journal of Statistical Software*, 96(1), 1-49. doi: https://doi.org/10.18637/jss.v096.i01

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 95–110
DOI: 10.46265/genresj.ZRCI8675
https://www.genresj.org
ISSN: 2708-3764

Solanum wild relative species indicate varying ecological resilience to climate change in Benin (West Africa)

Ahuéfa Mauricel Kégbé[a,c], Rodrigue Idohou[b,c,*], Birane Dieng[d], Gafarou Agounde[c], Anthony Egeru[e,f], Kandioura Noba[d], Achille Ephrem Assogbadjo[a,c]

[a] *Laboratoire d'Ecologie Appliquée, Faculté des Sciences Agronomiques, Université d'Abomey-Calavi, 01 BP 526, Cotonou, Bénin*
[b] *Ecole de Gestion et de Production Végétale et Semencière, Université Nationale Agriculture, BP 43, Kétou, Bénin*
[c] *Laboratoire de Biomathématiques et d'Estimations Forestières, Faculté des Sciences Agronomiques, Université d'Abomey-Calavi, 04 BP 1525, Cotonou, Bénin*
[d] *Laboratory of Botany-Biodiversity, Faculty of Sciences and Techniques, Department of Plant Biology, Cheikh Anta Diop University, P.O Box 5005, Dakar, Fann, Senegal*
[e] *Department of Geography and Climatic Sciences, Makerere University, PO Box 7062, Kampala, Uganda*
[f] *Regional Universities Forum for Capacity Building in Agriculture (RUFORUM), P.O. Box 16811, Wandegeya, Kampala, Uganda*

Abstract: Crop wild relatives are rich reservoirs of valuable genes for improving crop yields, but they have long been underestimated and neglected. Unfortunately, these resources are severely threatened in their natural habitats due to increasing stress caused by climate change and human disturbance. Recently, these wild species began receiving increasing attention for their effective inventory and sustainable conservation and use for the benefit of humanity. This study investigated the current distribution and forecasted the potential future climate change impact on ten *Solanum* wild relative species in Benin, assessed the effectiveness of protected areas in maintaining viable populations, and evaluated their conservation status using the International Union for Conservation of Nature Categories and Criteria. We used species distribution models under two socioeconomic pathways SSP370 and SSP585 projecting species ranges for the 2055 and 2085-time horizons. The models demonstrated high accuracy with an average value of the Area Under the Curve and True Skill Statistic of 0.89 and 0.74, respectively. The most suitable areas were located in the Sudano-Guinean and Guineo-Congolian zones of Benin. Furthermore, a significant proportion of these suitable areas is projected to become unsuitable for most wild *Solanum* species. Surprisingly, most of the identified hotspots were poorly represented within the existing protected area network, which appears insufficient to provide long-term refugia for the species. Nevertheless, new suitable areas were identified outside the current protected zones. Coordinated efforts are urgently needed to sustainably manage the populations of target species to enhance their future persistence in Benin.

Keywords: Crop wild relatives, climate change, protected area networks, *Solanum*, species distribution modelling

Citation: Kégbé, A. M., Idohou, R., Dieng, B., Agounde, G., Egeru, A., Noba, K. and Assogbadjo, A. E. (2025) "*Solanum* wild relative species indicate varying ecological resilience to climate change in Benin (West Africa)", *Genetic Resources*, 6(12), pp. 95–110. doi: 10.46265/genresj.ZRCI8675.

Introduction

Crop wild relatives (CWR) are wild plant species closely related to cultivated crops and possess valuable traits that can contribute to crop improvement and breeding efforts. Found in natural habitats, these wild species have co-evolved with domesticated crops over millennia (Maxted *et al*, 2006; Maxted *et al*, 2012). The genetic diversity present in CWR is crucial for enhancing the resilience and adaptability of agricultural systems, particularly in response to changing environmental conditions and the emergence of new pests and diseases (Maxted and Magos Brehm, 2023). However, habitat loss, climate change, and other human activities threaten the survival of these species (Idohou *et al*, 2025), limiting their potential contribution to food security, nutrition and human health. Despite these challenges, CWR have received and continue to receive increasing interest from scientists worldwide (Pilling *et al*, 2020).

* Corresponding author: Rodrigue Idohou (rodrigidohou@gmail.com)

Received: 12.08.2024 | Accepted: 03.07.2025 | Published online: 03.10.2025

The Solanaceae family is highly diverse, comprising over 2,000 species across approximately 90 genera, including major crops, their wild relatives, and a wide range of perennial and herbaceous annual species (Samuels, 2015; Gebhardt, 2016). Many species provide a wide range of goods and services (Gebhardt, 2016), and are used for food, medicinal and ornamental purposes (Samuels, 2015). In Benin, wild progenitors of *Solanum* are distributed throughout the country but are more abundant in the Sudano-Guinean and the Guineo-Congolian zones (Akoègninou *et al*, 2006). In fact, wild *Solanum* species also play an important role in the diet of the local population (Sarma and Sarma, 2011). Several of such species are directly consumed as wild foods rich in micronutrients, fibre and antioxidants. They diversify diets and enhance local food security, especially in rural and marginalized communities.

Fruits and leaves are used for soups or sauces and have been shown to successfully contribute to improving human health (Sarma and Sarma, 2011; Okokon *et al*, 2017). For instance, *Solanum* species have been reported to be effective in the treatment of various diseases, including malaria, stomach aches, asthma and diabetes (Sarma and Sarma, 2011). Despite their importance, wild relatives of cultivated *Solanum* vegetables are currently facing severe threats (Syfert *et al*, 2016) in their natural habitats due to habitat loss and conversion primarily from agriculture and grazing (Idohou *et al*, 2017; Lala *et al*, 2018) and more critically, climate change. Climate change impacts all aspects of agriculture, which remains a key economic sector and a cornerstone of food security and nutrition in Africa. According to the Intergovernmental Panel on Climate Change (IPCC), changes in temperature, precipitation patterns, and the frequency of extreme weather events have been observed since the 1950s (Masson-Delmotte *et al*, 2022). Climate change is now considered one of the major threats to biodiversity and food production, causing widespread disruptions to human systems and indigenous livelihoods, particularly in West Africa (IPCC, 2021). As conservation efforts increasingly adopt a 'biodiversity for livelihoods' approach, accurate information on plant distributions and ecosystem shifts under environmental change is urgently needed.

In the face of this global change, it became urgent to assess the distribution patterns of *Solanum* wild relatives under climate change. Indeed, many wild relatives of *Solanum* are found in both agricultural and protected areas (PA), which are currently experiencing high levels of disturbance (Steffen *et al*, 2015). Given the current distribution of these species and the potential impact of climate change in many ecosystems across Africa, we hypothesized that such environmental changes would negatively affect the future distribution ranges of wild relatives of cultivated *Solanum* leafy vegetables in Benin. We also assumed that abiotic factors, namely rainfall, temperature and soil, are the primary drivers influencing both the current and future distribution of these species.

Ecological niche models (ENM), also known as species distribution models (SDM) are powerful tools widely used to predict the species' abiotic niches and forecast their future distributions under climate change scenarios (Feng *et al*, 2020). Various methods have been applied to assess species distributions (Hijmans and Elith, 2017); however, the choice of the method largely depends on the aim of the study and the availability of data. Maximum Entropy Modelling (MaxEnt) (Phillips *et al*, 2006) is among the most frequently used SDM algorithms. It requires only presence data as inputs and estimates species' relative occurrence rates (Yackulic *et al*, 2013) by minimizing the relative entropy between the probability distributions of species presence and the background environment. In Benin, several researchers have explored the effectiveness of MaxEnt in predicting the potential impacts of climate change on species distribution (Daï *et al*, 2023; Hounsou-Dindin *et al*, 2023). However, no studies have specifically focused on the wild relatives of cultivated *Solanum* leafy vegetables to assess their adaptive responses to predicted climate variations. Furthermore, the assessment of PA effectiveness in conserving genetic resources must be included in such studies, given their crucial role in safeguarding biodiversity from human disturbances (Stolton *et al*, 2006). In addition, PA serve as natural refuges for many wild plants, although these species have long been overlooked within these conservation frameworks (Vargas *et al*, 2004; Burgess *et al*, 2005). Given the limited occurrence of wild *Solanum* species in PA, we hypothesized that these ecosystems are unlikely to serve as hotspots for them.

This study examined the potential impact of climate change on wild relatives of cultivated leafy *Solanum* vegetables in Benin to inform appropriate conservation strategies. Specifically, it aimed to (1) identify the abiotic factors influencing the distribution of the species and their hotspots, (2) evaluate both current and future shifts in the species distribution, (3) determine the conservation status of the target *Solanum* species in Benin based on the International Union for Conservation of Nature (IUCN) Red List Categories and Criteria and (4) assess the effectiveness of PA in conserving species distributions.

The research questions for this study are as follows: Which factors influence the geographical distribution patterns of *Solanum* wild relatives? What conservation strategies could mitigate these impacts and enhance the resilience of the species to climate change?

Materials and methods

Study area

The study was carried out in the Republic of Benin located in West Africa between 6°0' and 12°50'N, and 1°0' and 3°40'E (Figure 1). Three biogeographical zones are distinguishable, namely Sudanian, Sudano-Guinean, and Guineo-Congolian (Adomou, 2005; Figure 1). In the Guineo-Congolian, the mean annual temperature is generally 25–29°C. The relative humidity is 69–97% with a bimodal annual rainfall of 1,200mm. The land cover types are mainly fallow land and agricultural fields with the highest human population density. As for the Sudanian zone, it is the driest in the country. However, the rainfall is also unimodal, sometimes reaching 1,000mm per year (Adomou, 2005). Temperatures range from 24 to 31 °C with a relative humidity between 18 and 99%. The vegetation is characterized by dry forests, riparian forests, tree, shrub and herb savannahs and woodlands. The transition zone of the Sudano-Guinean region is dominated by bimodal rainfall with a tendency towards unimodality varying between 900 and 1,110mm per year (Adomou, 2005). The annual temperature ranges from 25 to 29°C and the relative humidity varies from 31 to 98%. Here, the vegetation is characterized by riparian forests, woodlands and dry forests. The flora of Benin is estimated to comprise

2,807 plant species, grouped into 1,129 genera and 185 families (Akoègninou *et al*, 2006). The species-rich families are Leguminosae (14.8%), Poaceae (9.3%), Rubiaceae and Cyperaceae (5% each), Asteraceae (4.6%) and Euphorbiaceae (4.3%) (Akoègninou *et al*, 2006; Adomou, 2005). Agriculture contributes about 28,04% to the gross national product, and the main crops cultivated in Benin are cereals, legumes, tubers, vegetables and industrial crops (MAEP, 2020).

Solanum species selection

A previous study by Idohou *et al* (2013) generated a list of priority CWR for Benin for conservation. The list was used to extract the existing list of *Solanum* wild relatives' species in Benin. The species names were cross-checked with those existing in the Analytic flora of Benin (Akoègninou *et al*, 2006). The species considered in this study were: *Solanum anguivi* Herb.Lamb. ex Dunal, *S. anomalum* Thonn., *S. dasyphyllum* Schumach. & Thonn., *S. distichum* Schumach. & Thonn., *S. erianthum* D.Don, *S. incanum* Ruiz & Pav., wild forms of *S. melongena* Ruiz & Pav., *S. nigrum* Vell., *S. terminale* Forssk., and *S. torvum* Buch.-Ham. ex Wall. The description, distribution, ecology, uses, threats and conservation status of each species are summarized in Supplemental Table 1.

Occurrence data

The occurrence records for the ten species were initially downloaded from open-access platforms such as the Global Biodiversity Information Facility (www.gbif.org) and RAINBIO (http://rs.tdwg.org/dwc/terms/). Records with the same coordinates and those without identification information were removed with ENMTools (www.ENMTools.com) (Warren *et al*, 2010). To minimize temporal bias and ensure compatibility with the climate datasets, records prior to 2000 were excluded (Idohou *et al*, 2017). Additionally, further occurrence data were collected during a field expedition conducted between 2019 and 2021 across the Sudano-Guinean and Guineo-Congolian zones. These records encompass a diverse array of ecosystems including farms, fallows and semi-deciduous forests. To reduce spatial autocorrelation among occurrence records, we used the occurrence rarefy tool in ArcGIS v.10.8 Spatial Analyst tool (Brown, 2014) and retained only one record per 1×1km grid cell. Initially, 1,175 raw occurrences were compiled; after filtering for uniqueness and accuracy, 693 records remained, of which 636 were ultimately used in ecological niche modelling (Supplemental Table 2). The minimum number of georeferenced occurrence records was 20, aligning with the recommendations of Wisz *et al* (2008), who advised 25 records when performing SDM for African species.

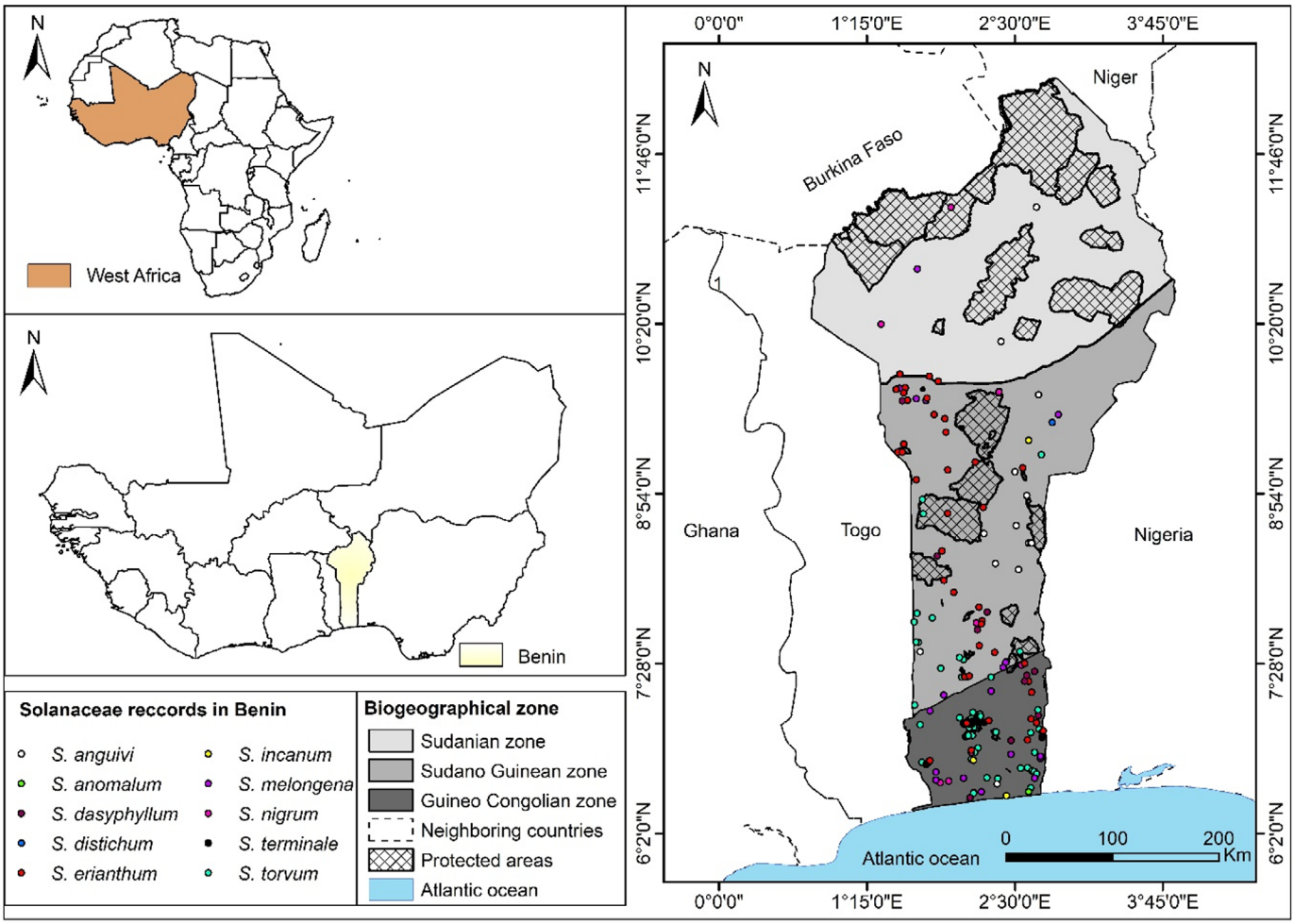


Figure 1. Location of the study area and occurrence records of *Solanum* wild relative species poplations across Benin

Table 1. Number of records of wild *Solanum* species and their sources. GBIF, Global Biodiversity Information Facility; RAINBIO, Mega-database of tropical African vascular plants distributions.

Species	Field	GBIF	RAINBIO	Raw data	Used data
Solanum anguivi Herb.Lamb. ex Dunal	30	98	0	128	20
Solanum anomalum Thonn.	10	22	9	41	25
Solanum dasyphyllum Schumach. & Thonn.	14	101	14	129	43
Solanum distichum Schumach. & Thonn.	52	52	11	115	22
Solanum erianthum D.Don	28	50	9	87	60
Solanum incanum Ruiz & Pav.	0	90	27	117	15
Solanum melongena Ruiz & Pav.	3	114	14	131	102
Solanum nigrum Vell.	30	106	6	142	115
Solanum terminale Forssk.	0	14	21	35	31
Solanum torvum Buch.-Ham. ex Wall.	45	205	0	250	203
Total	212	852	111	1,175	636

Environmental data

Current and future climatic data at 30 arcsec spatial resolution (~1km at the equator) were obtained from the Chelsa website (www.chelsa-climate.org) (Parviainen *et al*, 2008). The datasets included 19 bioclimatic variables (Supplemental Table 3) related to temperature and rainfall (Karger *et al*, 2017) and represented a baseline condition (1979 to 2013) along with two future periods: 2055 (2041–2070) and 2085 (2071–2100). Future climate conditions were projected using the IPSL-CM6A-LR climate model (Boucher *et al*, 2020), under two shared socioeconomic pathways (SSP): the middle-of-the-road scenario (SSP3-7.0) and the worst-case scenario (SSP5-8.5). These scenarios predict conditions likely for Africa by 2055 and 2085 (Williams and Jackson, 2007). SSP3-7.0 predicts an additional radiative forcing of 7W/m2 by 2100 while SSP5-8.5 forecasts an additional radiative forcing of 8.5W/m2 by 2100. The two scenarios were used to represent the moderate and the worst-case impacts of climate change for wild *Solanum* species in Benin. In addition, edaphic variables were downloaded from the Africa Soil Profiles Database (https://www.isric.org) and processed for Benin. The selected variables are summarized in Supplemental Table 3. Elevation data were obtained from the WorldClim database (www.worldclim.org). All these variables had a spatial resolution of 1km and were merged using the raster package in R 4.0.3 (R Core Team, 2023). To determine the environmental factors influencing the current distribution of each taxon, we followed a 3-step procedure. First, we excluded bio8, bio9, bio18, and bio19 because these variables frequently exhibited discontinuities between neighboring pixels across the African continent (Montoya-Jiménez *et al*, 2022). Second, we performed a Pearson correlation test using the usdm package (Naimi, 2017) to select variables with low correlation, retaining those with an absolute correlation coefficient ($|r|$) below 0.7 (Supplemental Table 4). Third, we conducted a jackknife test using the SDMtune package (Vignali *et al*, 2020) to identify variables that contributed significantly to the models. The final set of environmental factors used in modelling the ecological niche of each species was the least correlated, which made a substantial contribution to the models, and was ecologically relevant.

Modelling technique

Forty SDM were built per species (4 methods x 2 replication methods x 5 replicates). Four machine learning (ML) models available in the sdm package version 1.2-46 (Naimi and Araujo, 2016) were used: boosted regression trees (BRT) (Friedman *et al*, 2000), random forests (RF) (Breiman, 2001), support vector machines (Meyer and Wien, 2001) and MaxEnt (Phillips *et al*, 2006; Phillips *et al*, 2017). This ensemble modelling approach enhances model accuracy and robustness compared to single-algorithm methods (Ahmad *et al*, 2020). Due to the lack of true absence data, we generated 104 pseudo-absences as recommended (Barbet-Massin *et al*, 2012) using the sdm package. The dataset was split into 70% for model training and 30% for testing, and model fitting was expedited using parallel processing across four 'n-cores' (Naimi and Araujo, 2016). Model performance was evaluated using 5-fold cross-validation and bootstrapping replication methods. An ensemble function was applied to integrate predictions from all four machine learning models. Performance metrics included the area under the receiver operating characteristic (ROC) curve (AUC) and the true skill statistic (TSS). The AUC indicates the probability that the predictive power of a model is better than random prediction (AUC = 0.5) (Ramírez-Villegas *et al*, 2010; Castañeda-Álvarez *et al*, 2015); a model with an AUC value close to 1 (AUC $\geq$ 0.75) is considered to have a good fit. The TSS is a measure of the ability of the model to detect true presence (sensitivity) and true absence (specificity). It is expressed as sensitivity plus specificity -1, with TSS > 0.5 indicating a good predictive power (Zhang *et al*, 2015). Finally, suitability layers representing the current and future distribution of each species were exported in binary raster TIFF format using the average TSS value from the five replications, where a value of 1 denotes suitable habitat and 0 denotes unsuitable habitat.

Suitability habitat mapping, dynamic of the suitable areas, species richness and gap analysis

The output raster representing the habitat suitability of each species (i.e. suitable and unsuitable habitats) was imported into ArcGIS 10.8, and the current and future suitable areas were then mapped.

Habitat dynamics were quantified using the Spatial Analyst tool in GIS by counting the total number of pixels corresponding to the current and future distribution. The Rate of Change Index (RCI) was calculated using the following formula (Coulibaly *et al*, 2021):

$$\triangle = \frac{(FA_i - CA_i) * 100}{TA_i}$$

FA corresponds to the future area (e.g. suitable) of a species in the target horizon under scenario i (here horizon = 2055 and i = SSP3-7.0 and SSP5-8.5); CA corresponds to the current distribution area (e.g. suitable); Δ is the percentage of area gain (Δ > 0) or lost (Δ < 0) and stable Δ = 0. A pattern of change was qualified as minor when values varied between (0–5%), (6–15%) indicated moderate decrease and the upper 21% major decrease.

The dynamics of species richness over time were evaluated by summarizing the binary (1 = presence, 0 = absence) layers for the present and future distributions of the ten *Solanum* species. In this context, species richness referred to the number of *Solanum* species found in 1km² area, based on the resolution of environmental variables used in the models. Richness levels were then classified into five different categories based on the number of Solanaceae species modelled: high (8–10), moderate (5–7), low (2–4), very low (1) and no species (0).

Gap analysis was conducted by overlaying the shapefile of Benin's PA network (World Database on Protected Areas, UNEP-WCMC 2023) with the habitat suitability and species richness maps. This analysis aimed to evaluate the effectiveness in covering suitable habitats and to identify potential priority areas for future conservation efforts.

Conservation status assessment of the species

In this study, we assessed the conservation status of ten *Solanum* species in Benin using projections from SDM under future climate scenarios. Although many studies traditionally apply IUCN Criterion B by estimating the extent of occurrence (EOO) and area of occupancy (AOO) (e.g. Dassou *et al*, 2024), criterion A3(c) is more appropriate for SDM as it considers the expected reduction in population size inferred from the projected decline in habitat suitability. Criterion A3(c) relates to a reduction in EOO, AOO or habitat quality up to a maximum of 100 years (IUCN, 2024). According to IUCN guidelines, a species is classified as Extinct (EX) if it is projected to lose 100% of its suitable area, Critically Endangered (CR) with ≥ 80% reduction, Endangered (EN) with ≥ 50% and < 80%, Vulnerable (VU) with ≥ 30% and < 50%, Near Threatened (NT) with < 30%, and Least Concern (LC) if its suitable area is stable or increasing. The methodology for calculating the percentage change in the suitable area and for assigning IUCN categories was detailed in the previous section. In Benin, threats to CWR, including *Solanum* species, fall into three major categories: (1) agricultural expansion and urbanization, leading to habitat loss, (2) overharvesting, resulting in population decline, and (3) invasive species and climate change, also contributing to population decline (Idohou *et al*, 2013). Among these, agricultural encroachment has been identified as the most significant threat to *Solanum* species nationwide, severely affecting their suitable habitats. Consequently, the preliminary conservation status of each species was determined by combining the predicted reductions in suitable area from SDM with an understanding of ongoing threats.

Results

Variable contribution, model validation and performance evaluation

A total of 21 non-correlated variables (Figure 2) were identified as important in determining the modelled distribution of the ten *Solanum* species. Among these, temperature- and precipitation-related variables had a significantly greater influence on the SDM compared to soil and elevation variables. Individually, isothermality (bio3) emerged as the most influential variable shaping the distribution of the *Solanum* species. However, precipitation of the driest quarter (bio17) had the highest individual contribution, particularly for *S. nigrum* with contributions of 27.16% and 24.23%, respectively. *S. melongena* showed the highest dependence on bio3 (42%), followed by *S. distichum* (27.27%), *S. terminale* (20.87%), *S. nigrum* (20.11%), and *S. dasyphyllum, S. erianthum* and *S. incanum*, each with contributions around 16.5%. Annual temperature (bio1) moderately influenced the distribution of *S. anguivii* and *S. incanum*, each with a contribution of around 7%. The mean diurnal range (bio2) was a major predictor for *S. dasyphyllum* (38.17%) and *S. distichum* (20.66%). Temperature seasonality (bio4) and maximum temperature of the warmest month (bio5) were key variables for *S. torvum* (42.11%) and *S. anomalum* (16.04%), respectively. Precipitation seasonality of (bio15) was the most influential factor for *S. anguivii* (47.74%), followed by moderate to low contributions for *S. erianthum* (33.37%), *S. distichum* (24.39%) and *S. torvum* (16.84%). *S. terminale* strongly depended on precipitation of the driest month (43.10%) (bio14) while *S. anomalum* and *S. erianthum* also responded significantly (33.5% and 37.5%, respectively). For *S. incanum*, however, bio14 had a relatively low contribution (13.66%). Overall, soils characteristics showed a weak contribution for all the species. Among those, exchangeable acidity in subsoil (EACKCL_d1) and exchangeable magnesium (EMGX_d2) had some effect on *S. incanum, S. anomalum, S. nigrum, S. torvum* and *S. melongena*. Bulk density in topsoil (BLD_d1) contributed 10.05% for *S. distichum*, while subsoil bulk density (BLD_d5) contributed approximately 13% for both *S. anguivi* and *S. dasyphyllum*. Subsoil aluminium concentration (ALUM3S_d1) had contributions ranging from 13% to 18% for *S. anguivi, S. anomalum* and *S. nigrum*. Soil texture class at 45cm depth weakly affected *S. terminale* (19.64%), while deeper texture layers (TEXMHT_d5) impacted *S. dasyphyllum* (17.11%), *S. distichum* (17.64%) and *S. torvum* (3.68%).

Silt fraction (SLTPPT_d1) influenced *S. melongena* (18.7%) and *S. erianthum* (7.6%). Soil pH in the topsoil (PHIHOX_d1) slightly influenced S. terminale (7.73%), and subsoil pH (PHIHOX_d6) impacted *S. erianthum* (6.27%) and *S. melongena* (10.65%).

Elevation (DEM) was a significant predictor only for *S. incanum*, contributing 29.86% to its distribution model.

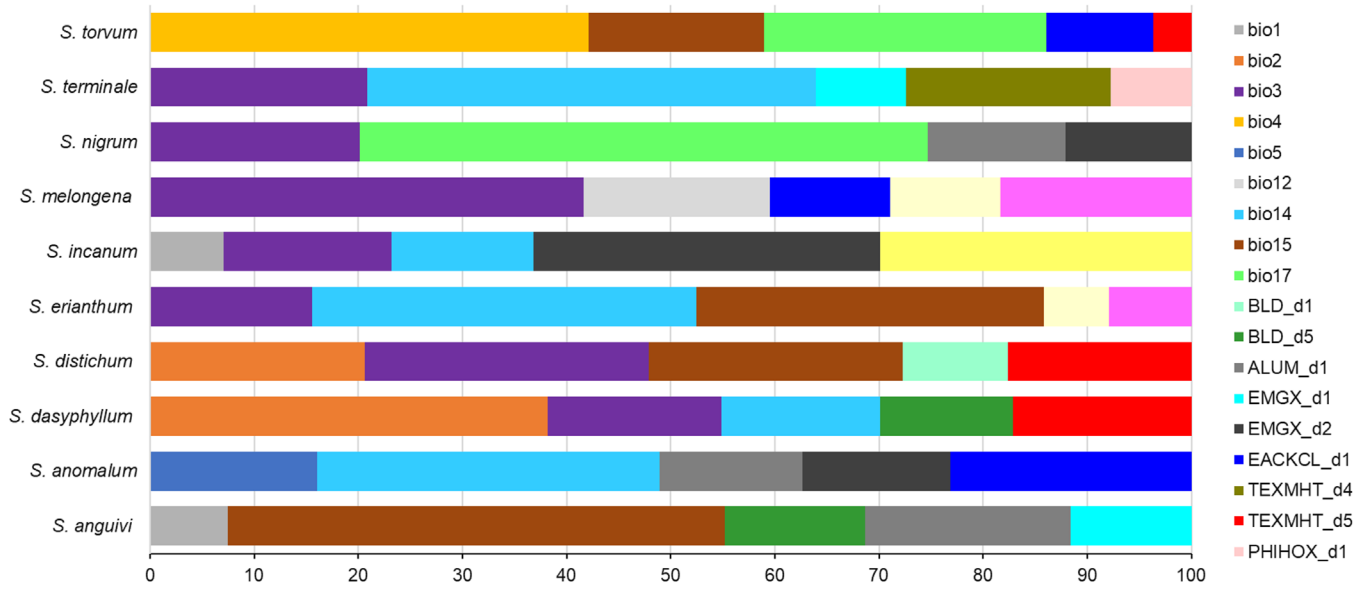


Figure 2. Contribution (%) of the variables to the models. The models were built using an ensemble of four algorithms boosted regression trees (BRT), random forests (RF), support vector machines (SVM), and MaxEnt with five replications each. Variable contributions were averaged across algorithms and replications to obtain final estimates.

The results showed an average AUC value of 0.89 across the ten *Solanum* species (Table 2), indicating that SDM performed well in predicting their potential distributions. Similarly, the average TSS value was 0.74, reflecting a strong agreement between observed and predicted occurrences.

Table 2. Model performance based on the two metrics. AUC, Area under the curve; TSS, True skill statistic.

Species	AUC	TSS
S. anguivi	0.9	0.8
S. anomalum	0.9	0.81
S. dasyphyllum	0.83	0.63
S. distichum	0.83	0.69
S. erianthum	0.96	0.85
S. incanum	0.8	0.5
S. melongena	0.87	0.69
S. nigrum	0.92	0.78
S. terminale	0.91	0.82
S. torvum	0.94	0.81

Geographic distribution patterns of the species under current and future climate conditions

The categorization of habitats under current conditions showed a diverse range of suitable areas for all species except for *S. anomalum* (Figure 3). The primarily suitable habitats for the studied species were predominantly located in the Sudano-Guinean and Guineo-Congolian zones. In contrast, a significant portion of the coastal region was unsuitable for some species, notably *S. anguivii* and *S. incanum*. For *S. anomalum*, suitable areas were limited and mostly concentrated along the eastern coast of the Guineo-Congolian zone. A similar distribution pattern was observed for *S. dasyphyllum*, although this also showed a broader extent of suitable habitat extending into the Sudanian zone.

Under future climate scenarios (SSP3-7.0 and SSP5-8.5) for the horizons 2055 and 2085, a general decline in suitable habitats was projected for many species. However, *S. melongena* and *S. dasyphyllum* showed a clear expansion of suitable habitats. The suitable areas for *S. distichum*, *S. nigrum* and *S. torvum* remained relatively stable across all scenarios and time horizons. In contrast, *S. anomalum*, *S. incanum* and *S. terminale* were projected to lose the suitable habitats by 2055. For *S. incanum*, the distribution remains narrow and relatively stable in the future compared to its initial distribution. *S. erianthum* showed a decline in suitability, with its distribution shifting southwards into the Guineo-Congolian zone – this spatial trend was consistent across both horizons and scenarios. When overlaying habitat suitability maps with the PA network, varied patterns emerged. In the Sudanian zone, parts of the suitable habitats for some species overlapped with existing PA except for *S. melongena* and *S. dasyphyllum*, which were largely unprotected. Under future climatic conditions, some PA emerged as important conservation zones. For instance, the classified forests of Kétou and Dogo are projected to harbour significant hotspots of *S. terminale*. Similarly, the Lama Forest appears to be a potential refuge for *S. incanum*.

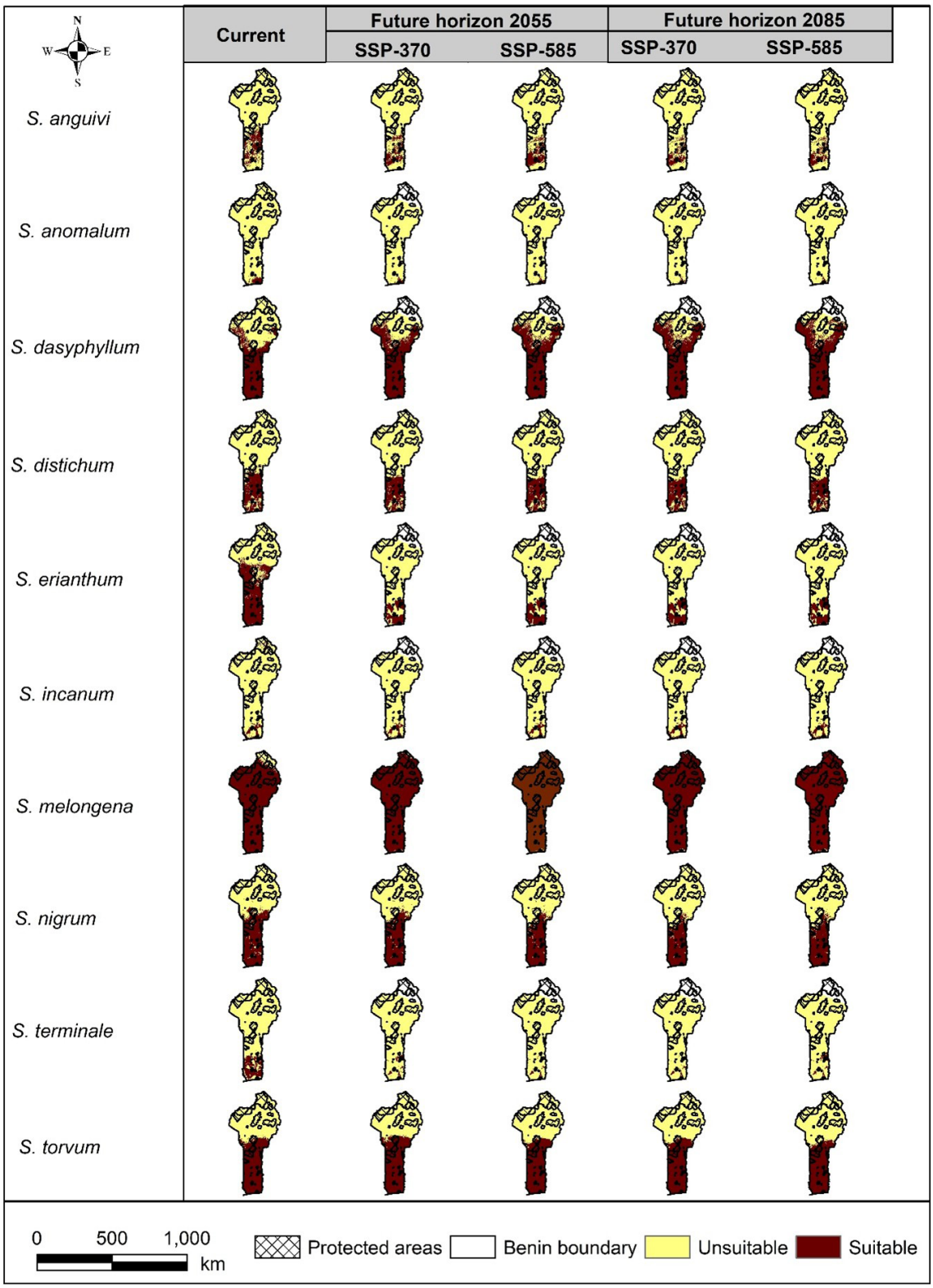


Figure 3. Suitable habitats for the ten *Solanum* species in the current and under future conditions accounted for socioeconomic pathways SSP3-7.0 and SSP5-8.5 by 2055 and 2085.

Dynamics of suitable habitats for *Solanum* species across Benin

The projected distribution of wild relatives of cultivated leafy vegetables of *Solanum* species, indicated that, under both present and future scenarios, the extent of suitable habitats will decrease for most species (Supplemental Table 5). Overall, eight out of ten species are expected to experience a reduction in suitable areas, while only two are projected to see an expansion.

Species such as *Solanum anguivii, S. anomalum, S. distichum, S. erianthum, S. incanum, S. nigrum*, *S. terminale* and *S. torvum* are projected to lose suitable habitats. Under current conditions, *S. torvum* occupies a suitable area of 4,560.47km². By 2055, this area is projected to increase by 7.06% under the SSP3-7.0 scenario but to decrease by up to

8.41% under SSP5-8.5. Other species, such as *S. distichum* and *S. incanum*, are predicted to experience relatively minor reductions in habitat suitability, with losses not exceeding 21%. The current suitable area of *S. distichum* is estimated at 2,601.14km², with projected reductions ranging from 11.09% to 20.49%, the largest occurring under SSP3-7.0 by 2085 (Supplemental Table 5). For *S. incanum*, the current suitable area is 743.28km², with projected losses of 17.74% under SSP3-7.0 and 14.06% under SSP5-8.5 by 2055. By 2085, this downward trend is expected to continue, with projected reductions of 14.06% and 14.06% under SSP3-7.0 and SSP5-8.5, respectively.

Moderate habitat losses, i.e. not exceeding 45%, were observed only for *S. anguivii* and *S. nigrum*. The current suitable areas for *S. anguivii* and *S. nigrum* were estimated at 1,583.58km² and 4,288.29km², respectively. However, the projected trends in habitat suitability are not stable for either species. For *S. anguivii*, a minor reduction of 15.09% is projected under the SSP5-8.5 scenario by 2055, whereas a more substantial reduction of 40.88% is expected under the same scenario by 2085. Similarly, under SSP3-7.0, reductions of 26.75% and 38.80% are projected for 2055 and 2085, respectively. *S. nigrum* exhibits the same pattern, with relatively minor reductions of 7.17% and 16.47% under SSP3-7.0 for 2055-time horizon and SSP5-8.5 time horizon, respectively, increasing to 21.91% under SSP3-7.0-2085 and 27.33% under SSP5-8.5-2085.

Species with major loss included *S. anomalum, S. erianthum,* and *S. terminale*, each experiencing projected reductions in suitable habitat exceeding 60% under all future scenarios (Supplemental Table 5). For instance, *S. anomalum*, which currently occupies 282.00km², is expected to lose up to 64.91% of its suitable area by 2085 under SSP5-8.5. Similarly, *S. erianthum*, with a current suitable area of 4,971.81km², is projected to decline by as much as 72.53% by 2055 under SSP3-7.0. *S. terminale*, currently distributed across 1381.48km², faces reductions up to 86.37% under the same scenario by 2085. These sharp contractions indicate a significant risk of habitat-driven population decline and potential local extirpation without targeted conservation efforts. In contrast, *S. dasyphyllum* and *S. melongena* are projected to expand their suitable habitats under all climate scenarios (Supplemental Table 5). *S. dasyphyllum*, currently found in 5,314.32km², may gain up to 57.15% by 2085 under SSP5-8.5, while *S. melongena*, with a current area of 10,717.61km², is projected to expand by 6.13% under the same scenario. Despite these positive trends, their future distributions remain susceptible to uncertainties inherent in climate projections, highlighting the need for continued monitoring. These contrasting patterns underscore the importance of adopting species-specific strategies for effective conservation planning in the face of climate change.

Dynamic of suitable areas of *Solanum* species within the protected areas network

Supplemental Table 6 presents the dynamics of suitable areas for *Solanum* species within the protected areas (PA) network. Overall, two main trends were observed, mirroring those at the national level. Eight out of ten species (*S. anguivi, S. anomalum, S. distichum, S. erianthum, S. incanum, S. nigrum, S. terminale,* and *S. torvum*) are projected to lose suitable area within PA, while only *S. dasyphyllum* and *S. melongena* are expected to gain.

S. anguivi currently occupies 1,790.92km² within PA and is projected to decline by up to 58.89% under SSP3-7.0 by 2085, reducing its extent to 736.27km². *S. anomalum*, with a current extent of 47.26km², shows consistent losses of 63.48% across all future scenarios, resulting in a stable yet severely reduced extent of 17.26km². *S. incanum* presently covers 3,158.83km² and is projected to decline by roughly 50% across all scenarios, reaching 1,561.9km² under SSP5-8.5 by 2085. *S. torvum* has a current distribution of 964.46km² and is projected to undergo losses under all scenarios, with the most severe reduction (62.66%) under SSP3-7.0 by 2085, dropping to 360.16km². While it displays a negligible gain of 0.55% in 2055 under SSP3-7.0, this is not sustained in later scenarios.

Moderate losses were observed for *S. distichum*, currently distributed over 3,472.39km². The most substantial reduction is projected under SSP5-8.5 by 2085, where the extent declines to 1,077.04km², accounting for a 68.98% reduction. *S. nigrum* shows contrasting dynamics, with a projected gain of 36.83% under SSP5-8.5 by 2055, but this is followed by a sharp decline (62.73%) under SSP5-8.5 by 2085, reducing its suitable area to 2,816.51km². *S. terminale* follows a similar trajectory, dropping from 7,557.58km² to 2,816.51km² under SSP5-8.5 by 2085 (62.73%). *S. erianthum* faces the most severe reductions, shrinking from 7,993.3km² to just 1,187.15km² by 2055 under SSP3-7.0 (85.15%) and showing similarly drastic losses under all scenarios.

By contrast, *S. dasyphyllum* and *S. melongena* are projected to expand within the PA network. *S. dasyphyllum* currently occupies 9,668.47km² and may increase its extent by up to 78.65% (17,272.45km²) by 2085 under SSP5-8.5. *S. melongena*, with the largest current extent (31,355.97km²), shows consistent expansion across all scenarios, peaking at 18.51% (37,159.88km²) under SSP5-8.5 by 2085. These contrasting dynamics emphasize the need for targeted conservation planning tailored to each species' future trajectory within the protected areas network.

Solanum richness dynamic accounted for current and future distribution

Under current conditions, SDM indicate a decline in *Solanum* species richness towards the southern regions (Figure 4). The highest richness category comprising eight to ten species was predominantly concentrated in the Guineo-Congolian and Sudano-Guinean zones. This richness class currently covers an estimated 681.11km*2* (5.94% of the total area). However, projections under future climate scenarios suggest a reduction of approximately 4.5% in this richness category (Supplemental Table 7). Moderate richness levels (5–7 species) were observed in the Ouémé-Boukou and Agoua PA, which are located in the Guineo-Congolian and Sudano-Guinean zones, respectively (Figure 4). Compared to the current condition, the future richness of the wild *Solanum* species is likely to decrease southwards (Figure 4). The current distribution area of *Solanum* species richness was estimated at 2,374.49km^2 (20.69%) with a projected reduction ranging from 3.18 to 5.34% under future climate scenarios. However, our findings indicated an increase in the species richness (2–4 species class) in the Sudanian zone. When we considered the richness class 2–4, an expansion towards the Guineo-Congolian and Sudano-Guinean zones

could occur. This class occupied 11.28% (1,294.75km^2) of the total area (Supplemental Table 7). The increase in area ranged from 3.56–10.54% with the lowest value under SSP3-7.0 by 2055 and the highest value at SSP5-8.5 in 2085. As for the richness class (1 species), it occupied 16.39% (1,880.91km^2). The increase ranged from 10.89–25.25% with the highest percentage under SSP5-85 in 2085 in terms of PA. The pattern of species richness shows that higher richness classes (5–7 and 8–10 species) are generally located outside the PA network. However, notable exceptions include the Kétou and Dogo PA, situated in the transition zones between the Sudano-Guinean and Guineo-Congolian regions. Under current conditions, richness classes with 0 and 1 species together account for approximately 66% of the richness within the PA network (Supplemental Table 8). Meanwhile, the 5–7 and 8–10 species richness classes make up only 11.37% of the total PA coverage.

The conservation status assessment of the *Solanum* genus

The conservation status of the ten *Solanum* species in Benin was based on the IUCN criteria. The results revealed varying levels of climate-induced vulnerability overall. One species was classified as CR, two as EN, three as VU, two as LC and two as NT. This was based on their major threats and distribution range (Supplemental Table 9), as well as the projected percentage changes in suitable habitat under the SSP3-7.0 and SSP5-8.5 scenarios for 2055 and 2085 (Supplemental Table 10).

S. terminale is the most severely affected species, with a projected habitat loss of over 70% under all future scenarios, reaching 85.86% under the SSP5-8.5 scenario by 2055. Coupled with its restricted distribution and high exposure to anthropogenic pressures, this supports its classification as CR. Similarly, *S. erianthum* and *S. anomalum* show consistent declines of between -53.06% and -69.42%, which justifies their classification as EN given their sensitivity to fire and habitat degradation. At the next level of severity, *S. incanum*, *S. distichum* and *S. anguivi* experience a moderate decline in suitable habitat, ranging from -14.10% to -40.88%. Despite having broader ecological amplitudes, these species face significant local threats, such as overgrazing, soil disturbance and agricultural encroachment. This warrants their classification as VU.

By contrast, *S. dasyphyllum* and *S. melongena* are projected to remain stable or even increase in number under future climate conditions, with changes ranging from -3.95% to +57.15%. Their wide ecological tolerance and lower exposure to critical threats mean they are designated as LC. However, ongoing monitoring is advised to detect any future shifts in vulnerability.

Finally, *S. nigrum* and *S. torvum* exhibit variable responses across scenarios, with projected changes ranging from a decline of 27.33% to an increase of 2.93%. While their overall decline is less severe, uncertainties surrounding their true wild distribution and local extinction risks justify their classification as NT.

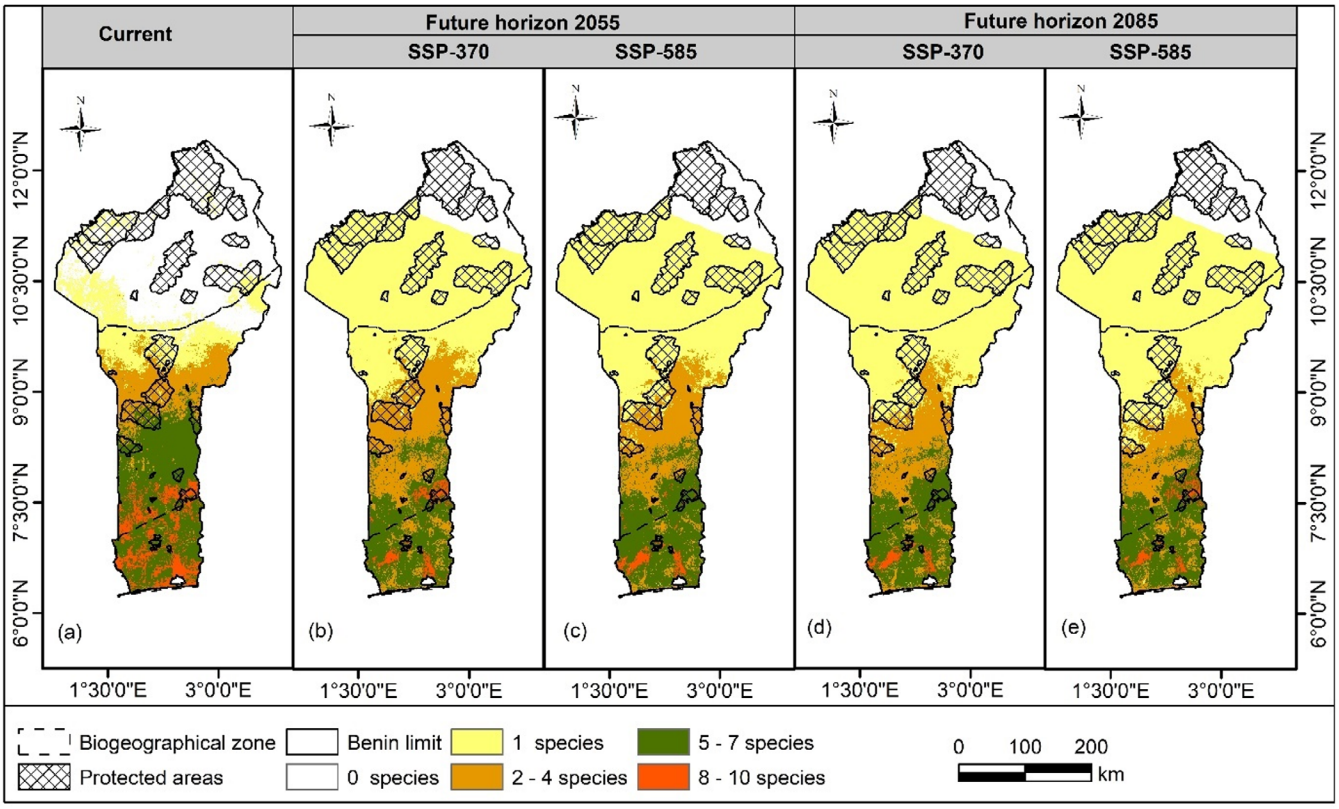


Figure 4. Species richness dynamics accounted for the current and future distribution.

Discussion

Factors determining the distribution of *Solanum* wild relatives

The distribution and suitable habitat of the species studied are largely influenced by key abiotic factors such as rainfall, temperature, and soil properties. Climate variables (e.g. temperature and precipitation) and edaphic conditions play a fundamental role in determining species presence in a given environment (Lewis *et al*, 2017). However, the ability of these species to persist in complex environment also depends on their interactions with biotic factors – competition, parasitism, commensalism – and on dispersal constraints that were not included in our models. Moreover, the models did not consider species' phenotypic plasticity and evolutionary adaptation to changing environments (Pidwirny, 2006). Integrating these aspects in future modelling efforts could substantially improve the predictive accuracy and ecological realism of species distribution models, especially under changing climate scenarios.

In this study, we found that the distribution patterns of the ten *Solanum* species were primarily influenced by bioclimatic variables, with isothermality emerging as the most dominant factor. Edaphic variables also played a role, albeit with a relatively smaller contribution to the distribution of each species.

Our findings align with those of Manda *et al* (2022) who identified isothermality as the second most important factor when modelling the potential impact of climate change on *Vigna* wild relatives. Similar patterns in the influence of bioclimatic and soil variables have been reported for tree species such as *Balanites aegyptiaca* (L.) Delile (Chérif *et al*, 2022) in Chad, *Fontainea* species (Brunton *et al*, 2023) in Australia and palm species in Benin (Idohou *et al*, 2017; Salako *et al*, 2019). In the case of wild relatives of cultivated *Solanum* leafy vegetables, their responses to soil characteristics varied among species. Indeed, several studies have highlighted the high nutritional value of African indigenous vegetables, including *Solanum* species, attributing their richness in nutrients to underlying edaphic conditions (Chinedu *et al*, 2011; Keatinge *et al*, 2011) . This relationship suggests that soil properties play a significant role in shaping the nutrient profile of these plants. Additionally, the observed variability among species likely reflects inherent physiological traits that influence water requirements and nutrient uptake. Since *Solanum* seed germination is sensitive to environmental stress, favourable soil and moisture conditions are essential for successful germination and, consequently, for the long-term persistence of the species in a given habitat (Stanton *et al*, 2012). These factors should be carefully considered when it comes to cultivation or conservation. Understanding their responses to environmental changes is therefore essential for designing effective conservation plans. As reported by Aksoy *et al* (2021), evaluating the impact of climate change is crucial for the conservation of *S. tuberosum*. Similarly, Ogundola *et al* (2023) reported that soil characteristics, particularly silty loam soils at a depth of 2cm, enhance the germination and viability of *S. nigrum* seeds. In our study, the most influential soil factors included exchangeable acidity, exchangeable magnesium, bulk density, silt content, and soil pH. However, further research is necessary to better understand how these specific edaphic parameters affect seed germination, viability and overall fitness of *Solanum* species under both current and future environmental conditions.

Suitable areas of *Solanum* wild relatives and climate change

The suitable habitats of wild relatives of cultivated *Solanum* leafy vegetables are likely to be largely unstable in the coming decades, reinforcing our first hypothesis that abiotic factors are the main drivers of the current and future distribution of wild *Solanum* species.

Climate change is predicted to increase the temperature by 1.5–2°C by 2100 in West Africa, and to cause an uneven distribution of precipitation (IPCC, 2021). Based on these projections, the suitable habitats of the ten wild *Solanum* species are likely to be severely altered by the 2055 and 2085time horizons under both SSP3-7.0 and SSP5-8.5 scenarios. As noted by Guisan and Thuiller (2005), several factors (population dynamics, migration, local adaptations, ecological interactions, disease prevalence and human intervention) can influence the distribution of a species. Our results indicate substantial habitat losses for nearly all species by both the 2055 and 2085 horizons under the two scenarios. Surprisingly, a rise in the number of suitable areas was observed for wild *S. melongena* and *S. dasyphyllum*. This suggests that climate change could have a positive impact on the potential distribution of these species. This expansion could imply an increase in the potential cultivation area for domesticated or improved forms of these species. However, further agronomic and socioeconomic studies would be necessary to confirm this. Under the SSP5-8.5, the situation appears particularly critical across both time horizons, with predicted high habitat losses. These findings align with numerous other studies which reported shifts in the distribution of many CWR. For instance, Manda *et al* (2022) reported major changes in the distribution of Vigna wild relatives in Benin. Similarly, findings in Southern Africa indicated that the majority of regionally priority CWR are expected to be negatively impacted by climate change, threatening their survival (Magos Brehm *et al*, 2022). Additionally, van Treuren *et al* (2020) demonstrated a reduction of suitable areas for CWR in the Netherlands. Seed dispersal is a crucial ecological mechanism that could either hinder or facilitate plant species' response to climate change (Öckinger *et al*, 2010). According to the same author, species with limited dispersal ability would be the most affected by the environmental changes compared to those with strong dispersal capacity. *Solanum* species are dispersed by mammals, bats, birds, human-induced actions, as well as wind and water flow (Roberts and Florentine, 2022), which are predicted to be highly disturbed by climate change, potentially disrupting natural dispersal processes. Besides, in the Sudano-Guinean zone, habitat fragmentation driven by agricultural expansion and rapid population growth has already impeded natural dispersal pathways (Neuenschwander & Adomou, 2017; Abdul Aziz *et al*, 2024). Consequently, natural dispersion may not be efficient with humans as a major dispersing agent. Compounding this issue, invasive species pose a significant threat to native biodiversity, challenging both ecological resilience and conservation efforts. In Benin, many studies showed the detrimental impacts of invasive species on native species particularly in PA (Gbètoho *et al*, 2017) and, at the same time, models predicted their capacity to thrive

in the changing climate (Fandohan *et al*, 2015). This is an important aspect to consider since most wild *Solanum* species are herbaceous and, thus, highly susceptible to displacement by aggressive invaders such as *Chromolaena odorata* (L.) R.M.King & H.Rob.

Solanum wild relatives and their conservation within protected areas

The results showed that, with a few exceptions in the Sudano-Guinean and Guineo-Congolian zones, most existing PA offer limited potential for conserving the target *Solanum* species. Future projections confirmed these patterns, although localised expansion or contraction may occur. The destruction of natural habitats on a large scale, combined with climate change, is one of the main threats to plant species (Hudson *et al*, 2014), and this includes CWR (Hunter *et al*, 2012; Magos Brehm *et al*, 2022; Maxted *et al*, 2012). Protected areas are often cited as refuges for threatened species (Le Saout *et al*, 2013); their ability to maintain populations within defined boundaries offers a valuable conservation mechanism, especially in regions undergoing rapid environmental change (Le Saout *et al*, 2013; Mao *et al*, 2020). However, the mere presence of a species within a PA does not guarantee its protection. More often than not, this represents passive *in situ* conservation, whereby species may persist without any specific monitoring or management actions. In the context of accelerating climate change, there is a critical need for active *in situ* conservation, which involves the direct management, monitoring and support of populations over time, to ensure long-term effectiveness. Our findings align with those of Manda *et al* (2022), who reported that the existing PA network was inefficient in conserving wild Vigna species, a trend also observed in several other countries (e.g. Davis *et al*, 2019; Ratnayake *et al*, 2021). Conversely, Idohou *et al* (2017) highlighted the potential of PA for conserving palm species in Benin. However, whether such potential persists under future climate scenarios remains uncertain and warrants further investigation. Despite being considered refuges for many plant species, our projections suggest that both the current and future distributions of the studied *Solanum* species may fall outside the existing PA network. This supports the hypothesis that current PA may not be sufficient to conserve wild *Solanum* habitats under climate change. Therefore, it is urgent to identify and prioritize additional conservation areas that are more likely to remain suitable in the future. Sub-Saharan Africa, particularly its arid and semi-arid zones, has been identified as one of the most climate-vulnerable regions in the world (Sintayehu, 2018), with native species facing multiple threats. Under such pressures, the distribution of most wild *Solanum* species is expected to become increasingly unstable in future climate scenarios. As a result, proactive conservation actions are essential (Schuster *et al*, 2023). We thus emphasize the need to conserve wild *Solanum* not only within PA networks but also beyond them, through mechanisms such as the IUCN's Other Effective Area-Based Conservation Measures (OECMs). In particular, agroforestry systems and home gardens represent promising complementary conservation environments. If properly designed and supported, these managed landscapes could host viable populations of wild relatives and provide additional resilience to climate stressors.

Implications for better management of *Solanum* wild relatives

In this study, we investigated the threats to the conservation of ten wild *Solanum* species under different future climate scenarios. Many studies demonstrated the great importance of CWR worldwide (e.g. Ng'uni *et al*, 2019; Tas *et al*, 2019; Maxted and Magos Brehm, 2023). Because of the key role they play in crop improvement and food production, there is a global call for the conservation and preservation of these resources (Maxted *et al*, 2010; Maxted *et al*, 2012). Assessment of the current and future distribution of *Solanum* species under climate change revealed a decline in suitable habitats for the majority of the species. Furthermore, the assessment based on IUCN Red List Categories and Criteria indicated that most species in Benin are threatened, underscoring their high risk of genetic erosion, a pattern similarly observed among numerous CWR taxa worldwide. Our study identified priority areas for target *Solanum* species conservation as suggested by Maxted *et al* (2009). The target species to be prioritized for conservation include *S. anguivi, S. erianthum, S. anomalum* and *S. terminale* (VU, EN and CR) since they are likely to continue losing suitable areas of distribution. Monitoring population status and habitat conditions, restoring degraded ecosystems, and analyzing genetic diversity are all essential steps for the effective conservation of *Solanum* species. Monitoring enables the early detection of population declines, while restoration efforts help ensure that species can persist in their natural habitats. Preserving a broad spectrum of genetic diversity is crucial for the long-term adaptability and survival of species. It also ensures that these genetic resources remain available to plant breeders, helping to improve crops and ensure food security. Ecogeographic diversity, which integrates environmental and geographical variation, can serve as a useful proxy for capturing genetic variation (Parra-Quijano *et al*, 2012). These species could also be protected in several sacred forests throughout the country. Although they are limited in size, sacred forests are often better conserved than most classified forests (Adomou *et al*, 2007), making them suitable refuges for the conservation of *Solanum* taxa through introduction or reintroduction programmes. To minimize ecological risks, such introductions should occur within each species' historical distribution range. In parallel, *ex situ* conservation is a valuable complementary strategy, particularly for species with critically reduced habitats. However, the long-term success of these measures depends on a better understanding of the ecological requirements of each species, including their capacity to adapt and their productivity under changing environmental conditions.

Conservation in the Wari-Maro and Belefoungou forests could be more effective than in the semi-deciduous forests of southern Benin, such as the Lama reserve or sacred forests, primarily because many of the *Solanum* species studied are heliophilous, meaning they thrive in environments with high light availability. Unlike in dense, closed-canopy forests, where light penetration is limited, the drier, more open structure of the Wari-Maro and Belefoungou forests provides favourable microhabitats for these species that demand light. These conditions support germination, growth and reproductive success, thereby enhancing their long-term survival prospects in these forest systems. Furthermore, classified forest reserves, including sacred forests, have decreased in recent decades,

implying the loss and fragmentation of suitable habitats for many taxa in Benin (Alohou *et al*, 2017). According to Pinto *et al* (2024), the continued loss of suitable areas may result in significant genomic erosion for metapopulation taxa. The impact of these losses on *Solanum* species should also be investigated. Finally, the occurrence of invasive species in natural habitats needs to be considered when defining these strategies.

Limitations of the approach

In this study, we investigated the current distribution of ten wild *Solanum* species in Benin and made predictions for their future distribution. The model accuracy and robustness were assessed by computing the AUC and TSS values of the final MaxEnt outputs (Allouche *et al*, 2006). However, we did not include seed dispersal limitations or constraints that could hinder the spread of species across landscapes. Furthermore, we did not account for physiological, phenological and morphological traits in our models (Koebsch *et al*, 2019; Agounde *et al*, 2025).

Although the SDM-based projections offer valuable insights into the potential impact of climate change on the distribution of wild *Solanum* species in Benin, it is important to acknowledge that our assessment is still preliminary. The absence of data on demographic trends and population viability may result in certain taxa being underestimated as vulnerable. Consequently, these results should be interpreted as indicative rather than definitive. Future assessments will require the integration of ecological and population-level data in order to better reflect extinction risk and guide conservation priorities.

Conclusion

Developing an effective conservation strategy for a given species requires accurate information on the species' distribution range as a fundamental basis. Using SDM tools and IUCN criteria, we assessed the potential distribution, habitat suitability and preliminary conservation status of ten wild *Solanum* species under current and projected climatic conditions. Overall, most of the species currently have a wide distribution in the Sudano-Guinean and Guineo-Congolian zones of Benin. Significant declines in potentially suitable habitats were found in the two scenarios for 2055 and 2085. However, for more accurate projections, we recommend additional distribution studies taking into account physiological, phenological and morphological data.

The study demonstrated that the current protected area network is ineffective in ensuring the long-term survival of these species in the face of changing climatic conditions. Therefore, conservation efforts should prioritize active *in situ* strategies, such as restoring degraded habitats, reinforcing natural populations within their historical ranges and integrating climate-smart management plans. Given the limited coverage and effectiveness of existing mechanisms in Benin, such actions are urgently needed. In parallel, *ex situ* conservation measures, including seedbanks and living collections, should complement *in situ* efforts to preserve the genetic diversity of these species and facilitate their potential use in future food security initiatives.

Acknowledgements

This research work was supported by European Commission Directorate-General for International Partnerships through the Regional Academic Exchange for Enhanced Skills in Fragile Ecosystems Management in Africa (Grant Number Nr 2017-2861/001-001) and the PhD fellowship from the Organization for Women in Science for the Developing World (Grant Number: 3240314469) provided to Ahuéfa Mauricel Kégbé. Ahuéfa Mauricel Kégbé is very grateful to Dr Valère Salako and Dr Ablaye Ngom for their encouragement and support during the implementation of this research. Rodrigue Idohou acknowledges support for BITC research and training visits from the John Roger Seamans Biodiversity Foundation, which enabled collaboration on this paper.

Author contributions

Ahuéfa Mauricel Kégbé, Rodrigue Idohou: conceptualization, methodology, visualization, software, resources, formal analysis, writing original draft, writing, review and editing; Birane Dieng: writing original first draft; Gafarou Agounde: software, formal analysis, resources, writing original first draft; Anthony Egeru: review and editing; Kandioura Noba: review and editing; Achille Ephrem Assogbadjo: supervision, software, writing, review and editing. All authors read and approved the final manuscript.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Supplemental data

Supplemental Table 1. The ecology, potential use and conservation status of the *Solanum* species

Supplemental Table 2: Species occurrences used in ecology niche modelling

Supplemental Table 3. Description of the environmental variables

Supplemental Table 4. Correlation analysis

Supplemental Table 5. Dynamic of the distribution of *Solanum* species in the whole study area

Supplemental Table 6. Dynamic of the distribution of *Solanum* species in protected areas

Supplemental Table 7. Spatial reduction of extent of richness class in the whole study area

Supplemental Table 8. Spatial reduction of extent of richness class in the protected areas

Supplemental Table 9. Threat and distribution range of the ten *Solanum* genus in Benin

Supplemental Table 10. Conservation status of *Solanum* in Benin

References

Abdul Aziz S., Émeline Sêssi Pelagie A., Séverin B., Ogoulonou Rodrigue B., Bertrand A., Samadori Sorotori Honoré B. (2024). Land use/land cover and plant community dynamics in the Benin's forest reserves: The effectiveness of participatory forest management *Trees, Forests and People* 16: 100543. https://doi.org/10.1016/j.tfp.2024.10054

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin: implications for biogeography and

conservation. Ph.D. Thesis, Wageningen University, The Netherlands.

Adomou, A. C., Yedomonhan, H., Sinsin, B., and Van der Maesen, L. J. G. (2007). Distribution des aires protégées et conservation de la flore en république du Bénin: N*otulae Florae Beninensis* 11.

Agounde, G., Salako, K.V., Idohou, R.A., Sode, A.I., Mensah, S., Dimobe, K., Assogbadjo, A.E., Glèlè Kakaï, R. (2025). Climate change may shift diet of the African savanna elephant: Preliminary results for 14 food tree and shrub species in the WAPOK transboundary ecosystem, West-Africa. *Global Ecology and Conservation*, 58, p.e03468. https://doi.org/10.1016/j.gecco.2025.e03468

Ahmad, S., Yang, L., Khan, T. U., Wanghe, K., Li, M. and Luan, X. (2020). Using an ensemble modelling approach to predict the potential distribution of Himalayan gray goral (*Naemorhedus goral bedfordi*) in Pakistan. *Glob. Ecol. Conserv.*, 21, e00845. doi: https://doi.org/10.1016/j.gecco.2019.e00845

Akoègninou, A., van der Burg, W. and van der Maesen, L. (2006). Flore Analytique du Bénin: Wageningen: Backhuys Publishers

Aksoy, E., Demirel, U., Bakhsh, A., Zia, M. A. B., Naeem, M., Saeed, F., ... and Çalışkan, M. E. (2021). Recent advances in potato (*Solanum tuberosum* L.) breeding. Advances in Plant Breeding Strategies: Vegetable Crops: Volume 8: Bulbs, Roots and Tubers, 409-487. https://doi.org/10.1007/978-3-030-66965-2_10

Allouche, O., Tsoar, A. and Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol., 43(6), 1223-1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Alohou, E. C., Gbemavo, D. S. J. C., Mensah, S., and Ouinsavi, C. (2017). Fragmentation of forest ecosystems and connectivity between Sacred Groves and Forest Reserves in southeastern Benin, West Africa. *Tropical Conservation Science*, 10, 1940082917731730.

Barbet-Massin, M., Jiguet, F., Albert, C.H. and Thuiller, W. (2012), Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in Ecology and Evolution,* 3: 327-338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S. and Bopp, L. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *J. Adv. Model. Earth Syst.*, 12(7), e2019MS002010. doi: https://doi.org/10.1029/2019MS002010

Brehm, J., Gaisberger, H., Kell, S., Parra-Quijano, M., Thormann, I., Dulloo, M. E., and Maxted, N. (2022). Planning complementary conservation of crop wild relative diversity in southern Africa. *Diversity and Distributions,* 28, 1358–1372. https://doi.org/10.1111/ddi.13512

Breiman, L. (2001). Random forests. Machine learning, 45, 5-32. doi:https://doi.org/10.1023/A:1010933404324

Brooks, T. M., Pimm, S. L., Akçakaya, H. R., Buchanan, G. M., Butchart, S. H., Foden, W., Hilton-Taylor, C., Hoffmann, M., Jenkins, C. N. and Joppa, L. (2019). Measuring terrestrial area of habitat (AOH) and its utility for the IUCN Red List. *Trends Ecol. Evol.*, 34(11), 977-986. doi: https://doi.org/10.1016/j.tree.2019.06.009

Brown, J. L. (2014). SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol. Evol.*, 5(7), 694-700. https://doi.org/10.1111/2041-210X.12200

Brunton, A. J., Conroy, G. C., Schoeman, D. S., Rossetto, M. and Ogbourne, S. M. (2023). Seeing the forest through the trees: Applications of species distribution models across an Australian biodiversity hotspot for threatened rainforest species of Fontainea. *Glob. Ecol. Conserv.*, 42, e02376. doi:https://doi.org/10.1016/j.gecco.2023.e02376

Burgess, N., Küper, W., Mutke, J., Brown, J., Westaway, S., Turpie, S., Meshack, C., Taplin, J., McClean, C. and Lovett, J. C. (2005). Major gaps in the distribution of protected areas for threatened and narrow range Afrotropical plants. *Biodivers. Conserv.*, 14, 1877-1894. doi:https://doi.org/10.1007/S40531-004-1299-2

Bussmann, R.W., Paniagua-Zambrana, N.Y., and Njoroge, G.N. (2021). *Solanum aculeastrum* Dunal *Solanum anguivi* Lam. *Solanum incanum* L. *Solanum nigrum* L. Solanaceae. In: Bussmann, R.W. (eds) Ethnobotany of the Mountain Regions of Africa. Ethnobotany of Mountain Regions. Springer, Cham. https://doi.org/10.1007/978-3-030-38386-2

Castañeda-Álvarez, N. P., De Haan, S., Juárez, H., Khoury, C. K., Achicanoy, H. A., Sosa, C. C., ... and Spooner, D. M. (2015). *Ex situ* conservation priorities for the wild relatives of potato (*Solanum* L. section Petota). *PLoS One*, 10(4), e0122599. https://doi.org/10.1371/journal.pone.0122599

Chérif, A., Sodé, A., Houndonougbo, J., Idohou, R., Fandohan, A., Kakaï, R. G. and Assogbadjo, A. (2022). Habitat suitability modeling for the conservation and cultivation of the multipurpose fruit tree, *Balanites aegyptiaca* L., in the Republic of Chad, Sahel. *Model. Earth Syst. Environ.,* 8(4), 4953-4963. doi:https://doi.org/10.1007/s40808-022-01416-4

Chinedu, S. N., Olasumbo, A. C., Eboji, O. K., Emiloju, O. C., Arinola, O. K. and Dania, D. I. (2011). Proximate and phytochemical analyses of *Solanum aethiopicum* L. and *Solanum macrocarpon* L. fruits. *Research Journal of Chemical Sciences*, 1(3), 63-71.

Coulibaly M., Idohou R., Akohoue F., Peterson A.T., Sawadogo M., and Achigan-Dako E.G. (2021). (2022). Coupling genetic structure analysis and ecological-niche modeling in Kersting's groundnut in West Africa. *Scientific reports*, 12(1), 5590.

Daï, E. H., Houndonougbo, J. S. H., Idohou, R., Ouédraogo, A., Kakaï, R. G., Hotes, S. and Assogbadjo, A. E. (2023). Modeling current and future distribution patterns of Uvaria chamae in Benin (West Africa): Challenges and opportunities for its sustainable management. *Heliyon*, 9(2). doi:https://doi.org/10.1016/j.heliyon.2023.e13658

Dassou G.H, Agoundé G, Akouété P, Favi GA, Kpétikou GC, Salako K.V, Ouachinou J, Makponsè J, Kouyaté A.M, Sari I, Glèlè Kakaï R.L, Yédomonhan H, and Adomou A.C (2024) Past, present, and future potential distributions of the African multipurpose tree Detarium senegalense (Fabaceae). *Plant Ecology and Evolution* 157(3): 343-357. https://doi.org/10.5091/plecevo.122470

Davis, A. P., Chadburn, H., Moat, J., O'Sullivan, R., Hargreaves, S. and Nic Lughadha, E. (2019). High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci. Adv.*, 5(1), eaav3473. doi:https://doi.org/10.1126/sciadv.aav3473

Fandohan, A. B., Oduor, A. M. O., Sodé, A. I., Wu, L., Cuni-Sanchez, A., Assédé, E. and Gouwakinnou, G. N. (2015). Modeling vulnerability of protected areas to invasion by Chromolaena odorata under current and future climates. *Ecosyst. Health Sustain.*, 1(6), 1-12. doi:https://doi.org/10.1890/EHS45-0003.1

Feng, X., Liang, Y., Gallardo, B. and Papeş, M. (2020). Physiology in ecological niche modeling: using zebra mussel's upper thermal tolerance to refine model predictions through Bayesian analysis. *Ecography*, 43(2), 270-282. doi: https://doi.org/10.1111/ecog.04627

Friedman, J., Hastie, T. and Tibshirani, R. (2000). ADDITIVE LOGISTIC REGRESSION: A STATISTICAL VIEW OF BOOSTING. *Ann. Stat.*, 28(2), 337-407. doi:https://doi.org/10.1214/aos/1016218223

Gbètoho, A. J., Aoudji, A. K., Roxburgh, L. and Ganglo, J. C. (2017). Assessing the suitability of pioneer species for secondary forest restoration in Benin in the context of global climate change. *Bois for. trop.*, 332, 43-55. doi:https://doi.org/10.19182/bft2017.332.a31332

Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theor. Appl. Genet.*, 129, 2281-2294. doi:https://doi.org/10.1007/s00122-016-2804-1

Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, 8(9), 993-1009.

Hijmans, R. J. and Elith, J. (2017). Species distribution modeling with R. R Cran Project.

Hounsou-Dindin, G., Idohou, R., Agre, P., Hounkpèvi, A., Adomou, A. C., Assogbadjo, A. E. and Kakaï, R. G. (2023). Habitat range shift and prediction of the potential future distribution of Ricinodendron heudelotii (Baill.) Heckel in Benin (West Africa). *Heliyon*, 9(9). doi:https://doi.org/10.1016/j.heliyon.2023.e20199

Hudson, L. N., Newbold, T., Contu, S., Hill, S. L., Lysenko, I., De Palma, A., Phillips, H. R., Senior, R. A., Bennett, D. J. and Booth, H. (2014). The PREDICTS database: a global database of how local terrestrial biodiversity responds to human impacts. *Ecol. Evol.*, 4(24), 4701-4735. doi: https://doi.org/10.1002/ece3.1303

Hunter, D., Maxted, N., Heywood, V., Kell, S. and Borelli, T. (2012). Protected areas and the challenge of conserving crop wild relatives. *Parks*, 18(1), 87.

Idohou, R., Odoulami, R., Houehanou, T., & Assogbadjo, A. (2025). Top priority crop wild relatives exhibit different resilience responses to climate change in Benin (West Africa). *Journal for Nature Conservation*, 83, 126769.

Idohou, R., Assogbadjo, A. E., Fandohan, B., Gouwakinnou, G. N., Glele Kakai, R. L., Sinsin, B., & Maxted, N. (2013). National inventory and prioritization of crop wild relatives: case study for Benin. *Genetic Resources and Crop Evolution*, 60, 1337-1352.

Idohou, R., Assogbadjo, A. E., Kakai, R. G. and Peterson, A. T. (2017). Spatio-temporal dynamic of suitable areas for species conservation in West Africa: eight economically important wild palms under present and future climates. *Agrofor. Syst.*, 91, 527-540. doi:https://doi.org/10.1007/S40457-016-9955-6

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp. doi:https://doi.org/10.1017/9781009157896.

IUCN (2024) The IUCN Red List of Threatened Species. Version 2024-2. https://www.iucnredlist.org ISSN 2307-8235

Jarvis, A., Lane, A. and Hijmans, R. J. (2008). The effect of climate change on crop wild relatives. *Agric. Ecosyst. Environ.*, 126(1-2), 13-23. doi:https://doi.org/10.1016/j.agee.2008.01.013

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P. and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Sci. Data*, 4(1), 1-20. doi:https://doi.org/10.1038/sdata.2017.122

Keatinge, J., Yang, R.-Y., Hughes, J. d. A., Easdown, W. and Holmer, R. (2011). The importance of vegetables in ensuring both food and nutritional security in attainment of the Millennium Development Goals. *Food Secur.*, 3(4), 491-501. doi:https://doi.org/10.1007/S42571-011-0150-3

Koebsch, F., Winkel, M., Liebner, S., Liu, B., Westphal, J., Schmiedinger, I., Spitzy, A., Gehre, M., Jurasinski, G. and Köhler, S. (2019). Sulfate deprivation triggers high methane production in a disturbed and rewetted coastal peatland. *Biogeosciences*, 16(9), 1937-1953. doi:https://doi.org/10.5194/bg-16-1937-2019

Lala, S., Amri, A. and Maxted, N. (2018). Towards the conservation of crop wild relative diversity in North Africa: checklist, prioritisation and inventory. *Genet. Resour. Crop Evol.*, 65(1), 113-124. doi:https://doi.org/10.1007/S40722-017-0513-5

Le Saout, S., Hoffmann, M., Shi, Y., Hughes, A., Bernard, C., Brooks, T. M., Bertzky, B., Butchart, S. H., Stuart, S. N. and Badman, T. (2013). Protected areas and effective biodiversity conservation. *Science*, 342(6160), 803-805. doi:https://doi.org/10.1126/science.1239268

Lewis, J. S., Farnsworth, M. L., Burdett, C. L., Theobald, D. M., Gray, M. and Miller, R. S. (2017). Biotic and abiotic factors predicting the global distribution and population density of an invasive large mammal. *Sci. Rep.*, 7(1), 44152. doi:https://doi.org/10.1038/srep44152

MAEP. (2020). Indicateurs macroeconomiques sur le secteur agricole au Benin: Direction de la Statisque Agricole, Republique du Bénin.

Magioli, C., and Mansur, E. (2005). Eggplant (*Solanum melongena* L.): tissue culture, genetic transformation and use as an alternative model plant. *Acta Botanica Brasilica*, 19, 139-148. https://doi.org/10.1590/S0102-33062005000100013

Manda, L., Idohou, R., Assogbadjo, A. E. and Agbangla, C. (2022). Climate change reveals contractions and expansions in the distribution of suitable habitats for the neglected crop wild relatives of the Genus *Vigna* (Savi) in Benin. *Front. conserv. sci.,* 3, 870041. doi:https://doi.org/10.3389/fcosc.2022.870041

Mao, L., Li, M. and Shen, W. (2020). Remote sensing applications for monitoring terrestrial protected areas: Progress in the last decade. *Sustainability*, 12(12), 5016. doi:https://doi.org/10.3390/su12125016

Masson-Delmotte, V., Zhai, P., Pörtner, H., Roberts, D., Skea, J. and Shukla, P. R. (2022). Global Warming of 1.5 C: IPCC special report on impacts of global warming of 1.5 C above pre-industrial levels in context of strengthening response to climate change, sustainable development, and efforts to eradicate poverty (https://doi.org/10.1017/9781009157940): Cambridge University Press.

Maxted, N., Ford-Lloyd, B. V., Jury, S., Kell, S. and Scholten, M. (2006). Towards a definition of a crop wild relative. *Biodivers. Conserv.*, 15(8), 2673-2685. doi:https://doi.org/10.1007/S40531-005-5409-6

Maxted, N., Kell, S. and Brehm, J. M. (2009). Commission on genetic resources for food and agriculture. Establishment of a global network for the in-situ conservation of crop wild relatives: status and needs. Background study paper(39), 212.

Maxted, N., Kell, S., Ford-Lloyd, B., Dulloo, E. and Toledo, Á. (2012). Toward the systematic conservation of global crop wild relative diversity. *Crop Sci.*, 52(2), 774-785. doi:https://doi.org/10.2135/cropsci2011.08.0415

Maxted, N., Kell, S., Toledo, Á., Dulloo, E., Heywood, V., Hodgkin, T., Hunter, D., Guarino, L., Jarvis, A. and Ford-Lloyd, B. (2010). A global approach to crop wild relative conservation: securing the gene pool for food and agriculture. *Kew Bulletin*, 65(4), 561-576. doi:https://doi.org/10.1007/S42225-011-9253-4

Maxted, N. and Magos Brehm, J. (2023). Maximizing the crop wild relative resources available to plant breeders for crop improvement. *Front. sustain. food syst.*, 7, 1010204. doi:https://doi.org/10.3389/fsufs.2023.1010204

Meyer, D. and Wien, F. (2001). Support vector machines. R News, 1(3), 23-26.

Montoya-Jiménez JC, Valdez-Lazalde JR, Ángeles-Perez G, De Los Santos-Posadas HM, and Cruz-Cárdenas G (2022). Predictive capacity of nine algorithms and an ensemble model to determine the geographic distribution of tree species. *iForest-Biogeosciences and Forestry* 15(5): 363. https://doi.org/10.3832/ifor4084-015

Naimi, B. (2017). Package 'usdm'. Uncertainty analysis for species distribution models. Wien: www.cran.r-project.org.

Neuenschwander P., Adomou A.C. (2017). Reconstituting a rainforest patch in southern Benin for the protection of threatened plants *Nature Conservation* 21: 57-82.https://doi.org/10.3897/natureconservation.21.13906

Ng'uni, D., Munkombwe, G., Mwila, G., Gaisberger, H., Brehm, J. M., Maxted, N., Kell, S. and Thormann, I. (2019). Spatial analyses of occurrence data of crop wild relatives (CWR) taxa as tools for selection of sites for conservation of priority CWR in Zambia. *Plant Genetic Resources*, 1-12. doi:https://doi.org/10.1017/S4479262118000497

Öckinger, E., Schweiger, O., Crist, T. O., Debinski, D. M., Krauss, J., Kuussaari, M., Petersen, J. D., Pöyry, J., Settele, J. and Summerville, K. S. (2010). Life-history traits predict species responses to habitat area and isolation: a cross-continental synthesis. *Ecol. Lett.*, 13(8), 969-979. doi:https://doi.org/10.1111/j.1461-0248.2010.01487.x

Ogundola, A. F., Afolayan, A. J., and Bvenura, C. (2023). *Solanum nigrum* Seed viability and germination, and soil modulation effect on seedling emergence. In Sustainable Uses and Prospects of Medicinal Plants (pp. 133-143). CRC Press.

Okokon, J. E., Davies, K. O., Amazu, L. U. and Umoh, E. E. (2017). Anti-inflammatory activity of leaf extract of Solanum anomalum. *Journal of Medicinal Herbs*, 7(4), 243-249.

Oyinloye, O.E., Ajayi, A.M. and Ademowo O.G. (2022). *Solanum dasyphyllum* leaf extract reduces inflammation in carrageenan-induced air pouch in rats by inhibition of cyclooxygenase-2 and inducible nitric oxide synthase. *Nutrire* 47, 24 https://doi.org/10.1186/s41110-022-00175-7

Parra-Quijano, M., Iriondo, J.M. & Torres, E. Ecogeographical land characterization maps as a tool for assessing plant adaptation and their implications in agrobiodiversity studies. *Genetic Resources Crop Evolution* 59, 205–217 (2012). https://doi.org/10.1007/s10722-011-9676-7

Parviainen, M., Luoto, M., Ryttäri, T. and Heikkinen, R. K. (2008). Modelling the occurrence of threatened plant species in taiga landscapes: methodological and ecological perspectives. *J. Biogeogr.*, 35(10), 1888-1905. doi:https://doi.org/10.1111/j.1365-2699.2008.01922.x

Phillips, S. J., Anderson, R. P., Dudik, M., Schapire, R. E. and Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, 40(7), 887-893. doi:https://doi.org/10.1111/ecog.03049

Phillips, S. J., Anderson, R. P. and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Modell.*, 190(3-4), 231-259. doi:https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pidwirny, M. (2006). Abiotic factors and the distribution of species. Fundamentals of Physical Geography.

Pilling, D., Bélanger, J., Diulgheroff, S., Koskela, J., Leroy, G., Mair, G., and Hoffmann, I. (2020). Global status of genetic resources for food and agriculture: challenges and research needs. *Genetic Resources* 1 (1), 4-16. doi: https://doi.org/10.46265/genresj.2020.1.4-16.

Pinto, A. V., Hansson, B., Patramanis, I., Morales, H. E., and van Oosterhout, C. (2024). The impact of habitat loss and population fragmentation on genomic erosion. *Conservation Genetics*, 25(1), 49-57. https://doi.org/10.1007/s10592-023-01548-9

Ramírez-Villegas, J., Khoury, C., Jarvis, A., Debouck, D. G., and Guarino, L. (2010). A gap analysis methodology for collecting crop genepools: a case study with *Phaseolus* beans. *PloS one*, 5(10), e13497. https://doi.org/10.1371/journal.pone.0013497

Ratnayake, S. S., Kariyawasam, C. S., Kumar, L., Hunter, D. and Liyanage, A. (2021). Potential distribution of crop wild relatives under climate change in Sri Lanka: implications for conservation of agricultural biodiversity. *Current Research in Environmental Sustainability,* 3, 100092. doi:https://doi.org/10.1016/j.crsust.2021.100092

Roberts, J. and Florentine, S. (2022). Biology, distribution and management of the globally invasive weed *Solanum elaeagnifolium* Cav (silverleaf nightshade): A global review of current and future management challenges. Weed Res., 62(6), 393-403. doi:https://doi.org/10.1111/wre.12556

R Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (Version 4.0.3.). Vienna, Austria. Retrieved from https://www.R-project.org/

Salako, V. K., Vihotogbé, R., Houéhanou, T., Sodé, I. A. and Glèlè Kakaï, R. (2019). Predicting the potential impact of climate change on the declining agroforestry species *Borassus aethiopum* Mart. in Benin: a mixture of geostatistical and SDM approach. *Agrofor. Syst.*, 93, 1513-1530. doi:https://doi.org/10.1007/S40457-018-0262-2

Samuels, J. (2015). Biodiversity of food species of the Solanaceae family: a preliminary taxonomic inventory of subfamily Solanoideae. *Resources*, 4(2), 277-322. doi:https://doi.org/10.3390/resources4020277

Sarma, H. and Sarma, A. (2011). *Solanum nigrum* L., a nutraceutical enriched herb or invasive weed? Paper presented at the International Conference on Environment and BioScience IPCBEE.

Schuster, R., Buxton, R., Hanson, J. O., Binley, A. D., Pittman, J., Tulloch, V., La Sorte, F. A., Roehrdanz, P. R., Verburg, P. H. and Rodewald, A. D. (2023). Protected area planning to conserve biodiversity in an uncertain future. *Conserv. Biol.*, 37(3), e14048. doi:https://doi.org/10.1111/cobi.14048

Sintayehu, D. W. (2018). Impact of climate change on biodiversity and associated key ecosystem services in Africa: a systematic review. *Ecosyst. Health Sustain.*, 4(9), 225-239. doi:https://doi.org/10.1080/20964129.2018.1530054

Stanton, R., Wu, H. and Lemerle, D. (2012). Factors affecting silverleaf nightshade (*Solanum elaeagnifolium*) germination. *Weed Sci*., 60(1), 42-47. doi:https://doi.org/10.1614/WS-D-11-00105.1

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O. and Ludwig, C. (2015). The trajectory of the Anthropocene: the great acceleration. *Anthr. Rev.,* 2(1), 81-98. doi:https://doi.org/10.1177/2053019614564785

Stolton, S., Maxted, N., Ford-Lloyd, B., Kell, S. and Dudley, N. (2006). Food Stores: Using Protected Areas to Secure Crop Genetic Diversity; Worldwide Fund for Nature: Woking, UK.

Syfert, M. M., Castañeda-Álvarez, N. P., Khoury, C. K., Särkinen, T., Sosa, C. C., Achicanoy, H. A., Bernau, V., Prohens, J., Daunay, M. C. and Knapp, S. (2016). Crop wild relatives of the brinjal eggplant (*Solanum melongena*): Poorly represented in genebanks and many species at risk of extinction. *Am. J. Bot.*, 103(4), 635-651. doi:https://doi.org/10.3732/ajb.1500539

Tas, N., West, G., Kircalioglu, G., Topaloglu, S. B., Phillips, J., Kell, S. and Maxted, N. (2019). Conservation gap analysis of crop wild relatives in Turkey. *Plant Genetic Resources,* 1-10. doi:https://doi.org/10.1017/S4479262118000564

Toffa, Y., Idohou, R. and Fandohan, A. B. (2022). Modélisation de la distribution des espèces en Afrique: état de l'art et perspectives. *Physio-Géo. Géographie physique et environnement*, 17, 43-65. doi:https://doi.org/10.4000/physio-geo.13738

UNEP-WCMC (2023) Protected areas map of the world, https://www.protectedplanet.net/en

van Treuren, R., Hoekstra, R., Wehrens, R. and van Hintum, T. (2020). Effects of climate change on the distribution of crop wild relatives in the Netherlands in relation to conservation status and ecotope variation. *Glob. Ecol. Conserv.*, 23, e01054. doi:https://doi.org/10.1016/j.gecco.2020.e01054

Vargas, J. H., Consiglio, T., Jørgensen, P. M. and Croat, T. B. (2004). Modelling distribution patterns in a species-rich plant genus, *Anthurium* (Araceae), in Ecuador. *Divers. distrib.*, 10(3), 211-216. doi:https://doi.org/10.1111/j.1366-9516.2004.00081.x

Vignali, S., Barras, A. G., Arlettaz, R. and Braunisch, V. (2020). SDMtune: An R package to tune and evaluate species distribution models. *Ecol. Evol.*, 10(20), 11488-11506. doi:https://doi.org/10.1002/ece3.6786

Warren, D. L., Glor, R. E. and Turelli, M. (2010). ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, 33(3), 607-611. doi:https://doi.org/10.1111/j.1600-0587.2009.06142.x

Williams, J. W. and Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological surprises. *Front. Ecol. Environ.*, 5(9), 475-482. doi:https://doi.org/10.1890/070037

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A. and NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Divers Distrib*., 14(5), 763-773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Campbell Grant, E. H. and Veran, S. (2013). Presence-only modelling using MAXENT: when can we trust the inferences? *Methods Ecol. Evol.*, 4(3), 236-243. doi:https://doi.org/10.1111/2041-210x.12004

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., and Wang, L. (2015). Consensus forecasting of species distributions: the effects of niche model performance and niche properties. *PloS one*, 10(3), e0120056. https://doi.org/10.1371/journal.pone.0120056

Genetic Resources

SHORT COMMUNICATION

Genetic Resources (2025), 6 (12), 111–119
DOI: 10.46265/genresj.SKPZ5680
https://www.genresj.org
ISSN: 2708-3764

A public mid-density genotyping platform for North American Atlantic salmon (*Salmo salar* L.)

Dongyan Zhao[a,†], Josue Chinchilla-Vargas[a,†], Alexander M. Sandercock[a], Cristiane H. Taniguti[a], Roseanna Long[b], Yniv Palti[b], Guangtu Gao[b], Michael Pietrak[c], Samuel A. May[c‡], Kasia Heller-Uszynska[d], Craig T. Beil[a], Moira J. Sheehan[a,*]

[a]*Breeding Insight, Cornell University, NY, 14853, Ithaca, USA*
[b]*National Center for Cool and Cold Water Aquaculture Research, USDA-ARS, WV, Leetown, 25430, USA*
[c]*National Cold Water Marine Aquaculture Center, USDA-ARS, ME, 04634, Franklin, USA*
[d]*Diversity Arrays Technology, ACT 2617, Bruce, Australia*
‡*Current address: The Center for Aquaculture Technologies, 8445 Camino Santa Fe Suite 104, San Diego, CA, USA*

Abstract: Genomics-enabled selective animal breeding has become common in recent years, prompting a growing need for diverse genotyping tools that facilitate collaboration among research groups while meeting specific programme needs and objectives. Here, we report the development of a medium-density amplicon panel (DArTag) of 2,950 loci for North American Atlantic salmon. It includes loci distributed across the genome and loci useful for distinguishing the continent-of-origin, parentage, and sex determination. This mid-density panel offers more cost-effective and rapid genotyping capabilities for Atlantic salmon researchers and breeders. The open access provided by this platform facilitates comparisons and enhances data reusability across projects, institutions and countries that use different genomic tools for genotyping. This genotyping panel can make routine genotyping a viable tool for breeding and research programmes.

Keywords: Salmon, aquaculture, amplicon-sequencing, selective breeding, DArTag genotyping

Citation: Zhao, D., Chinchilla-Vargas, J., Sandercock, A. M., Taniguti, C. H., Long, R., Palti, Y., Gao, G., Pietrak, M., May, S. A., Heller-Uszynska, K., Beil, C. T. and Sheehan, M. J. (2025) "A public mid-density genotyping platform for North American Atlantic salmon (Salmo salar L.)", *Genetic Resources*, 6(12), pp. 111–119. doi: 10.46265/genresj.SKPZ5680.

Introduction

Commercial aquaculture of Atlantic salmon (*Salmo salar* L.) accounts for approximately 70% of the global total salmon production (Pandey *et al*, 2023). Atlantic salmon has long been the focus of selective breeding programmes aimed at improving production traits related to growth, disease resistance and fillet quality (Vallejo *et al*, 2024; Kristjánsson *et al*, 2020). The past two decades have witnessed a rapid adoption of molecular methods to enhance breeding programmes, particularly the application of genomic selection methods (Meuwissen *et al*, 2001). These molecular advancements require developing genotyping tools for various applications, most often in the form of panels of preidentified Single Nucleotide Polymorphisms (SNPs). In plant breeding, these targeted genotyping technologies can be characterized into low density (i.e. hundreds of loci), medium density (i.e. hundreds to thousands of loci) and high density (i.e. tens of thousands to millions of loci). This classification differs notably from livestock standards, where medium-density arrays typically contain around 50K markers and high-density arrays can exceed 700K markers. For aquaculture study and breeding, until recently, the majority of these resources were designed and owned by private companies (Gao *et al*, 2023; Kijas *et al*, 2017), slowing the efficiency of Atlantic salmon genomic research. Therefore, publicly accessible genomic resources are needed to facilitate open, reproducible research for Atlantic salmon, with applications in aquaculture, conservation and fisheries management.

* Corresponding author: Moira Sheehan (moirasheehan@cornell.edu)
†These authors have contributed equally to this work

Received: 14.04.2025 | Accepted: 23.08.2025 | Published online: 08.10.2025

North American Atlantic (NAA) salmon are genetically distinct from European and Baltic Atlantic strains. The common ancestor of today's salmonids underwent a lineage-specific whole genome duplication event at ~96 million years ago (Allendorf and Thorgaard, 1984; Danzmann *et al*, 2008; Berthelot *et al*, 2014). Since that time, the re-diploidization of salmonids has proceeded independently in the Old World and the New World. Today, the genetic divergence of European and Baltic Atlantic salmon (2n = 2x = 58) from NAA salmon (2n = 2x = 54) is characterized by large structural changes and unique karyotypes that designate them as subspecies (de Boer *et al*, 2007; Brenna-Hansen *et al*, 2012). Despite high syntenic conservation, NAA salmon exhibited significant genomic differences compared to its European counterpart (Brenna-Hansen *et al*, 2012; Gao *et al*, 2020). Atlantic salmon farming in the eastern United States and Canada is limited to North American (NA) genetic stocks due to ecological and conservation concerns. Therefore, genotyping panels developed for European Atlantic salmon are less effective when applied to NA populations, underscoring the need for origin-specific genomic tools (Yáñez *et al*, 2016).

In the United States, the only remaining wild populations of Atlantic salmon are found in the Gulf of Maine and are listed as endangered under the Endangered Species Act (https://www.fisheries.noaa.gov/species/atlantic-salmon/protected). The USDA-ARS National Cold Water Marine Aquaculture Center (NCWMAC) has operated a selective breeding programme for the St. John River (SJR) strain of NAA salmon since 2003 for traits such as growth, fillet quality and resistance to sea lice (a major pathogen in marine aquaculture) (Peterson *et al*, 2020; Vallejo *et al*, 2024). The SJR strain, chosen for its rapid growth and suitability for captive aquaculture, undergoes a 4-year lifecycle involving specialized systems for egg incubation, fry growth, and maturation, culminating in spawning mature broodstock weighing approximately 3–8kg. Up to 150 families are cultured annually, with fish evaluated in biosecure tanks and commercial net pens for performance. In 2022, NCWMAC adopted a genomic selection index weighted 70% for growth and 30% for sea lice resistance, supported by a 50K SNP chip developed for the NA salmon genome (Gao *et al*, 2023; Vallejo *et al*, 2024).

The cost of using the 50K SNP chip remains a barrier to widespread usage. In contrast, low- or medium-density genotyping panels may meet many of the same goals at a reduced price with reduced lab equipment requirements and lower overhead. Here, we developed a multi-purpose medium-density DArTag panel with 3K markers by subsetting markers already included in the 50K NAA salmon SNP array (Gao *et al*, 2023). DArTag is a targeted amplicon sequencing platform developed by Diversity Arrays Technology, LLC, which provides low-cost and reproducible genotyping results across sequencing projects (Blyton *et al*, 2023; "DArTag," n.d.). We validated the DArTag panel by genotyping 3,710 NAA salmon from the United States Department of Agriculture (USDA) aquaculture stock and show that (1) the 3K DArTag panel can be used to obtain high-quality SNPs across genotyped individuals, (2) the panel accurately identifies relationships between individuals, and (3) the 3K panel can be effectively used for linkage analysis comparable to the high-density array. This open-source 3K DArTag panel can increase the accessibility of genotyping for programmes without access to in-house genotyping technology or specialized labour. It may also reduce the cost of genotyping by lowering the marker density without sacrificing much information, thereby increasing breeder access to genotyping services to allow for more intensive, routine and effective usage of genomic resources in NAA salmon breeding.

Materials and methods

Selection of 3K marker loci for building the DArTag genotyping panel

We previously published the results of the 50K SNP Affymetrix array (Gao *et al*, 2023) developed based on the alignment of whole-genome re-sequencing of 80 NAA salmon fish from three distinct aquaculture stocks to the NAA salmon reference genome (GenBank Accession GCA_021399835.1). From the 50K array, 10,353 SNPs were selected for their even genome distribution and functional annotations, including 8,803 SNPs from the NAA-based SNP dataset (Gao *et al*, 2023), 1,462 highly informative SNPs from a European-based SNP array (Houston *et al*, 2014), 64 SNPs for distinguishing the continent of origin (COO), 20 mitochondrial SNPs, and four sex determination SNPs. The evenly distributed 10K SNP set was submitted to Diversity Arrays Technology (DArT) for proprietary *in silico* quality control. DArT recommended the loci that passed quality control to produce the final 3K SNP panel. Upon initial testing of the 3K panel, it was observed that the 20 mitochondrial markers consumed ~40% of the reads per sample, indicating the abundance of mitochondrial DNA and undesirable preferential amplification. Therefore, these 20 markers were removed, leaving a final set of 2,980 genomic loci (Figure 1).

Notably, the 50K SNP dataset was initially developed based on the NAA salmon's contig-level assembly, and Gao *et al*, 2023 mapped these SNPs to the final chromosome-level assembly. Among the 2,980 DArTag markers, 2,911 were assigned chromosome coordinates by Gao *et al*, 2023. For the remaining 69 markers, we used BLAST to align the 180bp flanking sequences to the reference genome. We confirmed the positions of 49 of the 69 markers, bringing the number of markers with known pseudomolecule physical locations in the chromosomes up to 2,950 (Supplemental Table 1). The unmapped SNPs are likely due to several factors, such as their contigs not being included in the final assembly, contig splits due to Hi-C or Bionano scaffolding, and/or error corrections at the scaffold level that altered the reference sequences.

To compare the 50K SNP array with the 3K DArTag panel, we aligned the two datasets into a consistent genomic framework, including matching target SNP positions and reference and alternative base calls. The Axiom and DArTag arrays can include probes on the plus or minus strand. Thus, correct inference of reference and alternative alleles depends upon the oligo orientation, especially for A/T and C/G (i.e. SNPs ambiguous to DNA strand) SNPs. We established contig orientation through BLAST alignment of the 180bp flanking sequences of the 3K SNPs against the reference genome. Reference and alternative bases were designated based on the contig orientations and the Axiom and DArTag probe orientations.

Figure 1. A) Filters and criteria applied to produce the 3K DArTag marker panel from the 50K North American Atlantic (NAA) salmon SNP array. Abbreviations: K is thousands. B) The distribution of 2,950 SNPs on the 3K DArTag panel across the 27 chromosomes of the NAA salmon genome.

Methodology and procedure of the DArTag genotyping platform

The DArTag genotyping assay consists of four steps based on principles described in Krishnakumar *et al*, 2008 and implemented as described (Zhao *et al*, 2023; Sandercock *et al*, 2025a). Briefly, the pool of 3K NAA salmon oligos, each targeting one genetic variant plus adjacent flanking sequence, are hybridized to denatured gDNA in step 1, followed by SNP/INDEL copying into DArTag molecules by DNA polymerase in step 2. Also in step 2, amplicons are ligated to create circularized molecules. In step 3, the reaction content is treated with nucleases to remove any un-circularized molecules. DArTag products are subsequently amplified in step 4 with the simultaneous addition of sample-unique barcodes used downstream for demultiplexing. The products of the DArTag assay, after purification and quantification, are sequenced on NGS platforms (e.g. NovaSeq 6000, Illumina) with a depth of around 200x. Sequences are demultiplexed, quality controlled, and the genetic variants are detected using DArT's proprietary analytical pipelines.

Selection of samples for validating the DArTag panel and genotyping results

The salmon 3K marker panel was tested using a set of 3,710 fish from the SJR strain reared at NCWMAC (Supplemental Table 2). This set included 118 fish from year classes (YC) fertilized in YC2009 and YC2010, 1,099 fish from YC2014, 2,487 fish from YC2018, and two individuals lacking year class information. The dataset represented three generations of fish from the SJR strain of NAA salmon. Genotyping was processed in two batches: (1) 1,105 fish from 2014 and (2) the remaining 2,605 fish. DArT provided genotypes in VCF format, read counts for all markers, dosage calls, and missing allele discovery counts (MADC) for the second batch, which contained the read count for all 54bp microhaplotype alleles discovered in the samples. Read count data for reference (Ref) and alternative (Alt) alleles from both batches underwent a 2-step quality control and filtering process. First, we removed samples with high missing data rates (≥ 95%), where a marker was considered missing if it had fewer than 10 reads. Subsequently, we filtered out marker loci that were present in < 10 samples. To enable accurate comparative analysis between the two batches, we concatenated read count data into a single file and conducted dosage calls using the updog R package (Gerard *et al*, 2018). The original DArTag marker IDs were converted to the chromosome-level marker IDs in the 50K array and lookup table of the DArTag vs. Axiom IDs is provided (Supplemental Table 1).

Pedigree verification and parentage testing

Before verifying parentage, duplicated IDs, individuals appearing both as male and female, and any circular dependencies in the pedigree were removed using the clean_pedigree() function from BIGr (RRID: SCR_026677; v0.3.4) (Sandercock *et al*, 2025a; Sandercock *et al*, 2025b). Parentage testing was performed with the SEEKPARENTSF90 module in BLUPF90 (Misztal *et al*, 2014) with an allowed maximum threshold of 1% of markers showing Mendelian errors between parent-offspring pairs proposed by pedigree. Percentages over the threshold were flagged as a pedigree error. To identify potential parents in the genotyped set of individuals, the –seektype 2 flag was used in the analysis. Additionally, within-family clustering of individuals was performed with a principal component analysis (PCA) via the Breeding Insight Genomics App (RRID v0.6.2) (Beygelzimer *et al*, 2019; Sandercock *et al*, 2025a).

Supervised clustering and K nearest neighbor (KNN)

An initial parentage verification analysis of the DArTag datasets found discordance between pedigree records and genotypes, suggesting that the individuals in batch 2 were not labelled with the correct sample IDs. A 2-step approach was implemented to estimate the correct IDs: (1) match the sample IDs between the Axiom 50K and 3K DArTag datasets based on genetic similarity, and (2) validate the estimated sample IDs through a second parentage verification analysis.

Before matching, missing genotypic data in batch 2 samples were imputed using Beagle v5.4 with the default parameters (Browning, Zhou, and Browning 2018). The 50K dataset and the 3K dataset of fish born in 2018 were filtered to retain only shared loci in both datasets. To estimate potential matches (step 1), we used the K-nearest neighbors (KNN) algorithm as implemented in the Fast Nearest Neighbor Search Algorithms and Applications (FNN) R package (v1.1.4.1) (Beygelzimer *et al*, 2019). The individual samples on the 50K panel were paired with the sample exhibiting the smallest genetic distance (Euclidean distance) in the 2018 DArTag dataset. This was accomplished using the knn.dist() function of the FNN package (parameter k = 1), effectively assigning the single, most likely sample ID from the 50K dataset to its counterpart in the 2018 3K dataset. To assess the accuracy of this first step, we performed the same steps above with the 50K and 3K genotype data for the 2014 salmon samples (batch 1). We found 99.8% agreement between the putative sample ID and the matched sample ID using only this initial KNN match.

Despite the assessed accuracy of step 1, several individuals in the 50K dataset did not pair with a unique sample in the 2018 3K dataset. In these cases, only the match with the lowest genetic distance was retained for step 2. Finally, a second parentage analysis (step 2) was performed using the sample IDs estimated by KNN for mislabelled individuals in the 2018 DArTag dataset. The sample IDs that passed this verification confirmed that the revised IDs were consistent with Mendelian expectations, enabling the accurate identification of the parents of the mislabelled samples. Only validated samples that passed both steps were included for genetic map construction.

Genetic map construction

To evaluate the utility of the DArTag marker panel for closed-population marker-assisted selection (MAS), a linkage map was generated using Lep-MAP3 (v0.5.0; Rastas, 2017) from the validated salmon samples. Samples were retained if they belonged to a family with at least ten individuals, resulting in 1,035 samples, 55 families, and 2,806 informative SNP loci. First, the ParentCall2 function was used to call missing parental genotypes, with halfSibs = 1 to include half-sib information. Then, Filtering2 checked SNPs for non-informative markers or non-Mendelian markers (i.e. segregation distortion), although no additional SNPs were removed due to the previous, more stringent filtering in Plink 1.9 (Purcell *et al*, 2007). Markers were categorized into 33 linkage groups (LGs) using SeparateChromosome2, with lodLimit = 27 set as the expected number of haploid chromosomes in NAA salmon (1n = 1x = 27). Of the 33 LGs, 27 LGs contained markers aligned with their expected physical chromosomes. The six remaining LGs contained four or fewer markers, so the markers from these six LGs were categorized as 'single' markers for the next step. The 'single' markers were added to one of the 27 LGs with a more relaxed lodLimit = 10 and lodDifference = 3. Lastly, the genomic positions of the markers within each LG were ordered with OrderMarkers2. Additional filtering was performed to remove markers with (1) a physical position that deviated significantly from the other markers in the LG and (2) a lower pairwise LOD score with closely positioned markers (Supplemental Figure 1).

Results

Creation of the 3K NAA salmon DArTag panel

The 3K DArTag panel (Salmon DArTag3K BI Cornell University (1.0)) is comprised of 514 SNPs from a European-based SNP array (Houston *et al*, 2014), 64 SNPs for identifying COO, 4 SNPs from the sex determination (*sd*) locus, and 2,418 genic SNPs from the NAA salmon SNP dataset. To enhance comparability, the 3K SNPs were mapped to the NAA salmon chromosomes, and 2,950 SNPs were assigned unambiguously to physical positions (Supplemental Table 1). The 514 European-based markers were mapped across the 27 chromosomes, with an additional small fraction (0.6%) remaining in unplaced sequences based on the NAA salmon reference genome (Supplemental Table 3). The majority of chromosomes (18/27; 66.7%) maintained moderate to high marker coverage, containing between 15 and 30 markers each, indicating robust coverage across most of the genome.

Validation of the 3K salmon DArTag panel and genotyping results

To assess the 3K panel, a validation set of 3,710 samples was genotyped in two batches using the 3K DArTag panel to: (1) construct a genetic linkage map and (2) evaluate the usefulness of the DArTag panel for downstream genetic analyses. We established a minimum threshold of ten read counts for a marker locus to be considered valid. Under this criterion, 1,077 (97%) of 1,105 samples from batch 1 and 2,470 (95%) of the 2,605 samples from batch 2 retained data for 75% of the total markers. Batch 2 showed particularly robust performance, with 2,181 (84%) samples containing data for ≥ 90% of the total markers (Supplemental Table 4).

Of the 2,950 markers, 2,495 (85%) and 2,827 (96%) were present in ≥ 50% of the samples from batches 1 and 2, respectively, suggesting they are highly conserved sequences within the NAA salmon population. Batch 2 demonstrated superior marker performance, with 2,410 (82%) markers present in ≥ 90% of the samples compared to 1,857 (63%) in batch 1 (Supplemental Table 4). This disparity in missing data rates between batches was likely due to lower DNA quality in the batch 1 samples. Overall, the panel demonstrated the robustness and applicability of the panel for high-throughput genotyping in NAA salmon populations.

For comprehensive analyses, we merge read count data from both batches. Of the 2,950 SNPs assigned to physical positions, 2,278 markers were successfully genotyped in ≥ 85% of samples. Additionally, 2,493 samples retained

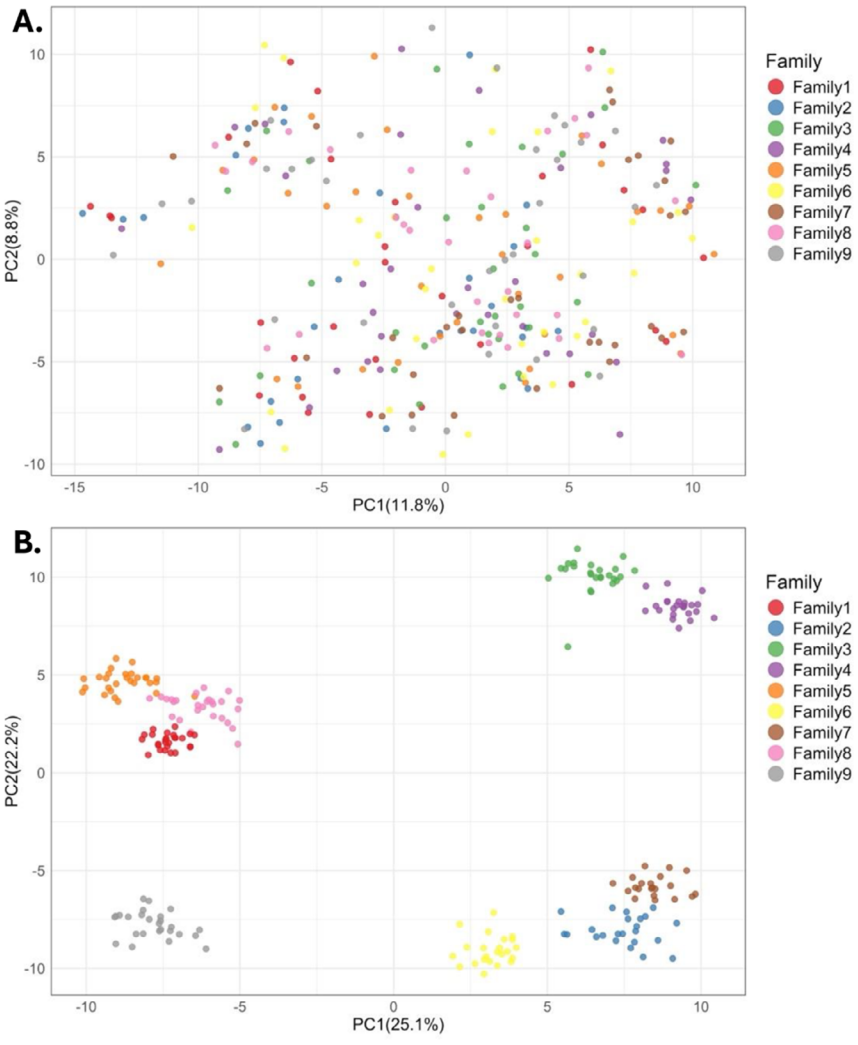


Figure 2. Principle Component Analysis (PCA) plots in the first two dimensions of the validation families. A) PCA plot of the nine full-sibling families with the largest members of 2018-born fish before parentage testing and KNN analysis. B) PCA plot of the same nine full-sibling families of 2018-born fish after pedigree correction with KNN.

genotype data for ≥ 85% of the markers. The merged dataset showed lower data rates compared to batch 2 independently, reflecting the impact of batch 1's lower performance. After filtering for missing data and concatenating the allele dosage results from both batches, 2,806 markers were retained for downstream analyses.

Pedigree verification, parentage testing, and correction via KNN analysis

An initial parentage testing run to verify pedigree accuracy showed that 99.6% of offspring-parent pairs in the dataset were erroneous. To better visualize the within-family clustering, the principal components of genotypes of 276 fish belonging to the nine families with the most siblings were plotted. No identifiable clustering by family was found (Figure 2A). This led us to identify widespread mislabeling of samples from fish born in 2018.

To identify the best-matching ID in the DArTag genotyping results, 3K genotype calls were compared to the same markers in the 50K array using KNN analysis. Genetic distances between matched samples ranged from 27.2 to 53.63, with a maximum value of 45 selected as a filtering threshold that limited the number of samples with multiple matches, retaining 1,493 fish. Parentage testing of the KNN-informed parent-offspring pairs found 992 samples (66.5%) fulfilled Mendelian expectations with their proposed parents. This two-step approach produced a set of 1,013 individuals, composed of 55 full-sib families of at least ten individuals and their respective parents, which were then used to generate the linkage map. Figure 2B shows the clustering of the nine families with the most individuals, as assigned by KNN.

Creation of a linkage map

The final salmon DArTag linkage map (Figure 3) consisted of 27 LGs with 2,642 markers and a total length of 1,983.81cM for the female map, and 927.8cM for the male map (with an average density of 1.33 markers/cM and 2.85 markers/cM, respectively).

LG length from the female map ranged from 52.44cM to 101.02cM, with an average of 73.5cM. The male map linkage group length ranged from 2.42cM to 75.3cM, with an average length of 34.4cM.

Consistent with findings from the same fish tested on the 50K marker panel (Gao *et al*, 2023), paternal and maternal recombination patterns differed (Figure 3B). In paternal chromosomes, recombination was elevated at the telomeres with strong interference near the centromere. In contrast, maternal chromosomes exhibited distinct patterns based on chromosome type: in acrocentric chromosomes, recombination was elevated around the centromere and decreased toward the telomeres, whereas in metacentric chromosomes, interference was pronounced at the centromere with comparable recombination patterns extending toward both telomeric ends. Markers were generally well distributed across the 27 LGs, with ~50% of the markers located within the first 10 LGs (Table 1). Additional mapping details are summarized in Supplemental Table 5.

Table 1. Linkage map from 1,035 fish spanning 55 families and 2,642 uniquely mapped SNPs by chromosome on the male and female maps (in cM).

Chromosome	Marker Count	Male (cM)	Female (cM)
Chr01	148	56.8	78.7
Chr02	55	10.1	92.6
Chr03	107	50.4	99.1
Chr04	90	26.7	93.7
Chr05	75	52.0	96.8
Chr06	80	20.8	99.4
Chr07	54	21.4	85.7
Chr08	75	61.8	87.2
Chr09	163	17.1	83.2
Chr10	151	43.6	75.7
Chr11	103	47.7	67.4
Chr12	114	29.6	67.3
Chr13	129	36.4	70.0
Chr14	131	22.8	60.8
Chr15	146	42.7	62.2
Chr16	93	5.6	56.6
Chr17	51	2.4	58.0
Chr18	98	36.3	68.9
Chr19	103	44.7	57.2
Chr20	113	17.1	58.5
Chr21	86	47.5	53.5
Chr22	79	55.3	53.6
Chr23	131	75.3	101.0
Chr24	64	34.8	56.2
Chr25	54	6.4	52.6
Chr26	90	49.2	95.6
Chr27	59	13.5	52.4
Min	51	2.4	52.4
Max	163	75.3	101.0
Average	97.9	34.4	73.5
Total	2,642	927.8	1,983.8

Discussion and conclusion

The NAA salmon 3K DArTag panel serves as a robust and versatile tool for genetic applications, providing reliable data for pedigree verification, parentage assignment and linkage map construction. Its mid-density design fills a gap in community resources between the high-density 50K array and the low-density option of 384 SNPs (Center for Aquaculture Technologies, personal communication). The 3K panel achieves comparable genome coverage to the 50K

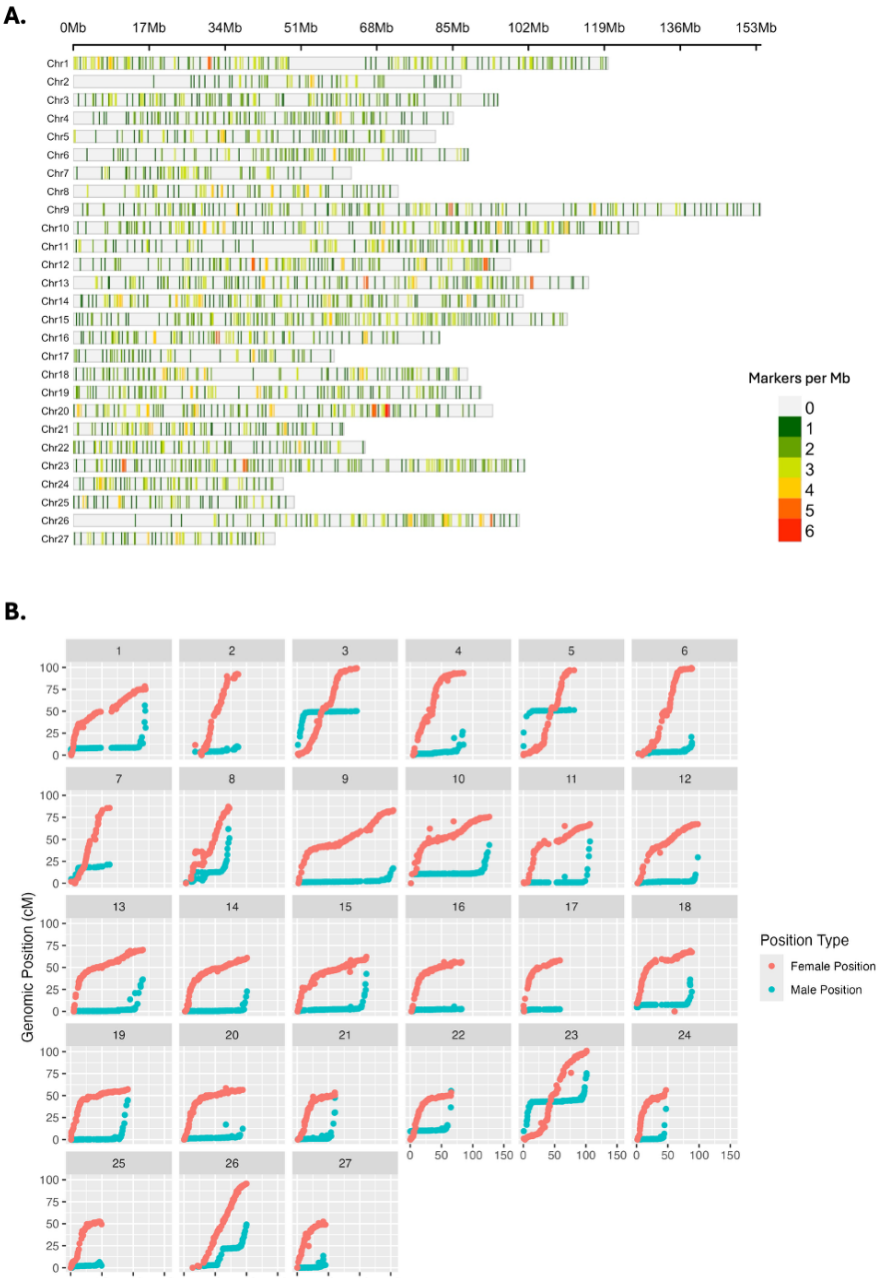


Figure 3. Genetic map of NAA salmon constructed from 1,035 individuals from 55 families. A) Distribution of 2,642 SNPs across 27 linkage groups of the North American Atlantic salmon linkage map. B) Relationship plots of physical map distance (Mb; x-axes) to genetic map distance (cM; y-axes) for each of the 27 chromosomes in the male and female genetic maps.

panel described by Gao *et al* (2023). Because it is a subset of loci on the 50K, it could be used for sparse testing plus imputation on higher numbers of progeny when parents are genotyped on the 50K in parent-progeny studies. The inclusion of four sex-linked markers and 68 COO markers further enhances its applicability across diverse research and breeding scenarios. We acknowledge that the fewer markers in the mid-density panel relative to the 50K array may result in a substantial loss of resolution for fine-scale mapping applications such as genome-wide association studies (GWAS), as QTL detection and mapping accuracy are highly dependent on marker density. However, the reduction in marker density is expected to have only a modest impact on genomic selection applications, where prediction accuracy may show only a slight decrease.

While the 3K panel was developed and validated for NAA salmon, its transferability to European salmon populations would need empirical validation. Importantly, the inclusion of 514 markers from a European salmon-based Affymetrix SNP array could potentially be useful for European salmon populations. The potential utility of this panel might be particularly relevant for comparative genomic studies, population structure analyses, or preliminary screening purposes for both North American and European Atlantic salmon. However, users should consider possible limitations when applying it to European populations, including: (1) potentially reduced marker polymorphism in European populations, (2) different linkage disequilibrium patterns

that might affect marker informativeness, and (3) possible ascertainment bias due to the North American-focused marker selection.

The NAA salmon DArTag panel is publicly available and open for any researcher or breeder to order through DArT (https://www.diversityarrays.com), with a cost midway between the 50K high-density array and the 384-SNP low-density options. The high detection rate and repeatability make this panel suitable for genetic map construction, marker-assisted selection, whole-genome association mapping, reconstruction of recombination patterns, allele dosage estimation, and parental confirmation in NAA salmon from the Northeast US. The panel's efficacy on breeding materials or populations outside the northeast US has not been tested.

One benefit DArTag has over fixed array platforms is the ability to update and improve the marker panel as needed. The panel is a pool of 2,950 oligos, one per locus, which are used to generate sequencing libraries from assayed material. Because the pool is created from individual oligo stocks, removing suboptimal loci or adding new loci can be quickly done by creating a new pool. Independently, as new significant trait markers and/or markers specific to other germplasm are detected, they can be included in the original pool in the panel's next version(s).

Due to our budgetary restrictions, we created a panel of 3,000 loci; however, smaller, complementary panels can be made at lower up-front and downstream usage costs. Sub-panels of a few hundred loci may also be developed using other amplicon techniques, such as Genotyping by Thousands (GTseq), for lower genotyping costs (Campbell *et al*, 2015). The practical upper limit for the number of probes on a DArTag panel is 7,000 loci. However, the optimal maximum may differ by species and genome complexity, and read depth required to sufficiently call genotypes (Andrzej Kilian DArT, personal communication).

Supplemental data

Supplemental File 1. Genotypic data in VCF format for the 1,013 individuals used to produce the linkage map

Supplemental Figure 1. Relationship plots of physical map distance (Mb; x-axes) to genetic map distance (cM; y-axes).

Supplemental Table 1. Physical position and identification of the 2950 SNPs included in the DArTag panel.

Supplemental Table 2. Accessions used in the testing of the salmon 3K DArTag panel and construction of genetic map.

Supplemental Table 3. Distribution of European Atlantic salmon array-based markers on salmon genome

Supplemental Table 4. Sample and marker missing data from two batches of DArTag genotyping.

Supplemental Table 5. Physical position, genetic distance and identification of the 2,642 SNPs included in the linkage map.

Author contributions

DZ, GG, YP and MJS contributed to experimental design and planning. GG and YP selected the diversity panel for WGS. SM contributed to writing and editing the paper. RL and MP collected and prepared all fin clip materials used in the study. DZ performed all the SNP database creation, filtering pipelines, and quality control analyses to create the 3K panel. KHU managed the panel creation at Diversity Arrays Technology. DZ, AS, JCV, AMS and CHT executed the data analyses and genetic mapping. DZ, AMS, JCV and MJS wrote the initial draft of the manuscript. CB managed experiments and communication among all authors involved. All authors contributed to reviewing the manuscript.

Acknowledgments and funding

The authors extend their thoughts to their late colleague, Guangtu Gao, with whom they had the pleasure of working on this study, and who passed away too soon. The authors also thank Dr Alex Casa for her careful review of this manuscript and her valuable feedback. Breeding Insight (RRID:SCR_026645) was funded through Cooperative Agreements between USDA-ARS and Cornell (project numbers: 8062-21000-043-004-A, 8062-21000-052-002-A, and 8062-21000-052-003-A). This study was also supported by the USDA Agricultural Research Service in-house project numbers 8030-31000-004/005 and 8082-31000-012/013.

Conflict of interest statement

The authors declare no conflicts of interest.

Data availability statement

The genotypic data in VCF format for the 1,013 individuals used to produce the linkage map are available as Supplemental File 1.

References

Allendorf, Fred W., and Gary H. Thorgaard. 1984. "Tetraploidy and the Evolution of Salmonid Fishes." In *Evolutionary Genetics of Fishes*, 1–53. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4684-4652-4_1.

Berthelot, Camille, Frédéric Brunet, Domitille Chalopin, Amélie Juanchich, Maria Bernard, Benjamin Noël, Pascal Bento, *et al*, 2014. "The Rainbow Trout Genome Provides Novel Insights into Evolution after Whole-Genome Duplication in Vertebrates." *Nature Communications* 5 (1): 3657. https://doi.org/10.1038/ncomms4657.

Beygelzimer, Alina, Sham Kakadet, John Langford, Sunil Arya, David Mount, and Shengqiao Li. 2019. "FNN: Fast Nearest Neighbor Search Algorithms and Applications." Comprehensive R Archive Network (CRAN). 2019. https://CRAN.R-project.org/package=FNN.

Blyton, Michaela D. J., Kylie L. Brice, Katarzyna Heller-Uszynska, Jack Pascoe, Damian Jaccoud, Kellie A. Leigh, and Ben D. Moore. 2023. "A New Genetic Method for Diet Determination from Faeces That Provides Species Level Resolution in the Koala," February. https://doi.org/10.1101/2023.02.12.528172.

Boer, Johan G. de, Ryosuke Yazawa, William S. Davidson, and Ben F. Koop. 2007. "Bursts and Horizontal Evolution of DNA Transposons in the Speciation of Pseudotetraploid Salmonids." *BMC Genomics* 8 (1): 422. https://doi.org/10.1186/1471-2164-8-422.

Brenna-Hansen, Silje, Jieying Li, Matthew P. Kent, Elizabeth G. Boulding, Sonja Dominik, William S. Davidson, and Sigbjørn Lien. 2012. "Chromosomal Differences between European and North American Atlantic Salmon Discovered by Linkage Mapping and Supported by Fluorescence in Situ Hybridization Analysis." *BMC Genomics* 13 (1): 432.

https://doi.org/10.1186/1471-2164-13-432.

Browning, Brian L., Ying Zhou, and Sharon R. Browning. 2018. "A One-Penny Imputed Genome from next-Generation Reference Panels." *The American Journal of Human Genetics* 103 (3): 338–48. https://doi.org/10.1016/j.ajhg.2018.07.015.

Campbell, Nathan R., Stephanie A. Harmon, and Shawn R. Narum. 2015. "Genotyping-in-Thousands by Sequencing (GT-Seq): A Cost Effective SNP Genotyping Method Based on Custom Amplicon Sequencing." *Molecular Ecology Resources* 15 (4): 855–67. https://doi.org/10.1111/1755-0998.12357.

Danzmann, Roy G., Evelyn A. Davidson, Moira M. Ferguson, Karim Gharbi, Ben F. Koop, Bjorn Hoyheim, Sigbjorn Lien, *et al*, 2008. "Distribution of Ancestral Proto-Actinopterygian Chromosome Arms within the Genomes of 4R-Derivative Salmonid Fishes (Rainbow Trout and Atlantic Salmon)." *BMC Genomics* 9 (1): 557. https://doi.org/10.1186/1471-2164-9-557.

"DArTag." n.d. Diversity Arrays Technology. https://www.diversityarrays.com/services/targeted-genotying/.

Gao, Guangtu, Michael R. Pietrak, Gary S. Burr, Caird E. Rexroad 3rd, Brian C. Peterson, and Yniv Palti. 2020. "A New Single Nucleotide Polymorphism Database for North American Atlantic Salmon Generated through Whole Genome Resequencing." *Frontiers in Genetics* 11 (February): 85. https://doi.org/10.3389/fgene.2020.00085.

Gao, Guangtu, Geoffrey C. Waldbieser, Ramey C. Youngblood, Dongyan Zhao, Michael R. Pietrak, Melissa S. Allen, Jason A. Stannard, *et al*, 2023. "The Generation of the First Chromosome-Level de Novo Genome Assembly and the Development and Validation of a 50K SNP Array for the St. John River Aquaculture Strain of North American Atlantic Salmon." *G3: Genes|Genomes|Genetics* 13 (9). https://doi.org/10.1093/G3JOURNAL/JKAD138.

Gerard, David, Luis Felipe Ventorim Ferrão, Antonio Augusto Franco Garcia, and Matthew Stephens. 2018. "Genotyping Polyploids from Messy Sequencing Data." *Genetics* 210 (3): 789–807. https://doi.org/10.1534/genetics.118.301468.

Houston, Ross D., John B. Taggart, Timothé Cézard, Michaël Bekaert, Natalie R. Lowe, Alison Downing, Richard Talbot, *et al*, 2014. "Development and Validation of a High Density SNP Genotyping Array for Atlantic Salmon (Salmo Salar)." *BMC Genomics* 15 (February): 90. https://doi.org/10.1186/1471-2164-15-90.

Kijas, J., N. Elliot, P. Kube, B. Evans, N. Botwright, H. King, C. R. Primmer, and K. Verbyla. 2017. "Diversity and Linkage Disequilibrium in Farmed Tasmanian Atlantic Salmon." *Animal Genetics* 48 (2): 237–41. https://doi.org/10.1111/age.12513.

Krishnakumar, Sujatha, Jianbiao Zheng, Julie Wilhelmy, Malek Faham, Michael Mindrinos, and Ronald Davis. 2008. "A Comprehensive Assay for Targeted Multiplex Amplification of Human DNA Sequences." *Proceedings of the National Academy of Sciences* 105 (27): 9296–9301. https://doi.org/10.1073/pnas.0803240105.

Kristjánsson, Ólafur H., Bjarne Gjerde, Jørgen Ødegård, and Marie Lillehammer. 2020. "Quantitative Genetics of Growth Rate and Filet Quality Traits in Atlantic Salmon Inferred from a Longitudinal Bayesian Model for the Left-Censored Gaussian Trait Growth Rate." *Frontiers in Genetics* 11 (November). https://doi.org/10.3389/fgene.2020.573265.

Meuwissen, T. H., B. J. Hayes, and M. E. Goddard. 2001. "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps." *Genetics* 157 (4): 1819–29. https://doi.org/10.1093/genetics/157.4.1819.

Misztal, Ignacy, Shogo Tsuruta, Daniela Lourenco, Yutaka Masuda, Ignacio Aguilar, Andres Legarra, and Zulma Vitezica. 2014. "Manual for BLUPF90 Family of Programmes."

Pandey, Rudresh, Frank Asche, Bård Misund, Rune Nygaard, Olugbenga Michael Adewumi, Hans-Martin Straume, and Dengjun Zhang. 2023. "Production Growth, Company Size, and Concentration: The Case of Salmon." Aquaculture (Amsterdam, Netherlands) 577 (739972): 739972. https://doi.org/10.1016/j.aquaculture.2023.739972.

Peterson, Brian C., Gary S. Burr, Michael R. Pietrak, and Dina A. Proestou. 2020. "Genetic Improvement of North American Atlantic Salmon and the Eastern Oyster Crassostrea Virginica at the U.s. Department of Agriculture–Agricultural Research Service National Cold Water Marine Aquaculture Center." *North American Journal of Aquaculture* 82 (3): 321–30. https://doi.org/10.1002/naaq.10144.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, *et al*, 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81 (3): 559–75. https://doi.org/10.1086/519795.

Rastas, Pasi. 2017. "Lep-MAP3: Robust Linkage Mapping Even for Low-Coverage Whole Genome Sequencing Data." Bioinformatics (Oxford, England) 33 (23): 3726–32. https://doi.org/10.1093/bioinformatics/btx494.

Sandercock, Alexander M., Michael D. Peel, Cristiane H. Taniguti, Josué Chinchilla-Vargas, Shufen Chen, Manoj Sapkota, Meng Lin, *et al*, 2025a. "BIGapp: A User-Friendly Genomic Tool Kit Identified Quantitative Trait Loci for Creeping Rootedness in Alfalfa (*Medicago Sativa* L.)." *The Plant Genome* 18 (3): e70067. https://doi.org/10.1002/tpg2.70067.

Sandercock, Alexander M., Cristiane H. Taniguti, Josue Chinchilla-Vargas, Dongyan Zhao, Shufen Chen, Meng Lin, Manoj Sapkota, and Team Breeding Insight. 2025b. "Breeding Insight Genomics Functions for Polypoid and Diploid Species." GitHub. https://github.com/Breeding-Insight/BIGr.

Vallejo, Roger L., Michael R. Pietrak, Melissa M. Milligan, Guangtu Gao, Shogo Tsuruta, Breno O. Fragomeni, Roseanna L. Long, Brian C. Peterson, and Yniv Palti. 2024. "Genetic Architecture and Accuracy of Predicted Genomic Breeding Values for Sea Lice Resistance in the St John River Aquaculture Strain of North American Atlantic Salmon." *Aquaculture* 586 (May): 740819. https://doi.org/10.1016/j.aquaculture.2024.740819.

Yáñez, J. M., S. Naswa, M. E. López, L. Bassini, K. Correa, J. Gilbey, L. Bernatchez, *et al,* 2016. "Genomewide Single Nucleotide Polymorphism Discovery in Atlantic Salmon (*Salmo Salar*): Validation in Wild and Farmed American and European Populations." *Molecular Ecology Resources* 16 (4): 1002–11. https://doi.org/10.1111/1755-0998.12503.

Zhao, Dongyan, Katherine Maria Mejia-Guerra, Marcelo Mollinari, Deborah Samac, Brian Irish, Katarzyna Heller-Uszynska, Craig Thomas Beil, and Moira Jane Sheehan. 2023. "A Public Mid-Density Genotyping Platform for Alfalfa (*Medicago Sativa* L.)." *Genetic Resources* 4 (8): 55–63. https://doi.org/10.46265/genresj.EMOR6509.

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 120–129
DOI: 10.46265/genresj.QNIV2785
https://www.genresj.org
ISSN: 2708-3764

Egg production characteristics of several Bulgarian chicken breeds

Hristo Lukanov[a], Ivelina Pavlova[b,*], Atanas Genchev[a], Todor Petrov[a]

[a]*Trakia University, Faculty of Agriculture, Department of Animal Husbandry - Non-ruminant animals and special industries, Student campus, Stara Zagora, Bulgaria*
[b]*Trakia University, Faculty of Veterinary Medicine, Department of General Livestock Breeding, Student campus, Stara Zagora, Bulgaria*

Abstract: This study aimed to investigate and analyze the egg productivity of four Bulgarian chicken breeds, namely: Rhodope painted chicken (RPCh), Southwest Bulgarian chicken (SWBCh), Bulgarian longcrower (BL), and Struma chicken (SCh). The following traits were analyzed: age of sexual maturity (at 20% egg-laying intensity) (days); average daily feed intake (g); daily egg production and culled eggs (number); daily egg weight (g); livability (%). The following productive parameters were calculated: egg number per hen-housed; egg-laying intensity; feed conversion ratio (per kg of eggs); feed conversion per egg (g feed per egg); Egg Production Efficiency Index (EPEI). RPCh was identified as the earliest maturing group, reaching 20% egg-laying intensity at 157 days of age with the highest hen-housed egg production (223.9 eggs), whereas SCh exhibited the latest maturity, reaching this stage at 250 days of age and the lowest productivity (123.9 eggs). The highest average egg weight was recorded in the SCh group (58.0±0.54g), followed by the RPCh group (56.7±1.84g) and the BL group (56.0±1.68g), while the lowest average values were observed in the SWBCh group (50.9±0.68g). Based on the findings of this study, we can conclude that among all the tested Bulgarian chicken breeds, RPCh demonstrate the highest egg-laying potential. When compared to other purebred chickens, which are part of European genetic diversity, RPCh show superior performance in terms of age at sexual maturity, number of eggs produced per productive period, and egg weight.

Keywords: Rhodope Painted Chicken, Southwest Bulgarian Chicken, Bulgarian Longcrower, Struma Chicken

Citation: Lukanov, H., Pavlova, I., Genchev, A. and Petrov, T. (2025) "Egg production characteristics of several Bulgarian chicken breeds", *Genetic Resources*, 6(12), pp. 120–129. doi: 10.46265/genresj.QNIV2785.

Introduction

Global poultry production has steadily increased, driven by rising demand for affordable and accessible animal protein sources (OECD-FAO, 2017). According to OECD-FAO (2024), of the approximately 354 million tonnes of meat produced worldwide in 2023, around 139 million tonnes were from poultry, making it the leading category with nearly 40% of the total production. The majority of this poultry meat production is chicken, with about 103.5 million tonnes produced in 2023 (USDA FAS, 2024). A similar trend is observed in egg production, with approximately 97 million tonnes produced in 2023, of which around 94% were chicken eggs (FAO, 2024). Regionally, the EU produced about 13.3 million tonnes of poultry meat (Eurostat, 2024) and 6.7 million tonnes of eggs (EC, 2023a) in 2023. This highlights the dominant role of domestic chickens in both global and regional poultry farming. In Bulgaria, poultry farming is one of the most advanced livestock sectors, with poultry meat accounting for over half of the country's total meat production (Genchev and Lukanov, 2025).

Parallel to the positive development of the poultry sector, a negative trend is observed regarding the preservation of genetic diversity in domestic chickens (Malomane *et al*, 2019). In modern industrial poultry farming, highly productive lines from just a few chicken breeds are used (Teneva *et al*, 2015; Preisinger, 2021) out of the vast number of breeds known worldwide. In addition to productive purposes, the domestic chicken serves a variety of other functions in human life (ornamental, exhibition, sporting, etc.), which form the foundation for the breed diversity observed within the species

* Corresponding author: Ivelina Pavlova (ivelina.hristova@trakia-uni.bg)

Received: 25.04.2025 | Accepted: 02.09.2025 | Published online: 30.10.2025

(Lukanov, 2017a). Historically, numerous attempts have been made to classify chicken breeds based on factors such as origin, purpose, plumage, body size and other characteristics (BDRG, 2006; Roberts, 2008; APA, 2023; Kochish *et al*, 2023). Among these, the combined classification appears to be the most comprehensive. It categorizes chicken breeds into the following groups: meat breeds, egg-laying breeds, dual-purpose breeds, fighting breeds and ornamental breeds (including long-tailed, long-crowing, true bantams, miniature breeds and other ornamental varieties)(Lukanov, 2017a). The preservation of breed and genetic diversity in domestic chickens globally is predominantly attributed to the efforts of hobbyist poultry breeders, in conjunction with national poultry genetic centres, where such institutions are present (Teneva *et al*, 2015; Pavlova and Lukanov, 2024).

In Bulgaria, a total of ten chicken breeds are recognized, seven of which are of standard body size, and three are true bantams (Lukanov, 2023). Birds that do not exhibit signs of dwarfism are considered standard breeds, whereas those that do are classified as bantams. These include the Katunitsa chicken, Black Shumen chicken, Stara Zagora Red chicken, Struma chicken, Southwest Bulgarian chicken, Bulgarian longcrower, Rhodope painted chicken, Bregovska dzhinka, Struma bantam, and Southwest Bulgarian dzhinka (Lukanov and Pavlova, 2021; Pavlova and Lukanov, 2024). Among these, the Struma chicken (SCh), Rhodope painted chicken (RPCh), Southwest Bulgarian chicken (SWBCh) and Bulgarian longcrower (BL) share a similar geographical origin in Southwestern Bulgaria (Lukanov, 2023). The referenced study has investigated the incubation characteristics of eggs from these four breeds, while another has examined the exterior traits of these local Bulgarian chicken breeds (Pavlova and Lukanov, 2023). All four breeds are of standard size (standard chicken breeds), with three of them (SWBCh, SCh and BL) being typical ornamental breeds, while RPCh can be classified as a dual-purpose breed. Egg production is one of the most important economic factors in the poultry industry (El-Sabrout *et al*, 2022), as it is essential for both table egg production and hatching eggs. In this context, the traits that characterize egg productivity are significant for various branches of poultry farming, including backyard and ornamental poultry. To date, there has been no assessment of RPCh, SWBCh, BL and SCh egg-laying productivity, which would reveal their potential in this area. In this context, the aim of the present study was to investigate and analyze the egg productivity of RPCh, SWBCh, BL and SCh breeds.

Material and methods

Experimental design

The study was conducted from September 2022 to November 2023 at the experimental station of the Poultry Science Section, Faculty of Agriculture, Trakia University, Stara Zagora, Bulgaria. For the purposes of the study, hatching eggs were collected for incubation from various breeders of the tested breeds as follows: 276 hatching eggs from the RPCh breed (6 farms, 6 breeding groups), 366 from the BL breed (4 farms, 7 breeding groups), 254 from the SWBCh breed (5 farms, 5 breeding groups), and 180 from the SCh breed (2 farms, 5 breeding groups). The resulting chicks were reared at the same experimental station until the beginning of the trial. The study included typical representatives of the Rhodope painted chicken (RPCh), Southwest Bulgarian chicken (SWBCh), Bulgarian longcrower (BL), and Struma chicken (SCh) breeds, all at the same initial age of 140 days. Four groups of 25 pullets each were formed, corresponding to the four breeds and designated as RPCh, SWBCh, BL and SCh. The study covered one laying cycle, from the onset of egg production (20% laying rate) to the start of natural molting, lasting 52 weeks in RPCh, 49 weeks in SWBCh, 52 weeks in BL, and 40 weeks in SCh. The differences in the test period are due to variations in age at sexual maturity among the breeds.

Experimental bird management

The birds were housed on a deep litter system in a semi-enclosed facility divided into four pens, each measuring 2.5 × 4m, with an initial density of 2.5 birds per m². Each group's housing was fitted with natural light openings of identical size, providing an approximate individual area of 3m² and allowing continuous exposure to diffused natural daylight throughout the full natural photoperiod. Perches were provided at one end of the pen, ensuring a minimum perch space of 20cm per bird. A nest was provided for every five hens (a module of five individual nests in each pen). The facilities were equipped with manually refillable pan feeders and automatic cup drinkers, appropriately adjusted to the number of birds in each pen (Genchev and Lukanov, 2025). Feeding was ad libitum with a balanced compound feed in two phases: pre-laying and laying phase (Table 1). Feed consumption was recorded daily by weighing the feed residue remaining 24 hours after the feed was supplied, and was calculated as the average daily feed intake (ADFI) per bird. Eggs were collected regularly throughout the period of highest laying activity, from morning until early afternoon, with an additional collection in the late afternoon. The total number of eggs collected per day was considered the daily yield and was used to calculate daily egg production. Following collection, the eggs were weighed to calculate the average daily egg weight. Climate control was not implemented due to the facility's specifics and the birds' management system. Temperature (instantaneous, minimum and maximum) was monitored using a digital thermometer (TFA Dostmann Ltd.) installed in the 'birds' room', away from direct sunlight. The minimum recorded temperature during the entire period was -6.6°C on 11 February 2023, while the maximum was 38.5°C on 4 August 2023. The presented data on ambient temperature refer to the average daily values recorded by the Stara Zagora meteorological station.

For the purposes of the study, a lighting programme with additional artificial lighting was used, similar to those applied in intensive poultry farming in open-house systems. In our conditions, the birds were housed with a natural day length of approximately 12 hours. Additional artificial lighting was applied for ten days (until they reached 150 days of age) to gradually extend the day length to 14 hours by the end of these ten days. At 157 days of age, the day length was further increased by one hour, reaching 15 hours, with nine hours of darkness. Two weeks later (171 days of age), the day length was increased by one more hour, reaching 16 hours of daylight and eight hours of darkness. By the end of the test period, a day length of 16 hours was maintained. The extension of the photoperiod was accomplished by delaying the onset of the dark phase, with artificial lighting provided following the end of the natural daylight period.

Table 1. Nutritional composition of the compound feed used. *, time of first egg production in the group

Component	Pre-lay phase (140 days of age – maturity*)	Laying phase (whole egg-laying period)
Metabolized energy, MJ/kg	11.6	11.5
Crude protein, %	17.5	17.0
Lysine, %	0.75	0.8
Methionine, %	0.36	0.35
Calcium, %	2.0	3.8
av. Phosphorus	0.43	0.38

Egg production data collection

The following traits were analyzed: age of sexual maturity (at 20% laying rate) (days); average daily feed intake (g); daily egg production and culled eggs, number; daily egg weight (g); livability for the entire production period (including culled birds) (%). The following productive parameters were calculated: egg number per hen-housed; egg-laying intensity (laying rate) (%); feed conversion ratio (per kg of eggs); feed conversion per egg (g feed per egg). The egg production efficiency index (EPEI) was calculated by using the formula (Lukanov *et al*, 2023):

$$EPEI = [(L \times DEMP)/FCR)] \times 100,$$

where: L is livability for the period (%), DEMP is daily egg mass produced (kg), and FCR is the feed conversion ratio (kg/kg egg mass).

$$DEMP = (CHDEP \times AEW)/t,$$

where: CHDEP is the cumulative hen-day egg production for the period (number), AEW is the average egg weight (kg), and t is the period (days).

Statistical analysis

Statistical analyses were conducted using the IBM® SPSS® Statistics software package (version 26). A one-way analysis of variance (one-way ANOVA) was applied to assess inter-group differences. The following statistical parameters were calculated for data analysis and interpretation: mean value ($\bar{x}$) and standard error of the mean (SEM). Data are expressed as mean ± SEM.

Inter-group differences were considered statistically significant at $P < 0.05$, based on the LSD post hoc test, provided that the assumptions of normality (Shapiro–Wilk test; $n < 50$) were met and the ANOVA model was significant (F-test, $P < 0.05$). Microsoft Excel 16.0 (2018, Windows version) was used for the graphical presentation of the results.

Results

Figure 1 illustrates the changes in egg-laying intensity over the entire productive period for the tested groups of hens. Peak values of 77.1% laying intensity for RPCh group were recorded during the 24th productive week, with an average weekly ambient temperature of 10.4°C.

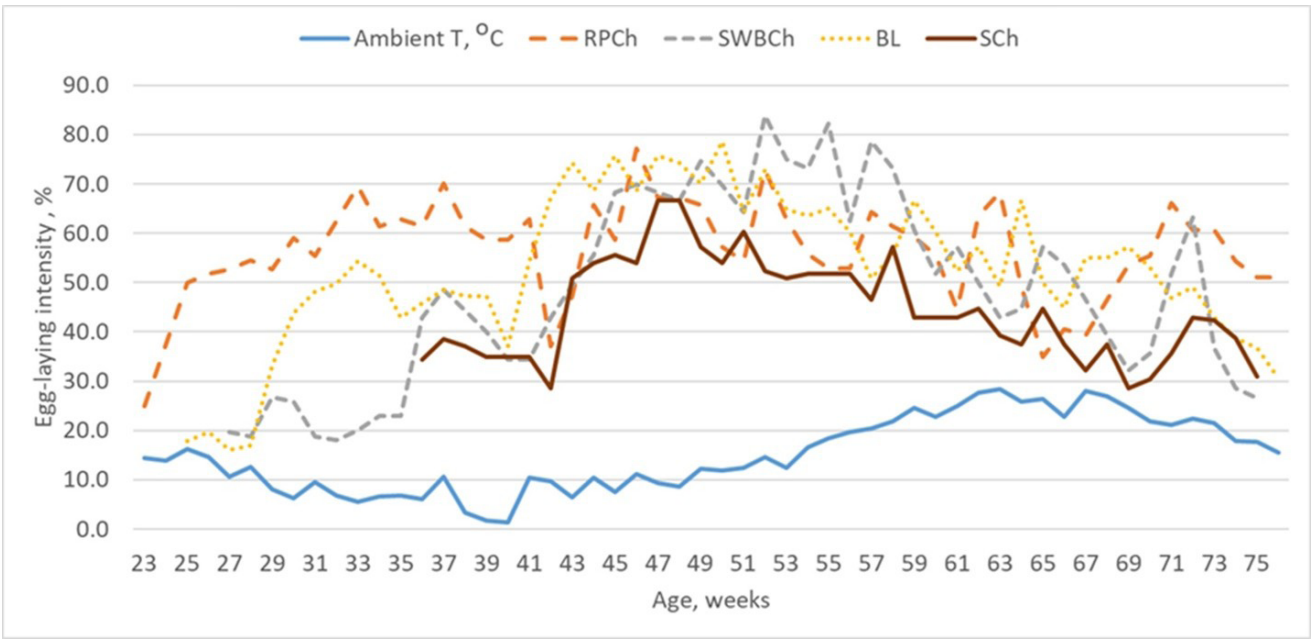


Figure 1. Egg-laying intensity recorded throughout the entire productive period. RPCh, Rhodope painted chicken; SWBCh, Southwest Bulgarian chicken; BL, Bulgarian longcrower; SCh, Struma chicken.

In contrast to RPCh, the SWBCh group showed a significantly delayed onset of productive maturity, reaching 20% laying rate at 179 days of age. This breed was characterized by a slow increase in laying performance, with the 50% threshold commonly used in industrial poultry production being reached only at 217 days of age. In this group, a sharp increase in laying intensity was observed after the 15th production week, reaching a peak value of 83.9% in the 26th week. A high laying rate was maintained until approximately the 32nd production week, after which the curve showed a marked decline. The SWBCh group also exhibited a shorter productive period. By week 49, the average weekly egg production had declined to 26.5%, with the majority of birds already undergoing molt.

In the BL group, 20% laying intensity was reached at 179 days of age, while the threshold of 50% was attained at 215 days. As with the other studied breeds, a typical laying curve characteristic of intensive poultry systems was not observed. A laying intensity of approximately 70% or higher was maintained between the 18th and 28th productive weeks, corresponding to ambient temperatures favourable for the species. Peak average weekly laying performance reached 78.6% during the 26th productive week. Unlike SWBCh birds, BL hens showed no sharp temperature-induced decline. A more substantial decrease was recorded only after the 48th productive week, likely linked to the onset of molting in some individuals and a gradual reduction in egg production within the group, declining to 30.6% by the 52nd productive week.

The SCh group was identified as the slowest-maturing among all the studied breeds, reaching a 20% laying rate at 250 days of age, with the first egg being laid slightly earlier, at 232 days. The genetic background, combined with the natural rearing conditions, is directly associated with the observed short productive period in the SCh group, which lasted for 40 weeks. When analyzing the dynamics of the trait change over the testing period, no significant differences in the laying curve are observed compared to the RPCh and BL groups. The later sexual maturity of the birds is also linked to the fact that the first half of the productive period occurred under more favourable ambient temperatures, which likely contributed to the relatively rapid achievement of peak production. For the SCh group, peak egg production values can be considered those above 50%, sustained between the 8th and 23rd productive weeks. The highest weekly average values for this trait were recorded in the 13th productive week (66.7%). Similar to the other groups, a decline in productivity is observed in parallel with the increasing age of the birds and ambient temperatures, with the dynamics of this decline being comparable to those of RPCh and BL groups.

The change in egg weight during the productive period in the RPCh group followed a typical growth curve with increasing age (Figure 2). At the start of the productive period, the average egg weight was 41.8g, gradually rising to the threshold of 53g by the 13th week. As the ambient temperature increased, a significant decline in laying intensity was observed, which coincided with the final third of the productive period. In terms of changes in SWBCh egg weight, no significant variation was observed throughout the entire laying period. In comparison to RPCh, which exhibited a 33.2% difference between the minimum and maximum average weekly egg weight, the variation in SWBCh was considerably lower, amounting to only 17.9%. At the beginning of the production cycle, the average weekly egg weight was 45.5g, reaching approximately 50g within 5–6 weeks. Peak values were recorded in the middle of the laying period (weeks 17–29), coinciding with favourable ambient temperature conditions. Unlike RPCh, the SWBCh group was characterized by a significantly lower egg weight, almost entirely falling within the S size category (< 53g) (EC, 2023b).

The egg weight in the BL group showed a rate of change throughout the productive period similar to the RPCh group, with the difference between the minimum and maximum weekly averages being 32.3%. In this breed, the optimal egg weight of 53g was reached by the 12th productive week and was maintained within the range of 53–62g until the end of the testing period. The highest weekly average egg weight was 62.2g, recorded in the 50th productive week.

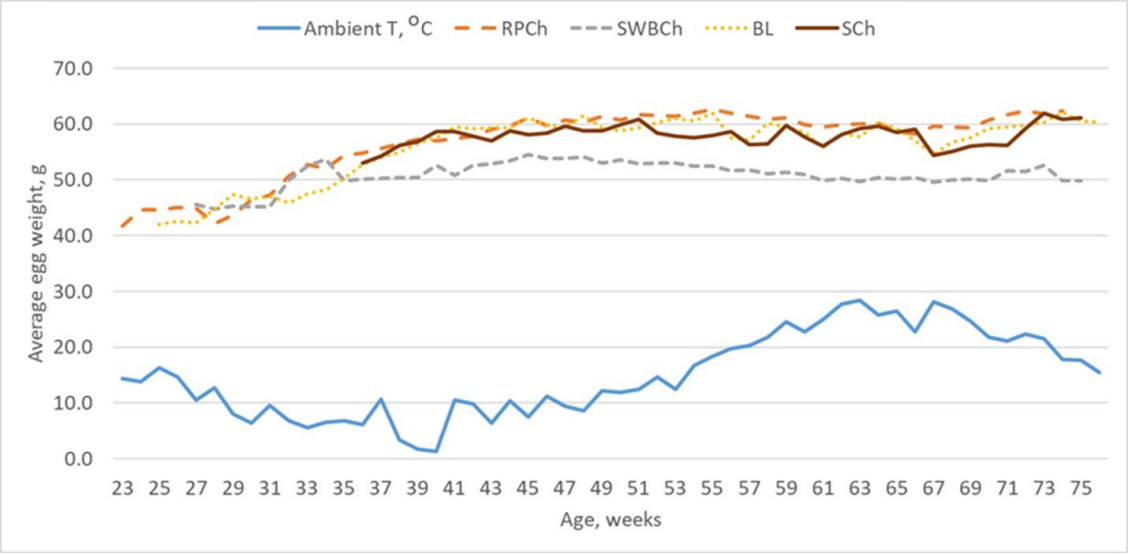


Figure 2. Average egg weight recorded throughout the entire productive period. RPCh, Rhodope painted chicken; SWBCh, Southwest Bulgarian chicken; BL, Bulgarian longcrower; SCh, Struma chicken.

The SCh egg weight exhibited the least fluctuation when compared to the reported minimum and maximum weekly average values throughout the entire productive period across all tested groups, with a difference of only 14.33%. This breed also showed the highest initial weekly average weight, which was 53.04g in the first productive week. The maximum weekly average egg weight was recorded in the 38th productive week, at 61.9g. In contrast to all other breeds included in the experiment, only the Struma chicken maintained egg weights throughout the entire productive period that consistently fell within the M weight category (53–62g) (EC, 2023b), based on the weekly average.

Patterns of average daily feed intake were generally similar among the four experimental groups, although RPCh exhibited more pronounced fluctuations up to approximately mid-lay (Figure 3). The highest intakes were recorded during the first half of the productive cycle, with RPCh consistently showing the greatest values (peaking above 160 g/day), followed by SWBCh, BL and SCh. As average daily ambient temperatures increased in late spring and summer, feed intake gradually declined across all groups, remaining at lower levels until the onset of molting or the end of the experimental period. The SCh group maintained the most stable intake pattern throughout the cycle, with only minor variations relative to seasonal changes in temperature.

The calculated EPEI for each of the four breeds included in the study is presented in Figure 4. This index reflects the efficiency of producing both table and hatching eggs. Lower EPEI values indicate less efficient production. Notably, the RPCh group recorded the highest efficiency (EPEI = 65.08), clearly distinguishing itself from the other breeds. The second-highest value was observed in the BL group, although it was 32.6% lower than that of RPCh. The least efficient performance was observed in the SWBCh and SCh groups, with EPEI values of 28.84 and 29.89, respectively. Liveability is one of the main parameters influencing the EPEI. Over the entire production period, the overall liveability of hens from the four groups was identical at 80%, with five birds per group lost due to mortality or culling, predominantly during the final stage of the laying period.

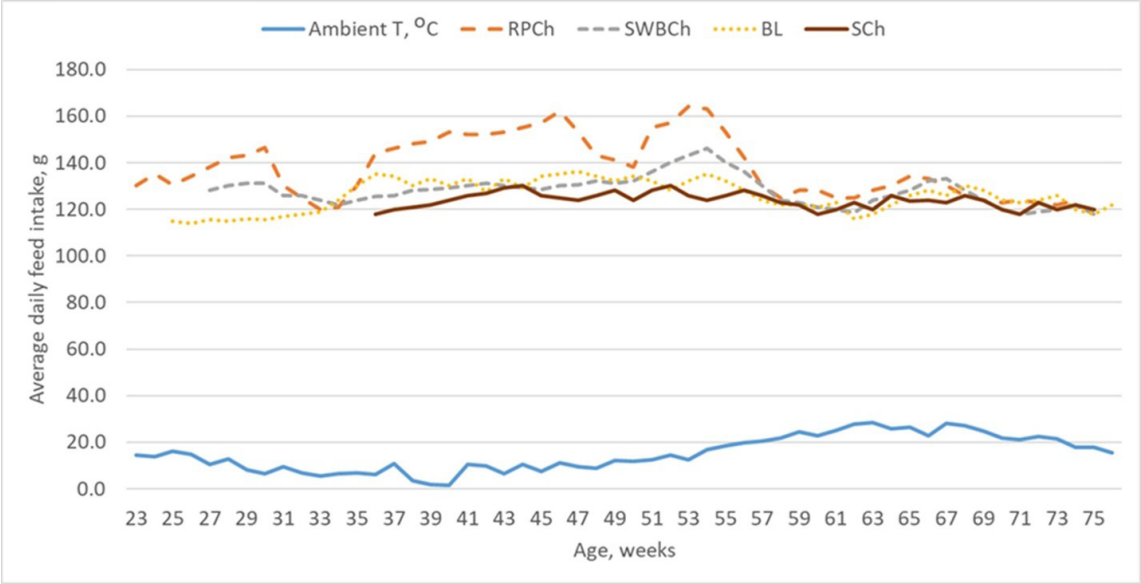


Figure 3. Average daily feed intake recorded throughout the entire productive period. RPCh, Rhodope painted chicken; SWBCh, Southwest Bulgarian chicken; BL, Bulgarian longcrower; SCh, Struma chicken.

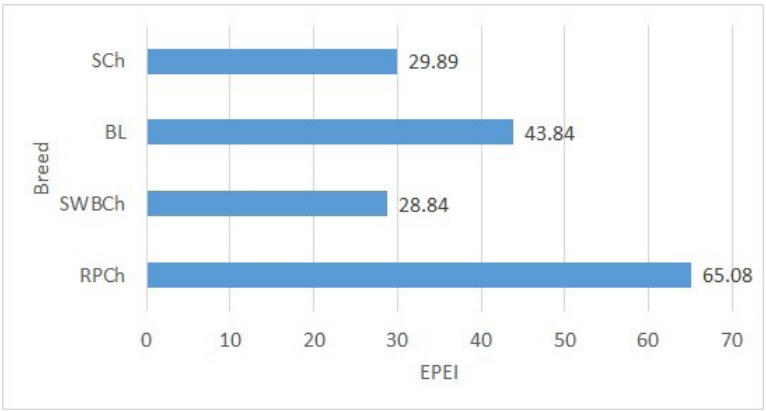


Figure 4. Egg production efficiency index (EPEI) calculated for each tested breed. SCh, Struma chicken; BL, Bulgarian longcrower; SWBCh, Southwest Bulgarian chicken; RPCh, Rhodope painted chicken.

Table 2 summarizes the results regarding the key parameters related to egg production in the four studied breeds. As previously noted, the RPCh group exhibited the earliest onset of sexual maturity, while the SCh group was the latest. The remaining two groups (BL and SWBCh) can be classified as intermediate in terms of sexual maturity.

Among the studied groups, RPCh exhibited the highest productivity, with a total of 223.9 eggs laid over a 52-week production period, corresponding to a laying rate of 57.4%. In contrast, the SCh group showed the lowest performance, producing 123.9 eggs – 44.7% less than RPCh – which equates to a laying rate of 33.9%. The remaining two groups, BL and SWBCh, displayed intermediate levels of productivity, with the BL group showing a higher production potential, recording 191.6 eggs during the same 52-week period and a laying rate of 52.5%.

According to the data presented in Table 2, the breed with the highest egg weight was SCh (58.0± 0.54g), while the lowest was observed in SWBCh (50.9 ± 0.68g) ($P < 0.001$). The other two groups (BL and RPCh) showed mean values similar to the SCh group, and their differences from the SWBCh group were also statistically significant ($P < 0.05$)

Considering the average egg weight and the number of eggs produced by the hen-housed, it can be summarized that the highest total egg mass per productive cycle was achieved by the RPCh group (approximately 12.7kg), followed by the BL group (10.73kg), SWBCh (8.49kg) and SCh (7.18kg).

Table 2 presents two expressions of feed converting efficiency, represented as the feed conversion ratio for producing one kilogram of egg mass or the feed required to produce a single egg. In this study, the most cost-effective feed conversion was observed in the RPCh group (244.1±9.2g of feed required to produce one egg and 4.34±0.18kg of feed required to yield one kilogram of eggs). Conversely, the least efficient feed conversion was observed in the SWBCh group (328.3±47.95g of feed required to produce one egg and 6.55±1.04kg of feed required to yield one kilogram of eggs).

Discussion

Age at sexual maturity is a major factor considered in selection for egg-laying poultry, as it represents an important reproductive trait (Xu *et al*, 2011; Liu *et al*, 2019; Genchev and Lukanov, 2025). Giesbrecht and Nordskog (1963) used the 20% level as the lowest point with reliable data when estimating age at sexual maturity, while suggesting 50% as the optimal threshold. Some authors propose an even lower threshold – such as a 10% laying rate – when evaluating the onset of maturity in indigenous chicken breeds (Schreiter and Freick, 2023). Hens from the RPCh group reached 20% laying rate at 157 days of age, with the age for reaching 50% laying rate, considered a benchmark in productive poultry farming (Genchev and Lukanov, 2025), being 162 days. Compared to the age of sexual maturity in modern high-performance laying hens (around 140–150 days), RPCh is relatively close, differing by approximately two weeks. When compared to purebred chickens that have not undergone targeted, scientifically based selection, RPCh ranks among early-maturing breeds, reaching sexual maturity at approximately 4.5 to 5.5 months (Lukanov, 2017a). The Spanish breed Asturian painted chicken, which is somewhat similar in exterior to RPCh but slightly larger and more massive (Pavlova, 2024), shows official data indicating a later maturation age of around 7 months (MAPA, 2025).

According to the 20% laying intensity and the threshold of 50% in the BL group, they are classified as a typical medium-maturing breed (Lukanov, 2017a). The late onset of sexual maturity, like in the SCh group, is characteristic of many large chicken breeds (Lukanov, 2017a).

The curve representing the laying intensity of the tested groups of hens does not follow the typical shape observed in hens raised under controlled microclimatic conditions. This is related to the rearing method, where the birds are kept in natural temperature conditions. In adult birds, the thermoneutral zone can be broadly defined, starting from

Table 2. Egg production parameters of the tested Bulgarian chicken breeds. Means ± SEM followed by the same letter are not significantly different at $P < 0.05$. RPCh, Rhodope Painted chicken; SWBCh, Southwest Bulgarian chicken; BL, Bulgarian Longcrower: SCh, Struma chicken; SM, Sexual maturity (at 20% egg-laying intensity); ADFI, average daily feed intake; AMEPR, average monthly egg production rate; HHEP, hen-housed egg production; AELI, average egg-laying intensity; AEW, average egg weight; FCR, feed conversion ratio; FCE, feed conversion per egg; SEM, standard error of the mean.

Breed	SM, days	ADFI, g	AMEPR, eggs	HHEP, eggs	AELI, %	AEW, g	FCR, kg/kg	FCE, g/egg
RPCh n = 25	157	138.5±3.67 abc	17.2±0.62 a	223.9	57.4±2.08 a	56.7±1.84 a	4.34±0.18 ab	244.1±9.2 a
SWBCh n = 25	191	129.0±1.67 ad	13.9±1.56	166.6	46.3±5.19 ab	50.9±0.68 abc	6.55±1.04 a	328.3±47.95
BL n = 25	179	127.1±2.85 b	14.7±1.22	191.6	52.5±4.08 c	56.0±1.68 b	5.45±0.98	291.2±38.56
SCh n = 25	250	122.1±2.52 cd	12.4±0.89 a	123.9	33.9±2.98 abc	58.0±0.54 c	5.35±0.41 b	308.8±22.49 a
ANOVA (P-value)		< 0.001	< 0.05		< 0.05	< 0.01	< 0.05	< 0.05

16°C (Poku *et al*, 2024) and reaching up to 29.9°C (Ribeiro *et al*, 2020), with an optimal range of 18°C to 22°C at a relative humidity of 50–75% (Kamanli *et al*, 2015). Temperature fluctuations have a significant impact on hens raised under such conditions, reflected in substantial variations in the laying curve (Gerzilov, 2011). A similarly negative effect of high daily temperatures on egg production has been reported by other authors (Gerzilov, 2011; Yoshida *et al*, 2011; Kim *et al*, 2024). This is mainly explained by heat stress, mediated by the reduced feed consumption of the birds (Getabalew and Negash, 2020). Under natural rearing conditions typical of the temperate climate in Bulgaria, birds are exposed not only to heat stress, mainly during the summer months (July, August and June), but also to cold stress, especially during the winter months (December, January and February). Cold stress is recognized as an environmental and managemental challenge, particularly in regions where temperatures regularly fall below 18°C (Kim *et al*, 2023). Similar to heat stress, birds exposed to temperatures lower than the thermoneutral zone exhibit negative parameters related to egg production (Torki et al, 2015; Li et al, 2020; Kim et al, 2023).

The number of eggs produced by an individual hen is a critical parameter in the selection process in modern poultry farming, while hen-housed egg production serves as a key indicator of the laying performance at the group level (Liu *et al*, 2019). The egg production capacity of local breeds, compared to modern results from high-productivity strains used in industrial poultry farming, shows a striking difference, especially over an extended productive period, which is commonly applied to modern egg-laying hens (El-Sabrou *et al*, 2022). It should be noted that modern egg-laying hybrids demonstrate this capacity under optimal rearing and feeding conditions. In contrast, local breeds are better adapted to the environment, i.e. when raised under uncontrolled conditions with limitations in optimal nutrition. This makes them valuable as a genetic reserve, including for potential inclusion in future breeding programmes (Chebo *et al*, 2022). The results obtained in this study position RPCh and BL as breeds with high genetic potential for egg production, when compared to other purebred chickens (BDRG, 2006; Henning *et al*, 2017; Lukanov, 2017a; Schreiter and Freick, 2023). Comparing RPCh with the data presented for the Asturian Painted Chicken (average of 160 eggs; MAPA, 2025), it can be said that the former shows significantly higher potential in terms of laying capacity. The egg production recorded for SCh over the productive period is comparable to that of other large ornamental chicken breeds, such as Brahma, Cochin and Orpington (Hrnčár *et al*, 2015).

Modern commercial laying hens produce eggs with an average weight of 62–65g over the entire laying period (Genchev and Lukanov, 2025). Globally, however, average egg weight tends to be slightly lower, around 60–61g, with regional preferences influencing egg size (Thiruvenkadan *et al*, 2010). The detrimental impact of elevated ambient temperatures on egg weight in domestic hens has been thoroughly documented by Bennion and Warren (1933). Similar to egg production, egg weight is adversely affected by heat stress and reduced feed intake of the birds (Kilic and Simsek, 2013).

The RPCh breed, identified in the study as having the highest laying performance, produced eggs with a lower average weight compared to the larger Asturian Painted Chicken, which reaches an average of 65g (MAPA, 2025). The low egg weight observed in SWBCh corresponds to the lower range of the trait reported for the breed by Pavlova (2024), namely 50–55g. Regarding egg weight, the Bulgarian longcrower exhibits typical values reported for other Balkan long-crowing chicken breeds (Lukanov, 2012; Różewicz and Kaszperuk, 2018). The Turkish breed Denizli, which is believed to be close to the Balkan long-crowing breeds (Lukanov, 2017b), is reported in various sources to have a lower average egg weight of 50–52g (Fidan and Nazlıgül, 2012; Özdemir *et al*, 2013) to values similar to those of BL (BDRG, 2006; Özdoğan *et al*, 2007; Kaya and Yıldız, 2014). In a review focused on Bulgarian chicken breeds, a lower egg weight range (50–55g) was suggested for BL; however, this estimate was based on preliminary assumptions rather than comprehensive research (Lukanov *et al*, 2021). According to Lukanov (2023), both RPCh and BL exhibited slightly greater egg weights, likely attributable to the advanced age of the hens, including those in their second productive cycle.

The egg weight reported by Hrnčár *et al*, (2015) for three large ornamental breeds (Brahma, Cochin and Orpington) is significantly lower than that of SCh. The recorded egg weight in the Struma chicken breed is consistent with the findings of a more recent study (Lukanov, 2023) and higher than the average values reported for the breed prior to these detailed investigations (Lukanov, 2012; Teneva *et al*, 2015; Lukanov, 2017a). Comparison with the other large native Bulgarian breed – the Katunitsa chicken – shows that the latter has higher egg weight (Gerzilov *et al*, 2015) and reaches sexual maturity significantly earlier (Nikolov and Gerzilov, 2011).

The average egg weight observed in all four tested Bulgarian chicken breeds is lower compared to that of commercial layers subjected to targeted selection for traits related to egg production, including egg weight. Nevertheless, the two breeds demonstrating the highest laying potential – RPCh and BL – exhibited relatively high egg weights when compared to many other local breeds. As no focused selection for this trait has been applied to these populations, their current performance suggests a promising potential for further genetic improvement in this direction.

The egg production efficiency index (EPEI) is a dimensionless indicator combining liveability, egg production and feed conversion into a single score, with higher values indicating better overall production efficiency. The calculated EPEI is significantly lower in all four studied breeds compared to the modern laying hybrids, in which the reference value is approximately 230 (Lukanov *et al*, 2023). These differences can be explained by the lower egg production parameters observed in the local chicken breeds included in the present study, as well as by the higher proportion of culls (which negatively affects liveability), compared with the optimal values reported for commercial hybrids in the cited study. The reduced efficiency in SCh and SWBCh groups can be largely explained by the limited laying performance in the SCh group and the lower average egg weight recorded in the SWBCh.

Feed conversion is an important economic trait that reflects the efficiency of converting feed into finished production (eggs), primarily determined by the feed conversion ratio (FCR) (Li *et al*, 2024). It is well known that feed constitutes a significant portion of the production cost in egg-laying poultry farming (Farooq *et al*, 2002; Thiruvenkadan *et al*, 2010), with considerable variation observed depending on the farming practices applied (Kato *et al*, 2022). The results

obtained regarding feed conversion are comparable to those reported for EPEI. When comparing the feed transformation efficiency of the four experimental groups of hens with the current performance levels of modern white- and brown-egg laying hybrids, it is evident that the experimental groups lag considerably behind. Even the best-performing group, RPCh, exhibited approximately twice the values for both types of FCR compared with those of modern commercial laying (Churchil and Suresh, 2021; Genchev and Lukanov, 2025). Studies involving various non-commercial chicken breeds report variable FCR values, ranging from those similar to modern laying hybrids (Besari *et al*, 2017) to higher values similar to our results (Lukanov *et al*, 2016; Phuong and Nha, 2024), or even more striking differences in some low-performing breeds (Nguyen Van *et al*, 2020).

Conclusion

The results of the study indicate that the Rhodope painted chicken (RPCh) breed demonstrates the highest potential in terms of egg production characteristics, followed by the Bulgarian longcrower (BL). Both breeds are distinguished by early maturity, with this trait being particularly pronounced in RPCh. Egg production in these two indigenous breeds is above average compared to other purebred chickens that are not part of industrial poultry farming. Due to their high egg production, egg weight and low feed consumption compared to the other three tested breeds, the RPCh group demonstrates the most efficient egg production. The two other tested breeds, the Struma chicken and the Southwest Bulgarian chicken, display less favourable characteristics in terms of egg productivity. The good egg production and attractive exterior of the RPCh and BL breeds provide strong grounds for their significant potential in amateur and backyard poultry farming.

Acknowledgements

The authors would like to express their sincere gratitude to all those who contributed to this work, especially the breeders of the studied chicken breeds. We also wish to acknowledge the financial support from Project 4AF/22 ‘Study on the Egg Productivity of Some Bulgarian Chicken Breeds’.

Author contributions

Hristo Lukanov and Ivelina Pavlova contributed equally to the conception and design of the study, the execution of the experimental work, data analysis, and manuscript preparation. Atanas Genchev participated in the development and optimization of the experimental methodology and provided technical support during data collection. Todor Petrov was responsible for monitoring the productivity parameters.

Conflict of interest statement

The authors declare that there is no conflict of interest.

Ethics statement

All experimental procedures involving animals were carried out in accordance with the relevant institutional and national regulations for the ethical treatment of animals. Since the study involved only routine productivity assessment under normal housing conditions, no specific ethical approval was necessary.

References

APA. (2023). American Standard of Perfection: 45th edition (American Poultry Association), 406 p.

BDRG. (2006). Der Rassegeflügel-Standard für Europa in Farbe (Bund Deutscher Rassegeflügelzüchter e.V., Fürth), 848 p. (In German)

Bennion, N. L. and Warren, D. C. (1933). Temperature and its effect on egg size in the domestic fowl. *Poultry Science*, 12(2), 69-82. doi: https://doi.org/10.3382/ps.0120069

Besari, F., Pandi, J., and Dom, M. (2017). Feed conversion efficiency and egg production of village chickens under improved feeding and management practices. In Paper presented at the 8th HUON Seminar: Celebrating 52 years of Nation Building – Embracing Challenges beyond 2017, University of Technology, Papua New Guinea.

Chebo, C., Betsha, S., and Melesse, A. (2022). Chicken genetic diversity, improvement strategies and impacts on egg productivity in Ethiopia: a review. *World’s Poultry Science Journal*, 78(3), 803–821. doi: https://doi.org/10.1080/00439339.2022.2067020

Churchil, R. R. and Suresh. (2021). Current concepts in nutrition and feeding of hybrid layer chicken. *Indian Journal of Veterinary and Animal Sciences Research*, 50(6), 1–16.

EC (2023a). European Commision. Eggs - Market Situation Dashboard (Directorate-General for Agriculture and Rural Development), 16 p. Retrieved from https://agriculture.ec.europa.eu/system/files/2024-02/eggs-dashboard_en.pdf

EC (2023b). European Commission Delegated Regulation (EU) 2023/2465 of 17 August 2023 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards marketing standards for eggs, and repealing Commission Regulation (EC) No 589/2008. Official Journal of the European Union, Document 32023R2465.

El-Sabrout, K., Aggag, S., and Mishra, B. (2022). Advanced practical strategies to enhance table egg production. *Scientifica*, 1393392. doi: https://doi.org/10.1155/2022/1393392

Eurostat. (2024). Agricultural production - livestock and meat. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agricultural_production_-_livestock_and_meat

FAO. (2024). Agricultural production statistics 2010–2023 (FAOSTAT Analytical Briefs, No. 96. Rome), 16 p. https://openknowledge.fao.org/handle/20.500.14283/cd3755en

Farooq, M., Mian, M. A., Durrani, F. R., and Syed, M. (2002). Feed consumption and efficiency of feed utilization by egg type layers for egg production. *Livestock Research for Rural Development*, 14(1).

Fidan, E. D., and Nazlıgül, A. (2012). The effect of cage position and density on some production traits in Denizli chickens. *Animal Health, Production and Hygiene*, 1, 31–37. (In Turkish)

Genchev, A. and Lukanov, H. (2025). Poultry farming (Trakia University Academic Press, Stara Zagora, Bulgaria), 278 p. (In Bulgarian)

Gerzilov, V. (2011). Egg productivity in some fowl strains from the National gene pool reared under bio-friendly conditions. *Agricultural Sciences*, 3(6), 105-112.

Gerzilov, V., Nikolov, A., Petrov, P., Bozakova, N., Penchev, G., & Bochukov, A. (2015). Effect of a dietary herbal mixture supplement on the growth performance, egg production and health status in chickens. *Journal of Central European Agriculture*, 16(2), 10–27. doi: https://doi.org/10.5513/JCEA01/16.2.1580

Getabalew, M. and Negash, A. (2020). Effect of high temperature on body weight gain, egg production, and egg shell formation process in laying hen: A review. *British Journal of Poultry Sciences*, 9(2), 42-47. url: https://idosi.org/bjps/9(2)20/3.pdf

Giesbrecht, F. G. and Nordskog, A. W. (1963). Estimating age at sexual maturity from flock records. *Poultry Science*, 42(1), 83–87. doi: https://doi.org/10.3382/ps.0420083

Henning, M., Ehling, C., Weigend, S., Fellmin, M., Feldmann, A., and Tiemann, I. (2017). Cryo reserve in the chicken (Project No. 10BM016 and -027). Final Report of a Model and Demonstration Project in the Field of Biological Diversity. Bund Dt. Rassegeflügelzüchter e.V./ Bruno-Dürigen-Institut bzw. Wissenschaftlicher Geflügelhof des BDRG. (In German)

Hrnčár, C., Gašparovič, M., Gálik, B., & Bujko, J. (2015). Egg traits, fertility and hatchability of Brahma, Cochin and Orpington chicken breeds. *Scientific Papers: Animal Science and Biotechnologies*, 48(2), 137–141.

Kamanli, S., Durmuş, I., Yalçın, S., Yıldırım, U. and Meral, Ö. (2015). Effect of prenatal temperature conditioning of laying hen embryos: Hatching, live performance and response to heat and cold stress during laying period. *Journal of Thermal Biology,* 51, 96–104. doi: https://doi.org/10.1016/j.jtherbio.2015.04.001

Kato, H., Shimizuike, Y., Yasuda, K., Yoshimatsu, R., Yasuda, K. T., Imamura, Y., and Imai, R. (2022). Estimating production costs and retail prices in different poultry housing systems: Conventional, enriched cage, aviary, and barn in Japan. *Poultry Science*, 101(12), 102194. doi: https://doi.org/10.1016/j.psj.2022.102194

Kaya, M. and Yıldız, M. A. (2014). Tavuğun evcilleştirilmesi ve Türkiye yerli tavuk ırkları. Tavukçuluk Araştırma Dergisi, 11(2), 21–28. (In Turkish)

Kilic, I. and Simsek, E. (2013). The effects of heat stress on egg production and quality of laying hens. *Journal of Animal and Veterinary Advances*, 12(1), 42-47. url: https://scispace.com/papers/the-effects-of-heat-stress-on-egg-production-and-quality-of-37n0064xod

Kim, D. H., Song, J. Y., Park, J., Kwon, B. Y. and Lee, K. W. (2023). The effect of low temperature on laying performance and physiological stress responses in laying hens. *Animals*, 13(24), 3824. doi: https://doi.org/10.3390/ani13243824

Kim, H.-R., Ryu, C., Lee, S.-D., Cho, J.-H., and Kang, H. (2024). Effects of Heat Stress on the Laying Performance, Egg Quality, and Physiological Response of Laying Hens. *Animals*, 14(7), 1076. doi: https://doi.org/10.3390/ani14071076

Kochish, I. I., Titov, V. Y., Nikonov, I. N., Brazhnik, E. A., Vorobyov, N. I., Korenyuga, M. V., Myasnikova, O. V., Dolgorukova, A. M., Griffin, D. K., and Romanov, M. N. (2023). Unraveling signatures of chicken genetic diversity and divergent selection in breed-specific patterns of early myogenesis, nitric oxide metabolism and post-hatch growth. *Frontiers in Genetics*, 13. doi: https://doi.org/10.3389/fgene.2022.1092242

Li, D., Tong, Q., Shi, Z., Zheng, W., Wang, Y., Li, B. and Yan, G. (2020). Effects of cold stress and ammonia concentration on productive performance and egg quality traits of laying hens. *Animals*, 10(12), 2252. doi: https://doi.org/10.3390/ani10122252

Li, Y., Ma, R., Qi, R., Li, H., Li, J., Liu, W., Wan, Y., Li, S., Sun, Z., Xu, J., and Zhan, K. (2024). Novel insight into the feed conversion ratio in laying hens and construction of its prediction model. *Poultry Science*, 103(10), 104013. doi: https://doi.org/10.1016/j.psj.2024.104013

Liu, Z., Yang, N., Yan, Y., Li, G., Liu, A., Wu, G., and Sun, C. (2019). Genome-wide association analysis of egg production performance in chickens across the whole laying period. *BMC Genetics*, 20, 67. doi: https://doi.org/10.1186/s12863-019-0771-7

Lukanov, H. (2012). Balkan breeds and breed groups (Parts I & II). *Aviculture Europe*, 8(6), 1–16.

Lukanov, H. (2017a). Exhibition and ornamental poultry breeding (Kota Publ. House, Stara Zagora, Bulgaria), 528 p. (In Bulgarian)

Lukanov, H. (2017b). Balkan longcrowing chicken breeds. *Aviculture Europe*, 13(3), 1-7.

Lukanov, H. & Pavlova, I. (2021). Morphological and morphometric characterization of Bulgarian local chicken breed - Southwest Bulgarian dzinka. *Agricultural Science and Technology*, 13(2), 147-151. doi: https://doi.org/10.15547/ast.2021.02.024

Lukanov, H. (2023). Incubation characteristics of some Bulgarian chicken breeds. *Bulgarian Journal of Animal Husbandry*, 60(6), 44-52. doi: https://doi.org/10.61308/TZDV8636 (In Bulgarian)

Lukanov, H., Genchev, A., and Petrov, T. (2023). The Egg Production Efficiency Index (EPEI) as an economic indicator for measuring poultry egg production. *Bulgarian Journal of Agricultural Science*, 29(4), 747–751.

Lukanov, H., Pavlova, I., and Genchev, A. (2021). Bulgarian chicken breeds - part of the world's genetic diversity: I. Standard breeds. In International scientific and practical conference "Innovative approaches to increasing the productivity of farm animals", Kuban State Agrarian University named after I. T. Trubilin, 359–365. (In Russian).

Lukanov, H., Petrov, P., Genchev, A., Halil, E. and Ismail, N. (2016). Productive performance of Easter egger crosses of Araucana and Schijndelaar roosters with white Leghorn hens. *Trakia Journal of Sciences*, 1, 72–79. url: http://tru.uni-sz.bg/tsj/Vol.14,%20N%201,%202016/H.Lukanov%20(1).pdf

Malomane, D. K., Simianer, H., Weigend, A., Reimer, C., Schmitt, A. O., and Weigend, S. (2019). The SYNBREED chicken diversity panel: A global resource to assess chicken diversity at high genomic resolution. *BMC Genomics*, 20, 345. doi: https://doi.org/10.1186/s12864-019-5727-9

MAPA. (2025). Pita Pinta: Production data (Ministerio de Agricultura, Pesca y Alimentación). (In Spanish) Retrieved from https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/razas-ganaderas/razas/catalogo-razas/aviar/pita-pinta/datos_productivos.aspx

Nguyen Van, D., Moula, N., Moyse, E., Do Duc, L., Vu Dinh, T., and Farnir, F. (2020). Productive performance and egg and meat quality of two indigenous poultry breeds in Vietnam, Ho and Dong Tao, fed on commercial feed. *Animals*, 10(3), 408. doi: https://doi.org/10.3390/ani10030408

Nikolov, A. and Gerzilov, V. (2011). Productivity of newly selected AN heavy chicken line. *Agricultural Sciences*, 3(6), 99–104.

OECD-FAO. (2017). Meat: OECD-FAO agricultural outlook 2017-2026 (Organisation for Economic Co-operation and Development & Food and Agriculture Organization of the United Nations), 142 p. doi: https://doi.org/10.1787/agr_outlook-2017-en

OECD-FAO. (2024). OECD-FAO Agricultural Outlook 2024-2033 (OECD Publishing, Paris/FAO, Rome), 335 p. doi: https://doi.org/10.1787/4c5d2cfb-en

Özdemir, D., Özdemir, E. D., De Marchi, M., and Cassandro, M. (2013). Conservation of local Turkish and Italian chicken breeds: A case study. *Italian Journal of Animal Science*, 12(2), e49. doi: https://doi.org/10.4081/ijas.2013.e49

Özdoğan, N., Gürcan, İ. S., and Bilgen, A. (2007). Egg weight and egg weight repeatability of Denizli and Gerze local hen breeds. Lalahan Hayvancılık Araştırma Enstitüsü Dergisi, 47(1), 21–28. (In Turkish)

Pavlova, I. (2024). Bulgaria as a part of the world's poultry genetic resources – Bulgarian chicken breeds. *Danubian Animal Genetic Resources,* 9(1), 16. doi: https://doi.org/10.59913/dagr.2024.17293

Pavlova, I. and Lukanov, H. (2023). Exterior characteristics of some Bulgarian chicken breeds. In Proceedings of the Scientific Conference with International Participation Animal Science – Challenges and Innovations, Sofia, Bulgaria. 118–127.

Pavlova, I. and Lukanov, H. (2024). Population status and distribution of Bulgarian indigenous chicken breeds. *Journal of Central European Agriculture*, 25(2), 313–324. doi: https://doi.org/10.5513/JCEA01/25.2.4204

Phuong, L. T. and Nha, P. T. (2024). Reproductive performance of local Noi chicken over two generations in the Mekong Delta of Vietnam. *Advances in Animal and Veterinary Sciences*, 12(8), 1596–1603.

Poku, R. A., Agyemang-Duah, E., Donkor, S., Ayizanga, R. A., Osei-Amponsah, R., Rekaya, R. and Aggrey, S. E. (2024). Changes in rectal temperature as a means of assessing heat tolerance and sensitivity in chickens. *Tropical Animal Health and Production*, 56(9), 391. doi: https://doi.org/10.1007/s11250-024-04242-1

Preisinger, R. (2021). Commercial layer breeding: Review and forecast. Züchtungskunde, 93(3), 210–228. (In German)

Ribeiro, B. P. V., Yanagi Junior, T., Duarte de Oliveira, D., Ribeiro de Lima, R. and Zangeronimo, M. G. (2020). Thermoneutral zone for laying hens based on environmental conditions, enthalpy and thermal comfort indexes. *Journal of Thermal Biology*, 93, 102678. doi: https://doi.org/10.1016/j.jtherbio.2020.102678

Roberts, V. (2008). British poultry standards: 6th edition (Wiley-Blackwell), 480 p.

Różewicz, M. and Kaszperuk, K. (2018). Long-crowing and long-tailed chickens: Characteristics of the breeds. *Wiadomości Zootechniczne*, 56(4), 181–199.

Schreiter, R. and Freick, M. (2023). Laying performance characteristics, egg quality, and integument condition of Saxonian chickens and German Langshan bantams in a free-range system. *Journal of Applied Poultry Research*, 32(3), 100359. doi: https://doi.org/10.1016/j.japr.2023.100359

Teneva, A., Gerzilov, V., Lalev, M., Lukanov, H., Mincheva, N., Oblakova, M., Petrov, P., Hristakieva, P., Dimitrova, I., and Periasamy, K. (2015). Current status and phenotypic characteristics of Bulgarian poultry genetic resources. *Animal Genetic Resources*, 56, 19-27. doi: https://doi.org/10.1017/S2078633615000016

Thiruvenkadan, A. K., Panneerselvam, S., and Rabakaran, R. (2010). Layer breeding strategies: An overview. *World's Poultry Science Journal*, 66(2), 477–502. doi: https://doi.org/10.1017/S0043933910000553

Torki, M., Akbari, M. and Kaviani, K. (2015). Single and combined effects of zinc and cinnamon essential oil in diet on productive performance, egg quality traits, and blood parameters of laying hens reared under cold stress condition. *International Journal of Biometeorology,* 59(9), 1169–1177. doi: https://doi.org/10.1007/s00484-014-0928-z

USDA FAS. (2024). Livestock and poultry: World markets and trade (United States Department of Agriculture, Foreign Agricultural Service), 15 p. https://www.fas.usda.gov/sites/default/files/2024-07/Livestock_poultry.pdf

Xu, H., Zeng, H., Luo, C., Zhang, D., Wang, Q., Sun, L., Yang, L., Zhou, M., Nie, Q., and Zhang, X. (2011). Genetic effects of polymorphisms in candidate genes and the QTL region on chicken age at first egg. *BMC Genetics*, 12, 33. doi: https://doi.org/10.1186/1471-2156-12-33

Yoshida, N., Fujita, M., Nakahara, M., Kuwahara, T., Kawakami, S.-I., and Bungo, T. (2011). Effect of high environmental temperature on egg production, serum lipoproteins, and follicle steroid hormones in laying hens. *The Journal of Poultry Science,* 48(3), 207–211. doi: https://doi.org/10.2141/jpsa.010126

Genetic Resources

CORRECTION

DOI: 10.46265/genresj.JZQN2033
https://www.genresj.org
ISSN: 2708-3764

CORRECTION: Organization of plant Biological Resource Centers for research in France: History, evolution and current status

Valérie Bergheaud*[,a], Jean-Marc Audergon[b], Arnaud Bellec[c], Anne Delaunay[1], Jérôme Duminil[d], Stéphane Dussert[d], Florence Esnault[e], Emmanuel Geoffriau[f], Brigitte Gouesnard[g], Christophe Jenny[g,h], Alain Label[i], Philippe Lashermes[d], Najate Maghnaoui[g,h], Cécile Marchal[b], Franciane Nuissier[j], Nilda Paulo-de-la-Réberdiere[g,h], Aurélia Priet[k], Valérie Rieucau[7,8], Paule Térès[g,h], Anne-Françoise Adam-Blondon[l]

[a]*Université Paris-Saclay, INRAE, BAP, France*
[b]*INRAE, PACA Center, UGAFL, France*
[c]*INRAE, CNRGV French Plant Genomic Resource Center, France*
[d]*DIADE, IRD, Université Montpellier, Cirad, France*
[e]*INRAE, Institut Agro, Université de Rennes, IGEPP, France*
[f]*Institut Agro, Université d'Angers, INRAE, IRHS, SFR Quasav, France*
[g]*UMR AGAP Institut, Université Montpellier, CIRAD, INRAE, Institut Agro, France*
[h]*CIRAD, UMR AGAP institut, Cirad, Montpellier, France*
[i]*Diagonal, INRAE, France*
[j]*INRAE, UR ASTRO, Guadeloupe, France*
[k]*INRAE, UR4 (UR P3F), Nouvelle-Aquitaine-Poitiers Center, France*
[l]*Université Paris-Saclay, INRAE, URGI, France*

**Corresponding author: Valérie Bergheaud (valerie.bergheaud@inrae.fr)*

Correction to: *Genetic Resources* https://doi.org/10.46265/genresj.ASZO2413; published online 04 March 2025

Incorrect affiliation
In the published article, there was an error in the affiliations of Nilda Paulo-de-la-Réberdiere. Instead of "[h,g]", it should have been "[g,h]".

Incorrect author names
In the published article, two author names were incorrectly written:
- Cristophe Jenny (correct: Christophe Jenny)
- Françoise Nuissier (correct: Franciane Nuissier)

Error in Table 1
In the published article, there was an error in Table 1:
- [Perennial plants in Guyana (PPG)]: Instead of "Guyana", it should have been "French Guiana".
- [*Coffea* spp., *Theobroma* spp., *Hevea* spp., *Dalbergia* spp.]: Instead of *Dalbergia*, it should have been *Aniba rosodora*.

The authors apologize for these errors and state that they do not affect the scientific conclusions of the article in any way.

The PDF and HTML versions of the original article have been updated.

Published online: 06 November 2025

Received: 05.11.2025 | Accepted: 05.11.2025 | Published online: 06.11.2025

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 131–152
DOI: 10.46265/genresj.LSBJ4572
https://www.genresj.org
ISSN: 2708-3764

Novel germplasm of tepary and other *Phaseolus* bean wild relatives from dry areas of southwestern USA

Daniel G. Debouck[a,*], Richard C. Pratt[b], Sarah Dohle[c], Timothy Porch[d], Marcela Santaella[a], Luis Guillermo Santos[a] and Milan O. Urban[e]

[a]*Genetic Resources Program, International Center for Tropical Agriculture (CIAT), Palmira, Colombia*
[b]*New Mexico State University, Department of Plant and Environmental Sciences, Las Cruces, New Mexico, USA*
[c]*United States Department of Agriculture, Agriculture Research Service (USDA-ARS), Pullman, Washington, USA*
[d]*United States Department of Agriculture, Agriculture Research Service (USDA-ARS), Mayagüez, Puerto Rico*
[e]*Bean Program, International Center for Tropical Agriculture (CIAT), Palmira, Colombia now at International Center for Biosaline Agriculture (ICBA), Dubai, United Arab Emirates*

Abstract: Heat and drought stresses threaten global bean production. Additional genetic resources are needed in genebanks for future improvement of bean crops through breeding for tolerance. The USA southwestern Sky Island mountains contain such genetic resources that have not been adequately collected nor characterized. Continuing the work done in 2023 (during which 14 populations were identified and described), a 9-day exploration in 2024 in southern New Mexico and Arizona for wild teparies and other *Phaseolus* species resulted in the collection of herbarium and seed samples of 18 populations of *P. acutifolius*, one each of *P. angustissimus* and *P. filiformis*, two of *P. grayanus*, three of *P. maculatus* and possibly three of *P. montanus*, or 28 populations in total. Samples of nodules and soil of rhizosphere were also collected. Outcomes and ways to improve these exploration endeavours are discussed.

Keywords: Crop wild relatives, drought stress, germplasm exploration, high-temperature stress, tepary, *Phaseolus acutifolius*

Citation: Debouck, D. G., Pratt, R. C., Dohle, S., Porch, T., Santaella, M., Santos, L. G. and Urban, M. O. (2025) "Novel germplasm of tepary and other *Phaseolus* bean wild relatives from dry areas of southwestern USA", *Genetic Resources*, 6(12), pp. 131–152. doi: 10.46265/genresj.LSBJ4572.

Introduction

The megadrought of western North America with effects extending south to the Central American Dry Corridor is a current and historic climatic feature (Williams *et al*, 2020; McKinnon *et al*, 2021; IPCC, 2023; Chen *et al*, 2025). That water-limited area is a huge arc extending in the north from Saskatchewan (Canada) to North Dakota and Nebraska south to New Mexico (USA), Chihuahua and Zacatecas (Mexico) and ending in eastern Guatemala. The present combination of record high temperatures, prolonged drought and limited water resources can have profound implications for agricultural systems. Increasingly, growing urban areas and agriculture will compete for ever scarcer fresh-water resources. Farmers in remote areas will likely seek grains with sufficient value (e.g. barley for breweries, beans for export to Mexico and Central America, quinoa for specialty markets) to cover production and postharvest costs such as transportation. High-value grains capable of providing more stable income with a lower water requirement will be essential. Crops and cultivars highly resistant to drought and heat stresses are now high on the agenda of agronomists and breeders in that area (Parker *et al*, 2023; Silber-Coats *et al*, 2025), but also in other arid regions of the world e.g. Africa (Assefa *et al*, 2019).

While the vast American arc aforementioned currently has many introduced crops, tepary bean (*Phaseolus acutifolius* Asa Gray) has long been known as a native drought and heat tolerant crop (Freeman, 1912) as it was grown by the pre-Columbian peoples of the Southwest (Carter, 1945; Kaplan, 1956); seeds have been dated by accelerator mass spectrometry to at least 2,000 years before present (Kaplan and Lynch, 1999). In addition to drought

* Corresponding author: Daniel Debouck (d.debouck@cgiar.org)

Received: 28.03.2025 | Accepted: 27.08.2025 | Published online: 06.11.2025

(Parsons and Howe, 1984; Barrera *et al*, 2024) and high temperature tolerance (Lin and Markhart III, 1996; Cruz *et al*, 2023), tepary is resistant to several diseases (ashy stem blight: Miklas *et al*, 1998; common bacterial blight: Coyne *et al*, 1963; Bean Golden Mosaic Virus (later shown to be Bean Golden Yellow Mosaic Virus): Miklas and Santiago, 1996; Bean Golden Yellow Mosaic Virus: Porch et al, 2021; Bean Common Mosaic Necrosis Virus: Bornowski *et al*, 2023; rust: Miklas and Stavely, 1998) and pests (bruchids: Shade *et al*, 1987; Jiménez *et al*, 2017; thrip and leafhopper: Porch and Estévez de Jensen, 2024). Further, some accessions of *P. acutifolius* are tolerant to low temperatures (Souter *et al*, 2017) and salinity (Bayuelo-Jiménez *et al*, 2002).

Not surprisingly, there have been many attempts to transfer the useful traits of tepary into the common bean through interspecific hybridization (Pratt and Nabhan, 1988), but with limited success because of the genetic distance between them (Debouck, 1999; Barrera *et al*, 2022). However, the technological context of bean breeding is changing, given the development of the genome sequences for both species, marker assisted selection, and genomic technologies (Schmutz *et al*, 2014; Moghaddham *et al*, 2021; Parker *et al*, 2023; Wang *et al*, 2024). Genome editing through the CRISPR-Cas 9 of 2012 and subsequent improvements are also quickly changing the field of potential pathways for cultivar development (Bandyopadhyay *et al*, 2020; Jha *et al*, 2022; Singh *et al*, 2024). There are a lot of possibilities given the high level of synteny between tepary and common bean (Gujaria-Verma *et al*, 2016; Moghaddham *et al*, 2021).

Transfer of candidate genes from common bean for highly heritable traits such as seed size, seed colour, growth habit, and disease and pest resistance, may prove more expedient than attempting to transfer complex polygenic traits such as heat and/or drought tolerance from tepary into common bean. Incidentally, the approach to breed tepary itself is increasingly being considered in its traditional areas of cultivation (Pratt *et al*, 2023) but also in sub-Saharan Africa (Mwale *et al*, 2020) and in the tropics (Porch *et al*, 2024). So, it might be faster to breed tepary itself with molecular markers developed using the reference genome sequences of both species, based on the high level of synteny between them.

Further, the food processing industry may offer opportunities for bean seed types outside the current market classes (Voysest and Dessert, 1991) and the traditional ways of cooking. These factors are likely contributing to the increasing trend in *P. acutifolius* germplasm requests from the United States Department of Agriculture (USDA) genebank (Dohle, 2024).

The success of these new approaches depends heavily on the availability of genetic variability in tepary bean. Unfortunately, a lot of landraces went extinct, first when the Spaniards introduced new irrigation techniques into Mesoamerica in the 16th century onwards, and second in the 1930s onwards when water pumps with fuel-based engines changed the watering systems in the Southwest and other parts of Aridoamerica (Nabhan and Felger, 1978; Pratt *et al*, 2023). Some cultivated germplasm has been collected from the historic range of tepary cultivation from northern Arizona (Carter, 1945; Nabhan and Felger, 1978) to Guanacaste in Costa Rica (Debouck, 1992), and it is conserved in several genebanks. Once internal duplicates are identified, there are about 100 different landraces of tepary bean in the major germplasm collections for that crop: USDA, the International Center for Tropical Agriculture (CIAT) and INIFAP (Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias of Mexico).

Given this inadequate representation of the intraspecific genetic variation in the genebanks and their obligation to anticipate breeders' needs instead of reacting to them, it is imperative to expand the reservoir of available genetic diversity. The greatest benefit will come from the two wild forms of tepary (often named var. *acutifolius* Asa Gray and var. *tenuifolius* Asa Gray: Delgado-Salinas, 1985). The distribution of these wild forms extends from the southwestern USA (central Arizona, southern New Mexico, and the Trans-Pecos region of Texas) to Michoacán, Mexico (Nabhan and Felger, 1978; Debouck, 2021). While there might be ecological reasons for the recognition of these two varieties (Pratt and Nabhan, 1988), genetic evidence indicates that it is still an open question (Muñoz-Florez *et al*, 2006; Blair *et al*, 2012). A sister species and wild relative of tepary, *P. montanus* Brandegee (synonym *P. parvifolius* Freytag), strikingly similar in appearance to, but separated from var. *tenuifolius* based on biochemical (Florez-Ramos *et al*, 2004) and molecular evidence (Zink and Nagl, 1998a; Muñoz-Florez *et al*, 2006; Blair *et al*, 2012), extends from southeastern Arizona down to Jalapa, Guatemala through several regions of the Pacific side of Mexico (Debouck, 2021). Because of its presence in the Chiricahua Mountains (Debouck, 2019), an additional question is whether *P. montanus* is present in western New Mexico (the Arizona-New Mexico state border line likely not being an ecological barrier, at least up to the continental divide).

Another group of wild beans and with some genetic relationship with common bean is section Rugosi (Zink and Nagl, 1998b; Delgado-Salinas *et al*, 2006). It includes *P. angustissimus* Asa Gray, *P. carterae* Freytag & Debouck and *P. filiformis* Bentham (Freytag and Debouck, 2002; Dohle *et al*, 2019). Gene transfer to common bean from these relatives seems very difficult (Maréchal and Baudoin, 1978). Which traits could be of interest for introgression into tepary or into common bean? Wild teparies and these Rugosi species might be tolerant to drought, salinity and freezing temperatures (Bayuelo-Jiménez *et al*, 2002; Balasubramanian *et al*, 2004). In conclusion, the main purpose of this project is to increase the representation of wild teparies and species of the section Rugosi from New Mexico, where collection to date has been inadequate, in the USDA collection and later that of CIAT. Further, given the rapid progress in comparative pangenome analysis (e.g. Khan *et al*, 2024), it might be good foresight to collect other *Phaseolus* species from New Mexico as the opportunity arises in the field.

At the start of this project in mid-2023 there were two accessions of wild *P. acutifolius* (PI 640990, a collection by Oliver Wendel Norvell of November 1969 and PI 702622, a collection by Richard Pratt of October 2017; see also Figure 1) and two accessions of *P. angustissimus* (PI 535272, a collection by Russ Buhrow and PI 535273, a collection by George Frederick Freytag) from New Mexico in the genebanks of USDA-Pullman and CIAT-Palmira. Accordingly, a joint exploration was carried out in fall 2023. Unfortunately, as it often happens in desert areas, 2023 was a special year with below normal, erratic and late rainfalls in the counties of interest in southern New Mexico. However, several populations of wild teparies yielded some seeds from dried

plants of the previous year, as did return visits to three sites (namely Big Burro Mountains and two sites in the Organ Mountains) with delayed flowering (Debouck *et al*, 2023). Those collections have been successfully increased in the glasshouse of the USDA Pullman during the spring of 2024 (S. Dohle, personal communication, 2024). Finally, the finding of a new species of rhizobium in nodules of *P. filiformis* tolerant to salinity and high temperatures in Baja California (Rocha *et al*, 2020) justifies the continued sampling (initiated in 2023) of rhizosphere microorganisms in wild teparies and Rugosi species to capture effective nitrogen fixation under these abiotic stresses.

Materials and methods

Populations of target species

The team used two primary kinds of information to decide where to sample: the study of herbaria (identified by their acronyms: Thiers, 2023) keeping samples of *Phaseolus* sensu stricto and the sightings posted on iNaturalist (https://www.inaturalist.org/). The former source by personal visits to 86 herbaria collections since 1978 (Debouck, 2021, p. 102-103) and the study of 16 herbaria through the Southwestern Environmental Information Network (SEINet) portal (https://swbiodiversity.org/seinet/collections/list.php/) in April 2023 (Supplemental Material 1) gave a total of 189 populations for the state of New Mexico for six species (*P. acutifolius* (29 populations), *P. angustissimus* (67), *P. filiformis* (9), *P. grayanus* (37), *P. maculatus* (40) and *P. parvulus* (7)) covering a period of field collecting between 1849–2014 (Supplemental Table 1). For the 29 populations of wild tepary, the three coordinates were provided by the collector(s) only for three locations (one by global positioning system (GPS)). The second source of information provided GPS coordinates and a colour picture collected by citizens in the year 2024, and some in 2023. In addition, photos and descriptions of habitat preferences of wild teparies were provided to several area hikers. They subsequently reported to us possible sightings during their hikes. These multiple strategies allowed us to find concordance between prospective collection sites and areas where (more) favourable seasonal rainfall patterns had occurred. On the other hand, germplasm of the type localities from original descriptions was also taken into account, given its importance for future taxonomic and genomic studies. For New Mexico, this representation of type materials in genebanks refers to *P. acutifolius* var. *tenuifolius* ('near the copper mines, New Mexico' in October 1851), *P. angustissimus* ('hill-sides above Doña Ana' in July 1851), *P. grayanus* ('San Luís Mountains' in September 1893) and *P. parvulus* ('Pinos Altos Mountains' in August 1880) (Wooton and Standley, 1915; Freytag and Debouck, 2002).

Timing of collection

In collecting in the northern Chihuahuan Desert and surrounding dry areas (Dick-Peddie, 1993; Cornett, 2013), it was critically important to know how the populations revealed by the two approaches aforementioned had their flowering and pod setting affected by the North American monsoon, in essence variable from year to year (Nolin and Hall-McKim 2006; Reichenbacher and Peachey, 2025). Based on the experience of 2023, monitoring of rain accumulation and distribution was started in July 2024 through late September 2024 for an area covering southwestern New Mexico (roughly south of 33° 15' latitude North and from 106° 27' longitude West) extending beyond the state border with Arizona to include Cochise and Greenlee counties. That area was the one containing most populations of target species and the object of the grants. Rainfall information provided by the US Drought Monitor (https://www.droughtmonitor.unl.edu) (Supplemental Figure 1) was complemented with invaluable field visits and consultations with staff of the US Forest Service (FS) and the Bureau of Land Management (BLM). Such field visits indicated that the Peloncillo, the Florida and the southern Black Range Mountains (almost no rain) would not be fruitful in terms of seed production, and to concentrate exploration instead in 2024 in the Gila region (see Figure 1). For example, information provided by a ranger from the Southwestern Research Station (SRS) near Portal, Arizona, on 23 September 2024 that water was still running in Turkey Creek while not in Cave Creek in the Chiricahua Mountains was key for the team to decide to cross the border into Arizona in search of *P. montanus* (see Discussion).

Implements used during field work were indicated elsewhere (Debouck, 1988; Moss and Guarino, 1995). GPS coordinates were obtained from a Garmin GPSmap 62S receiver and checked against a second GPS receiver (Garmin Etrex 32X) for any significant deviation (that did not happen: see Table 3 in Debouck *et al*, 2025), while primary elevation readings were provided by a barometric altimeter Thommen 3D-16 (0-5000). The application OnX Backcountry (OnX Maps Inc., Missoula, Montana) installed on the smartphone of one participant gave further validation to these direct readings, as well as valuable data such as offline maps, names of places and landmarks, and property ownership. Trying to get the right geographic coordinates was doubly important: first, to have the possibility of getting back to the same population for any additional sampling in the same or subsequent season, and second, to monitor the fate of these populations over time. Many collections by Edward Lee Greene, Henry Hurd Rusby, Elmer Ottis Wooton and Charles Wright in the 1850s–1900s lack accurate data about location, making our ability to monitor changes since these early observations difficult, if not impossible. Further, ensuring the exact coordinates is critically important in relation to future germplasm evaluation for stresses related to location such as drought, extreme temperatures or soil limitations (salinity, nutrient deficiency or excess in minor elements). For managing the time at the collection sites and to improve sampling, communication was critical, and two pairs of radios ('walkie-talkies') were found useful to link members of the team looking for plants in different parts of a single location.

Sampling and data collection

It was a deliberate strategy to sample as many plants as possible at a single site and therefore the team often worked in pairs or individually for half an hour (often longer) in search of additional plants in a population. The reason for this strategy aimed at collecting genetic diversity lies in the cleistogamous reproduction of tepary (Lord and Kohorn, 1986). The sampling of the populations should also be targeted at the conservation of rare alleles. The example of resistance to bruchids in wild common bean (Osborn *et al*, 1986, 1988) – where plants having the right arcelin causing antibiosis were less than 20% in the original

populations – clearly advocates for thorough sampling. The data taken at collection sites follow the guidelines proposed elsewhere (Debouck, 1988; Moss and Guarino, 1995), and were reflected in the labels of the herbarium vouchers (Supplemental Material 2). Vegetation types were described based on the classification of Castetter (1956), Dick-Peddie (1993) and O'Kane (2025). Vernacular names of plants followed Hitchcock (1971a,b) and Dodson's guide (2012). GPS coordinates were checked against the atlas and gazetteer of Arizona (DeLorme, 2008) and New Mexico (DeLorme, 2012) and web-based topographic maps provided by CalTopo – Backcountry Mapping (https://caltopo.com/) and the US Geological Survey (https://www.usgs.gov/programs/national-geospatial-program/topographic-maps). These sources and the comprehensive reference guide of place names in New Mexico produced by Julyan (1998) enable the checking of place names.

Results

General

In 2024, a total of 28 populations were found, including two disclosed in 2023 (Table 1 and detailed information about each population in Supplemental Material 2) for six species (*Phaseolus acutifolius* and its two variants, *P. angustissimus*, *P. filiformis*, *P. grayanus*, *P. maculatus* and *P. montanus*, the latter in southeastern Arizona) during a 9-day exploration. Two populations of wild tepary disclosed in 2023 (#3387, #3390) were revisited in 2024 (successfully) to collect more seed for germplasm conservation; for all other populations found in 2023, prior scouting visits indicated no plant development due to lack of rains (and verified for populations #3392 and #3396). Seeds were collected for conservation for all populations except #3407 (too early), and herbarium specimens were collected for all except #3420 and #3421 (too late) (see Supplemental Table 2).

Table 1. Populations found in chronological order (those highlighted in grey refer to populations found in 2023 for which seeds were sought and found in 2024).

Collection No.	Species	Latitude N	Longitude W	Elevation (masl)	Date
3406	*acutifolius*	32° 17' 48.4"	106° 36' 38.4"	1,565	3-Oct-2024
3407	*acutifolius*	32° 18' 19.0"	106° 35' 32.0"	1,740	3-Oct-2024
3408	*acutifolius*	32° 18' 44.4"	106° 34' 43.2"	1,859	3-Oct-2024
3387	*acutifolius*	32° 22' 08.6"	106° 33' 34.7"	1,750	4-Oct-2024
3409	*acutifolius*	32° 21' 50.8"	106° 33' 55.5"	1,893	4-Oct-2024
3390	*acutifolius*	32° 20' 16.6"	106° 35' 10.7"	1,757	5-Oct-2024
3410	*acutifolius*	32° 17' 33.5"	106° 35' 40.2"	1,639	5-Oct-2024
3411	*filiformis*	32° 17' 34.0"	106° 35' 39.1"	1,647	5-Oct-2024
3412	*acutifolius*	32° 02' 18.8"	106° 57' 23.0"	1,288	6-Oct-2024
3413	*acutifolius*	31° 53' 13.6"	109° 12' 30.6"	1,663	7-Oct-2024
3414	*montanus*	31° 53' 13.6"	109° 12' 30.6"	1,663	7-Oct-2024
3415	*grayanus*	31° 54' 33.2"	109° 15' 09.3"	1,967	7-Oct-2024
3416	*montanus*	31° 54' 33.1"	109° 15' 09.7"	1,964	7-Oct-2024
3417	*montanus*	31° 55' 44.6"	109° 13' 10.4"	1,693	7-Oct-2024
3418	*acutifolius*	31° 55' 44.6"	109° 13' 10.4"	1,693	7-Oct-2024
3419	*maculatus*	32° 39' 08.3"	108° 31' 56.2"	1,781	8-Oct-2024
3420	*acutifolius*	32° 39' 01.3"	108° 31' 56.8"	1,790	8-Oct-2024
3421	*acutifolius*	32° 47' 17.8"	108° 29' 38.2"	1,518	8-Oct-2024
3422	*acutifolius*	32° 51' 04.8"	108° 35' 26.0"	1,327	8-Oct-2024
3423	*acutifolius*	33° 02' 59.0"	108° 30' 03.4"	1,535	9-Oct-2024
3424	*angustissimus*	32° 57' 57.6"	108° 33' 48.2"	1,412	9-Oct-2024
3425	*maculatus*	32° 53' 32.7"	108° 14' 06.4"	1,991	10-Oct-2024
3426	*grayanus*	33° 07' 1.3"	108° 11' 57.5"	2,248	10-Oct-2024
3427	*maculatus*	33° 13' 46.2"	108° 15' 52.3"	1,744	10-Oct-2024
3428	*acutifolius*	33° 13' 35.9"	108° 16' 11.5"	1,795	10-Oct-2024
3429	*acutifolius*	33° 10' 45.0"	108° 12' 19.2"	1,698	10-Oct-2024
3430	*acutifolius*	33° 05' 16.6"	109° 05' 21.8"	1,827	11-Oct-2024
3431	*acutifolius*	32° 57' 04.6"	108° 57' 35.8"	1,918	11-Oct-2024

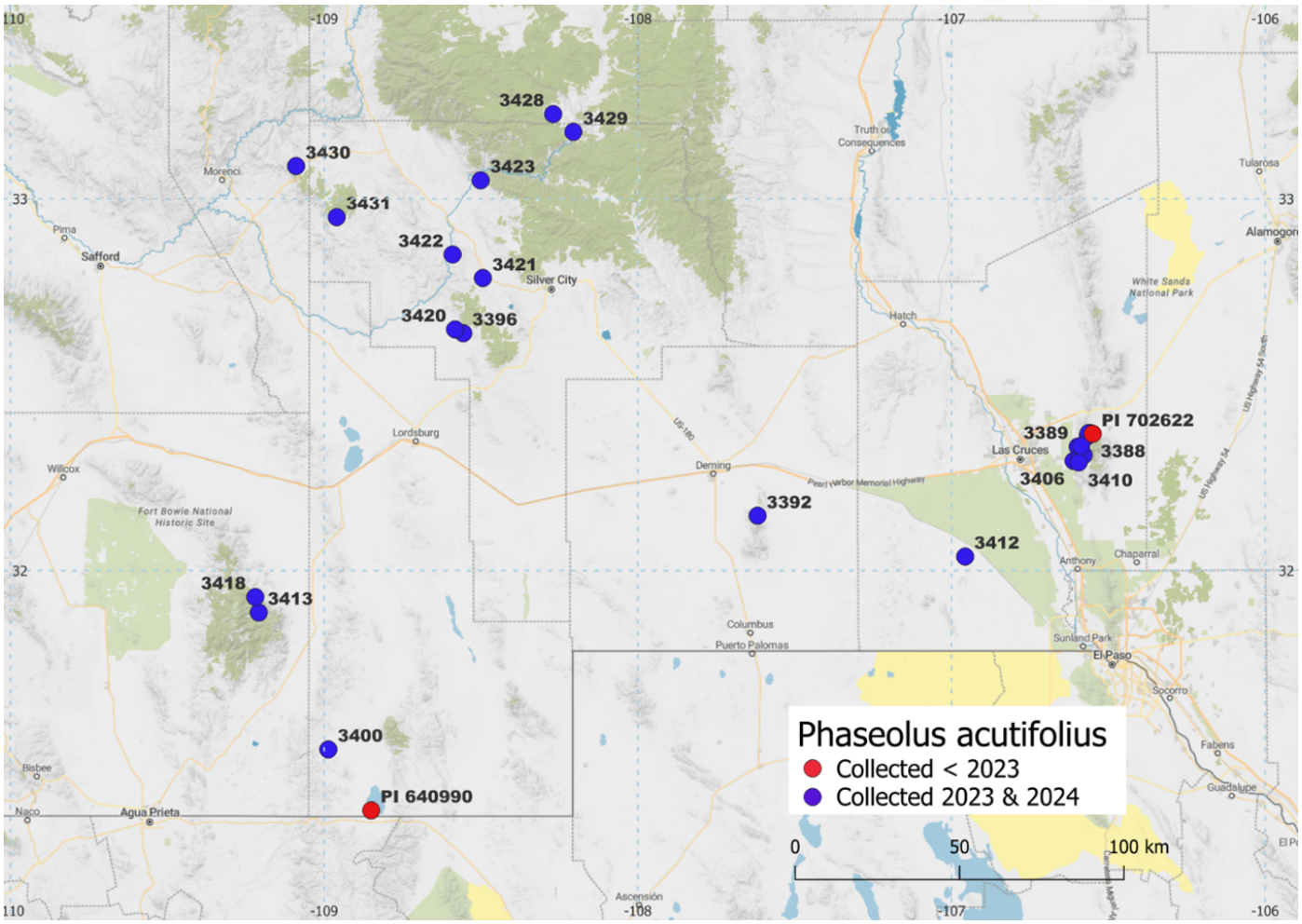


Figure 1. Map showing the number and location of accessions of wild tepary (*Phaseolus acutifolius*) in the USDA collection before (red dots) and after (blue dots) the explorations of 2023 and 2024. More information about the collections made east of Las Cruces, New Mexico, USA, can be found in Supplemental Figure 3.

Seeds were taken to the USDA ARS National Plant Germplasm System greenhouses located in Pullman, Washington, for increase, while 96 herbarium specimens were deposited at the New Mexico State University (NMSU) Biology Herbarium (NMC) for conservation and further distribution. Plants from seeds for all the annual species collected during the fall 2024 plant exploration are currently growing in the greenhouses at USDA Pullman (Supplemental Figure 2). Populations may be accessioned and available for distribution likely from 2026 onwards. As can be seen in Figure 1, the collections of 2023 and 2024 resulted in a significant increase in the USDA collection of wild tepary. Samples of nodules or soil around the rhizosphere of the collected plants were obtained for all populations (except two wild tepary populations #3407 and #3409; Supplemental Table 2). Soil and microbe samples are conserved at the New Mexico State University, Las Cruces, New Mexico, USA.

Per species

Phaseolus acutifolius

In 2024, 18 populations were found, most of them at seed dispersal stage. Some populations were noted with relatively broad leaflets (#3387, #3390, #3406, #3407, #3408, #3409, #3410, #3412), while others displayed very narrow ones (#3413, #3418, #3422, #3423, #3428, #3429, #3430, #3431). Two populations (#3420, #3421) were found with all leaves completely dry, making it impossible to prepare good herbarium specimens. In addition to variation in leaf shape, seed colour and size varied (Figure 2). The smallest 100-seed weight was for #3413 at 1.7g, and the largest was for #3408 and #3412, both at 3.5g. One population (#3428) was found on the slope close to the caves in the Gila Cliff Dwellings National Monument. This seems to be the northernmost latitude for *P. acutifolius* in New Mexico,

Figure 2. Closeups of seeds of *Phaseolus acutifolius* Asa Gray after field harvest. Top row: population #3406 (all scale bars in mm; note the difference in lower row), right #3410. Middle row: left #3422, right # 3428. Lower row: left # 3430, right #3431 (all photographs by SD).

and apparently, upon evidence available to us, the first record for Catron County. An iNaturalist report indicated the presence of wild tepary in the Aden Lava Flow Wilderness Study Area, Doña Ana County, New Mexico; before reaching the given coordinates, population #3412 (Figure 3 top) was identified and sampled for seed and herbarium specimens. This population is of particular interest because of its location being between the Organ and the Florida Mountains and its low elevation (1,210m barometric). For reasons related to proximity with Las Cruces and distance between transects, the distribution on the western slope of the Organ Mountains (Supplemental Figure 3) is better sampled as compared to other areas (Figure 1). Figure 1 also shows the positive return of scouting in August–September, in line with moisture level predicted in the Gila region by the US Drought Monitor (Supplemental Figure 1). Wild teparies were found in open, quite diverse habitats, from desert treeless grasslands (Figure 4 top) to pine juniper woodland (Figure 4 bottom). They were found in chaparral-like habitat (Figure 5 top) or in dry stream beds (Figure 5 bottom). These habitats match with those reported by Allred and Jercinovic (2020) and Alexander (2025), perhaps with the exception of the Ponderosa pine-oak community, because our sampling has not yet been targeted towards higher elevations (not enough rain in the Black Range in 2023 and 2024!).

Figure 3. Wild tepary *Phaseolus acutifolius* Asa Gray. Top: habitat of #3412, Aden Lava Flow Wilderness Area (photo DGD); in absence of cattle grazing, tepary plants (foreground and middle right, arrow) can reach significant development. Lower left: #3430 with narrow leaflets, the lateral ones lobed at base (arrow) (photo SD); lower right: #3406, dry pods as found in most populations at this time; the one to the left with eight seeds (photo SD).

Figure 4. Wild *Phaseolus acutifolius* habitats. Top: habitat of population #3406, a desert grassland at Sierra Vista Tank in Organ Mountains, Doña Ana County, New Mexico; a few stems can be seen on the *Opuntia* at foreground (arrow). The rough topography makes the entrance of cattle difficult, while big rocks reduce the effects of drying winds and help collect a bit of air moisture. Bottom: habitat of population #3423, a pine juniper woodland with scattered oaks in Turkey Creek in Brushy Canyon, Grant County, New Mexico (photos DGD).

Figure 5. Wild *Phaseolus acutifolius* habitats. Top: habitat of population #3387, open low oak chaparral, Aguirre Spring, Organ Mountains, Doña Ana County, New Mexico. Bottom: habitat of population #3396, a dry wash along Red Rock Road in Big Burro Mountains, Grant County, New Mexico. Arrows mark where plants were found in 2023 (photos DGD).

Phaseolus angustissimus

One population (#3424) was found east of Gila, Grant County, New Mexico (there was a previous collection in the area: 4 miles east of Gila, Bear Creek Canyon by *Bassett Maguire 11664*, 23 May 1935, kept at the herbarium of the New York Botanical Garden (NY) with sprawling stems and very narrow leaflets (Figure 6 lower right); several stems were cut to the base of the plant most likely due to grazing by cattle and/or nibbling by deer. The roadside location (Figure 6 top) perhaps saved this population from being completely wiped out by grazing. While this population was found thanks to an iNaturalist report of September 2024, another population of *P. angustissimus* similarly reported from a spot inside Silver City was not found. The habitat of population #3424 matches with the one reported by Jason Alexander (2025).

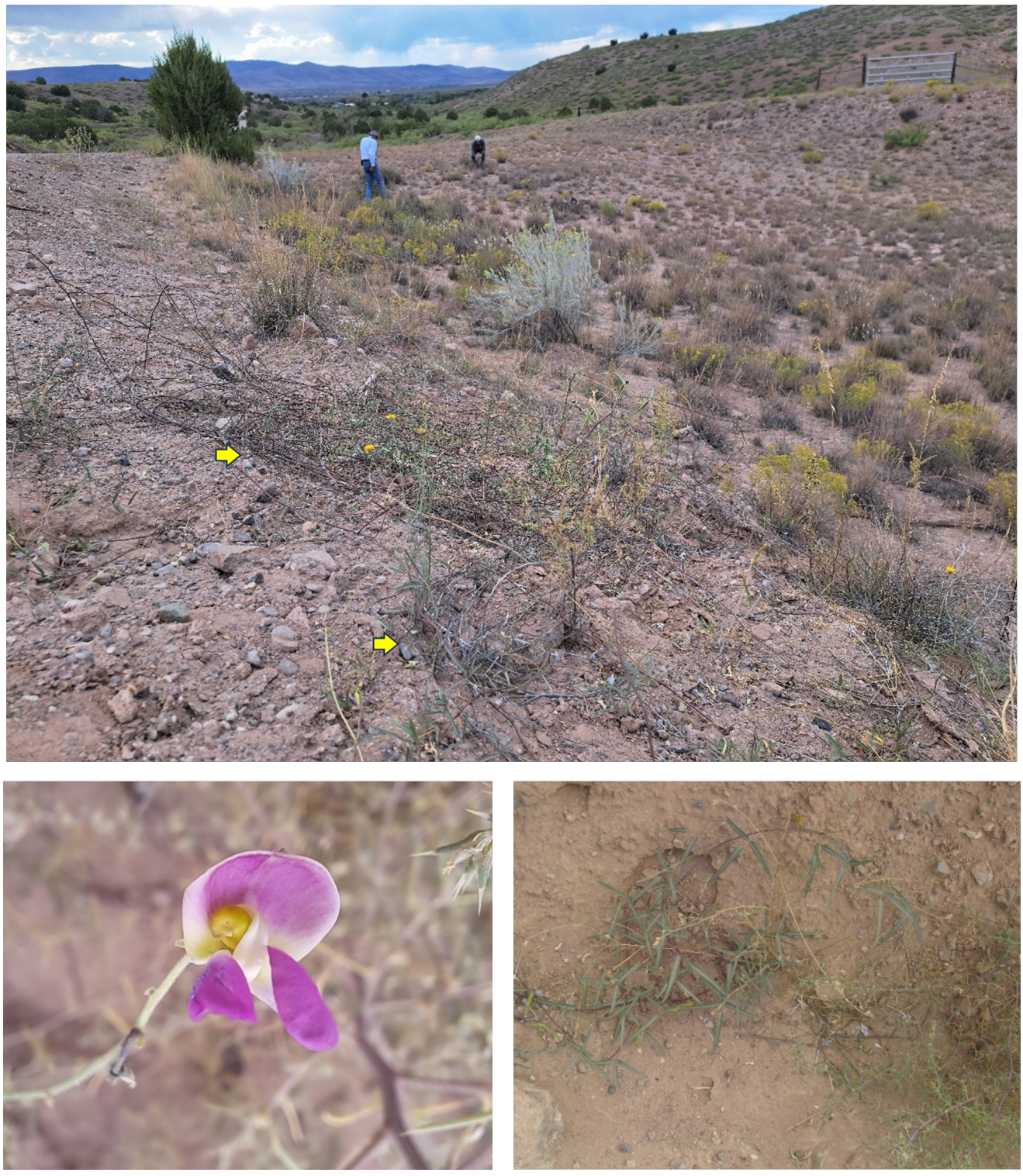


Figure 6. *Phaseolus angustissimus* Asa Gray, population #3424. Top: habitat, a desert scrubland with few scattered junipers and soil almost half bare; the yellow arrows mark where plants thrive (photo SD). Lower left: close-up of a late purple flower (photo SD). Lower right: a plant with sprawling stems in roadside gully (photo DGD).

Phaseolus filiformis

One small population (#3411; Figure 7) was found (very close to a wild tepary #3410) in the central-southern part of the Organ Mountains, Doña Ana County, New Mexico; because of the location (Cuates Canyon, after an unsuccessful search in Achenbach Canyon, same mountainous range), it might be a new record for the Organ Mountains. It was found at seed dispersal stage (Figure 7 lower right). The population #3393 found in 2023 in Rockhound State Park, Luna County, New Mexico, was visited again for seed in 2024, but because of lack of rains not a single plant was seen (nor for its sympatric wild tepary #3392). The two populations found so far (#3393 and #3411) fall within the diversity of dry and open habitats reported by Allred and Jercinovic (2020) and Alexander (2025).

Figure 7. *Phaseolus filiformis* Bentham, population #3411. Top: habitat in Cuates Canyon of Organ Mountains, the arrow marking where the plants thrive. Lower left: a late green branch, where all leaflets are quite parallel to sun rays because of very active pulvini, thus difficult to detect. Lower right: an almost dry stem on *Opuntia* with wrinkled leaflets, twisted pods (arrow) and seeds already dispersed (all photos SD).

Phaseolus grayanus

Two populations (#3415, #3426) were found at green and mature pod stage, often in altitude pine woodland (Figure 8 top). That habitat is one of those reported by Allred and Jercinovic (2020) and Alexander (2025). Many vines were seen without any raceme (Figure 8 lower left), and the low pod productivity may reflect the low amount of rain at these sites in 2024. If flowering is not triggered, more products of photosynthesis will go into the tuberous root (Figure 8 lower right) as the survival strategy of this pluriannual species.

Figure 8. *Phaseolus grayanus* Wooton & Standley. Top: habitat of population #3415, a pine forest in Cochise County, Arizona (photo DGD). Lower left: if left ungrazed, dense mats of sprawling vines can be seen as in population #3426 in pine woodland, Grant County, New Mexico (photo DGD). Lower right: a 4–5-year-old root (20cm long, diameter 15mm) of plant in population #3426 (photo SD).

Phaseolus maculatus

All three populations (#3419, #3425 and #3427) showed stems with completely dried, tan whitish leaflets, perhaps due to lack of rains in September or scarcity of water in the immediate rocky environment. *P. maculatus* is normally a pluriannual prostrate legume of the grasslands of the Chihuahuan Desert (Gentry, 1957). Its abundant biomass of palatable shoots explains its extinction in these flatlands due to cattle grazing, while it survives on steep rocky slopes (Figure 9 top). Roadsides (#3419) and pine-oak (#3425) communities were two of the habitats mentioned by Alexander (2025) and one by Allred and Jercinovic (2020). In contrast to wild teparies or *P. filiformis* (Figure 7 lower right), pods in *P. maculatus* dehisce less abruptly (Figure 9 lower right). Damage due to seed weevils (Coleoptera Brentidae subfamily Apioninae, J. King, personal communication, 2024) was seen in population #3427.

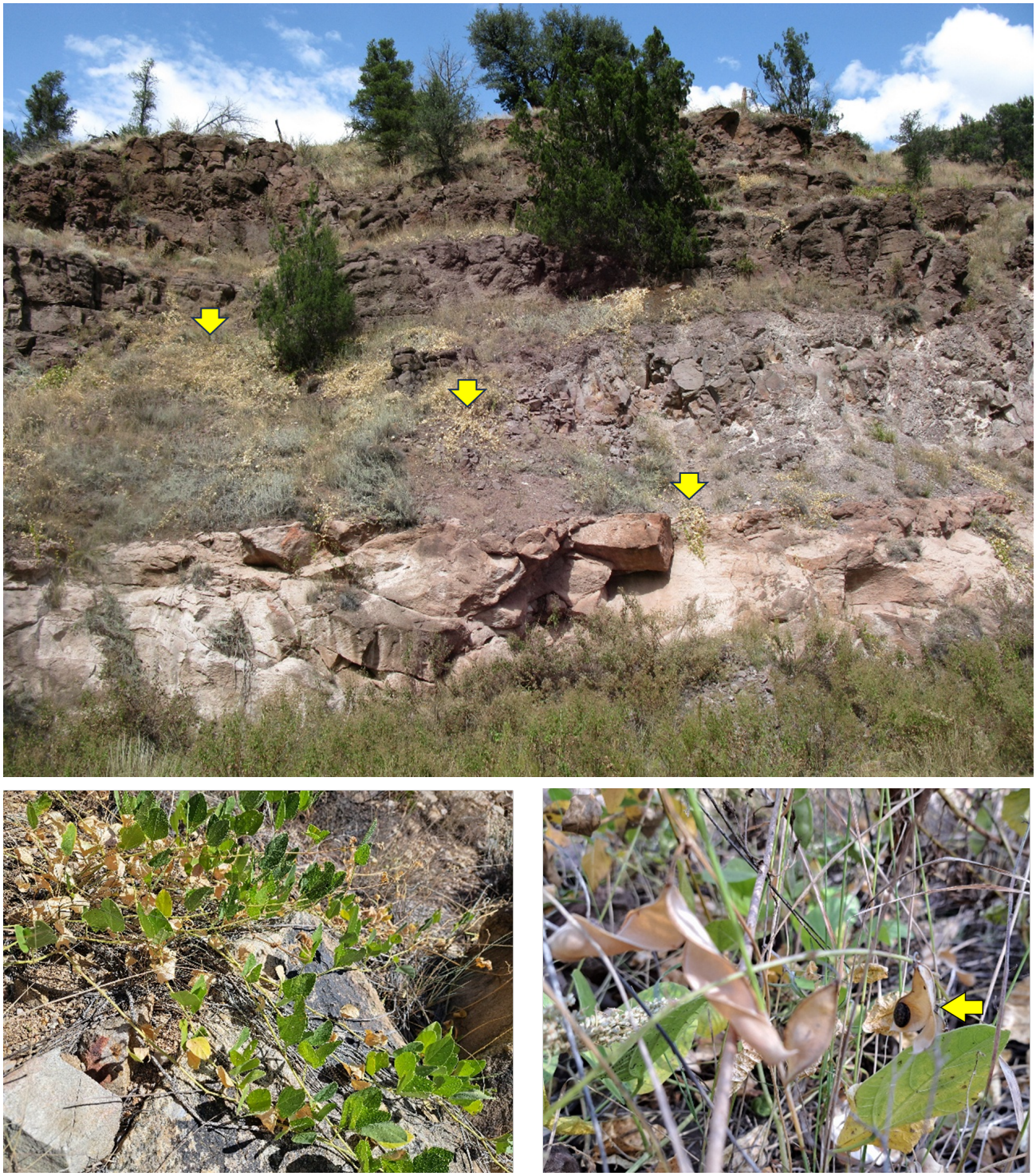


Figure 9. *Phaseolus maculatus* Scheele. Top: habitat of population #3427 at Gila Cliff Dwellings, Catron County, New Mexico, a rocky outcrop above a small riverine plain now converted into a parking lot; arrows mark plants with dried sprawling stems (photo DGD). Lower left: population #3419 on a cliff west of Red Rock, Grant County, New Mexico: note the active pulvini putting the leaflets in an upright position (photo SD). Lower right: population #3427 with pods at maturity with 1–3 globose seeds (arrow) (photo SD).

Phaseolus montanus

Three small populations (#3414, #3416, Figure 10, and #3417) were found in 2024, all in Cochise County in southeastern Arizona, while none were identified in New Mexico. As explained below, the team had to enter into Cochise Co., in the Chiricahua Mountains, to verify the presence of the species, and to investigate the species habitat in order to address the question about its presence in New Mexico. The plants were found at flowering and green pod stages, often intermixed with *P. acutifolius* var. *tenuifolius* with very narrow leaflets. It was thus important to verify some of the discriminant traits, namely in leaflets (Freytag and Debouck, 2002, page 174) and pods (Brandegee, 1893, page 130), as flowers were not present on all plants (Figure 11). Confirming field observations made in Durango in 1978, and in Guerrero and Jalapa both in 1987 (reported by Freytag and Debouck, 2002), the plants do not exceed 60cm in height and have narrow leaflets (Figure 11) with active pulvini that make their identification challenging in the field.

Figure 10. *Phaseolus montanus* Brandegee. Top: habitat of population #3414 (shared with *P. acutifolius* # 3413) (photo DGD). Bottom: leaflets of population #3416 in a shady spot; comparing with Figure 3 lower left, no lobed leaflets are present (photo SD).

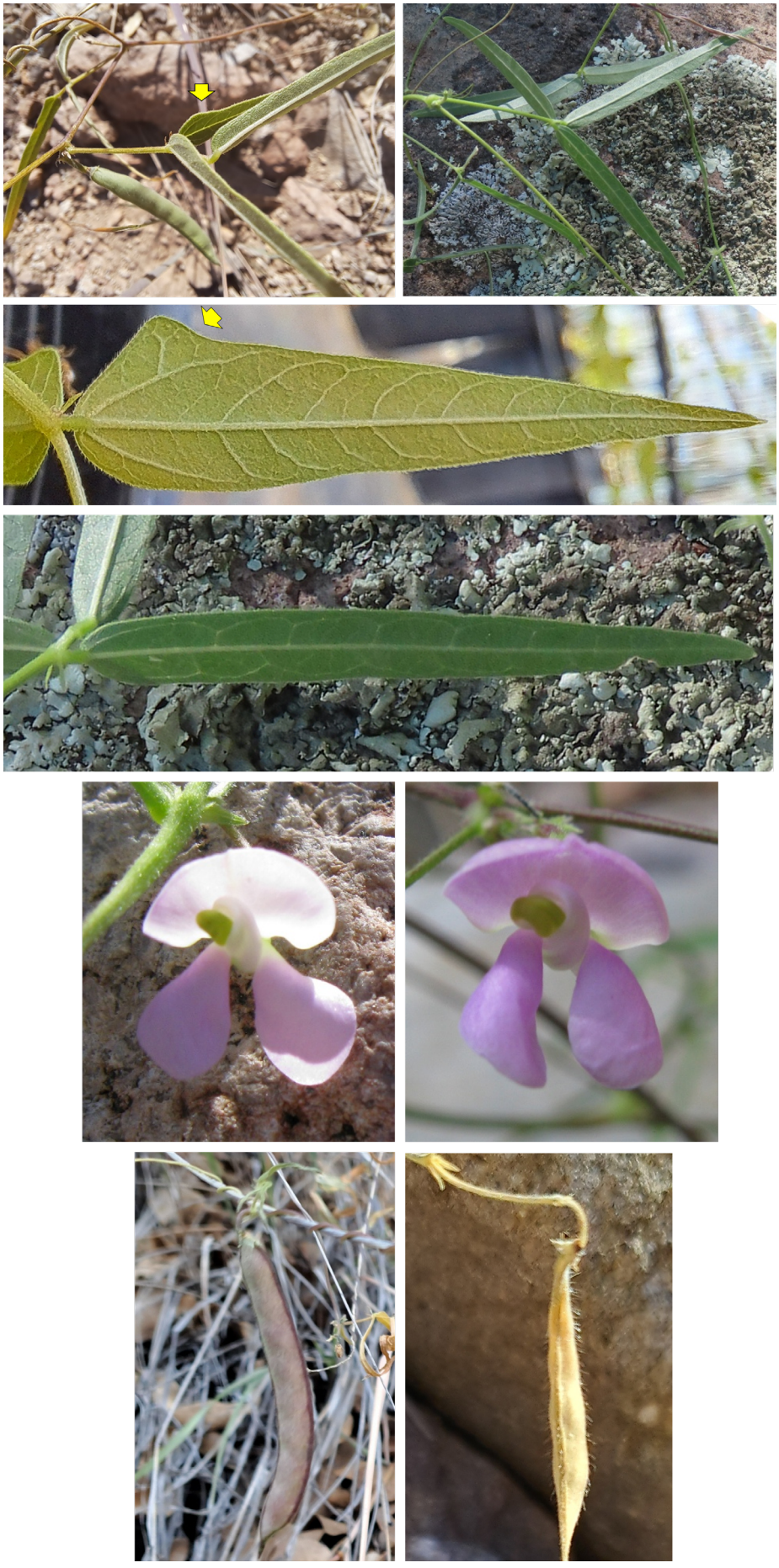


Figure 11. Traits used to help identify *Phaseolus acutifolius* var. *tenuifolius* (left) and *P. montanus* (right). Top row: trifoliolate leaves, #3423 (photo SD) and #3414 (photo LGS). Middle two rows: abaxial faces of lateral leaflets; upper: #3413 (photo SD), with arrow marking the formation of an external lobe, and lower: #3414 (photo LGS). Lower row: flowers, #3390 and #3416, the greenish tip of the keel serving as indication of scale (photos DGD). Bottom row: mature pods before shattering: #3428 left (photo SD) and #3417 right (photo SD) (note four developing seeds against eight in *P. acutifolius*, seven in photograph in upper corner left).

Phaseolus parvulus

One small population (#3395) was found in 2023, in the Pinos Altos Range NE of Silver City, Grant, NM. Given the lack of accuracy in the original species description ('in the Pinos Altos Mountains, New Mexico', Greene, 1881, page 217), it could be considered as falling within the range of the type specimen. Together with a few forbs of Compositae and scattered grasses, it thrives on organic soils in the undergrowth of old Ponderosa pine forest (Figure 12); this matches with the habitat mentioned by Allred and Jercinovic (2020) and one of the three habitats indicated by Alexander (2025).

Figure 12. *Phaseolus parvulus* Greene. Top: habitat of population #3395; bottom: three fully developed plants, the one in the centre with an open pod, scale 5cm long (photos DGD).

Discussion

These results suggest the following points for discussion, namely on purpose, newness, diversity and prospects. **First**, with the results of 2023, all six species (*P. acutifolius, P. angustissimus, P. filiformis, P. grayanus, P. maculatus* and *P. parvulus*) reported for the state of New Mexico (Wooton and Standley, 1915; Freytag and Debouck, 2002; Allred and Jercinovic, 2020; Alexander, 2025) have been found, and some germplasm has been secured for the USDA genebank. Importantly, the team has learned about vegetation type and microhabitats of each taxon to more readily find additional populations in the future. While the collecting priority was on wild teparies, some germplasm was found for the other species, and this is important for the genebanks serving broader interests in different disciplines (e.g. plant taxonomy, ecophysiology, genomics). In this regard, Arizona has the same six species and in addition, *P. montanus* and *P. ritensis* Jones. We concur with Allred and Jercinovic (2020) (page 452) that *P. ritensis* has not been found yet in New Mexico, and the same for *P. montanus*.

This observation links with a **second** point that goes beyond settling a floristic question, given the unique value of *P. montanus* for reciprocal breeding of common and tepary beans (Barrera *et al*, 2022). Finding this taxon was debated between the members of the team during preparation, and explained the brief entry into Arizona. We had one record, a collection by Jacob Corwin Blumer #1676 made in 1907 on Paradise slope in the Chiricahua Mountains in the northeastern extreme of the Madrean archipelago (Figure 1 in Van Devender *et al*, 2013). It was actually a mixed collection of *P. acutifolius* var. *tenuifolius* (specimens studied at the herbaria of ARIZ, F, ISC, MO, NMC and NY[1]) and of *P. montanus* (specimens studied at the herbaria of CAS, K, L and MIN[2]) (Debouck, 2019). Before arriving to the Paradise slope not far from the Southwestern Research Station, we found the two species (#3413 and #3414) almost growing side by side. This close proximity repeated itself eastwards from the locality of Paradise (with #3417 as *P. montanus* and #3418 as *P. acutifolius*), while population #3416 of *P. montanus* was found close to *P. grayanus* #3415. A collection made by Howard Scott Gentry #6472 in Sinaloa, Mexico in 1941 (annotated by one of us in NY; Debouck, 2019) also showed that the two species can be found at the same spot much further south. But the fact that *P. montanus* is found alone in many places from Guerrero, southern Mexico, south to Jalapa, eastern Guatemala (Freytag and Debouck, 2002; Debouck, 2021) would argue against it being a mere morphological segregant of *P. acutifolius* var. *tenuifolius*. Clearly, this close proximity requires further investigation.

A **third** point relates to sampling diversity of wild teparies in southern New Mexico, where a clearer picture starts to appear thanks to our field work. Extremes in elevation are so far: 1,288m for #3412 and 1,918m for #3431, and extremes in latitude: 31° 30' 56.8" for #3400 (from 2023 in the Peloncillo Mountains) and 33° 13' 35.9" for #3428 (from 2024 in the Gila Cliff Dwellings National Monument). The records from herbaria indicated the following range in elevation: the specimen collected by *Elmer Ottis Wooton 528* (kept at NY) was found at 1,370m and the specimen collected by *J Travis Columbus 1588* (kept at NMCR) was found at 1,980m. So, we have put the limit a bit further towards lower elevation (the collection #3412 of Aden Lava Flow Wilderness). Likewise, the study of herbarium specimens indicated a collection (NMC-15801) by *EO Wooton sn* in August 1902 'Mangas Springs; near Silver City' as the northernmost location (approxim. 32° 51'), so it seems that we have pushed the limit a bit further north. But the western slope of the Organ Mountains apart (Supplemental Figure 3), our sampling is still unequal and scanty (Figure 1), not because of lack of records – though variously documented (29 wild tepary populations out of 189 of *Phaseolus* records: Supplemental Table 1), but because of the heavy dependence on intensity and timing of the monsoon rainfall patterns. This uncertainty is common in North American deserts (Beck and Haase, 1969; Larson and Larson, 1997; Nolin and Hall-McKim, 2006; Reichenbacher and Peachey, 2025) and raises the point of how sampling could be improved. One can mention the critical importance of local scouting during August and early September, while the US Drought Monitor gave only a broad picture (Supplemental Figure 1). In this regard, given the lack of meteorological stations in the wilderness of New Mexico, the information provided by rangers of the Bureau of Land Management and of the US Forest Service was extremely valuable, while iNaturalists and informal hikers' observations gave a 50% chance of accurate data about wild *P. acutifolius*, the rest being other legumes such as *Galactia* (for example, for Gallinas canyon NE of San Lorenzo in the Mimbres watershed). The help by iNaturalists can perhaps be made more effective if genebanks put on their websites macrophotos of flowers and pods, cumulating many traits for an accurate identification of target species (Figure 11).

Finally, as noted in the first collection year (2023), grazing in protected areas may be a threat to wild teparies because the soil seed bank might not recover sufficiently to ensure the long-term survival of the populations in the context of the drying Southwest. As a suggestion, from our field work, populations of wild teparies might be selected for the Bureau of Land Management or the US Forest Service to launch a pilot project of *in situ* conservation, more precisely to address many questions related to the soil seed bank.

Concluding, the field work of 2023 and 2024 resulted in a significant increase of representation of wild tepary germplasm in genebanks, adding 22 new accessions from diverse habitats of the Southwest (New Mexico, Arizona). That ecological diversity may announce an important diversity to be disclosed at the genetic level, thus opening up new prospects for breeding of that crop.

Supplemental data

Supplemental Material 1. List of Museums of Natural History and Herbaria where specimens were annotated.

Supplemental Material 2. Detailed information about each population found.

Supplemental Table 1. Numbers of populations by species (verified) and by county of New Mexico.

1 ARIZ, University of Arizona, USA; F, Field Museum of Natural History, USA; ISC, Iowa State University, USA; MO, Missouri Botanical Garden, USA; NMC, New Mexico State University, USA; NY, New York Botanical Garden, USA

2 CAS, California Academy of Sciences, USA; K, Royal Botanic Garden, Kew, UK; L, Naturalis Biodiversity Center, The Netherlands; MIN, University of Minnesota, USA

Supplemental Table 2. Results about the collection of seed, herbarium specimens and nodules/ soil samples of the immediate rhizosphere.

Supplemental Figure 1. Map of New Mexico indicating drought prediction as compared to moisture expected across that state at that date.

Supplemental Figure 2. Photographs of seedlings during the seed increase process at USDA Pullman.

Supplemental Figure 3. Satellite map of the Organ Mountains showing the progress of sampling of wild tepary populations.

Author contributions

This germplasm exploration was originally conceived after a distance workshop during COVID19 lockdown in which participated scientists of the three institutions. For the first (2023) and second (2024) explorations, RP brought the experience and information about suitability of areas from the scouting prior to the field work. He provided many edits and several references about tepary research. SD did the soil sampling in 2024 and collected material of the rhizosphere as well. She added names of several locations of individual collections and provided many high-quality photographs. RP and SD rechecked names of places for all collections reported in Supplemental Material 2. MS and TP participated very actively in the seed collection in 2023 and 2024, respectively. TP provided several insights about tepary breeding and genomics and several references about tepary research. LGS prepared the herbarium specimens in 2023; in 2024 he provided the data about the second GPS recording, photographs and the handling of herbarium specimens. MOU did the soil sampling in 2023 and contributed the rhizosphere samples as well. DGD participated into the identification of populations and collection of data. He did the literature review, wrote the initial draft of the paper and integrated all edits by co-authors. All authors, who were in the field in 2023 and 2024, contributed to the population sampling, read, revised and approved the manuscript.

Acknowledgements

The field work in 2023 was possible thanks to an USDA/ ARS National Plant Germplasm System Plant Exploration Office funding for plant exploration in New Mexico for *Phaseolus* spp. And the field work in Arizona and New Mexico in 2024 was possible thanks to the project 'Plant Exploration in New Mexico to Collect Wild Species of *Phaseolus acutifolius* and *Phaseolus filiformis* Germplasm for Crop Improvement' sponsored by the USDA/ARS award no. 58-2090-4-042. The authors extend special thanks to Dr Anne Frances and the reviewers for considering the proposal, and the Phaseolus Crop Germplasm Committee for their letter of support. Additional funding was provided by the Plant and Environmental Sciences Department of New Mexico State University, the Genetic Resources Program of CIAT, and the Agriculture Research Service of USDA. Permits to collect herbarium specimens, seed for germplasm conservation and samples of microorganisms and soil were kindly and swiftly granted by the Forest Service of USDA, the Bureau of Land Management of the Department of Interior and the Land Trust of the State of New Mexico. The authors express deep gratitude to Britton Bourland (USDA) for help on the maps.

The help and interest of the following persons: Josh Bachman (NMSU), Stephen Beebe (CIAT), Geoff Bender (SRS), Esteban Bolaños (CIAT), Nury Escobar (CIAT), Sara Fuentes Soriano (NMSU), Lois Grant (NMSU-retired), Gabriela Guerrero Florez (WSU), Anowar Islam (NMSU), Joanie King (NMSU), Juan David Libreros (CIAT), Claudia Maldonado (CIAT), Carla Olson (USDA), Erin Riordan (Arizona-Sonora Desert Museum), Zachary Rogers (NMSU), Kirsten Romig (BLM), Fermin Salas (NPS), John Jairo Sánchez (CIAT), Joe Tohme (CIAT), Eliana Urquijo (CIAT), Carlos Urrea (UNL), Marilyn Warburton (USDA) and Peter Wenzl (CIAT) at different steps of these explorations are deeply acknowledged. The authors express gratitude to a regional director of USDA for interest and to the Editor for helping to improve the manuscript.

Conflict of interest statement

The authors are all interested in increasing knowledge about the native bean species of New Mexico, and adding genetic diversity into USDA and CIAT genebanks, specifically of wild teparies. SD is currently the Curator of the USDA Phaseolus collection and responsible for the bean genebank at the Western Plant Introduction Station of USDA, Pullman, Washington. RP is Professor at New Mexico State University and has investigated the agronomy and ecology of tepary and relatives since the 1980s. TP is a bean breeder and researcher of the Agricultural Research Service based at the Tropical Agricultural Research Station in Mayagüez, Puerto Rico, and has launched tepary breeding since the 2000s. MS is the Manager of the genebank of Future Seeds of the Alliance of Bioversity International and CIAT, based in Palmira, Colombia, and oversees all operations for the seed collections of *Phaseolus* beans and tropical forages. LGS is Curator of the bean collection kept in the genebank of Future Seeds in Palmira, Colombia. MOU, now at the International Center for Biosaline Agriculture, worked as the bean physiology leader in CIAT, with interest to develop new technologies to measure tolerance to heat, drought, low phosphorus and high aluminium in crop plants. DGD, CIAT Emeritus, has been responsible for CIAT genebank in 1996–2017; now retired he continues writing and sharing information about crops of neotropical origin.

Ethics statement

The primary objective of this collaborative project involving USDA, NMSU and CIAT (the concerned branch of the Alliance of Bioversity International and CIAT) was to increase the genetic diversity in the USDA germplasm collection. Therefore, it was of utmost importance that the different materials were collected in full knowledge of the authorities and with the appropriate permits, and the role of Sarah Dohle, being Staff member of the Agricultural Research Service of USDA, was key in this regard. Where permits were required on public land, they were obtained as follows:

- New Mexico Bureau of Land Management, reference no. 6850 (9300) Laura Hronec, Acting Deputy State Director Division of Lands and Resources (2023 and 2024)
- New Mexico National Forest, inter agency courtesy provided by Jessie Willett, New Mexico Zone Contracting Officer (2023 and 2024)

- New Mexico State Parks, Research Permit #017 (2023) and #027 (2024) provided by Robert Stokes, Program Support Bureau Chief
- Arizona Apache National Forest, inter agency courtesy permission provided by Trace Douglas Timber Management Officer for Apache-Sitgreaves NFs, Alpine and Springerville Ranger Districts (2024)
- Arizona Coronado National Forest, inter agency courtesy permission provided by Douglas Ruppel, District Ranger Douglas Ranger District (2024).

At the Gila Cliff Dwellings National Monument, an in-person permission was provided by Fermin Salas, Superintendent of Gila Cliff Dwellings National Monument, National Park Service, and the Staff of the National Monument very kindly accompanied the collecting team along the official path within the Monument.

References

Alexander, J. A. (2025). Fabaceae – Pea family. In *Vascular Plants of New Mexico*, ed. K. D. Heil and S. L. O'Kane. Monographs in Systematic Botany, Volume 140. (St. Louis, Missouri, USA: Missouri Botanical Garden Press), 529-598. ISBN: 978-1-935641-30-8.

Allred, K.W., Jercinovic, E.M. (illustr. R. DeWitt Ivey). (2020). Flora Neomexicana III: an illustrated identification manual. Part 2: dicotyledonous plants. 2nd edition. (Durham, North Carolina, USA: Published by the author), 451-455. ISBN: 979-8-651774-81-4.

Assefa, T., Assibi-Mahama, A., Brown, A. V., Cannon, E. K. S., Rubyogo, J. C., Rao, I. M., Blair, M. W., Cannon, S. B. (2019). A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.). *Mol. Breeding* 39, 1–23. https://doi.org/10.1007/s11032-018-0920-0.

Balasubramanian, P., Vandenberg, A., Hucl, P., Gusta, L. (2004). Resistance of *Phaseolus* species to ice crystallization at subzero temperature. *Physiol. Plant*. 120, 3, 451-457. https://doi.org/10.1111/j.0031-9317.2004.00257.x.

Bandyopadhyay, A., Kancharla, N., Javalkote, V. S., Dasgupta, S., Brutnell, T. P. (2020). CRISPR-Cas12a (Cpf1): a versatile tool in the plant genome editing tool box for agricultural advancement. *Front. Plant Sci.* 11, 1-17 (584151). https://doi.org/10.3389/fpls.2020.584151.

Barrera, S., Berny Mier y Teran, J. C., Lobaton, J. D., Escobar, R., Gepts, P., Beebe, S., Urrea, C. A. (2022). Large genomic introgression blocks of *Phaseolus parvifolius* Freytag bean into the common bean enhance the crossability between tepary and common beans. *Plant Direct 6 (12): 1-12 (e470)*. https://doi.org/10.1002/pld3.470.

Barrera, S., Berny Mier y Teran, J.C., Aparicio, J., Diaz, J., León, R., Beebe, S., Urrea, C. A., Gepts, P. (2024). Identification of drought and heat tolerant tepary beans in a multi-environment trial study. *Crop Sci.* 64, 6, 3399-3416. https://doi.org/10.1002/csc2.21354.

Bayuelo-Jiménez, J., Debouck, D. G., Lynch, J. (2002). Salinity tolerance in *Phaseolus* species during early vegetative growth. *Crop Sci.* 42, 6, 2184-2192. https://doi.org/10.2135/cropsci2002.2184.

Beck, W. A., Haase, Y. D. (1969). Historical atlas of New Mexico (Norman, Oklahoma, USA: University of Oklahoma Press), 151p. ISBN 978-0-8061-0817-9.

Blair, M. W., Pantoja, W., Muñoz, L. C. (2012). First use of microsatellite markers in a large collection of cultivated and wild accessions of tepary bean (*Phaseolus acutifolius* A. Gray). *Theor. Appl. Genet*. 125, 6, 1137-1147. https://doi.org/10.1007/s00122-012-1900-0.

Bornowski, N., Hart, J. P., Palacios, A. V., Ogg, B., Brick, M. A., Hamilton, J. P., Beaver, J. S., Buell, C. R., Porch, T. (2023). Genetic variation in a tepary bean (*Phaseolus acutifolius* A. Gray) diversity panel reveals loci associated with biotic stress resistance. *The Plant Genome* 16, e20363. https://doi.org/10.1002/tpg2.20363.

Brandegee, T. S. (1893). Flora of the Cape Region of Baja California. *Proc. Calif. Acad. Sci. ser.* 2, 3, 108-182.

Carter, G. F. (1945). Plant geography and culture history in the American Southwest. *Viking Fund Publ. Anthropol*. 5, 1-140.

Castetter, E. F. (1956). The vegetation of New Mexico. *New Mexico Quaterly* 26, 3, 255-288.

Chen, L., Brun, P., Buri, P., Fatichi, S., Gessler, A., McCarthy, M. J., Pellicciotti, F., Stocker, B., Karger, D. N. (2025). Global increase in the occurrence and impact of multiyear droughts. *Science* 387, 278-284. https://doi.org/10.1126/science.ado4245.

Cornett, J. W. (2013). The Chihuahuan Desert: a brief natural history (Palm Springs, California, USA: Nature Trails Press), 81p. ISBN 978-0-937794-46-3.

Coyne, D. P., Schuster, M. L., Al-Yasiri, S. (1963). Reaction studies of bean species and varieties to common blight and bacterial wilt. *Plant Dis. Reptr*. 47, 6, 534-537.

Cruz, S., Lobatón, J., Urban, M.O., Ariza-Suarez, D., Raatz, B., Aparicio, J., Mosquera, G., Beebe, S. (2023) Interspecific common bean population derived from *Phaseolus acutifolius* using a bridging genotype demonstrate useful adaptation to heat tolerance. *Front. Plant Sci*. 14, 1-15 (1145858). https://doi.org/10.3389/fpls.2023.1145858.

Debouck, D. G. (1988). *Phaseolus* germplasm exploration. In *Genetic resources of* Phaseolus *beans: their maintenance, domestication, evolution and utilization*, ed. P. Gepts (Dordrecht, Holland: Kluwer Academic Publishers), 3-29. ISBN: 90-247-3685-4.

Debouck, D. G. (1992). Frijoles, *Phaseolus spp*. In *Cultivos marginados: otra perspectiva de 1492, ed.* E. Hernández Bermejo and J. León (Rome, Italy: Food and Agriculture Organization of the United Nations), 45-60. ISBN: 92-5-303217-0.

Debouck, D. G. (1999). Diversity in *Phaseolus* species in relation to the common bean. In *Common bean improvement in the twentyfirst century*, ed. S. P. Singh (Dordrecht, The Netherlands: Kluwer Academic Publishers), 25-52. ISBN: 0-7923-5887-2.

Debouck, D. G. (2019). Cahiers de phaséologie – section *Acutifolii* Freytag. International Center for Tropical Agriculture, Cali, Colombia. 128p. Available from: https://ciat.cgiar.org/what-we-do/crop-conservation-and-use/ in program files. (accessed on 13 October 2023).

Debouck, D. G. (2021). *Phaseolus* beans (Leguminosae, Phaseoleae): a checklist and notes on their taxonomy and ecology. *J. Bot. Res. Inst. Texas* 15, 1, 73-111. https://doi.org/10.17348/jbrit.v15.i1.1052.

Debouck, D. G., Dohle, S., Marquez, D., Pratt, R., Santaella, M., Santos, L. G., Urban, M. (2023). Report on a *Phaseolus* germplasm exploration in New Mexico, USA, Sep 27 – Oct 8, 2023. New Mexico State University, Las Cruces, New Mexico, United States Department of Agriculture, Pullman, Washington, USA, and International Center for Tropical

Agriculture (CIAT), Cali, Colombia. Mimeographed. 26p. https://cgspace.cgiar.org/items/7767dfac-6d29-4d29-9597-61d8b234db6b.

Debouck, D. G., Dohle, S., Porch, T., Pratt, R., Santos, L. G. (2025). *Phaseolus* germplasm exploration in New Mexico, USA, Oct 3 – Oct 11, 2024, Report. United States Department of Agriculture, Pullman, Washington, New Mexico State University, Las Cruces, New Mexico, USA, and International Center for Tropical Agriculture (CIAT), Cali, Colombia. Mimeographed. 34p. https://cgspace.cgiar.org/items/6fb6d1aa-b4be-4db1-8885-eea3510872c6.

Delgado-Salinas, A. O. (1985). Systematics of the genus *Phaseolus* (Leguminosae) in North and Central America. Ph.D. Thesis, Univ. of Texas-Austin, Texas, USA. 363p.

Delgado-Salinas, A., Bibler, R., Lavin, M. (2006). Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Syst. Bot.* 31, 4, 779-791. https://doi.org/10.1600/036364406779695960.

DeLorme. (2008). Arizona Atlas and Gazetteer 7th edition (Yarmouth, Maine, USA: DeLorme), 68p.

DeLorme. (2012). New Mexico Atlas and Gazetteer 6th edition (Yarmouth, Maine, USA: DeLorme), 72p.

Dick-Peddie, W. A. (1993). New Mexico vegetation, past, present, and future (Albuquerque, New Mexico, USA: University of New Mexico Press), 244p. ISBN: 0-8263-2164-X.

Dodson, C. (2012). A guide to plants of the northern Chihuahuan Desert (Albuquerque, New Mexico, USA: The University of New Mexico Press), 194p. ISBN: 978-0-8263-5022-0.

Dohle, S. (2024). USDA National Plant Germplasm System, *Phaseolus* Crop Germplasm Committee, Curator report presented on August 20, 2024. https://www.ars-grin.gov/documents/cgc/committee/2024%20Phaseolus%20CGC%20Minutes%20Meeting.pdf.

Dohle, S., Berny Mier y Teran, J. C., Egan, A., Kisha, T., Khoury, C. K. (2019). Chapter 4. Wild beans (*Phaseolus* L.) of North America. In *North American Crop Wild Relatives*. Volume 2, ed. S. L. Greene, K. A. Williams, C. K. Khoury, M. B. Kantar, L. F. Marek (Berne, Switzerland: Springer Nature AG.), 99-127. https://doi.org/10.1007/978-3-319-97121-6_4.

Florez-Ramos, C. P., Ocampo-Nahar, C. H., Toro-Chica, O. (2004). A biochemical trait helps to recognize *Phaseolus parvifolius* Freytag in the gene pool of tepary bean. *Annu. Rep. Bean Improv. Coop. (USA)* 47, 163-164.

Freeman, G. F. (1912). Southwestern beans and teparies. University of Arizona Agricultural Experiment Station. Bulletin 68. Tucson, Arizona, USA. 55p.

Freytag, G. F., Debouck, D. G. (2002). Taxonomy, distribution, and ecology of the genus *Phaseolus* (Leguminosae-Papilionoideae) in North America, Mexico and Central America. *SIDA Bot. Misc.* 23, 1-300. https://doi.org/10568/54291.

Gentry, H. S. (1957). Los pastizales de Durango - Estudio ecológico, fisiográfico y florístico (México, D.F., México: Instituto Mexicano de Recursos Naturales Renovables), 361p.

Greene, E. L. (1881). New plants of New Mexico and Arizona. *Bot. Gaz.* 6, 6, 217-219.

Gujaria-Verma, N., Ramsay, L., Sharpe, A. G., Sanderson, L-A., Debouck, D. G., Tar'an, B., Bett, K. E. (2016). Gene-based SNP discovery in tepary bean (*Phaseolus acutifolius*) and common bean (*P. vulgaris*) for diversity analysis and comparative mapping. *BMC Genomics* 17, 239, 1-16. https://doi.org/10.1186/s12864-016-2499-3.

Hitchcock, A. S. (1950) (1971a). Manual of the grasses of the United States. Revised 2nd edition. USDA Miscel. Public. No. 200. Volume 1 (New York, New York, USA: Dover Publications, Inc.), 1-569. ISBN: 0-486-22717-0.

Hitchcock, A. S. (1971b). Manual of the grasses of the United States. 2nd edition. Volume 2 (New York, New York, USA: Dover Publications, Inc.), 570-1051. ISBN: 0-486-22718-9.

Intergovernmental Panel on Climate Change (IPCC). (2023). Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, H. Lee and J. Romero (eds). Intergovernmental Panel on Climate Change. Geneva, Switzerland. 184p. https://doi.org/10.59327/IPCC/AR6-9789291691647.

Jha, U. C., Nayyar, H., von Wettberg, E. J. B, Naik, Y. D., Thudi, M., Siddique, K. H. M. (2022). Legume pangenome: status and scope for crop improvement. *Plants* 11, 1-14 (3041). https://doi.org/10.3390/plants11223041.

Jiménez, J. C., de la Fuente, M., Ordás, B., García Domínguez, L. E., Malvar, R. A. (2017). Resistance categories to *Acanthoscelides obtectus* (Coleoptera: Bruchidae) in tepary bean (*Phaseolus acutifolius*), new sources of resistance for dry bean (*Phaseolus vulgaris*) breeding. *Crop Protection* 98, 255-266. https://doi.org/10.1016/j.cropro.2017.04.011.

Julyan, R. H. (1998). The place names of New Mexico. Revised edition (Albuquerque, New Mexico, USA: University of New Mexico Press), 385p. ISBN: 0-8263-1689-1.

Kaplan, L. (1956). The cultivated beans of the prehistoric Southwest. *Ann. Mo. Bot. Gard.* 43, 189-251.

Kaplan, L., Lynch, T. (1999). *Phaseolus* (Fabaceae) in archaeology: AMS radiocarbon dates and their significance for pre-Colombian agriculture. *Econ. Bot.* 53, 3, 261-272.

Khan, A. W., Garg, V., Sun, S., Gupta, S., Dudchenko, O., Roorkiwal, M., Chitikineni, A., Bayer, P. E., Shi, C., Upadhyaya, H. D., Bohra, A., Bharadwaj, C., Rouf Mir, R., Baruch, K., Yang, B., Coyne, C. J., Bansal, K. C., Nguyen, H. T., Ronen, G., Lieberman Aiden, E., Veneklaas, E., Siddique, K. H. M., Liu, X., Edwards, D., Varshney, R. K. (2024). *Cicer* super-pangenome provides insights into species evolution and agronomic trait loci for crop improvement in chickpea. *Nature Genetics*, 1-20. https://doi.org/10.1038/s41588-024-01760-4.

Larson, P., Larson, L. (1997). The deserts of the Southwest: a Sierra Club naturalist's guide. Second edition (San Francisco, California, USA: Sierra Club Books), 283p. ISBN: 1-57805-052-9.

Lin, T-Y., Markhart III, A. H. (1996). *Phaseolus acutifolius* A. Gray is more heat tolerant than *P. vulgaris* L. in the absence of water stress. *Crop Sci.* 36, 1, 110-114. https://doi.org/10.2135/cropsci1996.0011183X003600010020x.

Lord, E. M., Kohorn, L. U. (1986). Gynoecial development, pollination, and the path of pollen tube growth in the tepary bean, *Phaseolus acutifolius*. *Amer. J. Bot.* 73, 1, 70-78. https://doi.org/10.1002/j.1537-2197.1986.tb09682.x.

Maréchal, R., Baudoin, J.-P. (1978). Observations sur quelques hybrides dans le genre *Phaseolus*. IV. L'hybride *Phaseolus vulgaris* x *Phaseolus filiformis*. *Bull. Rech. Agron. Gembloux* 13, 3, 233-240.

McKinnon, K. A., Poppick, A., Simpson, I. R. (2021). Hot extremes have become drier in the United States Southwest. *Nature Climate Change* 11, 7, 598-604. https://doi.org/10.1038/s41558-021-01076-9.

Miklas, P. N., Santiago, J. (1996). Reaction of select tepary

bean to Bean Golden Mosaic Virus. *HortScience* 31, 3, 430-432. https://doi.org/10.21273/HORTSCI.31.3.430.

Miklas, P. N., Schwartz, H. F., Salgado, M. O., Nina, R., Beaver, J. S. (1998). Reaction of select tepary bean to ashy stem blight and *Fusarium* wilt. *HortScience* 33, 1, 136-139. https://doi.org/10.21273/HORTSCI.33.1.0136.

Miklas, P. N., Stavely, J. R. (1998). Incomplete dominance of rust resistance in tepary bean. *HortScience* 33, 1, 143-145. https://doi.org/10.21273/HORTSCI.33.1.0143.

Moghaddam, S. M., Oladzad, A., Koh, C., Ramsay, L., Hart, J. P., Mamidi, S., Hoopes, G., Sreedasyam A., Wiersma, A., Zhao, D., Grimwood, J., Hamilton, J. P., Jenkins, J., Vaillancourt, B., Wood, J. C., Schmutz, J., Kagale, S., Porch, T., Bett, K. E., Buell, C. R., McClean, P. E. (2021). The tepary bean genome provides insight into evolution and domestication under heat stress. *Nature Communic*. 12, 2638, 1-14. https://doi.org/10.1038/s41467-021-22858-x.

Moss, H., Guarino, L. (1995). Gathering and recording data in the field. In *Collecting plant genetic diversity – Technical guidelines*, ed. L. Guarino, V. R. Rao and R. Reid (Wallingford, United Kingdom: CAB International), 367-417. ISBN: 0-85198-964-0.

Muñoz-Florez, L. C., Duque, M. C., Debouck, D. G., Blair, M. W. (2006). Taxonomy of tepary bean and wild relatives as determined by amplified fragment length polymorphism (AFLP) markers. *Crop Sci*. 46, 4, 1744-1754. https://doi.org/10.2135/cropsci2005-12-0475.

Mwale, S. E., Shimelis, H., Mafongoya, P., Mashilo, J. (2020). Breeding tepary bean (*Phaseolus acutifolius*) for drought adaptation: A review. *Plant Breeding* 139, 5, 821-833. https://doi.org/10.1111/pbr.12806.

Nabhan, G. P., Felger, R. S. (1978). Teparies in southwestern North America. a biogeographical and ethnohistorical study of *Phaseolus acutifolius*. *Econ. Bot*. 32, 1, 3-19.

Nolin, A. W., Hall-McKim, E. A. (2006). Frequency modes of monsoon precipitation in Arizona and New Mexico. *Monthly Weather Rev*. 134, 12, 3774-3781.

O'Kane, S. L. (2025). Vegetation. In *Vascular Plants of New Mexico*, ed. K. D. Heil and S.L. O'Kane. Monographs in Systematic Botany, Volume 140 (St. Louis, Missouri, USA: Missouri Botanical Garden Press), 5-22. ISBN: 978-1-935641-30-8.

Osborn, T. C., Blake, T., Gepts, P., Bliss, F. A. (1986). Bean arcelin. 2. Genetic variation, inheritance and linkage relationships of a novel seed protein of *Phaseolus vulgaris* L. *Theor. Appl. Genet*. 71, 6, 847-855.

Osborn, T. C., Alexander, D. C., Sun, S. S. M., Cardona, C., Bliss, F. A. (1988). Insecticidal activity and lectin homology of arcelin seed protein. *Science* 240, 207-210. https://doi.org/10.1126/science.240.4849.207.

Parker, T. A, Acosta-Gallegos, J., Beaver, J., Brick, M., Brown, J. K., Cichy, K., Debouck, D. G., Delgado-Salinas, A., Dohle, S., Ernest, E., Estevez de Jensen, C., Gómez, F., Hellier, B., Karasev, A. V., Kelly, J. D., McClean, P., Miklas, P., Myers, J. R., Osorno, J. M., Pasche, J. S., Pastor-Corrales, M. A., Porch, T., Steadman, J. R., Urrea, C., Wallace, L., Diepenbrock, C. H., Gepts, P. (2023). Genetic resources and breeding priorities in *Phaseolus* beans: vulnerability, resilience and future challenges. *Plant Breeding Reviews* 46, 6, 289-417. https://doi.org/10.1002/9781119874157.ch6.

Parsons, L. R., Howe, T. K. (1984). Effects of water stress on the water relations of *Phaseolus vulgaris* and the drought resistant *Phaseolus acutifolius*. *Physiol. Plant*. 60, 2, 197-202.

Porch, T., Beaver, J., Arias, J., Godoy-Lutz, G. (2021). Response of tepary beans to Bean Golden Yellow Mosaic Virus and powdery mildew. *Annu. Rep. Bean Improvem. Coop. (USA)* 64, 73-74.

Porch, T., Estévez de Jensen, C. (2024). Response of the tepary diversity panel to combined Asian bean flower thrip and leafhopper pressure. *Annu. Rep. Bean Improvem. Coop. (USA)* 67, 117-118.

Porch, T. G., Rosas, J. C., Cichy, K., Godoy Lutz, G., Rodríguez, I., Colbert, R. W., Demosthene, G., Hernández, J. C., Winham, D. M., Beaver, J. S. (2024). Release of tepary bean cultivar 'USDA Fortuna' with improved disease and insect resistance, seed size, and culinary quality. *J. Plant Registrations* 18, 1, 42-51. https://doi.org/10.1002/plr2.20322.

Pratt, R. C., Grant, L., Velasco-Cruz, C., Lauriault, L. (2023). Field performance of selected and landrace tepary bean varieties in diverse southwestern USA irrigated production environments. *Legume Sci*. 5, e157, 1-8. http://doi.org/10.1002/leg3.157.

Pratt, R. C., Nabhan, G. P. (1988). Evolution and diversity of *Phaseolus acutifolius* genetic resources. In *Genetic resources of* Phaseolus *beans*, ed. P. Gepts (Dordrecht, Holland: Kluwer Academic Publishers), 409-440.

Reichenbacher, F. W., Peachey, W. D. (2025). Cyclic interannual variation in monsoon onset and rainfall in South Central Arizona, USA. *Climate* 13, 75, 1-21. https://doi.org/10.3390/cli13040075.

Rocha, G., Le Queré, A., Medina, A., Cuéllar, A., Contreras, J-L., Carreño, R., Bustillos, R., Muñoz-Rojas, J., Villegas, M. C., Chaintreuil, C., Dreyfus, B., Munive, J.-A. (2020). Diversity and phenotypic analyses of salt- and heat-tolerant wild bean *Phaseolus filiformis* native of a sand beach in Baja California and description of *Ensifer aridi* sp. nov. *Archiv. Microbiol*. 202, 2, 309-322. https://doi.org/10.1007/s00203-019-01744-7.

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S. M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., Gepts, P., Goodstein, D. M., Gonzales, M., Hellsten, U., Hyten, D. L., Jia, G., Kelly, J. D., Kudrna, D., Lee, R., Richard, M. M. S., Miklas, P. N., Osorno, J. M., Rodrigues, J., Thareau, V., Urrea, C. A., Wang, M., Yu, Y., Zhang, M., Wing, R. A., Cregan, P. B., Rokhsar, D. S., Jackson, S. A. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics* 46, 7, 707-713. https://doi.org/10.1038/ng.3008.

Shade, R. E., Pratt, R. C., Pomeroy, M. A. (1987). Development and mortality of the bean weevil, *Acanthoscelides obtectus* (Coleoptera: Bruchidae), on mature seeds of tepary beans, *Phaseolus acutifolius*, and common beans, *Phaseolus vulgaris*. *Environ. Entomol. 16, 5, 1067-1070*. https://doi.org/10.1093/ee/16.5.1067.

Silber-Coats, N., Elias, E., Fernald, K., Gagliardi, M. (2025). Evaluating alternative crops as a solution to water stress in the U. S. Southwest. *Agric. Water Manag*. 312, 1-15 (109439). https://doi.org/10.1016/j.agwat.2025.109439.

Singh, S., Mishra, V., Riyazzudin, R., Chugh, V., Upadhyay, S. K. (2024). CRISPR/Cas9-mediated genome editing for trait improvement and stress tolerance in Leguminosae (Legume family). *Plant Breeding* 0, 1-19. https://doi.org/10.1111/pbr.13234.

Souter, J. R., Gurusamy, V., Porch, T. G., Bett, K. E. (2017). Successful introgression of abiotic stress tolerance from wild tepary bean to common bean. *Crop Sci*. 57, 3, 1160-1171. https://doi.org/10.2135/cropsci2016.10.0851.

Thiers, B. M. (2023). [continuously updated]. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium, New York, USA. http://sweetgum.nybg.org/science/ih/. Accessed on 2 January 2023 and verified on 19 July 2023.

Van Devender, T. R., Avila-Villegas, S., Emerson, M., Turner, D., Flesch, A. D., Deyo, N. S. (2013). Biodiversity in the Madrean archipelago of Sonora, Mexico. In *Merging science and management in a rapidly changing world: biodiversity and management of the Madrean Archipelago III*, ed. G. J. Gottfried, P. F. Ffolliott, B. S. Gebow, L. G. Eskew and L. C. Collins (Fort Collins, Colorado, USA: USDA Forest Service, Rocky Mountain Research Station Proceedings) 67, 10-16.

Voysest, O., Dessert, M. (1991). Bean cultivars: classes and commercial seed types. In *Common beans: research for crop improvement*, ed. A. van Schoonhoven and O. Voysest (Wallingford, United Kingdom: Commonwealth Agricultural Bureaux International), 119-162. ISBN: 0-85198-679-X.

Wang, Y-W., Wood, J. C., Hamilton, J. P., Mailloux, K., Vaillancourt, B., Estévez de Jensen, C., Porch, T., Buell, C. R. (2024). Genome-enabled breeding across *Phaseolus* species. *Annu. Rep. Bean Improvem. Coop. (USA)* 67, 57-58.

Williams, A. P., E.R. Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles. K., Baek, S. H., Badger, A. M., Livneh, B. (2020). Large contribution from anthropogenic warming to an emerging North American megadrought. *Science* 368, 314-318. https://doi.org/10.1126/science.aaz9600.

Wooton, E. O., Standley, P. C. (1915). Flora of New Mexico. *Contr. US Natl. Herb*. 19, 9-794.

Zink, D., Nagl, W. (1998a). A taxon identified by microsatellite-primed PCR and Southern hybridization in the secondary gene pool of the tepary bean. *Annu. Rep. Bean Improv. Coop. (USA)* 41, 107-108.

Zink, D., Nagl, W. (1998b). Interspecific microsatellite-primed PCR analysis in 20 different *Phaseolus* species. *Annu. Rep. Bean Improv. Coop. (USA)* 41, 109-110.

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 153–159
DOI: 10.46265/genresj.NZSL1498
https://www.genresj.org
ISSN: 2708-3764

Genetic variation of Burgo chicken from Bengkulu, Indonesia, based on the *ND1*-mitochondrial DNA gene

Jarulis[a,*], Aceng Ruyani[b], Nurmeiliasari[c], Ahmat Fakhri Utama[a]

[a]*Magister Program of Biology, Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Bengkulu. Jl. W.R. Supratman, Kandang Limun, Bengkulu 38371, Bengkulu, Indonesia*
[b]*Bachelor Program of Biology, Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Bengkulu, Jl. W.R. Supratman, Kandang Limun, Bengkulu 38371, Bengkulu, Indonesia*
[c]*Department of Animal Science, Faculty of Agriculture, Universitas Bengkulu, Jalan W. R. Supratman, Kandang Limun, Bengkulu, Indonesia*

Abstract: Burgo chicken Burgo chicken is one of the domesticated red jungle chickens found in Bengkulu Province, Indonesia. Taxonomically, the position of Burgo chicken as a subspecies, species or breed remains unclear due to the lack of supporting data, highlighting the need for further taxonomic identification. We identified two specific sites, 52 and 375, representing single nucleotide polymorphisms in the *ND1* gene, with a gene sequence length of 450bp. Three haplotypes were detected in Burgo chickens, with haplotype 2 shared between Burgo chicken, *Gallus gallus* (Java) and *G. gallus* bankiva. The average genetic distance in the Burgo chicken population was 0.1%. When compared to other chicken populations, the average distance was 0.12%, while the distance to other *Gallus* spp. was 3.62%. All Burgo chickens formed the same clade in the phylogenetic tree, although two individuals (C2F3ND1 and K4F2ND1) showed slight differences. These two individuals were found in Rejang Lebong and Kepahiang, two nearby locations, indicating the possibility that a meeting occurred. Genetic differences within Burgo chickens from Bengkulu, and with other chickens in Indonesia and various parts of the world, were present but not significant. Our data show that Burgo chickens may exhibit differences from other chickens in Indonesia and globally. However, although the genetic data revealed some divergence in mitochondrial DNA, additional morphological and morphometric analyses are needed to provide supporting evidence.

Keywords: Burgo chickens, conservation, domestication, genetics, taxonomy

Citation: Jarulis, Ruyani, A., Nurmeiliasari and Utama, A. F. (2025) "Genetic variation of Burgo chicken from Bengkulu, Indonesia, based on the ND1-mitochondrial DNA gene", *Genetic Resources*, 6(12), pp. 153–159. doi: 10.46265/genresj.NZSL1498.

Introduction

The domestication of wild animals is part of the journey of human civilization. One of the most commonly domesticated animals is the chicken. Chickens are bred for egg and meat production. The red partridge is the first chicken that was successfully domesticated in Southeast Asia and Southwest China (Fumihito *et al*, 1994; Väisänen *et al*, (2005); Liu *et al*, 2006; Miao *et al*, 2013). Studies indicate that the domestication process of red partridges in Asia began around 3,000 years ago, leading to the species now known as the domestic chicken (*Gallus gallus domesticus*). Domestication of the red partridges (*Gallus* spp) in East Asia occurred in the mid-late Holocene. (Miao *et al*, 2013; Larson *et al*, 2014). Domestication has influenced changes in the behaviour, physiology and productivity of chickens; however, some similarities persist between domestic chickens and their ancestors, such as aggressive behaviour during mating and urinary protein excretion, which remain consistent with that of their wild counterparts. (Al-Nasser *et al*, 2007). Meanwhile, local chickens found in Indonesia have continued to develop since this successful domestication process.

This situation has led to Indonesian chickens forming a different genetic clade from other chickens in Asia. Therefore, Indonesia is considered one of the centres of chicken domestication in Asia (Sulandari *et al*, 2007). In

* Corresponding author: Jarulis (jarulis@unib.ac.id)

Received: 23.01.2025 | Accepted: 01.10.2025 | Published online: 13.11.2025

Indonesia, there are red partridges (*G. gallus bankiva* and *G. gallus spadiceus*) and green jungle fowl (*G. varius*) with a total of 31 strains spread across the regions of Sumatra, Java, Bali and Nusa Tenggara. (Sibley and Monroe, 1990; Nataannjaya, 2000). One of the local chicken breeds found in Indonesia is the Burgo chicken. Burgo chickens are fertile and can produce a high number of offspring, as well as five times more eggs than the red partridges, averaging 32 eggs per period (Sutriyono, 2016). In addition, Burgo chickens have a distinctive crowing sound and beautiful feather colours, which encourage people to raise them as ornamental animals and livestock. The Burgo chicken population is found in all districts of Bengkulu Province, Sumatra Island (Putranto *et al*, 2017). However, there has been no research into their genetic relationship and characteristics, so it remains unclear whether it is the result of inherited genetics or the impact of environmental factors. As a source of germplasm, Burgo chickens are threatened by various anthropogenic factors, including habitat fragmentation, which causes isolation in these species, further threatening their populations. Moreover, as one of the local chicken clades, the taxonomic position of Burgo chickens remains unknown. Taxonomic determination is generally based on morphological and genetic characteristics. Studies related to the morphology of Burgo chickens have been conducted previously (Rafian *et al*, 2017; Safitra *et al*, 2022). Mitochondrial DNA (mtDNA) has been widely used to analyze genetic variation between populations and species due to the high number of DNA copies, making it suitable for analysis with a limited amount of DNA or easily degraded DNA (Ni'mah *et al*, 2016). One of the mtDNA genes used for species identification is NADH Dehydrogenase subunit 1 (*ND1*) (Amin and Mushlih, 2020). The *ND1* gene is part of complex I, also known as NADH. Ubiquinone oxidoreductase is the first and largest enzyme complex in the mitochondrial respiratory chain, playing a role in oxidizing NADH to release electrons that assist in the translocation of protons to the inner membrane, producing proton gradients (Hirst, 2010).

The genetic diversity of the *Gallus* genus, based on the mitochondrial DNA COI gene, shows a genetic similarity of 98% between red partridges from Bengkulu and South Sumatra (Jarulis *et al*, 2022). Several previous studies have utilized the *ND1* gene. For instance, Bowles and Mcmanus (1993) revealed inter- and intraspecies variations in *Echinococcus* from 59 isolates; Raharjo *et al* (2018) detected rat meat contamination in meatballs using the *ND1* gene; and Widayanti *et al* (2022) successfully identified mutations at three sites within the 972-nucleotide sequence of the *ND1* gene of Indonesian catfish. Therefore, we investigated the potential of the *ND1* gene to determine the level of genetic similarity among Burgo chicken populations, other chickens in Indonesia, and other *Gallus* species. No comparative genetic study of Burgo chickens, particularly based on the mitochondrial DNA ND1 gene, has ever been conducted. Therefore, this research is essential to provide data on the genetic diversity and variation among Burgo chicken populations and between species of the *Gallus* genus in Indonesia. The findings will support the Bengkulu Provincial Government's efforts to identify and designate the Bengkulu Burgo chicken cluster for submission to the central government, as part of future conservation initiatives aimed at preserving the population's genetic diversity.

Materials and methods

Blood collection

Blood samples were collected from 28 Burgo roosters owned by members of the Bengkulu Burgo chicken hobbyists. There three locations where the Burgo chicken samples were taken are Bengkulu city, Kepahiang, and Rejang Lebong. Blood samples were drawn through the carpal joints and pectoralis veins. Preserved using EDTA tube according to Seutin *et al* (1991) and stored in a freezer at -20°C, before use. All blood samples were analyzed in the Molecular Biology Laboratory, Department of Biology, Universitas Bengkulu.

DNA extraction and purification

The blood samples (10-20μl) were preserved in EDTA tubes. The DNA was isolated using the Dneasy® Blood and Tissue Kit Cat. No. 69504 (50), following the Spin-Column Protocol Qiagen procedure with modification. In our research, the elution solution used was 50μl with three repetitions. The isolated DNA was observed on 1.2% agarose gel using electrophoresis and stored in a freezer at -20°C, before the amplification process.

Polymerase chain reaction (PCR) DNA

The *ND1* gene of Burgo chickens was replicated using a PCR technique with a DNA template derived from the total DNA product. The *ND1* gene sequence used to design the specific primer in this study was obtained from the complete genome of mitochondrial DNA from *G. gallus* from Kalimantan (GenBank accession number KY039421). *ND1*The primers were BR*ND1*F (5'CCCACCCTAACAAACCTTCTAATC-3') and BR*ND1*R (5'TAGGGTGACTTCGTAT GAGAT TGT-3'), which amplified a 450bp fragment of the 974 bp *ND1* sequence. All reaction mixtures followed the existing protocol Gotaq green. The reaction mixture contained 25μl Gotaq Green, 1.5 μL forward primer, 1.5μl reverse primer, 3μl DNA template, and 19μl nuclease-free water. PCR amplification was performed using a SimpliAmp Thermal Cycler with the following programme: denaturation at 94°C (1 minute), annealing at 55°C (45 seconds) and elongation at 72°C (1 minute) for 30 cycles. Furthermore, the successful amplification samples were sent to PT. Genetika Sains for sequencing.

Data analysis

The *BIOEDIT 7.0.9* software (Hall, 1999) was applied to edit the *ND1* gene sequence and visualize the electrograms and nucleotide base sequences.The nucleotide sequence (forward and reverse) products were aligned using *Clustal W* of the *MEGA 11.0* programme (Tamura *et al*, 2013). Each individual's gene sequence was compared with the *ND1* reference to determine the similarity level of the samples. The genetic distance between individuals was calculated using the 2-parameter Kimura (K2P) method (Kimura, 1980). The phylogeny tree was constructed using the neighbour-joining (NJ) method with 1,000 replications (Tamura *et al*, 2013). Additional *ND1 G. gallus* gene sequences found in GenBank were downloaded and included in the phylogenetic tree reconstruction analysis (see Table 1). Genetic diversity parameters, namely haplotype (Hd) and nucleotide (π) diversity were calculated using *DNASp v6.12.03* software

(Rozas *et al*, 2017). The haplotype analysis was presented in a sequence location distribution map/operational taxonomic unit (OTU) and haplotype network images to depict the latest connectivity and genetic distribution between populations using model median-joining by *Network v10.2.0.0* software (Bandelt *et al*, 1999).

Results

Single nucleotide polymorphism

The nucleotide sequence of the *ND1* gene observed in 450bp between Burgo chicken species from Bengkulu had two nucleotide polymorphisms (SNP) that differed among individuals at positions 52 and 375 (Table 1). Site 52 showed a transversion substitution in Burgo chicken individuals from Rejang Lebong Regency and Kepahiang Regency, namely from cytosine (C) to adenine (A), while a transition substitution at site 375 was found among Burgo chicken individuals from Central Bengkulu, namely from the nucleotide base adenine (A) to guanine (G).

Table 1. SNP between individuals of Burgo chickens from Bengkulu based on the *ND1* gene (450bp). Sample code indicates accession numbers of sequences sourced from GenBank. Dots (.) indicate identical nucleotide to the reference sequence for *ND1* (KY039420.1). A, adenine; C, cytosine; G, guanine.

No.	Sample code	Location/Source	Local name	Site number		Haplotype group
				52	375	
1	KY039420.1	GenBank	red junglefowl	C	A	Hap 2
2	KY039418.1	GenBank	red junglefowl	.	.	Hap 2
3	KY039422.1	GenBank	red junglefowl	.	.	Hap 4
4	KY039421.1	GenBank	Red junglefowl	.	.	Hap 5
5	AP003323.1	GenBank	Bankiva	.	.	Hap 2
6	NC007238.1	GenBank	Green junglefowl	.	.	Hap 6
7	NC007240.1	GenBank	Grey junglefowl	.	.	Hap 7
8	NC007239.1	GenBank	Ceylon junglefowl	.	.	Hap 8
9	BR1	Central Bengkulu	Burgo	.	G	Hap 1
10	BR2	Central Bengkulu	Burgo	.	.	Hap 2
11	BR3	Central Bengkulu	Burgo	.	.	Hap 2
12	BR4	Central Bengkulu	Burgo	.	.	Hap 2
13	BR5	Central Bengkulu	Burgo	.	G	Hap 1
14	BR6	Central Bengkulu	Burgo	.	.	Hap 2
15	BR7	Central Bengkulu	Burgo	.	.	Hap 2
16	BR8	Central Bengkulu	Burgo	.	.	Hap 2
17	BR9	Central Bengkulu	Burgo	.	.	Hap 2
18	BR10	Central Bengkulu	Burgo	.	.	Hap 2
19	BR11	Central Bengkulu	Burgo	.	.	Hap 2
20	BR12	Central Bengkulu	Burgo	.	.	Hap 2
21	BR13	Central Bengkulu	Burgo	.	.	Hap 2
22	BR14	Central Bengkulu	Burgo	.	G	Hap 1
23	BR15	Central Bengkulu	Burgo	.	.	Hap 2
24	C1F2	Rejang Lebong	Burgo	.	.	Hap 2
25	C2F3	Rejang Lebong	Burgo	A	.	Hap 3
26	C4F2	Rejang Lebong	Burgo	.	.	Hap 2
27	C5F2	Rejang Lebong	Burgo	.	.	Hap 2
28	K1F1	Kepahiang	Burgo	.	.	Hap 2
29	K2F1	Kepahiang	Burgo	.	.	Hap 2
30	K3F2	Kepahiang	Burgo	.	.	Hap 2
31	K4F2	Kepahiang	Burgo	A	.	Hap 3
32	K5F2	Kepahiang	Burgo	.	.	Hap 2
33	K10F2	Kepahiang	Burgo	.	.	Hap 2
34	K11F2	Kepahiang	Burgo	.	.	Hap 2
35	K12F2	Kepahiang	Burgo	.	.	Hap 2
36	K13F3	Kepahiang	Burgo	.	.	Hap 2

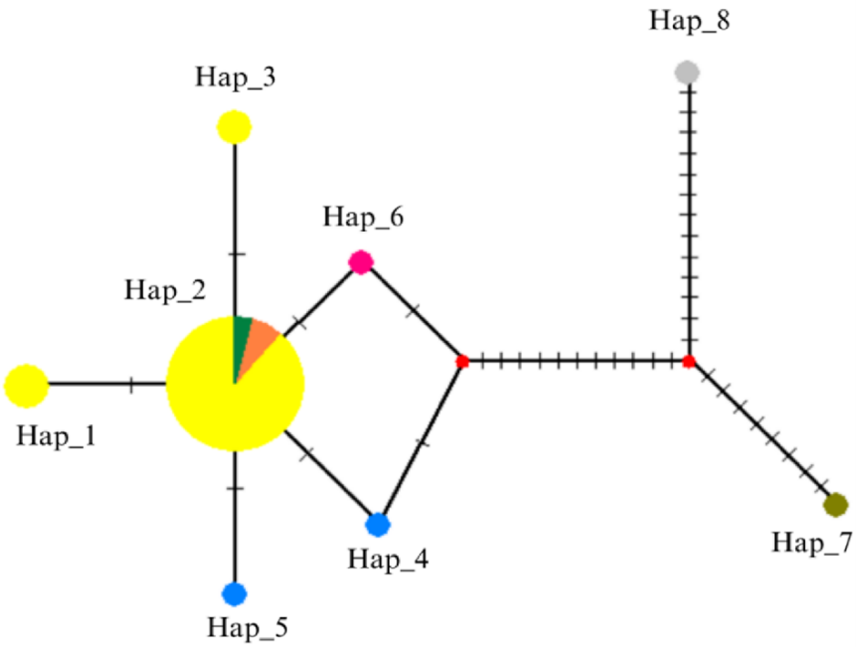


Figure 1. The haplotype network of the *Gallus* spp. population, based on the *ND1* gene alignment. Yellow, Burgo chicken (this study); blue, *G. gallus* from Kalimantan (KY039422.1; KY039421.1); green, *G. gallus bankiva* (AP003323.1); orange, *G. gallus* from Java (KY039420.1; KY039418.1); red, *G. varius* (NC007238.1); grey, *G. sonneratii* (NC007240.1); army green *G. lafeyetii* (NC007239.1).

Haplotype network

In this study, network reconstruction was performed using median-joining (Bandelt *et al*, 1999). Twenty-eight samples of Burgo chickens were complemented with additional genetic data of eight samples of the *Gallus* genus from Genbank, including *G. gallus* from Kendu (KY039420.1), *G. gallus* from Garut (KY039418.1), *G. gallus* from Nunukan (KY039422.1), *G. gallus* from Tarakan (KY039421.1), *G. gallus bankiva* (AP003323.1), *G. varius* (NC007238.1), *G. sonneratii* (NC007240.1), and *G. lafayetii* (NC007239.1). We succeeded in identifying eight haplotypes with a sequence length of 450bp. In Burgo chicken samples, three haplotypes were found: hap 1, hap 2 and hap 3 (Figure 1). Each haplotype is separated by a single nucleotide base, represented by a small horizontal line connecting the haplotypes.

Genetic distance

Genetic distances were analyzed using pairwise distances with the *MEGA 11* software (Table 2). In general, genetic distance is divided into three groups, namely genetic distance between individuals (intraspecific), genetic distance between *G. gallus* species, and genetic distance between species of the *Gallus* genus (interspecific). In this study, a slight change was observed in the interspecific genetic distance compared to all Burgo chicken samples, incorporating genetic data from red partridges and subspecies from GenBank. Meanwhile, the outgroup comparison involved all red partridges and their offspring with all other partridges. The intraspecific genetic distance among three districts in Bengkulu Province, based on the *ND1* gene, showed the lowest value of 0%, while the highest distance was 0.4%.

Table 2. Intra- and interspecific genetic distance in Burgo chickens based on the *ND1* gene (450bp)

Genetic Distance	Maximum	Minimum	Average
Intrapopulation of Burgo chicken	0.4%	0.0%	0.1%
Burgo chicken versus other *Gallus gallus*	0.4%	0.0%	0.12%
Interspecies of *Gallus* spp.	5.9%	0.2%	3.62%

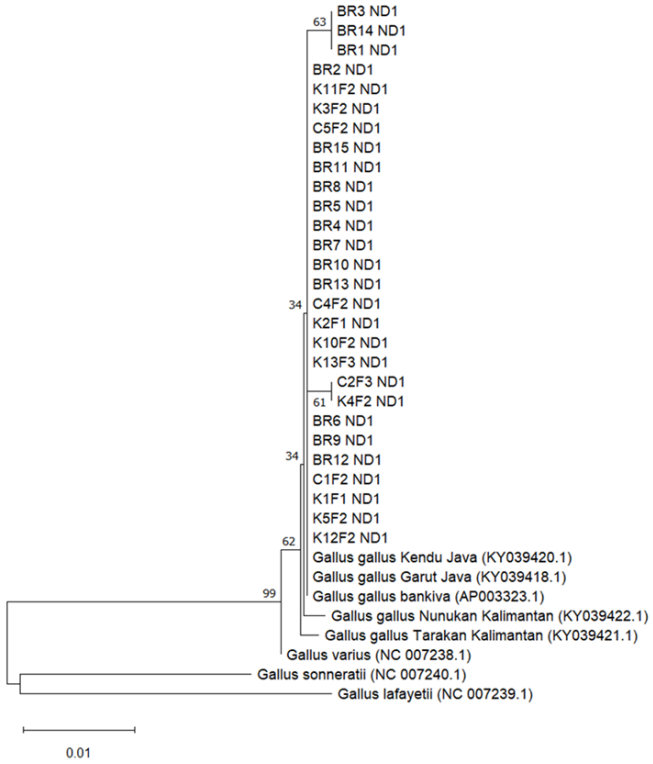
Phylogeny

Phylogenetic trees were reconstructed using the *ND1* gene to determine the taxonomic position of Burgo chickens in comparison with the available reference data. The phylogeny of the Burgo chicken sample was grouped into one clade with red partridge from Java, Kalimantan, and the subspecies of *G. gallus bankiva* (Figure 2). However, five samples formed a small group, namely BR3, BR14 and BR1 from Central Bengkulu, and C2F3 and K4F2 from Rejang Lebong and Kepahiang. This is due to the discovery of mutations in the sequence that caused a slight change; however, the shape of the phylogenetic tree remains stable.

Discussion

Human intervention drives the continuous domestication of chickens. Domestication has led to the development of many chicken breeds worldwide. Indonesia has 31 local chicken breeds that have adapted over tens to hundreds of years. Each local chicken breed has characteristics influenced by its specific region (Nataamijaya, 2010). These characteristics are intrinsically linked to their genetic foundation, such as SNPs. SNPs are often used to interpret variations and identify species or individuals (Torres, 2016). In our study, there were two differentiation sites, namely sites 52 and 375. At site 52, a change was present in nucleotide bases from C to A (C2F3 and K4F2), while a change from A to G was found at site 375 (BR1, BR5, and BR14). These changes are caused by mutations. According to Warmadewi *et al* (2020), mutations can enhance adaptability by eliminating original traits. Sometimes, the treatment such as maintaining high stocking density and implementing accelerated growth diets of domestic chickens has negative impacts such as health problems, brittle bones, and even sudden death (Hirsch, 2003; Meseret, 2016). However, it is not yet known for certain whether the changes that occur in Burgo chickens have a positive or negative impact on their ability to adapt, so more in-depth research is needed regarding the morphometry and morphology of Burgo chickens. The C2F3 and K4F2 samples were Burgo chickens obtained from Kepahiang and Rejang Lebong Districts, while the BR1, BR5 and BR14 samples were Burgo chickens from Central Bengkulu District. The landscapes in each location differ: Kepahiang and Rejang Lebong are highland areas, whereas Central Bengkulu is a lowland area, leading to different adaptation processes.

Figure 2. Phylogenetic tree construction with neighbour-joining (NJ) modelling of 28 Burgo chickens from Bengkulu using a K2P model and 1,000-time bootstrap, based on the *ND1* gene (450bp).



Based on the SNP data, Burgo chickens are grouped into three haplotypes according to their sequence similarity, namely hap 1, hap 2 and hap 3. The 450bp alignment of the *ND1* gene yielded eight haplotypes of the entire sequence (Figure 1). Similar genetic data is present from several species, including Burgo chicken, *G. gallus* from Java, and *G. gallus bankiva*. This is interesting because several Burgo chickens share the same genetic components as *G. gallus* (Java) and *G. gallus bankiva*; however, the results of the haplotype analysis may be influenced by the number of samples and population. Research by Wang *et al* (2020) using 863 native and domestic chicken genomes showed that crossbreeding occurred among red partridge subspecies. Therefore, it is possible that all three originated from the same ancestor. These data are strengthened by previous studies that revealed the origin of red partridge as the ancestors of local chickens worldwide. Sulandari *et al* (2008) found 69 haplotypes in the genetic characterization of local Indonesian chickens and local chickens outside Indonesia using D-loop, besides discovering the Indonesian chicken genes in other countries. Although our data only used three breed populations in Bengkulu Province, they revealed a direct relationship between Burgo chickens, red partridge, and their descendants, as indicated by haplotype 2.

Genetic distance is one of the tools used for species identification, alongside morphological and morphometric data. Lately, bird research has been relying on genetic data to facilitate the identification process. Each species has a threshold value for genetic distance; if the genetic distance is equal to or greater than 3%, species separation occurs (Fouquet *et al*, 2007). Based on genetic distance identification, the distance between Burgo chickens and red partridges (intersubspecific) is 0.1–0.4%. Meanwhile, the distance among Burgo chickens (intraspecific) ranges from 0 to 0.4%, indicating a close relationship at the species level. This suggests that they likely originate from a closely related or similar population. Therefore, Burgo chicken can be identified as a new breed of red partridge. However, further study on morphometry as well as sound identification is required to ensure this theory. Similarly, Utama *et al*, (2023) obtained the genetic distance between Burgo chickens and red partridges as 0–0.8% using the COI gene. In addition, Zein and Sulandari (2008) reported that the genetic distance between native chicken populations in Lombok, using the D-loop, ranged from 0.1–1.7%. The distance between red partridges from Bengkulu and South Sumatra based on the COI gene has also been confirmed to range from 0–1.4%

(Jarulis *et al*, 2022). The genetic distance between red partridges and domesticated individuals, as measured by the D-loop and COI gene, exhibits a divergence range of 0–1.7%, which is higher than the divergence observed in the *ND1* gene. Therefore, the *ND1* gene is more conserved.

The results of the phylogenetic tree reconstruction of 28 Burgo chickens from Bengkulu using the NJ model and 1,000 bootstraps are presented in Figure 2. NJ is one of the phylogenetic analysis methods based on the difference in the evolution rate of each branch. The components in NJ analysis are the operational taxonomic units and evolutionary distance. Based on the phylogenetic tree, all Burgo chickens form a large clade, joined by *G. gallus bankiva* and *G. gallus* from Kalimantan and Java. This suggests that, genetically, Burgo chicken still have a direct relationship as descendants of the red partridge. However, there are two small groups among the individual Burgo chickens, due to nucleotide base differences at sites 52 and 375. Several factors can cause differences in nucleotide bases, including geographical and environmental factors, as well as the duration of isolation, all of which can trigger mutations. In general, the genes used for identification are the COI gene and the non-coding region (D-loop). Many studies have focused on these two genes (Zein and Sulandari, 2012; Bilgin *et al*, 2016). However, several previous studies have stated that the *ND1* gene can also be used for identification because it contains conserved regions (Bowles and Mcmanus, 1993; Raharjo *et al*, 2018; Widayanti *et al*, 2022). Therefore, the use of the *ND1* gene in species identification can be applied as an alternative to the COI gene with more stable traits in the region. However, this method is not yet accurate in determining the taxonomic position or discovering species history because our study only used 450bp (± 50%) of the total length of the *ND1* gene (974bp).

Conclusion

The *ND1* gene sequence of mitochondrial DNA from the original Burgo chicken in Bengkulu has been successfully obtained. SNPs were identified at two sites of the *ND1* gene, with a sequence length of 450bp. The average genetic distance within the Burgo chicken population was 0.1%, while the distance between Burgo chicken to other chicken populations was 0.12%. All Burgo chickens formed the same clade in the phylogenetic tree, though two individuals (C2F3 and K4F2) showed slight differences, forming small groups based on variations in nucleotide bases. Genetic differences among Burgo chickens from Bengkulu, other chicken species in Indonesia, and several locations worldwide are present but non-significant. Our data show that Burgo chickens may be genetically distinct from other chickens found in Indonesia and globally. However, further research on the morphology and morphometrics of Burgo chickens is needed to confirm these findings.

Acknowledgements

We would like to thank the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia for research funding assistance through the Fundamental Research scheme in 2024 with contract number No. 3930/UN30.15/PT/2024. We also appreciate the Bengkulu Burgo chicken community for their assistance and all the people who have helped with the research.

Author contributions

Jarulis and Ahmat Fakhri Utama contributed to the study's conception and design. Data collection was done by Jarulis. Data analysis and writing of the first manuscript draft were performed by Jarulis and Ahmat Fakhri Utama. All authors commented on the various versions of the manuscript, and read and approved the final manuscript.

Ethics statement

All experiments were approved by the local ethics committee of the University of Bengkulu, Indonesia. Animal procedures were conducted in accordance with the ethical code guidelines No. 15/KER-LPPM/EC/2023.

Conflict of interest statement

The authors declare no conflicts of interest.

References

Al-Nasser, A. *et al.* (2007) 'Overview of chicken taxonomy and domestication', *World's Poultry Science Journal*, 63(2), pp. 285–300. doi: https://doi.org/10.1017/S004393390700147X.

Amin, H.S., and Mushlih, M (2020). Identification of the Mitochondrial *ND1* Gene Carrier of Diabetes Mellitus Type 2 with Blood Samples. *Medicra (Journal of Medical Laboratory Science/Technology)*, 3(2), pp. 48–53. doi: https://doi.org/10.21070/medicra.v3i2.873

Bandelt, H.J., Forster, P., and Röhl, A (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), pp. 37–48. doi: https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Bilgin, R. *et al.* (2016) 'DNA barcoding of birds at a migratory hotspot in eastern Turkey highlights continental phylogeographic relationships', *PLoS ONE*, 11(6), pp. 1–17. doi: https://doi.org/10.1371/journal.pone.0154454.

Bowles, J. and Mcmanus, D.P. (1993) 'NADH dehydrogenase 1 gene sequences compared for species and strains of the genus Echinococcus', *International Journal for Parasitology*, 23(7), pp. 969–972. doi: https://doi.org/10.1016/0020-7519(93)90065-7.

Fouquet, A., *et al* (2007). Underestimation of Species Richness in Neotropical Frogs Revealed by mtDNA Analyses. (10). doi: https://doi.org/10.1371/journal.pone.0001109.

Fumihito, A., *et al* (1994). One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proceedings of the National Academy of Sciences of the United States of America*, 91(26), pp. 12505–12509. doi: https://doi.org/10.1073/pnas.91.26.12505.

Hall, T.A (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp. 95–98.

Hirsch V. (2003) Legal protections of the domestic chicken in the United States and Europe. East Lansing: Michigan State University College of Law; 2003.

Hirst, J (2010). Towards the molecular mechanism of respiratory complex I. *Biochemical Journal*, 425(2), pp. 327–339. doi: https://doi.org/10.1042/BJ20091382.

Jarulis, J. *et al,* (2022). DNA Barcode of Red Junglefowl Gallus gallus L, 1958 (Aves: Phasianidae) of Sumatra Based on Mitochondrial COI DNA Gene. *Biosaintifika: Journal of Biology & Biology Education*, 14(2), pp. 200–210. doi:

https://doi.org/10.15294/biosaintifika.v14i2.36530.
Kimura, M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2), pp. 111–120. doi: https://doi.org/10.1007/BF01731581.
Larson, G. *et al.* (2014) 'Current perspectives and the future of domestication studies', *PNAS*, 111(17), pp. 6139–6146. doi: https://doi.org/10.1073/pnas.1323964111/-/DCSupplemental.
Meseret, S. (2016) 'A review of poultry welfare in conventional production system', *Live stock Research for Rural Development*, 12(28). Available at: http://www.lrrd.org/lrrd28/12/mese28234.html
Liu, Y.P. *et al,* (2006). Multiple maternal origins of chickens: Out of the Asian jungles. *Molecular Phylogenetics and Evolution*, 38(1), pp. 12–19. doi: https://doi.org/10.1016/j.ympev.2005.09.014.
Miao, Y.W. *et al.* (2013) 'Chicken domestication: An updated perspective based on mitochondrial genomes', *Heredity*, 110(3), pp. 277–282. doi: https://doi.org/10.1038/hdy.2012.83.
Nataamijaya, A.G. (2010). Developing the Potential of Local Chickens to Support Improved Farmer Welfare (in Indonesian), *Jurnal Litbang Peternakan*, 29(10), pp. 131–138.
Nataannjaya, A.G. (2000) 'The Native Chicken of Indonesia', *Buletin Plasma Nutfah*.
Ni'mah, A. *et al,* (2016). Detection of Pork Contaminantion in Fresh and Cooked Beef Using Geetic Marker Mitochondrial-DNA Cytochrome B by Duplex-PCR. *Journal of the Indonesia Tropical Animal Agriculture*, 4(1), pp. 88–100. doi: https://doi.org/10.14710/jitaa.41.1.7-12.
Putranto, H.D. *et al.* (2017) 'The estimation of dynamical distribution of domesticated Burgo chicken population in Bengkulu coastal area, Indonesia', *Biodiversitas*, 18(2), pp. 458–464. doi: https://doi.org/10.13057/biodiv/d180203.
Rafian, T., Jakaria, J., and Ulupi, N (2017). Phenotypic Diversity of Qualitative Traits of Burgo Chicken in Bengkulu Province, *Jurnal Sain Peternakan Indonesia*, 12(1), pp. 47–54. doi: https://doi.org/10.31186/jspi.id.12.1.47-54.
Raharjo, T.J. *et al.* (2018) Mitochondrial ND-1 gene-specific primer polymerase chain reaction to determine mice contamination in meatball. *International Food Research Journal*, 25(2): 638-642. url: http://www.ifrj.upm.edu.my/25%20(02)%202018/(26).pdf
Rozas, J. *et al,* (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34(12), pp. 3299–3302. doi: https://doi.org/10.1093/molbev/msx248.
Safitra, M.I., Putranto, H.D., and Brata, B (2022). Characteristics of the crowing sound of male Burgo chickens in Bengkulu City (in Indonesian), *Jurnal Peternakan*, 19(1), p. 64. doi : https://doi.org/10.24014/jupet.v19i1.15047.
Setianto, J., Prakoso, H., and Sutriyono (2013). Laporan penelitian. 22(2), pp. 184–206.
Seutin G, White BN, Boag PT (1991). Preservation of avian blood and tissue samples for DNA analysis. Can. J. Zool. 69:82-90. doi: https://doi.org/10.1139/z91-013
Sibley, C.G. & B.L. Monroe. 1990. Distribution and Taxonomy of Birds of the World. Yale University Press. New Haven & London. P 1111.
Sulandari, S., Syamsul Arifin Zein, M., and Sartika, T (2008). Molecular Characterization of Indonesian Indigenous Chickens based on Mitochondrial DNA Displacement (D)-loop Sequences. *HAYATI Journal of Biosciences*, 15(4), pp. 145–154. doi: https://doi.org/10.4308/hjb.15.4.145.
Sulandari, S., M.S.A. Zein, S. Paryanti & T. Sartika. 2007. Taksonomi dan asal usul ayam domestikasi. Dalam: K. Diwyanto & S.N. Prijono (Eds.). Keragaman Sumber Daya Hayati Ayam Lokal Indonesia: Manfaat dan Potensi. Pusat Penelitian Biologi, LIPI. ISBN 978-979-799-183-8. Edisi Pertama. Hal. 7-24.
Sutriyono, S. (2016) Production and population of domesticated red jungle fowl in North Bengkulu Regency and population development scenarios (in Indonesian), *Pros Sem NAs Masy Biodiv Indon*, 2, pp. 226–231. doi: https://doi.org/10.13057/psnmbi/m020218.
Tamura, K. *et al,* (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), pp. 2725–2729. doi: https://doi.org/10.1093/molbev/mst197.
Torres, J.B (2016). A history of you, me, and humanity: mitochondrial DNA in anthropological research. *AIMS Genetics*, 03(02), pp. 146–156. doi: https://doi.org/10.3934/genet.2016.2.146.
Utama, A.F. *et al,* (2023). DNA barcoding of Burgo chicken from Bengkulu, Indonesia, based on the cytochrome oxidase gene sub unit I mitochondria DNA. *Biodiversitas*, 24(11), pp. 6256–6263. doi: https://doi.org/10.13057/biodiv/d241148.
Väisänen, J., Håkansson, J. and Jensen, P. (2005) 'Social interactions in Red Junglefowl (*Gallus gallus*) and White Leghorn layers in stable groups and after re-grouping', *British Poultry Science*, 46(2), pp. 156–168. doi: https://doi.org/10.1080/00071660500062638.
Wang, M.S. *et al.* (2020) '863 genomes reveal the origin and domestication of chicken', *Cell Research*, 30(8), pp. 693–701. doi: https://doi.org/10.1038/s41422-020-0349-y.
Warmadewi, D.A. *et al,* (2020). The variation of phenotypics Bali cattle in Bali Province, Indonesia. *International Journal of Fauna and Biological Studies*, 7(4), pp. 44–47.
Widayanti, R. *et al.* (2022) 'Study of Genetic Diversity of Native Indonesian Catfish Based on Nucleotide Sequences of the *ND1* Gene', *Jurnal Sain Veteriner*, 40(3), p. 268. doi: https://doi.org/10.22146/jsv.62755.
Zein, M.S.A., and Sulandari, S (2008). Genetic diversity of Lombok chickens based on D-loop mitochondrial DNA sequences. *JITV* 13(4): 307-314.
Zein, M.S.A., and Sulandari, S. (2012) Genetic Diversity and Haplogroup Distribution of Native Chicken Using Hypervariable-I Control Region of Mitochondrial DNA (in Indonesian), *JITV,* 17(2): 120-131.

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 160–170
DOI: 10.46265/genresj.UNBA8473
https://www.genresj.org
ISSN: 2708-3764

Evaluation of red sanders (*Pterocarpus santalinus* L.f.) germplasm for conservation and breeding

Kumbarahally Murthigowda Shivaprasad[a,*], Pendela Surath Kumar[a], Avula Kishore[a], Dinesh Kumar[a], Swapnendu Pattanaik[a,*]

[a]*Genetics and Tree Improvement division, Indian Council of Forestry Research and Education (ICFRE)-Institute of Forest Biodiversity, Hyderabad, Telangana, India-500100*

Abstract: *Pterocarpus santalinus* L.f., an endangered species endemic to India's Eastern Ghats, faces threats from illegal trade due to high demand and limited legal supply. Field genebanks are essential for conserving genetic diversity and supporting sustainable use. This study evaluated a 12-year-old *P. santalinus* germplasm collection of 500 accessions, grown *ex situ,* for growth and heartwood traits. The survival rate was 80%, with notable variation in tree height (5.5–11.8m), girth at breast height (GBH, 26–78cm), clear bole height (0–6.2m), and heartwood core length (0–6.6cm), indicating substantial genetic diversity. Early heartwood formation (< 12 years) occurred in 18.50% of accessions, earlier than the typical 15 years. Superior accessions included S5R1-19 (6.6cm heartwood) and S7R1-4 (5.5cm), while S1R1-13, S5R4-20, and S1R3-18 exhibited desirable deep red heartwood. Accessions from Petbasheerabad showed high survival and heartwood yield, suggesting their value as elite seed sources. GBH positively correlated with heartwood length (r = 0.443), supporting its use as a selection trait. Principal component analysis and clustering grouped accessions into three clusters: Cluster 1 showed superior timber traits (high clear bole, low percentage of forking), whereas Cluster 3 displayed less desirable traits. These findings aid in identifying elite accessions and developing conservation and breeding strategies. Integration of molecular tools such as genome-wide association studies and transcriptomics is recommended to accelerate genetic improvement.

Keywords: Red sanders, heartwood formation, germplasm evaluation, tree improvement, field genebank

Citation: Shivaprasad, K. M., Kumar, P. S., Kishore, A., Kumar, D. and Pattanaik, S. (2025) "Evaluation of red sanders (*Pterocarpus santalinus* L.f.) germplasm for conservation and breeding", *Genetic Resources*, 6(12), pp. 160–170. doi: 10.46265/genresj.UNBA8473.

Introduction

Red sanders (*Pterocarpus santalinus* L. f.) is a highly valuable timber species endemic to the southern part of the Eastern Ghats of India and is classified as endangered on the International Union for the Conservation of Nature (IUCN) Red List of threatened species (Ahmedullah, 2021). It is known for its characteristic timber of exquisite colour, beauty, and superlative technical qualities (Arunakumara *et al*, 2011). The red-coloured heartwood is the economic part of the plant, and the wood has significant demand globally, primarily sought after for its application in crafting the traditional Japanese musical instrument known as the 'Shamisen'. Additionally, wood sourced from red sanders is utilized in the fabrication of name seals, frames, carvings, furniture, and various traditional artefacts (Arunakumara *et al*, 2011; Pattanaik, 2024). Further, red sanders wood possesses a significant insoluble or sparingly soluble red dye, consisting of approximately 16% of the santalin pigment, a principal colouring agent utilized notably in textile dyeing and in European medicine as a colouring agent. In the United States, it holds approval as a food dye for alcoholic beverages and is similarly sanctioned for use within Europe, and classified as a spice extract rather than a conventional food colourant (Arunakumara *et al*, 2011; New, 1981). Additionally,

* Corresponding authors: : Kumbarahally Murthigowda Shivaprasad (shivaprasadkm.ifb@gmail.com), Swapnendu Pattanaik (swapnen@yahoo.com)

Received: 09.06.2025 | Accepted: 08.10.2025 | Published online: 17.11.2025

Pterostilbene, a methyl ester derivative of resveratrol present in red sanders, offers a diverse array of promising pharmacological properties (Schmidlin *et al*, 2008; Seshadri, 1972). The exploitation of red sanders forests without commensurate restoration in the past has led to the present degraded state (Ahmed and Nayar, 1984). Overexploitation prompted the Union Government in the 1980s to recommend the inclusion of red sanders in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Subsequently, the species was listed in Appendix II of CITES in 1995, and export of red sanders was prohibited in 2004. Despite the complete ban on felling and trade of natural origin wood, the gap between demand and supply fuelled its illicit felling and smuggling (Ahmedullah *et al*, 2019). Sustained pressure on these natural resources has resulted in their widespread degradation, as evidenced by the fact that most of the standing crop is of coppice origin and that regeneration is poor or lacking at several locations of occurrence.

The Andhra Pradesh Forest Department has estimated a global annual demand of approximately 3,000 metric tonnes (MT) of red sanders wood (NBA, 2017), and reported substantial revenue from the auction of 5,489.51MT of seized wood between 2014 and 2019 (Ahmedullah *et al*, 2019). In response to increasing international interest, the Directorate General of Foreign Trade (Government of India) revised its export policy in 2019 to permit the export of red sanders wood derived from cultivated sources, underscoring the urgent need to expand plantation-based production systems. Despite red sanders being cultivated in various regions of Andhra Pradesh and other Indian states, the wood produced in these plantations is typically of inferior quality compared to wood from natural populations. This quality gap limits the economic returns for farmers, even as global demand for premium-grade red sanders remains high (Ahmedullah *et al*, 2019). The lack of genetically improved red sanders varieties forces farmers to rely on uncertified bulk seeds, resulting in considerable variability in key traits such as heartwood formation. Furthermore, red sanders typically require 15–25 years to form heartwood, delaying economic returns (NBA, 2017). To enhance productivity and profitability for farmers, it is essential to develop and supply superior planting material with improved wood traits. Evaluation of available genetic resources is a critical step in identifying elite germplasm suitable for breeding and large-scale propagation. Field genebanks play a vital role in the conservation and evaluation of such genetic resources, enabling long-term genetic improvement strategies. This study was undertaken to assess a red sanders germplasm assemblage for variation in growth and heartwood traits, with the goal of identifying superior accessions for use in future breeding and conservation programmes.

Materials and methods

Study area

The present evaluation study was carried out in the field genebank located at ICFRE-Institute of Forest Biodiversity (17.556911° N latitude, 78.446385° E longitude, and 542m above mean sea level). The original collection, consisting of 500 germplasm from eight different seed sources, was established in August 2012 for *ex situ* conservation of red sanders. The genebank was designed with a spacing of 5 × 4m, and the trees were maintained without pruning operations to allow natural branching patterns (Figure 1). The details of genotypic origin and sample sizes are provided in Table 1 and Figure 2.

Figure 1. View of the red sanders field genebank at ICFRE-Institute of Forest Biodiversity, Hyderabad, India.

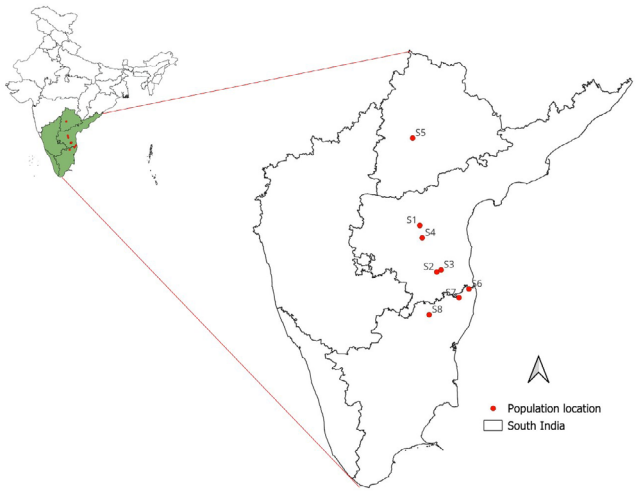


Figure 2. Map showing the red sanders seed sources used for the red sanders field genebank

Table 1. Geographic details of seed sources and corresponding number of plants established in the red sanders field genebank at ICFRE-Institute of Forest Biodiversity, Hyderabad, India. Longitude and latitude are expressed in decimal degrees.

Code	Seed source	Latitude	Longitude	No. of germplasms	No. of germplasm survived and evaluated
S1	Ahobilam	15.1408N	78.7344E	75	55
S2	K. V. Bhavi	13.8917N	79.2341E	100	82
S3	Red Wood Park	13.9462N	79.3528E	100	79
S4	Mudireddipally	14.8062N	78.8096E	100	62
S5	Petbasherabad	17.5132N	78.4736E	100	97
S6	Gumudipoondi	13.4466N	80.1306E	10	10
S7	Nayapakkam	13.2065N	79.8592E	10	10
S8	Amirthi	12.7305N	79.0485E	5	5

Evaluation of growth traits and heartwood formation

Germplasm evaluation was carried out in 2024 with ten morphological traits. Trait measurements included tree height (m), measured from the base of the tree to the highest point of the crown using a measuring pole, and clear bole height (m), measured from the base of the tree to the first branching point using a measuring tape. Girth at breast height (GBH) (cm) was recorded at 1.37m above the ground using a tape, while forking was assessed visually to record its presence or absence. The number of forked stems was counted, and the average GBH in forked trees (cm) was calculated by measuring the GBH of each forked stem and computing the average for trees with multiple stems. Heartwood core length (cm) and sapwood length (cm) were determined from wood core samples extracted using an increment borer. The distinct heartwood and sapwood regions of the core were measured with a scale and expressed in cm, while the total core sample length (cm) was recorded as the combined length of heartwood and sapwood. Heartwood colour was evaluated subjectively, and the wood samples were classified ocularly into three broad groups, *viz*., light red, medium red and deep red. These measurements were conducted systematically to ensure consistency and accuracy across all accessions, enabling detailed analysis of phenotypic variation.

Descriptive statistics

Descriptive statistical analysis was performed to summarize and compare the traits of 400 surviving germplasm accessions out of the 500 initially planted. Measures such as mean, minimum and maximum values were calculated using the Excel Data Analysis Toolpak for an initial overview. Frequency distributions and percentage pie charts were also plotted to illustrate key trends in survival rates and heartwood formation. Subsequently, descriptive statistics for each seed source were analyzed independently, including the calculation of mean and standard deviation (SD) for all traits. To enhance visualization and facilitate comparisons among seed sources, the R programming (R Core Team, 2018) environment was utilized. Packages such as dplyr and tidyr were employed for data preprocessing and aggregation, while ggplot2 was used to generate visual representations, including bar plots. These tools allowed for clear comparisons of trait variability and performance across seed sources.

Correlation

To explore relationships among the traits, a correlation analysis was conducted. Pearson correlation coefficients were computed for all trait combinations using R (R Core Team, 2018). The corrplot package was employed to visualize the correlation matrix, and RColorBrewer was used to enhance the colour scheme, facilitating easier interpretation of trait interdependencies. The correlation plot provided insights into the strength and direction of relationships among the traits across seed sources.

Principal component analysis and boxplot

Principal component analysis (PCA) was performed to reduce dimensionality while retaining maximum variance among tree growth traits. Numeric traits were standardized using z-score normalization, and PCA was conducted using the prcomp function in R (R Core Team, 2018). Traits with absolute loadings above 0.3 were identified as major contributors. To classify trees into distinct groups, K-means clustering was applied to PCA scores (PC1 and PC2), with the optimal number of clusters determined using the NbClust function based on the silhouette method. The resulting clusters were visualized using ggbiplot, with confidence ellipses highlighting group structures. All analyses were performed in R, utilizing packages such as ggplot2, ggbiplot, dplyr, NbClust, and RColorBrewer for statistical analysis and visualization.

Results

Morphological diversity in red sanders accessions

Out of the 500 established germplasm lines, 400 survived and were assessed using ten morphological traits. The evaluation of the 12-year-old germplasm revealed an average tree height of 9.04m, with the tallest tree (11.8m) observed in accession S2R4-10 and the shortest (5.5m) in S4R1-7. The average clear bole height was 1.05m, with the highest (6.2m) recorded in S1R2-13. The number of forked stems ranged from 0 to 6, with 176 accessions exhibiting no forking. The accession S8R4-4 recorded the maximum of six forked stems.

Figure 3 shows the frequency distribution of forked stems among red sanders accessions, illustrating the variation in forking observed across the population. The average GBH was 45.00cm, with the highest (78.00cm) observed in S2R1-23 and the lowest (26.00cm) in S3R2-2. Increment core samples had an average length of 7.58cm, with a maximum of 12.5cm in S2R1-23 and a minimum of 4.00cm in S3R2-2. The average sapwood length was 7.1cm, with the highest value (7.9cm) recorded in S5R2-2. Heartwood formation was observed in 74 accessions (18.50%), with an average heartwood core length of 2.53cm. The maximum heartwood core length (6.6cm) was recorded in S5R1-19, followed by S7R4-4 (5.5cm), S6R4-3 (5.2cm), S5R4-20 (5.0cm), S3R4-23 (5.0cm) and S5R4-25 (4.5cm). The lowest heartwood core length (0.5cm) was observed in S4R1-19 and S3R1-17 (Supplemental Table 1). Analysis of heartwood colour of 74 accessions revealed light red colour in 41 accessions (55.41%), a medium red colour in 19 accessions (25.68%), and a deep red colour in 14 accessions (18.92%) (Figure 4A and B). A detailed list of the 74 heartwood-forming accessions, along with their heartwood colour and heartwood core length, is provided in Table 2.

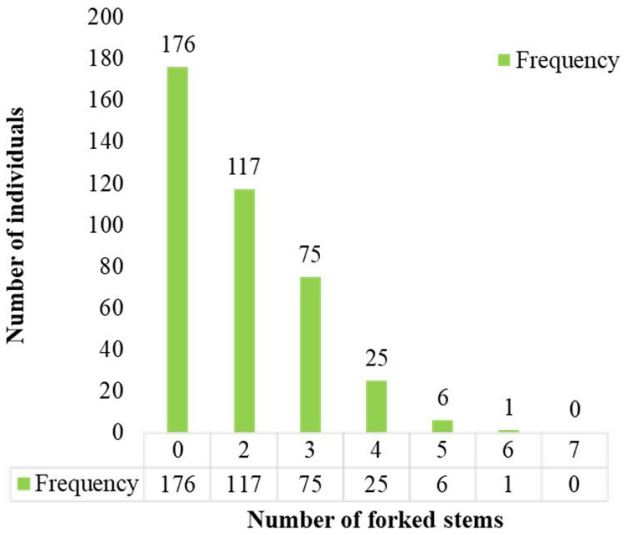


Figure 3. Frequency distribution of forked stems in red sanders accessions

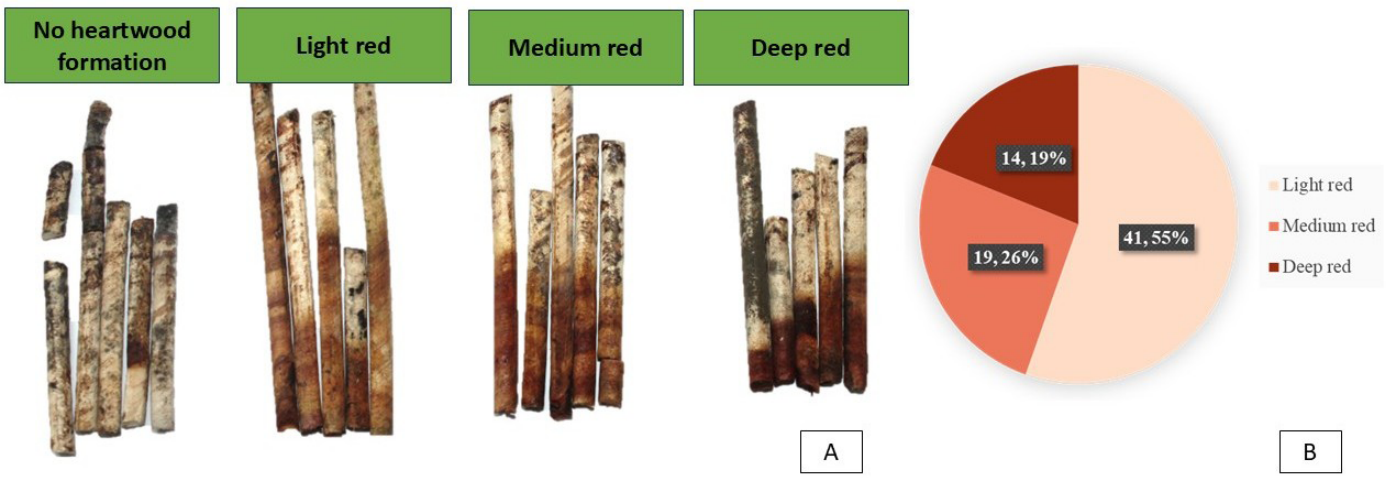


Figure 4. Heartwood colour variation in red sanders germplasm (A). Pie chart representing the distribution of different heartwood colour grades in the collection (B).

Table 2. List of the 74 heartwood-forming red sanders accessions and their heartwood colour and heartwood core length.

Accession	Heartwood core length (cm)	Heartwood colour	Accession	Heartwood core length (cm)	Heartwood colour
S5R1-19	6.6	Light red	S5R4-18	2.3	Deep red
S7R1-4	5.5	Medium red	S5R4-21	2.3	Light red
S6R1-3	5.2	Light red	S5R3-5	2.2	Light red
S5R4-20	5	Deep red	S5R2-13	2.1	Light red
S3R4-23	5	Light red	S1R3-7	2.1	Deep red
S5R4-25	4.5	Light red	S5R1-9	2	Light red
S4R3-22	4.3	Medium red	S5R2-14	2	Light red
S5R3-19	4.3	Light red	S1R2-23	2	Light red
S2R1-23	4	Medium red	S2R2-23	2	Light red
S5R2-4	4	Light red	S4R4-2	2	Medium red
S5R4-2	4	Medium red	S4R4-17	2	Medium red
S3R4-21	4	Medium red	S3R4-14	2	Light red
S5R4-4	3.9	Medium red	S3R4-15	2	Deep red
S5R4-15	3.8	Light red	S1R3-24	1.9	Light red
S3R2-20	3.6	Light red	S5R1-21	1.8	Light red
S3R2-22	3.5	Medium red	S1R3-2	1.8	Deep red
S5R3-4	3.4	Light red	S3R1-5	1.7	Deep red
S1R2-2	3.3	Medium red	S2R3-22	1.7	Light red
S5R1-2	3.2	Light red	S5R3-20	1.6	Medium red
S4R3-18	3.2	Medium red	S3R2-25	1.5	Light red
S4R3-24	3.1	Light red	S3R4-24	1.5	Medium red
S4R4-6	3	Medium red	S1R1-17	1.4	Deep red
S5R4-13	3	Medium red	S4R3-7	1.4	Light red
S5R4-9	2.9	Light red	S1R3-4	1.4	Medium red
S5R1-15	2.8	Light red	S5R2-17	1.3	Medium red
S5R1-23	2.8	Light red	S5R2-12	1.2	Light red
S5R2-5	2.7	Light red	S3R3-6	1.2	Light red
S3R3-4	2.7	Light red	S1R2-21	1.1	Medium red
S5R3-6	2.7	Light red	S3R4-4	1.1	Light red
S1R1-20	2.6	Light red	S1R1-13	1	Deep red
S3R3-14	2.5	Deep red	S2R2-4	1	Light red
S1R3-8	2.5	Light red	S3R3-2	1	Light red
S5R1-12	2.4	Light red	S1R3-1	1	Medium red
S3R2-23	2.3	Light red	S1R3-14	1	Deep red
S4R3-3	2.3	Light red	S1R3-18	0.8	Deep red
S5R3-25	2.3	Light red	S3R1-17	0.5	Deep red
S5R4-3	2.3	Deep red	S4R1-19	0.5	Deep red

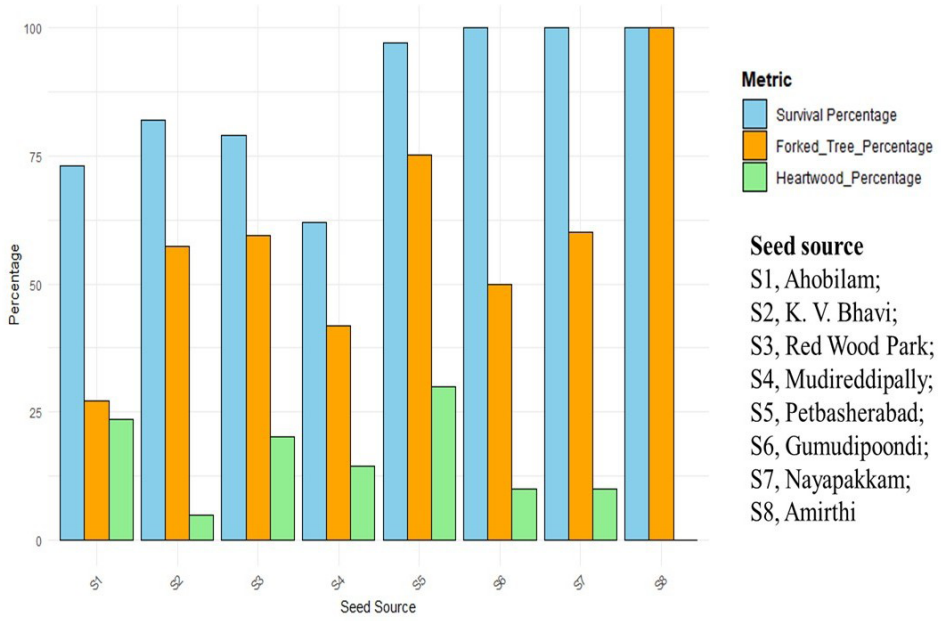


Figure 5. Frequency distribution of survival, forking and heartwood formation traits in red sanders seed sources

Morphological diversity in red sanders seed sources

The study evaluated eight seed sources based on survival rates, heartwood formation, and various growth and heartwood formation parameters in different accessions (Table 3). Survival rates varied significantly among sources, with Gumudipoondi (S6), Nayapakam (S7) and Amirthi (S8) achieving 100% survival, whereas Madireddipally (S4) exhibited the lowest survival rate (62%). Heartwood formation also varied considerably, with Petbasheerabad (S5) recording the highest heartwood formation (29.9%), followed by Ahobilam (S1) (23.64%) and Red Wood Park (S3) (20.25%), while K.V Bhavi (S2) had the lowest heartwood formation (4.88%), and no heartwood formation was observed in Amirthi (S8) (Figure 5).

Growth traits revealed that Amirthi accessions exhibited the greatest tree height (10.24m), while Madireddipally accessions recorded the shortest (8.40m). Petbasheerabad accessions demonstrated a favourable balance between growth and heartwood quality, with a high survival rate (97%) and consistent heartwood formation, indicating its potential as a reliable seed source. Other traits, including clear bole height, heartwood core length, sapwood length, and GBH, exhibited substantial variation across seed sources. Notably, Amirthi and Gumudipoondi accessions excelled in tree height and GBH. The observed variability highlights the importance of selecting genetically superior seed sources to enhance plantation productivity and optimize timber quality. These findings provide a scientific basis for identifying and promoting superior seed sources in breeding programmes aimed at improving plantation success and meeting commercial demands for high-quality wood products (Table 3, Figure 6).

Correlation analysis among various traits in red sanders

The correlation analysis revealed complex but statistically meaningful relationships among the recorded growth traits (Figure 7). Tree height showed a significant negative correlation with clear bole height ($r = -0.314$, $p < 0.05$) and significant positive correlations with the number of forking branches ($r = 0.278$, $p < 0.05$), average GBH ($r = 0.334$, $p < 0.05$), sapwood length ($r = 0.316$, $p < 0.05$), and total core sample length ($r = 0.435$, $p < 0.01$). Clear bole height exhibited a highly significant negative correlation with the number of forking stems ($r = -0.796$, $p < 0.001$) but only weak and mostly non-significant associations with other traits. The number of forking stems was negatively correlated with average GBH ($r = -0.326$, $p < 0.05$), while its correlations with heartwood core length ($r = -0.077$) and sapwood length ($r = 0.026$) were weak and non-significant. Average GBH showed strong and highly significant positive correlations with heartwood core length ($r = 0.443$, $p < 0.01$), sapwood length ($r = 0.630$, $p < 0.001$), and total core sample length ($r = 0.916$, $p < 0.001$). Heartwood core length exhibited a significant negative correlation with sapwood length ($r = -0.295$, $p < 0.05$) and a significant positive association with total core sample length ($r = 0.457$, $p < 0.01$), whereas sapwood length was strongly and significantly correlated with total core sample length ($r = 0.710$, $p < 0.001$) (Figure 7). These results underscore the role of clear bole height as an inverse determinant of branching, with its strongest relationship being the negative correlation with the number of forking stems. Average GBH emerged as the most integrative trait, showing consistently strong and highly significant associations with both wood anatomical traits and core sample length, highlighting its utility as a key indicator for overall growth performance in tree improvement and breeding programmes.

Table 3. Descriptive statistics of morphological and heartwood formation traits in red sanders

Code	Seed source	Accessions planted	Surviving accessions	Accessions with heartwood formation		Tree height (m)	Clear bole height (m)	No. of forked stems	Girth at breast height (cm)	Heartwood core length (cm)	Sapwoodlength (cm)	Total core sample length (cm)
S1	Ahobilum North, Rudravaram (A.P)	75.00	55 (73.33%)	14 (25.45%)	Mean	9.43	1.87	0.62	48.01	0.43	7.26	7.76
					SD	1.30	1.51	1.06	10.75	0.83	1.40	1.70
					Min	5.70	0.00	0.00	27.67	0.00	4.60	4.60
					Max	11.70	6.20	4.00	73.00	3.30	10.20	11.60
S2	K.V Bhavi Kodur (A.P)	100.00	82 (82%)	4 (4.88%)	Mean	8.69	1.13	1.54	43.96	0.11	7.32	7.42
					SD	1.21	1.47	1.47	8.74	0.53	1.39	1.52
					Min	w	0.00	0.00	27.00	0.00	4.20	4.20
					Max	11.80	4.80	5.00	78.00	4.00	11.50	12.50
S3	Red Wood Park (A.P)	100.00	79 (79%)	16 (20.25%)	Mean	8.90	1.04	1.53	45.74	0.46	7.22	7.68
					SD	1.15	1.42	1.44	9.80	1.06	1.34	1.58
					Min	6.20	0.00	0.00	26.00	0.00	4.00	4.00
					Max	11.50	5.00	5.00	74.00	5.00	10.50	11.80
S4	Madireddipally, Onipenta (A.P)	100.00	62 (62%)	9 (14.52%)	Mean	8.40	1.24	1.11	43.56	0.35	6.88	7.23
					SD	1.29	1.14	1.39	8.35	0.95	1.23	1.32
					Min	5.50	0.00	0.00	26.50	0.00	4.70	4.70
					Max	11.50	3.10	4.00	63.00	4.30	9.70	10.00
S5	Petbasheerabad, Hyderabad (T.S)	100.00	97 (97%)	29 (29.90%)	Mean	9.40	0.48	2.06	44.46	0.88	6.85	7.74
					SD	1.01	0.90	1.38	9.75	1.50	1.51	1.46
					Min	6.80	0.00	0.00	27.00	0.00	3.60	4.60
					Max	11.70	3.40	5.00	76.00	6.60	10.80	12.00
S6	Gumudipoondi (T.N)	10.00	10 (100%)	1 (10%)	Mean	9.97	1.04	1.30	47.73	0.52	7.63	8.15
					SD	0.55	1.12	1.49	6.17	1.64	1.45	0.78
					Min	8.60	0.00	0.00	38.75	0.00	4.00	7.10
					Max	10.50	2.40	4.00	58.00	5.20	8.90	9.20
S7	Nayapakam (T.N)	10.00	10 (100%)	1 (10%)	Mean	9.73	0.74	1.50	46.10	0.55	7.07	7.62
					SD	0.62	0.99	1.35	8.58	1.74	1.50	1.37
					Min	8.90	0.00	0.00	33.00	0.00	4.30	5.00
					Max	10.60	2.50	3.00	62.00	5.50	8.90	9.80
S8	Amirthi (T.N)	5.00	5 (100%)	-	Mean	10.24	0.00	3.40	37.93	0.00	7.14	7.14
					SD	0.30	0.00	1.52	6.91	0.00	0.97	0.97
					Min	9.80	0.00	2.00	26.00	0.00	5.70	5.70
					Max	10.50	0.00	6.00	44.00	0.00	8.40	8.40

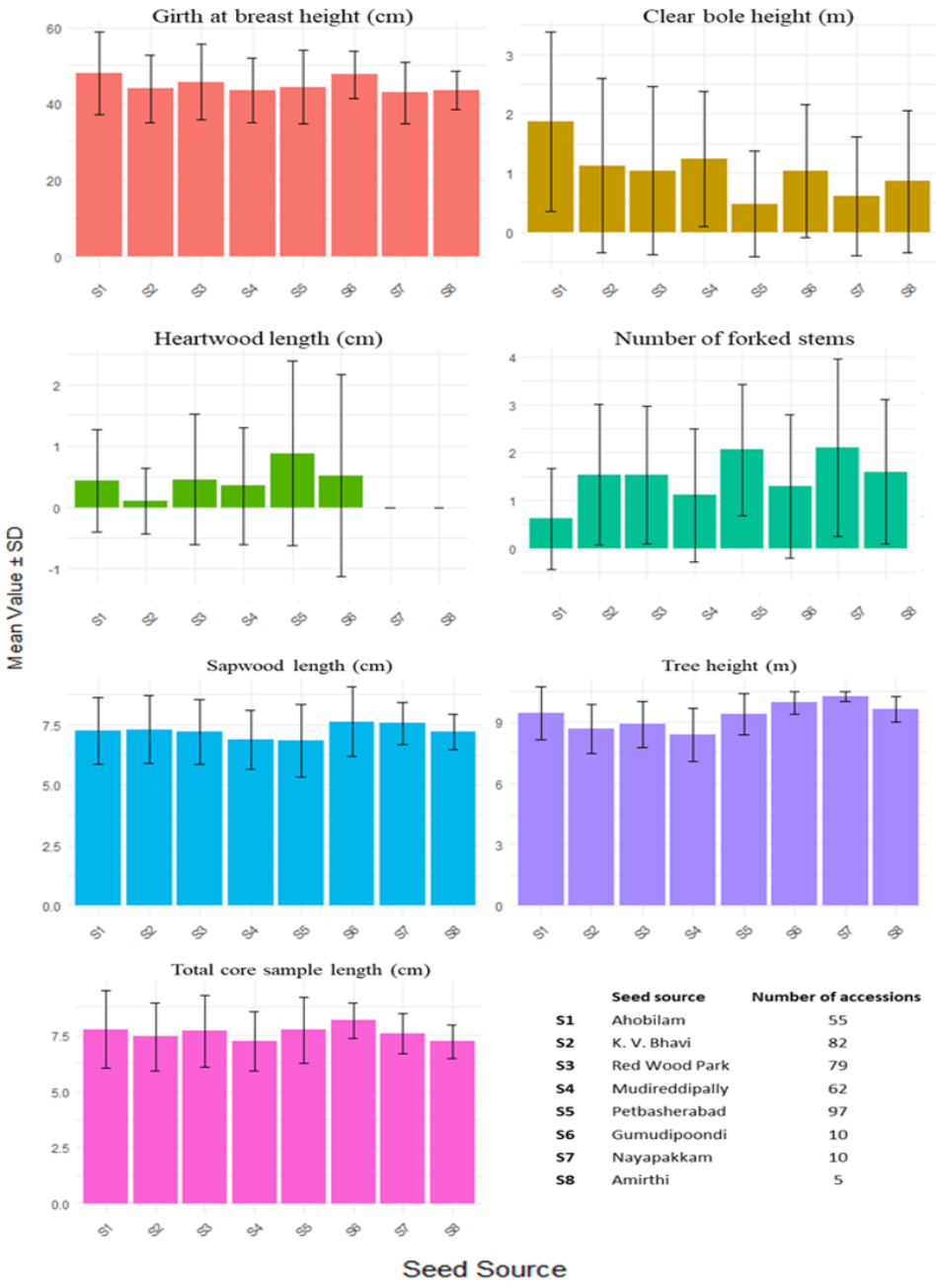


Figure 6. Bar plot representing mean and SD (standard deviation) for growth-related traits in red sanders seed sources

Principal component analysis among various traits in red sanders

PCA identified seven principal components (PCs) summarizing trait variance, with PC1 and PC2 explaining 40.85% and 28.86% of the variance, respectively, capturing a cumulative 69.71% of the total variation. PC3 contributed 18.20%, while PC4 explained 8.69%, and the remaining components (PC5, PC6 and PC7) collectively accounted for less than 4% of the variance (Supplemental Table 2). These results indicate that the first three components encapsulate most of the trait variation, underscoring the importance of total core sample length, GBH, forking, heartwood core length and sapwood core length in defining diversity. To determine the most influential traits for each principal component, a loading threshold of 0.3 was applied. Traits exceeding this threshold were considered significant contributors to the observed variation. PC1 was primarily influenced by total core sample length (0.58), average GBH (0.55) and sapwood length (0.42), indicating their role in explaining the primary axis of variation. PC2 was dominated by clear bole height (-0.63) and the number of forked stems (0.66), suggesting that this axis captures trait variation driven by contrasting growth patterns. PC3 was significantly influenced by heartwood core length (0.78) and inversely by sapwood length (-0.60), reflecting their negative correlation (Supplemental Table 3). Clustering analysis using the silhouette method in NbClust identified three optimal clusters. K-means clustering further revealed distinct groups: Cluster 1 comprised trees with shorter tree height, higher clear bole height, and minimal

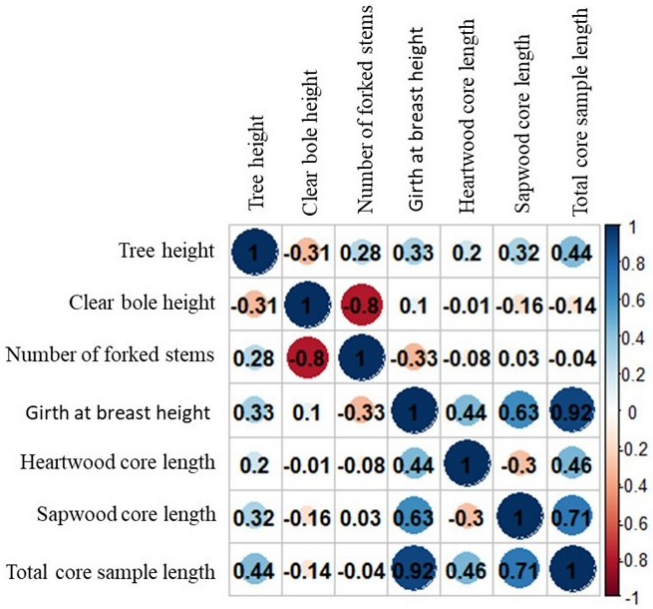


Figure 7. Correlation plot for various traits in red sanders. The circles represent the magnitude and direction of the Pearson correlation coefficients between trait pairs. The color scale shows correlation values, where blue indicates positive correlations and red indicates negative correlations. The intensity of the color and the size of the circles correspond to the strength of the correlation.

forking; Cluster 2 included taller trees with moderate clear bole height, higher GBH and moderate branching; while Cluster 3 consisted of trees with significantly more forked stems and negligible clear bole height. The PCA biplot further supported these clustering results. In the biplot, the black vectors represent the original quantitative traits, where the length and direction of each vector indicate the strength and direction of each trait's contribution to the principal components. The coloured points correspond to individual accessions grouped into the three clusters, while the elliptical circles represent the 68% confidence limits for each cluster. The distinct separation of clusters along PC1 and PC2 axes reflects trait-based differentiation among populations. These findings highlight the genetic and phenotypic diversity within the seed sources and provide insights for selecting superior accessions for tree improvement programmes (Figure 8).

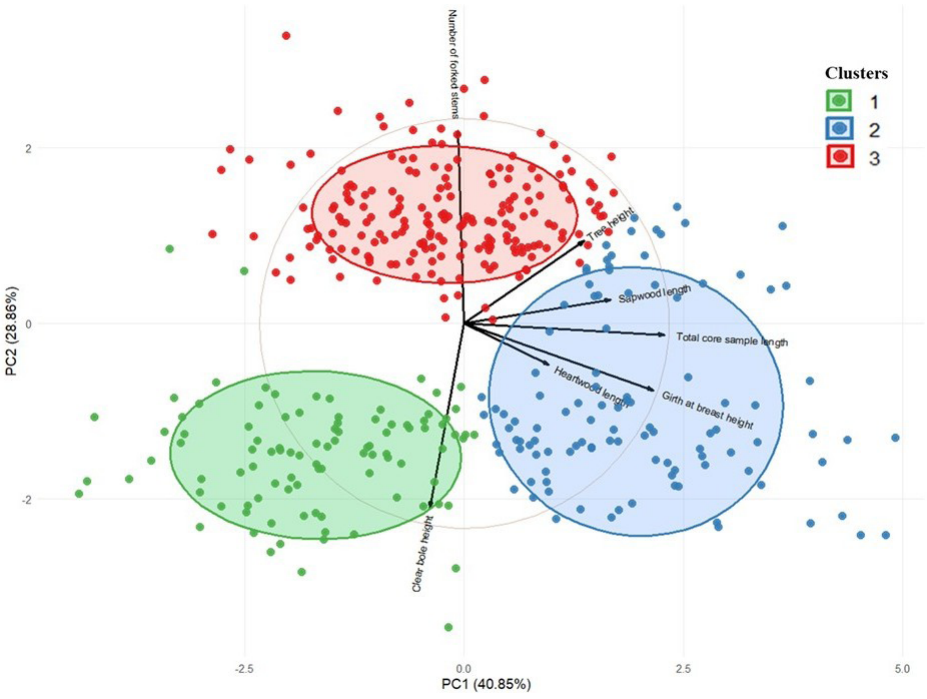


Figure 8. Principal component analysis biplot with clusters for various traits in red sanders. In the biplot, the black vectors represent the original traits, where the length and direction of each vector indicate the magnitude and direction of each trait's contribution to the principal components. The coloured circles (ellipses) represent the 68% confidence intervals for the clusters, while the colours of the points correspond to the three identified groups of accessions. The separation of clusters along PC1 and PC2 reflects trait-based differentiation among populations.

Discussion

The morphological and heartwood formation evaluation of red sanders germplasm provides valuable insights into the genetic variability and potential of this species for breeding and conservation. The high survival rate (80%) observed in the germplasm bank demonstrates the adaptability and resilience of red sanders to environments beyond its limited natural range, which is a crucial factor for its sustainable utilization in forestry programmes.

Diversity in growth morphological traits and heartwood formation parameters

The observed variability in growth traits, such as tree height, GBH and clear bole height, reflects significant genetic diversity. The tallest accession, S2R4-10 (11.8m), and the shortest, S4R1-7 (5.5m), along with the wide range in GBH from 26.00cm to 78.00cm, demonstrate the population's growth potential. These findings align with the natural red sanders populations in the Rajampet Forest Division, where tree heights range from 7.55m to 13.11m and GBH varies from 22.30cm to 81.82cm (Senthilkumar *et al*, 2015). Such diversity provides a valuable resource for selecting accessions for specific silvicultural and ecological objectives.

The average tree height of 9.03m, exceeding the 8.52m average reported for 15–25-year-old red sanders plantations in Gujarat (Hegde, 2023) highlights the superior performance of certain germplasm lines. However, the lower average GBH (45.00cm) compared to the 61.21cm reported in Gujarat's plantations and 49.84cm in a 20-year-old plantation (Arun Kumar *et al*, 2017) may be attributed to both genetic and environmental factors. The highest clear bole height, observed in S1R2-13 (6.2m), holds particular importance for timber production, where straight, knot-free logs are preferred.

Heartwood formation was observed in 18.25% of accessions in this 12-year-old field genebank, contrasting with previous reports that heartwood formation begins at 15 years (Hegde, 2023; Suresh *et al*, 2017). Heartwood content increases significantly with age, as observed in studies reporting 70% heartwood formation in 20-year-old plantations and up to 97% in 45-year-old plantations (Arunkumar, 2011). The occurrence of early heartwood formation in certain accessions highlights the strong genetic influence on this trait. Accessions such as S5R1-19 (6.6cm heartwood core length) and S678R4-9 (5.5cm) show potential for breeding programmes aimed at improving heartwood production, enabling higher economic returns in shorter cultivation cycles. Heartwood colour classification revealed a predominance of light red variants (54.79%), followed by medium red (26.03%) and deep red (19.18%) colours. This variation is crucial for meeting specific market demands and enhancing the economic value of red sanders plantations.

Seed source evaluation

The significant differences in survival rates and growth traits among seed sources underscore the influence of genetic and environmental factors. Superior survival rates in seed sources such as Gumudipoondi, Nayapakam and Amirthi (100% survival) indicate their adaptability to environmental conditions. Conversely, the lower survival rates observed in Madireddipally (62%) warrant further investigation. The superior performance of accessions from Amirthi (tree height of 10.24m) and Petbasheerabad (high heartwood formation and survival rates) highlights their potential as elite seed sources for plantation programmes.

Trait correlations

The correlation analysis provided critical insights into trait relationships, guiding breeding and management strategies. The negative correlation between clear bole height and the number of forking stems (-0.796) indicates trade-offs in stem architecture, where selecting for one trait may compromise the other. The strong positive correlations between average GBH and traits such as heartwood core length (0.443), sapwood length (0.630), and total core sample length (0.916) suggest that GBH can serve as an integrative indicator of overall tree performance. These findings align with previous studies reporting similar positive relationships between GBH and heartwood content in 20 and 45-year-old plantations (Arun Kumar *et al*, 2017). Given the significant influence of age on heartwood development, selecting accessions with early heartwood formation can enhance breeding efficiency. Such an approach maximizes genetic variability utilization at younger ages, increasing the likelihood of identifying superior accessions for economic gains.

Principal component analysis

The PCA results indicate that PC1 primarily represents growth-related traits, whereas PC2 reflects structural variations such as stem architecture and clear bole height. The clustering further highlights natural groupings among the trees, which could be linked to genetic and environmental factors influencing tree morphology. Cluster 3 trees, with excessive branching and forked stems, might be less desirable for commercial timber purposes but could be valuable for conservation and biodiversity. In contrast, Cluster 1 trees, which have minimal branching and a high clear bole, might be more suitable for timber production, as fewer knots improve wood quality. Cluster 2 represents an intermediate growth pattern, possibly offering a balance between commercial and ecological benefits. This analysis highlights the multifaceted nature of trait variation and provides a robust framework for prioritizing traits in breeding programmes.

Implications for breeding and conservation

The observed diversity in morphological and heartwood formation traits forms a robust foundation for breeding programmes targeting the improvement of red sanders. Superior accessions such as S2R4-10 (tree height), S1R2-13 (clear bole height) and S5R1-19 (heartwood core length) should be prioritized in future breeding efforts. Integrating morphological, genetic and environmental data will enhance the precision of selection strategies. The identified clusters provide valuable insights for breeding programmes, selection strategies and conservation planning. Trees in Cluster 1 may be prioritized for commercial plantations, while Cluster 3 trees could be conserved for genetic diversity. Future studies integrating genomic data with phenotypic clustering could further enhance the understanding of trait inheritance and selection strategies in red sanders breeding programmes. The application of molecular tools such as genome-wide

association studies (GWAS) and transcriptomics can further elucidate the genetic basis of key traits, enabling marker-assisted selection for accelerated genetic improvement. From a conservation perspective, maintaining the genetic diversity observed in the evaluated germplasm is crucial to ensure the long-term sustainability and productivity of red sanders plantations.

Supplemental data

Supplemental Table 1. Morphological and heartwood formation traits data of red sanders accessions

Supplemental Table 2. Summary of principal components showing eigenvalues and percentage of variance explained

Supplemental Table 3. Principal component loadings of morphological and heartwood formation traits showing the contribution of each trait to the respective components in red sanders accessions

Acknowledgments

The authors thank the Director, ICFRE-Institute of Forest Biodiversity, Hyderabad, for his support. Additionally, we sincerely acknowledge the technical assistance provided by Mr Sardar and Mr B. Suresh Kumar.

Author contributions

Swapnendu Pattanaik and Kumbarahally Murthigowda Shivaprasad contributed to the study conception and design. Material preparation, data collection and analysis were performed by Kumbarahally Murthigowda Shivaprasad, Pendela Surath Kumar, Avula Kishore and S. Dinesh Kumar. The first draft of the manuscript was written by Kumbarahally Murthigowda Shivaprasad and Swapnendu Pattanaik and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing of interest.

Ethics statement

This study did not involve any human participants or animals. The research was conducted on cultivated germplasm of *Pterocarpus santalinus* maintained at the ICFRE–Institute of Forest Biodiversity, Hyderabad, India. No field sampling from wild populations was carried out, and no collection of plant material from protected or endangered habitats was involved. Therefore, no specific permits or ethical committee approvals were required to conduct this research.

References

Ahmed, M., Nayar, M. (1984). Red sanders tree (*Pterocarpus santalinus* Linn. F.) on the verge of depletion. *Nelumbo* 26:142-143. doi: https://doi.org/10.20324/nelumbo/v26/1984/74871

Ahmedullah, M. (2021). *Pterocarpus santalinus*. The IUCN Red List of Threatened Species 2021: e.T32104A187622484. doi: https://dx.doi.org/10.2305/IUCN.UK.2021-1.RLTS.T32104A187622484.en

Ahmedullah, M., Rasingam, L., Swamy, J., Nagaraju, S., Shankara Rao, M. (2019). Non-detriment findings report on the red sanders tree (*Pterocarpus santalinus* Lf). Botanical Survey of India (Deccan Regional Centre), MoEFCC, Hyderabad.

Arun Kumar, A., Joshi, G., Manikandan, S. (2017). Variability for heartwood content in three commercially important tree species of Peninsular India—*Hardwickia binata*, *Pterocarpus santalinus* and *Santalum album*. In Wood is Good: Current Trends and Future Prospects in Wood Utilization. Springer, pp. 117-126. doi: https://doi.org/10.1007/978-981-10-3115-1_12

Arunakumara, K. K. I. U., Walpola, B. C., Subasinghe, S., Yoon, M-H. (2011). *Pterocarpus santalinus* Linn. f. (Rath handun): a review of its botany, uses, phytochemistry and pharmacology. *Journal of the Korean Society for Applied Biological Chemistry* 54:495-500. doi: https://doi.org/10.3839/jksabc.2011.076

Arunkumar, A. (2011). Variability studies in *Pterocarpus santalinus* in different aged plantations of Karnataka. *My Forest* 47:343-353.

Hegde, M. T. (2023). Non-Detriment Findings (NDF) Study For Cultivated Red Sanders (*Pterocarpus santalinus* L. f.) Trees in Gujarat State. ICFRE - Arid Forest Research Institute (AFRI) Jodhpur, Rajasthan, p. 31.

NBA (2017). Report of the Expert Committee on Red Sanders. A Comprehensive Policy for Conservation, Sustainable use and Fair and Equitable Sharing of Benefits Arising from Utilisation of Red Sanders under the Biological Diversity Act, 2002. National Biodiversity Authority, Chennai.

New, S. (1981). The use of stain by furniture makers 1660—1850. *Furniture History* 17:51-60.

Pattanaik, S. (2024). Improvement of Redsanders (*Pterocarpus santalinus* L. f.) through the All India Co-ordinated Research Project (AICRP). Wood is good. IWST, Bengaluru, pp. 151-154.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Schmidlin, L., Poutaraud, A., Claudel, P., Mestre, P., Prado, E., Santos-Rosa, M., Wiedemann-Merdinoglu, S., Karst, F., Merdinoglu, D., Hugueney, P. (2008). A stress-inducible resveratrol O-methyltransferase involved in the biosynthesis of pterostilbene in grapevine. *Plant Physiology* 148:1630-1639. doi: https://doi.org/10.1104/pp.108.126003

Senthilkumar, N., Mayavel, A., Subramani, S., Balaji, K., Deenathayalan, P. (2015). Red sanders, *Pterocarpus santalinus* L. in Rajampet forest range, Rajampet forest division, Andhra Pradesh, *India. Adv Appl Sci Res* 6:130-134.

Seshadri, T. (1972). Polyphenols of Pterocarpus and Dalbergia woods. *Phytochemistry* 11(3):881–898. doi: https://doi.org/10.1016/S0031-9422(00)88430-7

Suresh, K., Hegde, M., Deenathayalan, P., Karthick Kumar, P., Thangapandi, M., Gurudev Singh, B., Krishnakumar, N. (2017). Variation in heartwood formation and wood density in plantation-grown Red Sanders (*Pterocarpus santalinus*). In Wood is Good: Current Trends and Future Prospects in Wood Utilization. Springer, pp. 139-151. doi: https://doi.org/10.1007/978-981-10-3115-1_14

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 171–193
DOI: 10.46265/genresj.SLTA9371
https://www.genresj.org
ISSN: 2708-3764

Biochemical characteristics of bread wheat genotypes related to SSR markers in moisture stress conditions

Fatemeh Bavandpouri*[,a], Ezatollah Farshadfar[a], Kianoosh Cheghamirza[a] and Mohsen Farshadfar[b]

[a]*Department of Plant Production and Genetics Engineering, Faculty of Agricultural Sciences and Engineering, Razi University, Kermanshah, Iran*
[b]*Forests and Rangelands Research Department, Kermanshah Agricultural and Natural Resources Research and Education Center, (AREEO), Kermanshah, Iran*

Abstract: Wheat is one of the oldest and most important staple crops worldwide, facing various biotic and abiotic stresses that affect its productivity. This study examines microsatellite markers related to grain yield, biochemical traits and drought tolerance indices in 25 wheat genotypes. The experiment was set up based on the randomized complete block design with three replications under rainfed and irrigated conditions. Combined variance analysis revealed significant differences among genotypes. Principal component analysis identified drought-tolerant genotypes (6, 10, 15, 18, 13, Pishtaz) linked to superior yield, stress indices, and antioxidant activity under rainfed conditions. Polymorphic SSR markers revealed key associations: XCFD168 with catalase, XGWM350 with ascorbic peroxidase (both under rainfed conditions), and XGWM136 with yield in irrigated conditions and multiple stress indices. Marker XGWM410(a1) was associated with yield in both environments, catalase in irrigated conditions, and multiple indices. Marker XGWM2(a2) was linked to yield in irrigated conditions, ascorbic peroxidase in rainfed conditions, and abiotic tolerance index, while XGWM124(a2) was associated with yield, superoxide dismutase in rainfed conditions, and multiple indices. The study identifies these genotypes as top candidates for drought tolerance due to their high yield and optimal biochemical responses under stress. Furthermore, key molecular markers – XCFD168, XGWM350, XGWM136, XGWM124(a2), XGWM410(a1), and XGWM2(a2) – associated with biochemical and yield traits are prioritized for marker-assisted selection to enhance drought tolerance and yield stability in breeding programmes.

Keywords: Antioxidant enzymes, bread wheat, genetic variation, molecular marker

Citation: Bavandpouri, F., Farshadfar, E., Cheghamirza, K. and Farshadfar, M. (2025) "Biochemical characteristics of bread wheat genotypes related to SSR markers in moisture stress conditions", *Genetic Resources*, 6(12), pp. 171–193. doi: 10.46265/genresj.SLTA9371.

Introduction

Wheat *(Triticum aestivum* L.) contributes to approximately one-third of the global food supply. The Food and Agriculture Organization of the UN (FAO) estimates that by 2050, an annual production of around 840 million tonnes of wheat will be required (Ma *et al*, 2022). However, wheat production is increasingly affected by various biotic and abiotic stresses that reduce crop yield and productivity. Among these, drought stress stands out as a major abiotic challenge, posing a significant threat to global food security, especially in the context of climate change (Sunil kumar *et al*, 2023). As a result, there is a critical need to identify and cultivate drought-tolerant, high-yielding genotypes to ensure sustainable food production and meet the demands of a growing global population (Galal *et al*, 2023). Drought stress in wheat triggers morphological, physiological, biochemical and molecular changes (Gupta *et al*, 2024; Rashid *et al*, 2022). Utilizing selection factor indicators can significantly improve the identification of genotypes that perform well in both optimal and stress conditions. A promising strategy to enhance wheat drought tolerance is to improve its antioxidant defense mechanisms (Gupta *et al*, 2024). Antioxidant enzymes are critical in protecting plants from oxidative damage caused by various environmental stresses.

Molecular markers associated with biochemical parameters can significantly expedite the identification of tolerant genetic materials in breeding programmes. The

* Corresponding author: Fatemeh Bavandpouri (f.bavandpori@yahoo.com)

Received: 15.03.2025 | Accepted: 18.08.2025 | Published online: 24.11.2025

simple sequence repeat (SSR) marker system is highly effective for detecting significant marker–trait associations in wheat germplasm (Pour-Aboughadareh *et al*, 2022). SSRs, also known as microsatellites, are short, tandemly repeated DNA sequences (typically 1–6 nucleotides in length) that are distributed genome-wide, exhibiting high polymorphism due to replication slippage in non-coding regions (Ellegren, 2004). Because of their multi-allelic nature, co-dominant inheritance, uniform genomic distribution and simple detection methodology, these markers are widely favoured for assessing genetic variation and analyzing population structures (Jabari *et al*, 2023; Ahmed *et al*, 2024).

Exploring the genetic foundations of quantitative traits in crops and understanding the relationship between DNA polymorphisms and phenotypic variations are essential for plant breeding programmes. Identifying quantitative trait loci (QTLs) linked to drought tolerance through marker-assisted selection is particularly important for crop improvement and represents a valuable strategy for boosting wheat yield (Zhao *et al*, 2023). Multivariate regression analysis (MRA) offers a fast and effective approach for establishing the association between traits and markers. A significant benefit of MRA is its capacity to pinpoint loci associated with quantitative traits. Furthermore, this method is both time-efficient and cost-effective (Vaillancourt *et al*, 2008) and does not require the creation of specialized populations for mapping.

The genetic diversity of 18 wheat genotypes was evaluated for drought tolerance using 25 microsatellite markers alongside morpho-physiological traits. Findings revealed that integrating these two approaches enhanced the efficiency of the screening process and provided more reliable outcomes for improving drought tolerance in wheat (Ahmed *et al*, 2023). A study investigating morphological, biochemical and genetic diversity for diagnosing salt tolerance in 18 wheat genotypes using SSR markers highlighted significant findings. The stepwise regression analysis emphasized the importance of root dry matter, relative turgidity and their respective contributions to shoot dry matter. Out of 23 SSR primers analyzed, 17 exhibited polymorphisms (Al-Ashkar *et al*, 2020). An association analysis performed on wild relatives of wheat in drought stress conditions, using 24 SSR markers, identified eight and nine significant marker-trait associations (MTAs) in control and drought stress conditions, respectively. Notably, two MTAs were consistently observed in both growth conditions (Pour-Aboughadareh *et al*, 2022). A study on Iranian wheat varieties and landraces employed agronomic traits and drought tolerance indices to identify significant SNP loci associated with drought-tolerance characteristics. The findings revealed that association mapping based on multiple drought tolerance indices can be highly effective in identifying critical markers for drought tolerance and uncovering linked gene networks (Rabieyan *et al*, 2023). Additionally, a research evaluation combining genetic and phenotypic analyses was conducted to identify drought-tolerant bread wheat genotypes using multivariate analysis techniques, including stepwise multiple linear regression. The results demonstrated that SSR markers were associated with nine agro-physio-biochemical traits, highlighting their utility as a valuable tool in the selection process for drought tolerance (Sallam *et al*, 2024a).

Despite these advances, very few studies have explored the association between molecular markers and biochemical traits whose activity increases in drought stress. Biochemical traits, like the accumulation of proline or antioxidants, are the measurable physiological responses of a plant to stress. Molecular markers are DNA sequences that can pinpoint the specific genes or genomic regions controlling these biochemical pathways (Oguz *et al*, 2022). Therefore, this research aimed to: (1) characterize bread wheat genotypes in terms of biochemical traits, grain yield, and drought tolerance indices, (2) analyze the impact of drought stress on wheat traits to enhance yield and drought tolerance, (3) evaluate the genetic diversity of wheat genotypes for drought tolerance using studied traits and SSR markers, and (4) investigate the association between the studied traits and indices with SSR markers and identify informative markers associated with grain yield, biochemical traits, and drought tolerance indices in wheat in both rainfed and irrigated conditions.

Materials and methods

Field experiment

Twenty-five bread wheat genotypes were evaluated, including two cultivars (Pishtaz and Pishgam) as controls, and 23 accessions of bread wheat (Table 1). The genotypes were sourced from the genebank of Karaj Seedling and Seed Breeding Research Institute, Iran. Field experiments were conducted during the 2018–2019 growing season using a randomized complete block design with three replications under rainfed and irrigated conditions in a cold Mediterranean climate (34°21'N, 47°9'E; altitude: 1,319m; mean annual rainfall: 430–460mm) in Iran. Each experimental plot consisted of five rows, each with a length of 2m, row spacing of 23cm, and a planting density of 400 seeds per square metre. The planting date (14 November 2018) coincided with the first irrigation, but no irrigation was provided to the rainfed plots during the growth period. The total rainfall during the experimental year was 401.51mm. For the irrigated treatment, three additional irrigations were applied: The first on 15 May, at the heading stage (50% spike emergence). The second in late May, after full spike emergence. The third on 14 June, during the seed milking stage. No chemical fertilizers were applied during the experiment, and weeding was performed manually.

Molecular experiment

For molecular evaluation of the studied genotypes, 20 pairs of SSR markers were utilized. DNA was extracted from two- to three-week-old seedlings grown from seeds using the cetyltrimethylammonium bromide (CTAB) method, based on the modified protocol of Doyle and Doyle (1987), in bulk. Genomic DNA was extracted from 50 mg of cryogenically homogenized tissue. Samples were suspended in 800μl extraction buffer (100ml containing: 4g CTAB, 16.36g NaCl, 3.15g Tris-HCl, 1.48g EDTA, and 400μl β-mercaptoethanol; pH 8.0) and incubated at 65°C for 30 min. After adding 800μl chloroform-isoamyl alcohol (24:1), samples were vortexed for 60 min, centrifuged (13,000 × g, 15 min), and the aqueous phase was transferred to fresh tubes. This phase was mixed with 500μl cold isopropanol and held at -20°C for 2 hours. Subsequent centrifugation (13,000 × g, 15 min) yielded DNA pellets, which were washed twice with 500μl cold 80% ethanol (brief centrifugation, supernatant removal). Pellets were air-dried and resuspended in 100μl nuclease-free H_2O. Extracted genomic DNA was evaluated by 0.8% agarose gel electrophoresis.

Table 1. List of bread wheat planting materials used for the study. IR, Iran; US, United States of America.

Genotype number	Genotype number and name	Origin
1	WC-4924	Kalat, IR
2	WC-4582	Kermanshah, IR
3	WC-4592	Kermanshah, IR
4	WC-47341	Montana, US
5	WC-4965	Kashan, IR
6	WC-4840	Sarakhs, IR
7	WC-4958	Badranloo, IR
8	WC-47399	Bulgaria
9	WC-4600	Kermanshah, IR
10	WC-4987	Unknown, IR
11	WC-47615	Mexico
12	WC-4612	Kordestan Babrar, IR
13	WC-5001	Unknown, IR
14	WC-4994	Unknown, IR
15	WC-47638	Peru
16	WC-47583	Canada
17	WC-47522	Mexico
18	WC-47569	Minnesota, US
19	Pishtaz	Pishtaz, IR
20	Pishgam	Pishgam, IR
21	WC-47640	Minnesota, US
22	WC-47467	Mexico
23	WC-4553	Kerend, IR
24	WC-4583	Kermanshah, IR
25	WC-4554	Kerend, IR

Polymerase Chain Reaction (PCR) was conducted in three temperature-dependent steps. DNA samples diluted to 10ng/μl were amplified using 20 primer pairs (primer sequences are provided in Table 2). PCR was performed in 20μl reaction volumes using a Bio-Rad thermocycler.

The PCR products were electrophoresed on a 3% agarose gel in 1x TBE buffer, stained with 10μl safe stain. DNA bands were visualized using a Quantum ST4 Gel Documentation system. As not all samples were loaded on the same gel, a 100–1500bp DNA size marker (producing 11 bands) was included. Band presence was scored as '1' and absence as '0', compiling the data into a matrix. Alleles detected in genotypes were designated *a*, *b*, *c*, and *d* for each marker. This matrix served as the foundation for subsequent statistical analyses based on electrophoretic band pattern.

Biochemical enzyme assays

Measurement methods of the biochemical traits were carried out as follows. Extraction buffer preparation: A 200ml Tris-HCl extraction buffer (pH 8.0) was prepared by dissolving 2.428g of Tris and 0.2g of PVP in 40ml of distilled water. The solution's pH was then adjusted to 8.0 using HCl, and the final volume was brought to 200ml with distilled water. The prepared buffer was stored at 4°C, protected from light with aluminium foil.

Enzyme extraction: Flag leaf samples were ground into a fine powder in liquid nitrogen. Subsequently, 250mg of the homogenized powder was combined with 1ml of the pre-cooled extraction buffer in a 2ml microtube. The mixture was vortexed for 30 seconds and then incubated at 4°C for 12 hours. During this incubation period, the samples underwent two additional 30-second vortexing steps at 2-hour intervals. Following incubation, the homogenate was centrifuged at 13,000 × g for 15 minutes at 4°C. The resulting clear supernatant was carefully collected for the subsequent analysis of soluble protein content and antioxidant enzyme activities (Ramachandra Reddy *et al*, 2004). Meanwhile, the BioTek PowerWave XS2 microplate reader was used to measure biochemical traits.

Peroxidase Activity (POD) was assayed according to the method of Chance and Maehly (1995) with slight modifications by combining 6.6μl of diluted enzyme extract (1:4) with 200μl of substrate solution [408.71μl guaiacol + 78.3μl 0.9 M H_2O_2 in 50mM potassium phosphate buffer (pH 7.0)]. After a 15-minute incubation, absorbance was measured at 470nm every 30s.

Superoxide Dismutase Activity (SOD) was assayed following the method of Beauchamp and Fridovich (1971). The assay solution consisted of 50 mM potassium phosphate buffer (pH 7.8), 12.26mg nitroblue tetrazolium (NBT), 387.92mg L-methionine, 1mM EDTA, and 0.04mM riboflavin (stored in light-protected containers). For the assay, 196, 197, 198, and 199μl of extraction buffer were mixed with 4, 3, 2, and 1μl of diluted enzymatic extract (1:4), respectively, to achieve 200μl reaction mixtures. These mixtures were transferred to a 96-well microplate, followed by addition of 10μl riboflavin solution under dark conditions. After 30 min illumination in a light chamber, absorbance was measured at 560nm.

Catalase Activity (CAT) was assayed according to the method of Sinha (1972) with slight modifications. The assay was performed by combining 1.5μl of diluted enzymatic extract (1:4) with 150μl of 50 mM phosphate buffer (pH 7.0). The reaction was initiated by adding 75μl of 0.32 mM hydrogen peroxide solution. At timed intervals (2, 4, 6 and 8 min post-initiation), 62μl of dichromate reagent was rapidly added to each tube with immediate vortexing. Tubes were transferred to a preheated 95°C water bath for 10 min. After chromogenic development (green-to-yellow gradient), samples were centrifuged at (10,000g, 5 min), and the supernatant absorbance was measured at 570 nm.

Ascorbic Peroxidase Activity (APX) was assayed according to the method of Nakano and Asada (1981). The reaction was initiated by adding 50μl of the enzymatic extract to 1ml of an assay solution containing 50 mM potassium phosphate buffer (pH 7.0), 0.1 mM EDTA, 0.5 mM ascorbic acid (ASA), and 0.15 mM hydrogen peroxide (H_2O_2). Absorbance at 290 nm was recorded every 10 s for 1 min.

Soluble Protein Concentration (PROTEIN) was determined using the method of Bradford (1976). For the assay, 1μl of the extracted sample was mixed with 200μl of freshly prepared Coomassie Brilliant Blue G-250 dye reagent. After 15 min incubation, absorbance at 595nm was measured, with dye reagent as the blank. Protein concentration was calculated from a bovine serum albumin (BSA) standard curve (0-1,500μg/ml).

Table 2. SSR primers used to assess the genetic diversity of bread wheat genotypes.

No.	Name	(5'-3') Sequence	TM	GC%	Band size	References
1	XGWM350-7D-F	5' ACCTCATCCACATGTTCTACG 3'	57	47.6	150bp	Haque *et al* (2020); Bavandpouri *et al* (2025)
	XGWM350-7D-R	5' GCATGGATAGGACGCCC 3'		64.7		
2	XGWM334-6A-F	5' AATTTCAAAAAGGAGAGAGA 3'	50	30	100bp	Rosewarne *et al* (2013); Halder *et al* (2023)
	XGWM334-6A-R	5' AACATGTGTTTTTAGCTATC 3'		30		
3	XGWM155-3A-F	5' CAATCATTTCCCCCTCCC 3'	58	55.6	100bp	Ahmed *et al* (2020); El-Rawy and Hassan (2021)
	XGWM155-3A-R	5' AATCATTGGAAATCCATATGCC 3'		36.4		
4	XGWM577-7B-F	5' ATGGCATAATTTGGTGAAATTG 3'	56	31.8	150bp	El-Rawy and Hassan (2021); El-demery *et al* (2022); Firouzian *et al* (2023)
	XGWM577-7B-R	5' TGTTTCAAGCCCAACTTCTATT 3'		36.4		
5	XGWM70-6B-F	5' AGTGGCTGGGAGAGTGTCAT 3'	52.5	55	200bp	Batool *et al* (2018); Ilyas *et al* (2020)
	XGWM70-6B-R	5' GCCCATTACCGAGGACAC 3'		61.6		
6	XGWM642-1D-F	5' ACGGCGAGAAGGTGCTC 3'	58	45	180–200bp	Islam *et al* (2012); Ahmed *et al* (2024)
	XGWM642-1D-R	5' CATGAAAGGCAAGTTCGTCA 3'		64.7		
7	XGWM136-1A-F	5' GACAGCACCTTGCCCTTTG 3'	52	57.9	250bp	Budak *et al* (2013); Kaur *et al* (2016); Khan *et al* (2021); Bavandpouri *et al* (2025)
	XGWM136-1A-R	5' CATCGGCAACATGCTCAT 3'		52.6		
8	XGWM124-1B-F	5' GCCATGGCTATCACCCAG 3'	57.5	61.1	200bp	Amalova *et al* (2024); Bavandpouri *et al* (2025)
	XGWM124-1B-R	5' ACTGTTCGGTGCAATTTGAG 3'		45		
9	XGWM265-2A-F	5' TGTTGCGGATGGTCACTATT 3'	58.5	45	150bp	Choudhary *et al* (2016); Kumari *et al* (2025)
	XGWM265-2A-R	5' GAGTACACATTTGGCCTCTGC 3'		52.4		
10	XGWM410-2B-F	5' GCTTGAGACCGGCACAGT 3'	51	61.6	250bp	Maccaferri *et al* (2011); Naroui Rad *et al* (2012)
	XGWM410-2B-R	5' CGAGACCTTGAGGGTCTAGA 3'		55		
11	XGWM165-4B-F	5' TGCAGTGGTCAGATGTTTCC 3'	50.6	50	200bp	Ahmed *et al* (2020); El-Rawy and Hassan (2021)
	XGWM165-4B-R	5' CTTTTCTTTCAGATTGCGCC 3'		45		
12	XGWM4-4A-F	5' GCTGATGCATATAATGCTGT 3'	52.5	40	250bp	Mallick *et al* (2022a)
	XGWM4-4A-R	5' CACTGTCTGTATCACTCTGCT 3'		47.6		
13	XGWM192-5D-F	5' GGTTTTCTTTCAGATTGCGC 3'	50.7	45	100bp	Islam *et al* (2012); Heidari *et al* (2024)
	XGWM192-5D-R	5' CGTTGTCTAATCTTGCCTTGC 3'		47.6		
14	XGWM233-7A-F	5' TCAAAACATAAATGTTCATTGGA 3'	46.7	26.1	100bp	Mallick *et al* (2022a)
	XGWM233-7A-R	5' TCAACCGTGTGTAATTTTGTCC 3'		40.9		
15	XGWM2-3D-F	5' CTGCAAGCCTGTGATCAACT 3'	49.4	50	250bp	Ahmed *et al* (2020); Kumari *et al* (2025)
	XGWM2-3D-R	5' CATTCTCAAATGATCGAACA 3'		35		
16	XCFD5-5B-F	5' TGCCCTGTCCACAGTGAAG 3'	59.5	57.9	200bp	Ahmed *et al* (2020); Mallick *et al* (2022a)
	XCFD5-5B-R	5' TTGCCAGTTCCAAGGAGAAT 3'		45		
17	XGWM129-5A-F	5' TCAGTGGGCAAGCTACACAG 3'	50.6	55	250bp	Ahmed *et al* (2020); Mallick *et al* (2022a)
	XGWM129-5A-R	5' AAAACTTAGTAGCCGCGT 3'		44.4		
18	XCFD168-2D-F	5' CTTCGCAAATCGAGGATGAT 3'	56	45	250bp	Khan *et al* (2021); Bavandpouri *et al* (2025)
	XCFD168-2D-R	5' TTCACGCCCAGTATTAAGGC 3'		50		
19	XGWM234-5B-F	5' GAGTCCTGATGTGAAGCTGTTG 3'	54	50	220–230bp	Khan *et al* (2021); Mallick *et al* (2022b)
	XGWM234-5B-R	5' CTCATTGGGGTGTGTACGTG 3'		55		
20	XGWM33-1A-F	5' GGAGTCACACTTGTTTGTGCA 3'	59	47.6	100bp	Ahmed *et al* (2020); Mallick *et al* (2022a)
	XGWM33-1A-R	5' CACTGCACACCTAACTACCTGC 3'		45.5		

Malon-Dialdehyde (MDA) was determined according to the method of Heath and Packer (1968). Briefly, 0.25g of wheat leaves were homogenized in 500μl ice-cold 1.0% (w/v) trichloroacetic acid (TCA) using a porcelain mortar. The homogenate was centrifuged at 10,000 × g for 5 min at 4°C. Subsequently, 250μl of the supernatant was reacted with 1mL of thiobarbituric acid (TBA) reagent [0.5% (w/v) TBA in 20% (w/v) TCA]. The mixture was incubated at 95°C for 30 min in a water bath, then immediately cooled on ice and centrifuged again (10,000 × g, 10 min, 4°C). A 200μl aliquot of the resulting chromogenic supernatant was transferred to a 96-wall microplate. Absorbance was measured at 532 and 600nm.

Proline Concentration (PC) was determined according to the method of Bates *et al* (1973). Briefly, 0.05g of fresh leaf tissue was homogenized in 1mL of ice-cold 3% (w/v) sulfosalicylic acid using a pre-chilled mortar. The homogenate was centrifuged at 4,000 × g for 15 min (4°C). A 10μl aliquot of the resulting supernatant was then reacted with 200μl of acid-ninhydrin reagent [1.25 g ninhydrin in 30ml glacial acetic acid + 20ml 6 M phosphoric acid] and 200μl of glacial acetic acid. Tubes were incubated at 95°C for 60 min, immediately cooled on ice for 5 min, and then mixed with 400μl toluene via 30-second vortexing. After a 20-minute phase separation at 25°C, the upper toluene layer was transferred to a 96-well microplate. Absorbance was measured at 520nm with pathlength correction. Proline concentration was determined from a standard curve (0-20μg/ml).

Statistical analysis

Combined variance analysis based on the data obtained from the evaluation of 25 genotypes, including two cultivars and 23 accessions, was performed to determine the contribution of the main effects of genotype, irrigated conditions, and their interaction using SAS 9.1.3 software. A comparison of mean genotypes by the Least Significant Difference (LSD) test was performed. A bar graph related to the mean comparison was drawn in Excel. PCA was calculated based on the means of traits and genotypes. Principal components analysis was carried out using the Minitab16 software, and correlations between the studied traits and indicators were analyzed using the "corrplot" package in R-Studio version 4.5 (R Core Team, 2025). To analyze the differences among the studied genotypes using SSR molecular markers, analysis of molecular variance (AMOVA) was performed by GenAlex software version 6.502. The association between SSR markers, field-measured traits, and biochemical traits was analyzed using stepwise multiple regression in SPSS 26 software. Each quantitative trait was treated as a dependent variable, while the SSR markers served as independent variables. The studied traits and indices were measured in the field and molecular experiment section, as shown in Table 3.

Results

Analysis of combined variance and mean compression

The combined analysis of variance for grain yield and biochemical characteristics is presented in Table 4. Significant differences were observed across various irrigated conditions for all characteristics. Genotypes showed significant variation for all traits except soluble protein. Furthermore, the genotype-by-irrigated interaction effect was significant for most biochemical traits, except for grain yield and malon-dialdehyde.

The mean comparison (mean of three replications) of genotypes based on the studied traits in rainfed and irrigated conditions, presented in the form of a bar graph, is as follows. Genotype 10 showed the highest grain yield under rainfed and irrigated conditions (Figure 1, Chart GY) with values of 424.73 and 565.75, respectively. The maximum peroxidase (POD) activity in rainfed and irrigated conditions was observed in genotype 6 (0.49) and genotype 18 (0.34), respectively (Figure 1, Chart POD). For superoxide dismutase (SOD), the highest values in rainfed and irrigated conditions belonged to genotype 15 (1.02) and genotype 12 (0.64), respectively (Figure 1, Chart SOD). Catalase (CAT) activity was most significant in genotype 12 (3.01) in rainfed conditions and genotype 24 (1.56) in irrigated conditions (Figure 1, Chart CAT). The highest soluble protein content was found in genotype 14 (112.03) in rainfed conditions and genotype 12 (167.09) in irrigated conditions (Figure 1, Chart PROTEIN). Proline (PC) levels were highest in genotype 15 (10.14) in rainfed conditions and genotype 8 (7.24) in irrigated conditions (Figure 1, Chart PC). The maximum ascorbic peroxidase (APX) activity was recorded for genotype 6 (418.12) in rainfed conditions and genotype 15 (263.35) in irrigated conditions (Figure 1, Chart APX). Finally, the highest malondialdehyde (MDA) values in both conditions were observed in genotypes 23 and 24 (0.45) in rainfed conditions and genotype 23 (0.42) in irrigated conditions (Figure 1, Chart MDA). Complete information on the comparison of the mean genotypes for each trait is shown in Table 5.

Assessment of broad-sense heritability and genetic gain of studied traits in rainfed and irrigated conditions

The estimation of broad-sense heritability and genetic gain for grain yield and biochemical traits under rainfed conditions is summarized in Table 6. In rainfed conditions, the average broad-sense heritability and genetic gain for grain yield were 0.278 and 16.08%, respectively. Almost all biochemical traits exhibited heritability above 0.90, including PC (0.998), SOD (0.997), CAT (0.983), and APX (0.972). Among these, PC showed the highest heritability. For genetic gain, CAT (92.022%), SOD (89.91%), APX (67.62%), and PC (63.28%) were most significant, with CAT ranking highest. Under irrigated conditions, heritability and genetic gain for grain yield were 0.604 and 33.20%, respectively. The traits CAT (0.997), protein (0.989), PC (0.979), SOD (0.971), APX (0.966) and POD (0.929) all demonstrated high heritability (> 0.90), with CAT showing the highest value. Also, CAT exhibited the most significant genetic gain (133.7%), followed by SOD (86.9%), PC (83.19%), APX (80.39%). In both conditions, the MDA trait showed the lowest heritability and genetic gain.

Table 3. Measurement methods of the studied traits and indices. Yi: Yield of a genotype under irrigated conditions; Yr: Yield of a genotype under rainfed conditions; $\overline{Y}_i$: Mean yield of all genotypes under irrigated conditions; $\overline{Y}_r$: Mean yield of all genotypes under rainfed conditions; δ_g^2: Genotypic variance; δ_p^2: Phenotypic variance; $\overline{X}$: Overall mean of the trait; TCP: Trait Changes Percentage; MTIC: Mean of the trait under irrigated conditions; MTRC: Mean of the trait under rainfed conditions. The PIC index for SSR markers was calculated based on allele frequency at each locus across all genotypes. In the calculation of the RP index, Pi refers to the proportion of genotypes that possess a particular band. Cov (x1x2): Covariance between variables x1 and x2. V(x1): variance of one trait (x1). V(x2): variance of other trait (x2).

Traits and Indices	Measurement method and formulas
GY: Grain Yield	The grain weight from three 1m sections of the middle rows per plot.
ATI: Abiotic Tolerance Index (Moosavi *et al*, 2008)	$ATI = \left[\frac{Y_i - Y_r}{\frac{\overline{Y}_i}{\overline{Y}_r}}\right] \times \sqrt{(Y_r \times Y_i)}$
SSPI: Stress Susceptibility Percentage Index (Moosavi *et al*, 2008)	$SSPI = \left[\frac{Y_i - Y_r}{2\overline{Y}_i}\right] \times 100$
TOL: Tolerance (Rossielli and Hamblin, 1981)	$TOL = Y_i - Y_r$
MP: Mean Productivity (Rossielli and Hamblin, 1981)	$MP = (Y_r + Y_i)/2$
GMP: Geometric Mean Productivity (Fernandez, 1992)	$GMP = \sqrt{(Y_r \times Y_i)}$
HMP: Harmonic Mean Productivity (Fernandez, 1992)	$HMP = \frac{2Y_i \times Y_r}{Y_i + Y_r}$
STI: Stress Tolerance Index (Fernandez, 1992)	$STI = \frac{Y_i \times Y_r}{\overline{Y}_i^{\,2}}$
SSI: Stress Susceptibility Index (Fischer and Maurer, 1978)	$SSI = \left(1 - (Y_r / Y_i)\right) / \left(1 - (\overline{Y}_r / \overline{Y}_i)\right)$
PEV: Press Evaluation (Bouslama and Schapaugh, 1984)	$PEV = 1 - \left(\frac{Y_r}{Y_i}\right)$
RDY: Relative Decrease in Yield (Emre *et al*, 2011)	$RDY = 100 - (Y_r / 100 \times Y_i)$
h2b.s , GG: broad-sense Heritability and Genetic Gain (Kearsey and Pooni (1996) and the GLM MANOVA analysis in SAS 9.3.1 software)	$h^2_{b.s} = \frac{\delta_g^2}{\delta_p^2}$ $GG = \frac{\left(\left(2.06.\frac{\delta_G^2}{\sqrt{\delta_P^2}}\right).100\right)}{\overline{X}}$
Correlation (Miller *et al*, 1958)	$r(x_1 x_2) = \frac{Cov(x_1 x_2)}{\sqrt{V(x_1) V(x_2)}}$
TCP%: percentage of changes in the irrigated environment compared to rainfed for traits (Nourmand-moaied *et al*, 2001)	$TCP = \frac{MTIC - MTRC}{MTIC} \times 100$
Polymorphic percentage (Mohammadi and Prasanna, 2003)	The number of polymorphic bands is divided by the total number of amplified bands and multiplied by 100.
PIC: Polymorphic Information Content Index (Anderson *et al*, 1993)	$PIC = 1 - \sum pi^2$
MI: Marker Index (Kumar *et al*, 2009)	The number of polymorphic bands was multiplied with the PIC value.
EMR: Effective Multiplex Ratio Index (Kumar *et al*, 2009)	This index was obtained by multiplying the percentage of polymorphic loci by the number of polymorphic loci.
RP: Resolving Power (Altintas *et al*, 2008)	$RP = \sum IB$ $IB = 1 - \left[2 \times (0.5 - Pi)\right]$

Table 4. Analysis of combined variance in both rainfed and irrigated conditions for grain yield and biochemical characteristics in 25 bread wheat genotypes. ns, not significant; *, significant at 5% probability level; **, significant at 1% probability level; S.O.V, Source of variations; Error 1, is nesting the replication in the irrigated factor; Error 2, is the total error of the experiment.

S.O.V	df	Grain yield	Peroxidase activity	Superoxide dismutase activity	Catalase activity	Soluble protein	Proline content	Ascorbic peroxidase	Malon-dialdehyde
Irrigated	1	470283.30**	0.37**	2.58**	33.22**	32924.41**	272.49**	489561.60**	0.05**
Error 1	4	21490.83	0.0003	0.0002	0.0004	11.35	0.02	178.19	0.001
Genotype	24	28250.64**	0.02**	0.17**	1.42*	1416.56^{ns}	16.07**	21384.36*	0.005**
Genotype× Irrigated	24	6237.08^{ns}	0.006**	0.05**	0.63**	855.42**	5.44**	8793.39**	0.001^{ns}
Error 2	96	4496.36	0.001	0.0003	0.005	25.39	0.04	153.87	0.001
(C.V) %		19.74	8.89	4.08	6.28	5.07	3.44	6.39	10.07

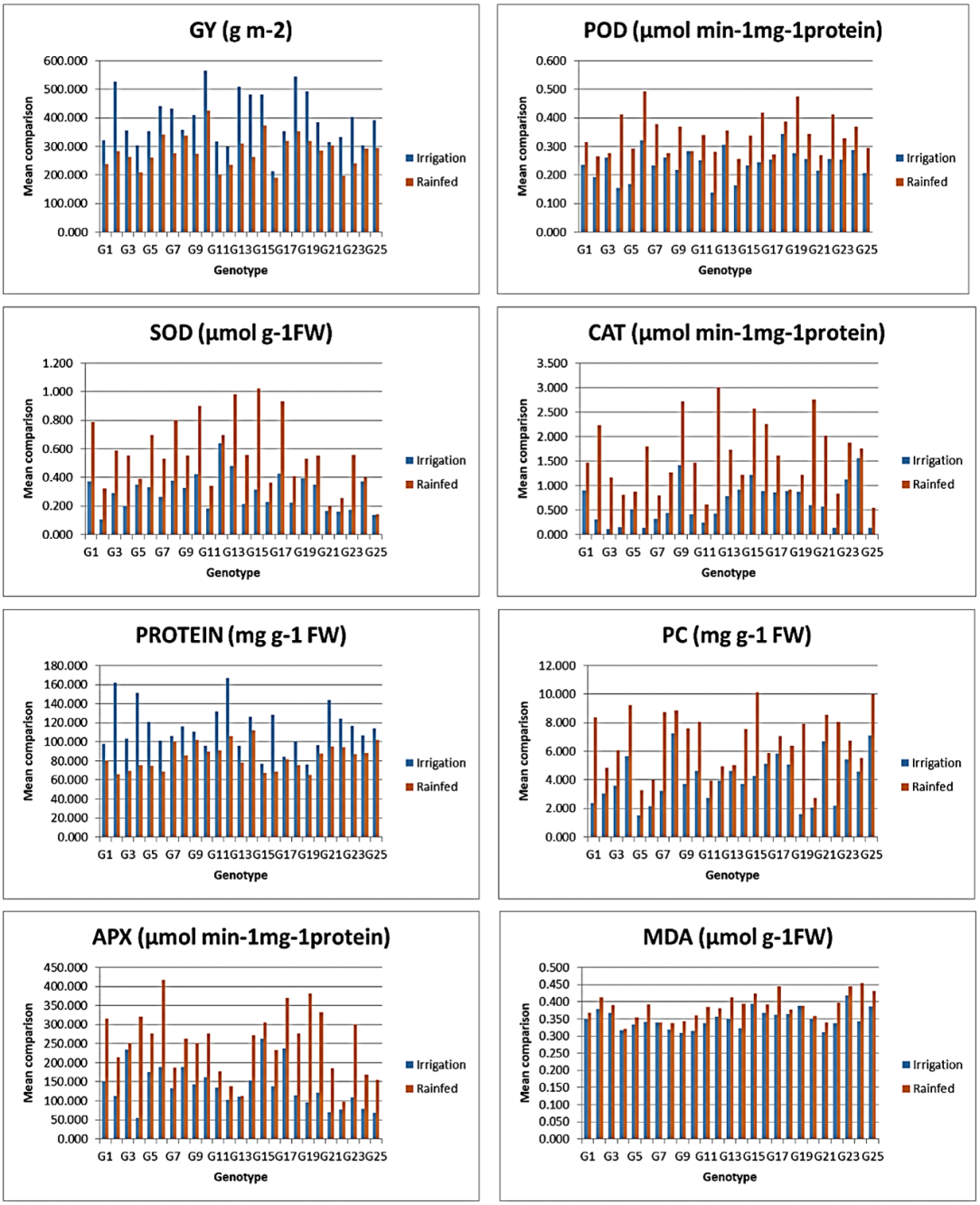


Figure 1. Bar graphs related to the comparison of mean genotypes in rainfed and irrigated conditions. GY, Grain Yield; POD, Peroxidase Activity; SOD, Superoxide Dismutase Activity; CAT, Catalase Activity; PROTEIN, Soluble Protein; PC, Proline Concentration; APX, Ascorbic Peroxidase Activity; MDA, Malon-dialdehyde.

Table 5. Mean comparison of bread wheat genotypes under rainfed and irrigated conditions for the studied traits. G, genotype; R, rainfed; I, irrigated.

G	Grain Yield (g m^2)		Peroxidase Activity (μmol min^{-1}mg^{-1}protein)		Superoxide Dismutase Activity (μmol g^{-1}FW)		Catalase Activity (μmol min^{-1}mg^{-1}protein)		Soluble Protein (mg g^{-1} FW)		Proline Content (mg g^{-1} FW)		Ascorbic Peroxidase (μmol min^{-1}mg^{-1}protein)		Malon-dialdehyde (μmol g^{-1}FW)	
	R	I	R	I	R	I	R	I	R	I	R	I	R	I	R	I
G1	239.38	322.20	0.32	0.24	0.79	0.37	1.48	0.90	80.06	98.06	8.38	2.39	316.01	151.37	0.37	0.35
G2	238.40	526.59	0.27	0.19	0.32	0.11	2.24	0.31	66.09	162.09	4.88	3.03	213.84	113.37	0.41	0.38
G3	262.82	355.60	0.28	0.26	0.59	0.29	1.17	0.12	69.76	103.06	6.07	3.60	251.30	234.07	0.39	0.37
G4	208.83	304.55	0.41	0.15	0.55	0.20	0.82	0.15	75.49	151.33	9.23	5.67	320.57	54.87	0.32	0.32
G5	261.51	353.34	0.29	0.17	0.39	0.35	0.88	0.53	75.09	120.79	3.27	1.53	276.49	175.16	0.36	0.33
G6	342.69	442.37	**0.49**	0.32	0.70	0.33	1.80	0.15	68.96	101.36	4.02	2.14	**418.12**	188.34	0.39	0.34
G7	277.26	431.71	0.38	0.24	0.53	0.27	0.80	0.33	100.46	106.23	8.72	3.23	187.76	133.83	0.34	0.34
G8	338.53	358.07	0.28	0.26	0.80	0.38	1.27	0.45	85.69	116.39	8.86	**7.24**	263.08	189.28	0.34	0.32
G9	275.12	410.24	0.37	0.22	0.55	0.33	2.71	1.41	101.79	110.73	7.59	3.71	250.79	143.65	0.34	0.31
G10	**424.73**	**565.75**	0.28	0.28	0.90	0.42	1.47	0.42	89.43	95.69	8.08	4.62	276.15	161.90	0.36	0.32
G11	201.89	317.21	0.34	0.25	0.34	0.19	0.61	0.25	90.99	131.69	3.93	2.75	176.85	133.93	0.39	0.34
G12	236.94	298.98	0.28	0.14	0.70	**0.64**	**3.01**	0.43	105.99	**167.09**	4.96	3.97	138.82	103.03	0.38	0.36
G13	309.49	508.45	0.36	0.31	0.98	0.48	1.73	0.79	78.27	95.79	5.05	4.65	113.31	110.31	0.41	0.35
G14	263.96	482.62	0.26	0.16	0.56	0.21	1.22	0.92	**112.03**	126.23	7.55	3.71	271.23	153.73	0.39	0.32
G15	372.95	482.01	0.34	0.23	**1.02**	0.32	2.58	1.22	67.43	76.73	**10.14**	4.28	304.86	**263.35**	0.42	0.40
G16	190.15	214.21	0.42	0.24	0.37	0.23	2.26	0.89	68.93	128.49	5.88	5.13	232.76	137.50	0.39	0.37
G17	319.10	354	0.27	0.25	0.93	0.43	1.61	0.86	81.96	84.19	7.09	5.86	370.74	238.03	0.44	0.36
G18	354.27	544.26	0.39	**0.34**	0.41	0.22	0.91	0.89	75.86	100.53	6.39	5.09	276.11	115.07	0.38	0.36
C19	318.70	492.24	0.48	0.28	0.53	0.39	1.22	0.88	65.59	76.06	7.91	1.63	382.05	96.46	0.39	0.39
C20	285.14	384.01	0.35	0.26	0.56	0.35	2.76	0.60	87.53	96.66	2.75	2.06	333.49	121.35	0.36	0.35
G21	303.47	313.97	0.27	0.22	0.20	0.17	2.03	0.58	94.93	144.13	8.55	6.72	184.70	71.20	0.34	0.31
G22	197.36	331.92	0.41	0.26	0.26	0.16	0.84	0.14	94.53	124.46	8.07	2.22	98.22	77.42	0.40	0.34
G23	240.30	403.74	0.33	0.26	0.56	0.17	1.87	1.13	87.19	116.59	6.76	5.43	300.27	109.35	**0.45**	**0.42**
G24	292.04	304.24	0.37	0.29	0.40	0.37	1.76	**1.56**	88.33	106.69	5.56	4.57	167.80	79.10	**0.45**	0.34
G25	293.72	391.08	0.30	0.21	0.14	0.14	0.55	0.14	101.81	113.89	10.02	7.09	155.52	68.74	0.43	0.39
LSD 5%	111.12	109.03	0.06	0.02	0.02	0.04	0.16	0.04	10.94	4.14	0.19	0.39	23.50	16.65	0.06	0.06

Analysis of trait-index correlations in wheat genotypes under rainfed and irrigated conditions

The correlation patterns between studied traits and drought tolerance indices revealed distinct profiles across conditions. Yield in irrigated conditions (Y_i) showed strong positive correlations ($p < 0.01$) with rainfed Yield (Yr; 0.71) and the indices STI (0.92), MP (0.96), GMP (0.93), HMP (0.91), SSI (0.52), TOL (0.78), ATI (0.90), SSPI (0.78), and PEV (0.52), but exhibited a significant negative association with RDY (-0.92). Similarly, Yr demonstrated strong positive correlations with STI (0.91), MP (0.88), GMP (0.91), and HMP (0.94) ($p < 0.01$), while having a negative correlation with RDY (-0.91).

In irrigated environments, POD activity correlated positively ($p < 0.05$) with its rainfed counterpart (0.44) and the indices STI (0.44), MP (0.41), GMP (0.43), and HMP (0.44), yet displayed a negative relationship with RDY (-0.44). SOD enzyme activity showed high consistency between the two conditions (irrigated vs. rainfed: 0.68; $p < 0.01$). Rainfed SOD further correlated positively with STI (0.43), MP (0.41), GMP (0.43), and HMP (0.44) ($p < 0.05$) but negatively with RDY (-0.43). CAT activity also followed this pattern with significant concordance between irrigated and rainfed conditions (0.44; $p < 0.05$).

For PC in irrigated conditions, inverse correlations appeared with STI (-0.50), MP (-0.48), GMP (-0.51), and HMP (-0.53) ($p < 0.05/p < 0.01$), in contrast to its positive linkage with RDY (0.50; $p < 0.05$). PC also aligned with its rainfed equivalent (0.51; $p < 0.01$) and showed negative associations with SSI (-0.47) and PEV (-0.47) ($p < 0.05$). APX and MDA activities maintained significant consistency between the two conditions (APX: 0.46, MDA: 0.69; $p < 0.05/p < 0.01$).

Inter-index correlations revealed tightly coupled networks: STI exhibited near-perfect positive alignment with MP (0.99), GMP (0.99), and HMP (0.99) ($p < 0.01$), moderate ties to ATI (0.69), TOL (0.49), and SSPI (0.49) ($p < 0.05$), and a complete inverse correlation with RDY (-1.00; $p < 0.01$). The MP, GMP, and HMP indices showed nearly identical mutual relationships (0.99–1.00; $p < 0.01$) and positive associations with ATI (0.64–0.74), TOL (0.44–0.56), and SSPI (0.44–0.56) ($p < 0.05/p < 0.01$), while uniformly opposing RDY (-0.99 to -1.00; $p < 0.01$).

SSI correlated strongly with PEV (1.00), TOL (0.93), SSPI (0.93), and ATI (0.80) ($p < 0.01$). TOL demonstrated positive linkages with SSPI (1.00), ATI (0.96), and PEV (0.93) ($p < 0.01$) but a negative correlation with RDY (-0.49; $p < 0.05$). ATI correlated positively with SSPI (0.96) and PEV (0.80) ($p < 0.01$) and negatively with RDY (-0.69; $p < 0.01$). Finally, SSPI and PEV shared a strong positive correlation (0.93; $p < 0.01$), while SSPI was inversely associated with RDY (-0.49; $p < 0.05$) (Figure 2).

Table 6. Estimation of broad-sense heritability and genetic gain for grain yield and biochemical characteristics in bread wheat genotypes in rainfed and irrigated conditions. GY, Grain Yield; POD, Peroxidase Activity; SOD, Superoxide Dismutase Activity; CAT, Catalase Activity; PROTEIN, Soluble Protein; PC, Proline Concentration; APX, Ascorbic Peroxidase Activity; MDA, Malon-dialdehyde.

Conditions	Traits	Mean	h^2_{bs}	GG
Rainfed	GY	283.75	0.278	16.08
	POD	0.340	0.750	28.74
	SOD	0.560	0.997	89.91
	CAT	1.58	0.983	92.022
	Protein	84.57	0.791	28.044
	PC	6.79	0.998	63.28
	APX	251.23	0.972	67.62
	MDA	0.390	0.250	6.82
Irrigated	GY	395.75	0.604	33.203
	POD	0.240	0.929	42.18
	SOD	0.300	0.971	86.9
	CAT	0.640	0.997	133.7
	Protein	114.2	0.989	42.932
	PC	4.09	0.979	83.19
	APX	136.98	0.966	80.39
	MDA	0.350	0.250	5.373

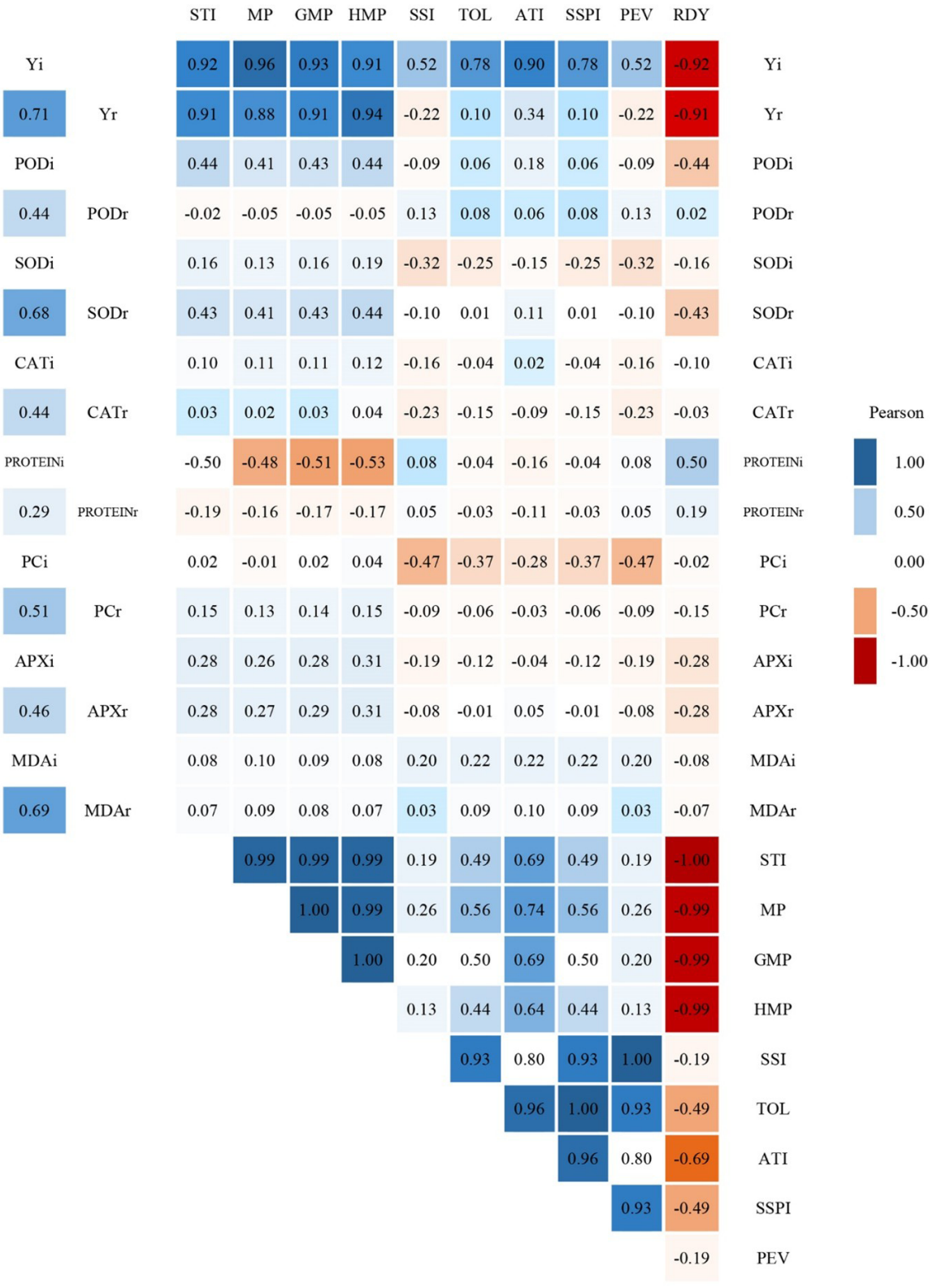


Figure 2. Heatmaps of Pearson's correlation coefficients between the studied characteristics and drought tolerance indices in 25 wheat genotypes in rainfed and irrigated conditions. Yi, Yield in irrigated conditions; Yr, Yield in rainfed conditions; i, irrigated; r, rainfed; POD, Peroxidase Activity; SOD, Superoxide Dismutase Activity; CAT, Catalase Activity; PROTEIN, Soluble Protein; PC, Proline Concentration; APX, Ascorbic Peroxidase Activity; MDA, Malon-dialdehyde; STI, Stress Tolerance Index; MP, Mean productivity; GMP, Geometric Mean Productivity; HMP, Harmonic Mean Productivity; SSI, Stress Susceptibility Index; TOL, Tolerance; ATI, Abiotic Tolerance Index; SSPI, Stress Susceptibility Percentage Index; PEV, Press Evaluation; RDY, Relative Decrease in Yield. Negative and positive correlations are indicated by red and blue cells, respectively. Color darkness scales with correlation strength (|*r*|) (Significance: *r* ≥ 0.40 at *p* < 0.05*; *r* > 0.50 at **p* < 0.01).

Principal components analysis and biplot graphic display based on drought tolerance indices and studied traits in rainfed conditions

PCA is calculated based on the mean of traits and genotypes. The results of both are shown in Figure 3. The PCA result for traits is presented in Table 7. It demonstrates that the first four components, with eigenvalues greater than one, contributed the most to explaining the variance in the dataset. Specifically, the first component explained 42.303% of the variance. The second component accounted for 21.78%. The third component contributed 9.612%. The fourth component explained 7.353%. Together, these four components explained 81.05% of the variance.

The first component was characterized by positive and high coefficients for the grain yield trait and the MP, GMP, HMP, STI, ATI, TOL and SSPI indices, as well as negative and high coefficients for the RDY index. This component was labelled the drought-tolerant PCA. The second component had positive and high coefficients for the SSI, PEV, TOL, SSPI and ATI indices, along with negative and high coefficients for grain yield and the superoxide dismutase enzyme. This component was referred to as the drought-stress PCA. The third component showed positive and high coefficients for the soluble protein and proline traits, while having negative and high coefficients for the peroxidase and ascorbic peroxidase enzymes. The fourth component was defined by positive and high coefficients for the catalase enzyme activity and malon-dialdehyde traits, and negative and high coefficients for proline and peroxidase enzyme.

According to the data, a biplot of the first two principal components was generated to analyze the traits and indicators under investigation. Based on the biplot (Figure 3), genotypes 10, 15, 6, 18, 13, and the Pishtaz cultivar, which were positioned near the vectors corresponding to the most effective drought tolerance indicators (MP, STI, GMP and HMP), demonstrated high yields in rainfed and irrigated conditions. Furthermore, in rainfed conditions, traits such as grain yield, superoxide dismutase activity, ascorbic peroxidase activity, proline content, malon-dialdehyde levels, and catalase enzyme activity were consistent with group A genotypes (those with high yield in both rainfed and irrigated conditions). Conversely, genotypes 22, 11, 4, 1, 5, 3, 12 and 16 exhibited the lowest levels of drought tolerance based on the selected indices, particularly the RDY index.

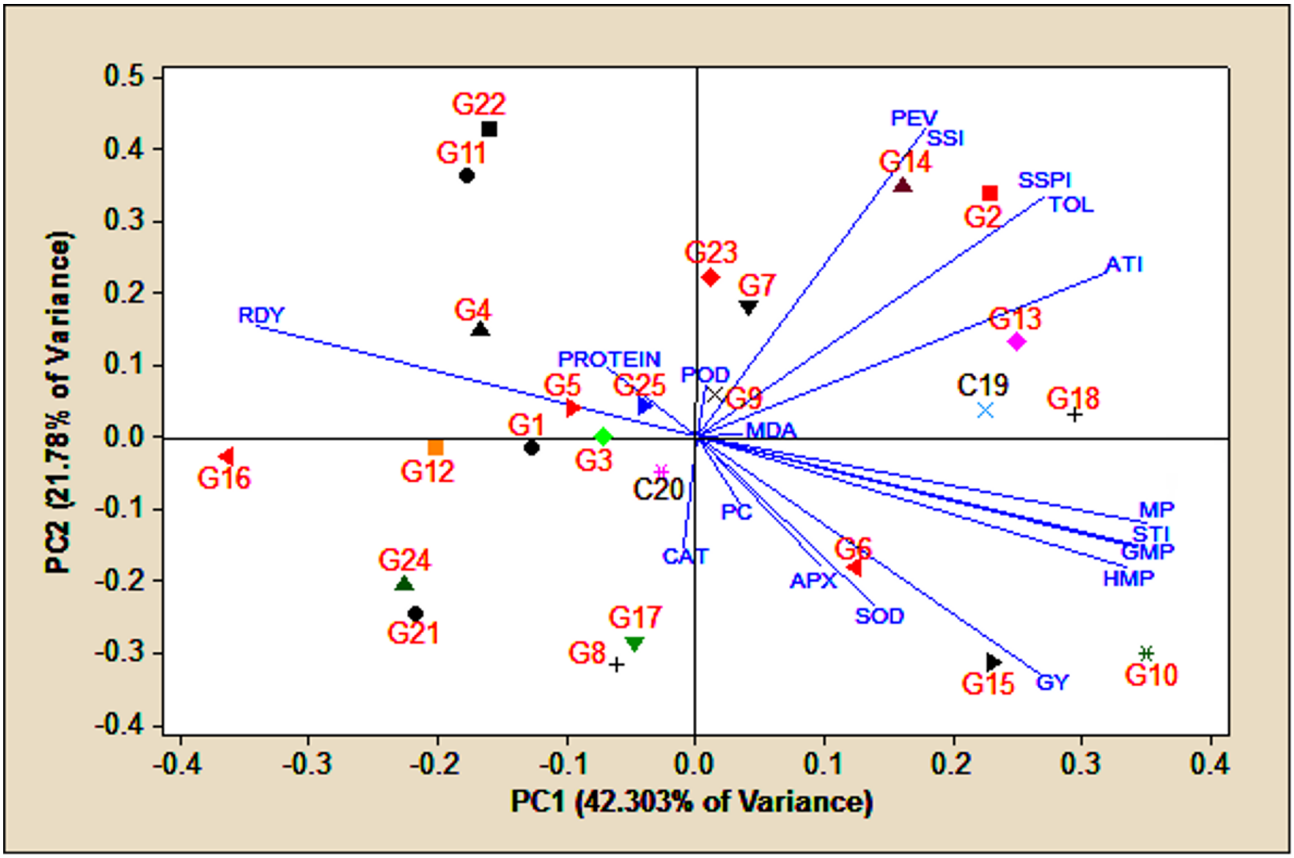


Figure 3. Biplot diagram of principal components analysis for drought tolerance indices and studied traits of wheat genotypes in rainfed conditions.

Table 7. Principal components analysis of 25 wheat genotypes in rainfed conditions. GY, Grain Yield; POD, Peroxidase Activity; SOD, Superoxide Dismutase Activity; CAT, Catalase Activity; PROTEIN, Soluble Protein; PC, Proline Concentration; APX, Ascorbic Peroxidase Activity; MDA, Malon-dialdehyde; STI, Stress Tolerance Index; MP, Mean productivity; GMP, Geometric Mean Productivity; HMP, Harmonic Mean Productivity; SSI, Stress Susceptibility Index; TOL, Tolerance; ATI, Abiotic Tolerance Index; SSPI, Stress Susceptibility Percentage Index; PEV, Press Evaluation; RDY, Relative Decrease in Yield.

Traits and indices	Component 1	Component 2	Component 3	Component 4
GY	0.266	-0.330	0.085	-0.030
POD	0.008	0.070	-0.494	-0.344
SOD	0.138	-0.234	-0.124	0.170
CAT	-0.010	-0.153	-0.135	0.604
Protein	-0.069	0.097	0.592	0.077
PC	0.034	-0.092	0.345	-0.489
APX	0.097	-0.178	-0.469	-0.197
MDA	0.036	0.005	-0.064	0.441
STI	0.341	-0.155	0.052	-0.023
MP	0.349	-0.120	0.064	-0.002
GMP	0.342	-0.151	0.064	-0.010
HMP	0.334	-0.182	0.063	-0.018
SSI	0.179	0.429	-0.044	-0.001
TOL	0.270	0.332	-0.016	0.050
ATI	0.318	0.227	-0.009	0.055
SSPI	0.270	0.332	-0.016	0.050
PEV	0.179	0.429	-0.044	-0.001
RDY	-0.341	0.155	-0.052	0.023
Eigenvalues	7.61	3.92	1.73	1.33
% of variance	42.303	21.78	9.612	7.353
Cumulative %	42.303	64.08	73.691	81.05

Determination of the genetic variability of wheat genotypes based on SSR markers

After evaluating 20 primer pairs across 25 bread wheat genotypes, 16 primers exhibiting high levels of polymorphism. 33 out of 35 total bands, showed high polymorphism, (93.75%). On average, each primer produced 2 bands, with a mean polymorphism of 2 bands per primer. The highest number of alleles was detected with primer XGWM136 (five). The primers XGWM155, XGWM234, XCFD168, XGWM577, XGWM642 and XCFD5 exhibited the highest polymorphic information content indices. Among the molecular indices assessed, the highest marker index values were identified for primers XGWM136, XCFD168 and XGWM350. The primers XGWM136, XGWM350, XCFD168 and XGWM165 recorded the highest Effective Multiplex Ratio. Regarding Resolving Power, the primers XGWM4, XCFD168 and XGWM350 showed the highest values (Table 8). The SSR markers banding pattern generated by the XGWM2, XGWM124, XGWM4, and XCFD5 primers for the wheat genotypes examined in this study is illustrated in Figure 4A–D.

Molecular variance analysis

The molecular variance analysis (AMOVA) for the SSR markers is presented in Table 9. Accordingly, a significant difference between the groups was observed at the 5% probability level. The proportion of variance attributed to intergroup differences was 10%, while intragroup variance accounted for 90%.

Investigating the relationship of studied characteristics and indices with SSR markers

The critical step in this process is assessing the efficiency of linkage markers associated with quantitative traits and identifying informative markers. To pinpoint alleles influencing grain yield, biochemical traits and drought tolerance indices in wheat genotypes under irrigated and rainfed conditions, an association analysis was conducted. This analysis examined the relationship between eight measured traits and ten indices (as dependent variables) and the molecular markers under study (as independent variables) using stepwise multiple regression analysis (Table 10, Table 11 and Table 12). The relationship with SSR markers was analyzed exclusively for characteristics that were statistically significant in the variance analysis.

Table 8. Molecular characteristics of more effective SSR primers in bread wheat genetic diversity in the present study.

Marker	No. of polymorphic bands	Polymorphic information content	Marker Index	Effective multiplex ratio	Resolving Power
XGWM350	3	0.337	1.011	3	3.12
XGWM155	1	0.499	0.499	1	0.96
XGWM577	2	0.467	0.934	2	1.60
XGWM642	1	0.461	0.23	0.5	0.72
XGWM136	5	0.352	1.76	5	2.32
XGWM165	3	0.324	0.973	3	2.96
XGWM4	2	0.211	0.422	2	3.52
XCFD5	2	0.442	0.883	2	1.76
XCFD168	3	0.489	1.466	3	3.20
XGWM234	1	0.493	0.493	1	0.88

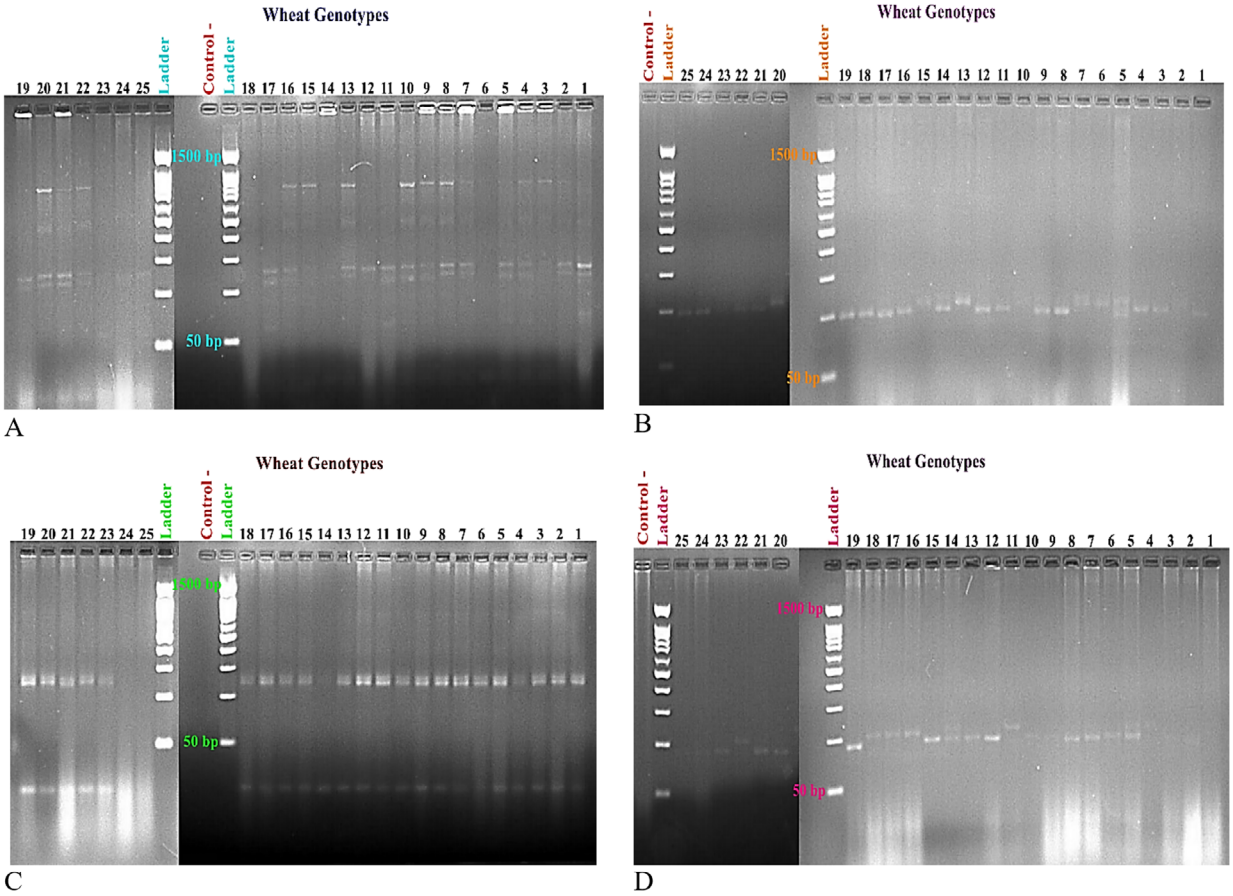


Figure 4. Patterns of some SSR markers used in the present study in wheat genotypes. A, XGWM2 primer; B, XGWM124 primer; C, XGWM4 primer; D, XCFD5 primer.

Table 9. Molecular variance analysis (AMOVA) of wheat genotypes. *, significant at 5% probability level.

Predicted Group	Source of variation	df	SS	MS	Estimated variance	Percentage of total variance	φ_{PT}
	Among Groups	3	25.85	8.62	0.68	10	0.039*
4	Within Groups	21	126.83	6.04	6.04	90	
	Total	24	152.68		6.72	100	

Grain yield and biochemical characteristics in rainfed conditions

The analysis identified two markers, **XGWM124(a2)** and **XGWM410(a1)**, as significantly related to yield in rainfed conditions, explaining 34% of the variation (Table 10). Additionally, the marker **XGWM124(a2)** showed a strong correlation with superoxide dismutase enzyme activity in rainfed conditions, accounting for 12% of the variation. The catalase enzyme activity was notably associated with the marker **XCFD168(a3)**, contributing 16% to the observed variance. Similarly, the ascorbic peroxidase enzyme activity displayed significant associations with three markers (**XGWM2(a1)**, **XGWM234(a1)**, and **XGWM350(a2)**), collectively explaining 52% of the variance. Moreover, a single locus amplified by the marker **XGWM155(a1)** was significantly associated with the malon-dialdehyde trait, accounting for 18% of the total variation. Overall, eight gene loci were identified as being associated with yield and biochemical characteristics in rainfed conditions. Notably, the **XGWM124(a2)** marker was shared between grain yield and the superoxide dismutase enzyme activity, highlighting its importance.

Grain yield and biochemical characteristics in irrigated conditions

The analysis revealed that grain yield was significantly correlated with seven amplified loci, including **XGWM577(a2, a1)**, **XGWM136(a3, a4)**, **XGWM265(a1)**, **XGWM410(a1)**, and **XGWM2(a2)** (Table 11). Among these, the loci **XGWM136(a3)**, **XGWM577(a2)**, and **XGWM410(a1)** demonstrated the most significant and positive effects. The marker **XCFD5(a2)** was significantly associated with the superoxide dismutase trait in irrigated conditions, explaining 17% of the variation. For the catalase enzyme activity, the marker **XGWM410(a1)** contributed 13% to the total variance. Altogether, nine gene loci were identified as being linked to yield and biochemical characteristics in irrigated conditions. Notably, the **XGWM410(a1)** marker was shared between grain yield and the catalase enzyme activity, underlining its importance.

Table 10. Markers association with grain yield and biochemical characteristics in rainfed conditions. *, significant at 5% probability level; **, significant at 1% probability level; † a1, a2, a3, a4, and a5 are the average alleles 1, 2, 3, 4, and 5, respectively.

Traits	Marker†	Regression coefficient (B)	Standard error (SE)	t-value	Significance level	R^2	Adjusted R^2
Grain Yield	Constant	320.44	17.871	17.931	**	0.397	0.342
	XGWM124(a2)	-68.482	20.944	-3.27	**		
	XGWM410(a1)	52.57	22.02	2.39	*		
Ascorbic Peroxidase	Constant	84.99	43.78	1.94	ns	0.583	0.523
	XGWM2(a1)	120.18	30.36	3.96	**		
	XGWM234(a1)	-80.304	24.72	-3.25	**		
	XGWM350(a2)	114.60	44.99	2.55	*		
Malon-dialdehyde	Constant	0.402	0.009	43.67	**	0.217	0.183
	XGWM155(a1)	-0.034	0.013	-2.53	*		
Catalase Activity	Constant	1.24	0.197	6.27	**	0.193	0.158
	XCFD168(a3)	0.62	0.264	2.35	*		
Superoxide Dismutase Activity	Constant	0.716	0.086	8.34	**	0.161	0.124
	XGWM124(a2)	-0.212	0.101	-2.1	*		

Table 11. Markers association with grain yield and biochemical characteristics in irrigated conditions. *, significant at 5% probability level; **, significant at 1% probability level; † a1, a2, a3, a4, and a5 are the average alleles 1, 2, 3, 4, and 5, respectively

Traits	Marker†	Regression coefficient (B)	Standard error (SE)	t-value	Significance level	R^2	Adjusted R^2
Grain Yield	Constant	374.72	12.76	29.37	**	0.885	0.838
	XGWM577(a2)	145.7	19.04	7.653	**		
	XGWM136(a3)	171.744	33.4	5.143	**		
	XGWM265(a1)	-122.541	26.084	-4.7	**		
	XGWM410(a1)	89.64	18.29	4.902	**		
	XGWM2(a2)	-62.8	18.43	-3.41	**		
	XGWM577(a1)	-58.85	16.294	-3.612	**		
	XGWM136(a4)	-79.04	29.13	-2.713	*		
Superoxide Dismutase Activity	Constant	0.233	0.036	6.513	**	0.206	0.171
	XCFD5(a2)	0.113	0.046	2.44	*		
Catalase Activity	Constant	0.55	0.089	6.13	**	0.163	0.126
	XGWM410(a1)	0.39	0.183	2.12	*		

Drought tolerance indices

The analysis identified a significant correlation between the ATI index and six amplified loci: **XGWM136(a_3, a4)**, **XGWM577(a2)**, **XGWM2(a2)**, **XGWM410(a1)** and **XGWM265(a1)**, collectively explaining 84% of the total variance (Table 12). The TOL and SSPI indices were significantly associated with the markers **XGWM136(a3, a4)**, **XGWM265(a2)**, **XGWM577(a2)** and **XGWM165(a2)**, accounting for 72% of the variation. Additionally, the MP, GMP and HMP indices demonstrated strong associations with three loci amplified by the markers **XGWM124(a2)**, **XGWM410(a1)** and **XGWM165(a1)**, explaining 57%, 57% and 56% of the total variation, respectively. The SSI and PEV indicators were significantly linked to the markers **XGWM136(a3, a4)** and **XGWM265(a2)**, accounting for 51% of the variance. Furthermore, the STI and RDY indices showed significant associations with two loci amplified by the markers **XGWM124(a2)** and **XGWM410(a1)**, each explaining 48% of the variation. Among these, the **XGWM410(a1)** marker exhibited the most substantial positive effect on the STI index, while the **XGWM124(a2)** marker had the strongest impact on the RDY index. Overall, 35 gene loci were identified for the drought tolerance indicators, with 10 gene loci being common across all measured indices.

Discussion

Significant differences in most of the studied characteristics highlighted the genetic diversity among wheat genotypes. This diversity suggests the potential to select superior cultivars based on grain yield and biochemical characteristics in rainfed and irrigated conditions. In addition, based on the percentage of changes in the irrigated environment compared to rainfed (TCP%), grain yield and soluble protein increased under irrigated conditions and decreased with stress. But on the other hand, the activity of peroxidase, superoxide dismutase, catalase, proline content, ascorbic peroxidase and malondialdehyde increased with stress, and the increase in the activity of these biochemical compounds aligns with enhanced stress resistance and reduced stress-induced damage. Therefore, the presence of a better antioxidant enzyme system, as evidenced by higher POD, SOD, CAT and APX activities in drought-tolerant wheat genotypes, could indicate that these genotypes are more efficient in removing superoxide anions produced in plants due to drought stress. Similarly, Saed-Moucheshi *et al* (2019) reported significant differences among genotypes for all yield and biochemical traits in triticale under regular irrigation and drought stress conditions. Furthermore, they observed significant increases in proline, malondialdehyde, protein content and antioxidant enzyme activities in response to drought stress, which aligns with the findings of this study. In a study by Pour-Aboughadareh *et al* (2022) evaluating biochemical traits in wild relatives of wheat under drought stress, ANOVA results revealed significant variations across growth conditions, except for dry matter in control and drought stress environments. Additionally, the activities of all antioxidant enzymes increased compared to the control conditions, which is consistent with current research.

Table 12. Markers association with drought tolerance indices of wheat genotypes. *, significant at 5% probability level; **, significant at 1% probability level; †, a1, a2, a3, a4, and a5 are the average alleles 1, 2, 3, 4, and 5, respectively. STI, Stress Tolerance Index; MP, Mean productivity; GMP, Geometric Mean Productivity; HMP, Harmonic Mean Productivity; SSI, Stress Susceptibility Index; TOL, Tolerance; ATI, Abiotic Tolerance Index; SSPI, Stress Susceptibility Percentage Index; PEV, Press Evaluation; RDY, Relative Decrease in Yield.

Indices	Marker†	Regression coefficient (B)	Standard error (SE)	t-value	Significance level	R^2	Adjusted R^2
ATI	Constant	21398.56	2357.34	9.08	**	0.881	0.841
	XGWM136(a3)	38568.94	6801.55	5.67	**		
	XGWM577(a2)	23254.62	3825.74	6.08	**		
	XGWM2(a2)	-18557.49	3870.85	-4.79	**		
	XGWM410(a1)	16573.71	3837.98	4.32	**		
	XGWM265(a1)	-17499.84	5480.78	-3.19	**		
	XGWM136(a4)	-17941.44	5977.63	-3	**		
SSPI	Constant	17.01	2.07	8.23	**	0.775	0.716
	XGWM136(a3)	26.8	4.25	6.31	**		
	XGWM136(a4)	-13.24	3.5	-3.79	**		
	XGWM265(a2)	-4.67	2.1	-2.22	*		
	XGWM577(a2)	5.81	2	2.91	**		
	XGWM165(a2)	-5.56	2.22	-2.5	*		
TOL	Constant	134.65	16.35	8.24	**	0.775	0.716
	XGWM136(a3)	212.14	33.63	6.31	**		
	XGWM136(a4)	-104.84	27.68	-3.79	**		
	XGWM265(a2)	-36.93	16.6	-2.23	*		
	XGWM577(a2)	46	15.81	2.91	**		
	XGWM165(a2)	-43.99	17.59	-2.502	*		
MP	Constant	393.77	17.393	22.64	**	0.619	0.565
	XGWM124(a2)	-87.581	20.85	-4.201	**		
	XGWM410(a1)	61.89	21.85	2.833	**		
	XGWM165(a1)	-72.78	34.633	-2.101	*		
GMP	Constant	386.27	16.91	22.85	**	0.619	0.565
	XGWM124(a2)	-84.36	20.27	-4.163	**		
	XGWM410(a1)	59.544	21.24	2.804	*		
	XGWM165(a1)	-73.03	33.67	-2.17	*		
HMP	Constant	378.97	16.64	22.78	**	0.613	0.558
	XGWM124(a2)	-81.242	19.94	-4.074	**		
	XGWM410(a1)	57.29	20.89	2.742	*		
	XGWM165(a1)	-73.24	33.13	-2.211	*		
SSI	Constant	1.11	0.121	9.16	**	0.575	0.514
	XGWM136(a3)	1.36	0.262	5.2	**		
	XGWM136(a4)	-0.824	0.231	-3.56	**		
	XGWM265(a2)	-0.309	0.147	-2.1	*		
PEV	Constant	0.314	0.034	9.16	**	0.575	0.514
	XGWM136(a3)	0.386	0.074	5.2	**		
	XGWM136(a4)	-0.233	0.065	-3.57	**		
	XGWM265(a2)	-0.087	0.042	-2.1	*		
STI	Constant	0.968	0.081	11.89	**	0.526	0.483
	XGWM124(a2)	-0.416	0.095	-4.36	**		
	XGWM410(a1)	0.294	0.1	2.93	**		
RDY	Constant	-1416.28	127.5	-11.11	**	0.526	0.483
	XGWM124(a2)	650.93	149.42	4.36	**		
	XGWM410(a1)	-460.43	157.09	-2.93	**		

In a study on 20 bread wheat cultivars under water stress and non-stress conditions, water stress caused a significant 54.9% reduction in grain yield and reductions in all studied traits except grain protein content (Al-Naggar *et al*, 2020), contrasting with the present study regarding protein content. Similarly, Firouzian *et al* (2023) reported that stress reduced yield components, physiological traits, and ultimately decreased grain yield by about 25% in bread wheat, while, in the present study, drought stress reduced grain yield by 28.3%. In a study investigating terminal heat stress effects on wheat cultivars, variance analysis of phenological traits, grain yield and biochemical traits showed significant variations in genotypes, environments, and genotype × environment interactions (for grain yield, SOD, POD, APX, CAT and proline) (Kumar *et al*, 2023a). In the present study, variance analysis also revealed significant variations for grain yield and all biochemical traits in the environment effect, for all traits except soluble protein in the genotype effect, and for all traits except grain yield and malon-dialdehyde in the interaction effects. Similarly, in a study by Mkhabela *et al* (2019) investigating drought-tolerant wheat genotypes under drought stress and non-stress conditions, the effects of genotype, stress condition, and genotype × stress condition interaction were significant for the tested traits, indicating differential genotypic responses to selection, in agreement with this study.

Estimating heritability helps plant breeders identify elite genotypes (Farshadfar, 2010). Likewise, genetic advancement reflects the mean genotypic value relative to the parental population and serves as an indicator of the genetic gain achieved through selection (Kumar *et al*, 2023b). High broad-sense heritability suggests that the trait is minimally influenced by environmental factors. However, modifying such a trait may be less beneficial, as broad-sense heritability encompasses the total genetic variance, including additive (fixable), dominance and epistasis (non-fixable) variances. On the other hand, high genetic advance or genetic gain indicates that the trait is primarily governed by additive genes, making selection a practical approach for improvement. Conversely, low genetic advance or gain suggests that the trait is controlled by non-additive genes, in which case heterosis breeding would be a more effective strategy (Farshadfar, 2010; Kaur *et al*, 2023). Traits with heritability ($h^2 > 60.0\%$) and genetic gain (GG > 20.0%) indicate that the observed variation is predominantly due to genetic factors, thereby making these traits reliable candidates for selection (Faysal *et al*, 2022; Kaur *et al*, 2023), which in the present study also included most biochemical traits with heritability above 90% and genetic gain above 30%. This indicates the high influence of genetic factors and aligns with the aforementioned findings. In summary, this study showed that broad-sense heritability and genetic gain for catalase, superoxide dismutase activity, proline content and ascorbic peroxidase activity were high under both irrigated and rainfed conditions, and were lowest for malon-dialdehyde. Therefore, it is recommended to use these traits as ideal criteria, along with yield, to select high-yielding genotypes in breeding programmes. In research on bread wheat genotypes, moderate heritability values and high genetic gain for grain yield were recorded, suggesting these traits are promising targets for improvement through favorable selection (Amare, 2023). In the present study, moderate broad-sense heritability and genetic gain were obtained for grain yield. Similarly, in the study by Saed-Moucheshi *et al* (2019), grain yield showed heritability values of 32.14% and 29.62% under normal irrigation and drought stress conditions, respectively, indicating environmental influence. Additionally, SOD and MDH showed the highest heritability under both conditions, while in the present study, CAT, SOD, PC and APX showed the highest heritability, and MDA the lowest heritability in both environmental conditions. The heritability of grain yield was 27.8% and 60.4% in rainfed and irrigated conditions, respectively. In studies by Shah *et al* (2019) on bread wheat under rainfed conditions and Sallam *et al* (2024b) on bread wheat under heat stress, the traits grain protein content, proline and catalase, respectively, showed high heritability and genetic gain, consistent with the study. Additionally, various researchers have utilized broad-sense heritability (Li *et al*, 2023; Sowadan *et al*, 2024) and genetic gain (Yusuf *et al*, 2021; Dukamo *et al*, 2023) to examine genetic variability and identify suitable traits for breeding programmes, aligning with the study's findings regarding the importance of these parameters.

These indices are used to calculate the level of drought tolerance in plants. Different indices are designed based on different traits that are related to grain yield. These indices are used in different agronomic, biochemical, molecular and even cytogenetic categories to select the best genotype. The correlation heatmaps were created to analyze the relationships between studied traits and drought tolerance indices in rainfed and irrigated conditions. Based on the nature of the indicators, it was observed that most of the studied traits showed a highly significant correlation with indices such as STI, MP, GMP, HMP and RDY. These indicators were identified as the most effective for selecting drought-tolerant and high-performing genotypes. Additionally, these indices had a strong influence on the first principal component, which was identified as the drought tolerance component. Similar to the present study, Reddy *et al* (2023), used correlation heatmaps to examine the relationship between phenotypic traits and drought tolerance indices such as STI, MP, and GMP. Also, in the study by Giovenali *et al* (2023), Pearson correlation coefficients were analyzed using heatmaps to investigate the relationships between yield-related traits, physiological parameters and biochemical parameters, and significant positive and negative correlations were obtained. In a related study, correlation heatmaps were employed to explore the relationships between phenological, physiological and biochemical variables in optimal conditions, heat stress conditions, prolonged heat stress conditions, and a combined environment. In optimal conditions, the correlation between seed yield and the APX and CAT traits was positive but not statistically significant. In heat stress conditions, a positive and significant correlation was observed between seed yield and the traits proline and SOD. When heat stress was prolonged, the correlation between seed yield and CAT became negative and was not significant. However, in high-temperature conditions, seed yield demonstrated a positive and significant correlation with proline, SOD and POD, while its relationship with APX remained positive but non-significant (Kumar *et al*, 2023a). In the present study, CAT and APX did not show statistically significant correlations with any of the drought tolerance indices. Under irrigation conditions, there was a significant negative correlation between PC and the SSI and PEV indices. There was also a significant positive correlation between POD and the STI, MP, GMP and HMP indices, and a

significant negative correlation with the RDY index. Under rainfed conditions, the correlation of SOD with the STI, MP, GMP and HMP indices was positive and significant, and with the RDY index was negative and significant.

Principal component analysis (PCA), a multivariate statistical method, serves as an efficient approach to data reduction by identifying strong correlations among variables to derive clear conclusions. In this study, PCA and biplot visualization were applied to analyze traits across 25 bread wheat genotypes under rainfed conditions. The first two principal components (PC1 and PC2) accounted for 64.08% of the total variation in drought tolerance indices and studied traits. Biplot visualization revealed considerable genetic diversity among genotypes in response to drought stress. These findings align with existing literature: Pour-Aboughadareh *et al* (2022) reported PC1 and PC2 explaining 64.52% of biochemical variation in wild wheat species under drought (PC1 = 47.86%; PC2 = 16.66%). Sallam *et al,* (2024a) identified four principal components (eigenvalues >1) capturing 89.79% of variance across 30 agro-physio-biochemical traits. PC1 correlated with 24 traits (e.g. grain yield, catalase, peroxidase, superoxide dismutase and proline), PC2 with five traits (e.g. soluble protein), PC3 showed no significant associations, and PC4 linked to glycine betaine. Similarly, in the present study, the first four principal components explained 81.05% of total variance: PC1 (grain yield and drought-tolerance indices), PC2 (drought-stress indices), PC3 (soluble protein and proline), and PC4 (catalase activity and malon-dialdehyde).

Of the 20 SSR markers tested, 16 showed significant polymorphism. Genetic diversity assessment of bread wheat genotypes utilizing SSR markers revealed XCFD168, XGWM350 and XGWM136 as fully polymorphic (100%). These markers demonstrated the highest allele counts and superior performance across key genetic indices: polymorphism information content (PIC), marker index (MI), effective multiplex ratio (EMR), and resolving power (RP) (Table 8). Consequently, these represent optimal candidates for advanced wheat genetic analyses. Notably, the significant discriminatory power achieved with limited primer sets confirms that highly polymorphic SSRs efficiently differentiate both individual accessions and population subgroups. These findings corroborate prior research identifying the same three markers as exceptionally informative. For instance, the marker XGWM136 was similarly highlighted by Budak *et al* (2013) and Kaur *et al* (2016), while Khan *et al* (2021) identified XGWM136 and XCFD168, and Haque *et al* (2020) emphasized XGWM350, collectively supporting their utility as reported in this study. In genomic diversity research on bread wheat using SSR markers, markers XGWM136, XCFD168, XGWM2, XGWM155, XCFD5, XGWM165, XGWM33 and XGWM129 were used (Ahmed *et al*, 2020), consistent with the marker selection in this study. Additionally, research on bread wheat genetic diversity revealed high PIC and marker index, showing greater diversity in the A and B genomes compared to the D genome (Feltaous, 2019). In the present study, more diversity was observed in the B genome, followed by D and A. In another study, 17 bread wheat genotypes evaluated with 16 SSR markers showed only 11 markers with high polymorphism and reproducibility (Kara *et al*, 2020). Similarly, a study assessing the genetic diversity and population structure of wheat genotypes employed ten SSR markers to characterize diversity across 22 genotypes (Hassan *et al*, 2025). Here, 25 bread wheat genotypes examined with 20 SSR markers revealed 16 with significant polymorphism. Regarding PIC values and marker utility, these findings align with prior reports. For example: in SSR evaluation of bread wheat, the PIC ranged from 0.276 to 0.541 (average: 0.384), using primers XGWM192 and XGWM642 (Islam *et al*, 2012). Three subsequent studies (El-Rawy and Hassan, 2021; Ahmed *et al*, 2024; Bavandpouri *et al*, 2025) reported a PIC range of 0.20–0.50 (average: 0.33). El-Rawy and Hassan (2021) utilized primers XGWM165, XGWM155, and XGWM577, with XGWM577 showing superior performance. Bavandpouri *et al* (2025) introduced three markers – namely XCFD168, XGWM350, and XGWM136 – as the most significant; while Ahmed *et al* (2024) highlighted XGWM642. In a separate analysis of ten bread wheat genotypes using ten SSR markers, Kumari *et al* (2025) detected 64 polymorphic bands, where alleles per locus ranged from 1 to 4 (highest for XGWM2 and XGWM265). In the present study, 33 out of 35 bands were polymorphic. The highest number of alleles (five) was observed for primer XGWM136, while the lowest number (two alleles) was recorded for primers XGWM155, XGWM410 and XGWM234. Collectively, these findings confirm SSR markers as reliable indirect selection tools for more efficient cultivar improvement.

Genetic structure within populations is commonly analyzed through variance analysis, where the variance between and within groups is determined based on the genetic distances among individuals. AMOVA is particularly effective in partitioning variance in wild species and among groups of cultivars originating from different regions (Farshadfar, 2023). The results of AMOVA revealed that the observed grouping of bread wheat genotypes could, to some extent, be explained by the diversity in SSR marker bands. The φ_{PT} statistic is employed as a criterion to test the assumption of population differentiation at the relevant level. In this experiment, AMOVA indicated that the φ_{PT} statistic was low due to the high genetic diversity observed within the populations. Similar findings were reported in a study investigating the genetic diversity of bread wheat using ISSR and SSR markers, where AMOVA for both types of markers revealed that genetic variation within species surpassed the genetic diversity among them (Jabari *et al*, 2023). Additionally, in another study, AMOVA results demonstrated that 19% of the total genetic variation occurred among subpopulations, while the remaining 81% was attributed to individual differences within each subpopulation (Sowadan *et al*, 2024).

In drought tolerance research, pinpointing QTLs linked to drought-responsive traits is pivotal for deciphering their genetic mechanisms (Sallam *et al*, 2019). To identify relevant SSR markers, regression analysis was conducted between grain yield and biochemical traits under rainfed and irrigated conditions, with ten drought tolerance indices as dependent variables and marker gene locations as independent variables. Results revealed significant trait–primer relationships. A key advantage of this multivariate regression approach is its efficiency in QTLs detection, reducing time and cost while eliminating the need for mapping populations (Ruan *et al*, 2009). This study specifically aimed to identify alleles correlated with grain yield and biochemical traits as informative markers. Outcomes supporting this objective are detailed in Table 10, Table 11 and Table 12. Critical associations include: XCFD168 marker showing strong correlation with catalase activity under rainfed conditions; XGWM350 linked

to ascorbic peroxidase activity (rainfed); and XGWM136 associated with yield under irrigated conditions and ATI, TOL, SSPI, SSI and PEV indices. Notably, XGWM410(a_1) correlated with yield in both environments, catalase activity under irrigation, and multiple indices; XGWM2(a2) tied to irrigated yield, rainfed ascorbic peroxidase activity, and ATI; while XGWM124(a2) demonstrated associations with yield, rainfed superoxide dismutase activity, and several indices. In a study investigating associations between biochemical traits and stress tolerance indices in wheat under drought using 24 SSR markers, two markers were linked to APX and three to CAT in control conditions, whereas under drought stress, two markers associated with APX and one with POD, alongside two markers significantly correlated with the STI index. These results suggest genomic regions governing growth and developmental characteristics across conditions (Pour-Aboughadareh *et al*, 2022). Comparatively, in the current study, biochemical trait analysis revealed one marker correlated with SOD and one with CAT under irrigation, while under rainfed conditions, one marker associated with SOD, one with CAT, three with APX and one with MDA. Regarding STI index, consistent with Pour-Aboughadareh *et al* (2022), two markers showed significant correlations. In a linkage mapping study for photosynthesis and yield traits under moisture stress and drought indices (SSI and STI) in winter bread wheat, 28 linkages were identified for drought tolerance indices, with one marker consistently associated across two seasons (Saeed *et al*, 2017). Here, 35 linkages emerged for drought indices, four of which were associated with both SSI and STI. El-Rawy and Hassan (2021) reported SSR markers XGWM260 and XGWM573 as specific to drought-tolerant bread wheat genotypes (low DSI values), suggesting marker-trait associations for drought tolerance. In contrast, our study identified XGWM136 and XGWM265 as significantly associated with SSI. Negisho *et al* (2022) detected 184 marker-trait associations (MTAs) for drought indices in Ethiopian durum wheat, with six MTAs (on chromosomes 2B, 3B, 4A, 5B and 6B) positively affecting GY-GMP. Notably, 41 MTAs (22.28%) associated with ≥ 2 indices, of which 16 (39.02%) linked to GMP and STI. Similarly, we identified an MTA positively affecting GMP on chromosome 2B, along with 35 gene loci for drought indices –10 (28.57%) common to all indices, with 30% of stable MTAs associated with GMP and STI. Across these studies, a positive regression coefficient indicates that selecting genotypes harbouring such alleles may enhance yield and drought tolerance.

The association between individual markers and multiple traits may arise from pleiotropic effects or overlapping QTLs influencing diverse characteristics. Primers such as XGWM136, XGWM234 and XCFD168 – previously used to investigate grain yield and agronomic traits relationships via SSR markers with potential for heat-tolerance breeding (Khan *et al*, 2021) – were similarly employed in this study, with XCFD168 and XGWM136 emerging as superior markers. Consistent with our findings, numerous studies report significant yield-marker associations: Maccaferri *et al* (2011) established XGWM410–yield relationships; Amalova *et al* (2024) documented correlations between grain yield and XGWM124. Bavandpouri *et al* (2025) reported significant relationships between XGWM265 and grain yield under irrigated conditions, and between XGWM410, XGWM577, and XGWM124 markers and grain yield under both rainfed and irrigated conditions, while Eldemery *et al* (2022) and Firouzian *et al* (2023) observed XGWM577–yield linkages under heat stress. Concurrently, XGWM165 – also utilized here – showed notable associations with SSPI, TOL, MP, GMP, and HMP indices, aligning with our results and collectively reinforcing these outcomes. Ultimately, this research confirms that molecular markers exhibiting strong regression coefficients for biochemical traits and drought tolerance indicators offer breeders actionable insights. Such markers enable selection of environmentally stable QTLs linked to yield and drought tolerance, accelerating the development of superior genotypes. Moreover, the identified MTAs hold direct utility in wheat breeding programmes targeting drought stress, particularly for marker-assisted selection and gene pyramiding strategies.

Conclusion

Molecular markers linked to biochemical traits can accelerate the identification of drought-tolerant germplasm, enhancing breeding efficiency. Significant genotypic variance confirmed substantial genetic diversity and differential drought stress responses. Key biochemical traits – catalase (CAT), superoxide dismutase (SOD) activity, proline content (PC), and ascorbate peroxidase (APX) activity – exhibited high heritability (> 90%) and genetic advance (> 30%) under both irrigated and rainfed conditions, unlike malondialdehyde (MDA). Thus, these traits are recommended for selecting high-yielding genotypes. Principal component analysis (PCA, 64.08% variance explained) and correlation identified stress tolerance index (STI), mean productivity (MP), geometric mean productivity (GMP), harmonic mean productivity (HMP), and relative decrease in yield (RDY) as the most effective drought tolerance indices, strongly correlated with grain yield and biochemical traits. Genotypes 6, 10, 15, 18, 13, and Pishtaz demonstrated superior drought tolerance, high yield potential, and optimal biochemical performance (SOD, APX, PC, MDA, CAT) under stress. Among 20 SSR markers, 16 showed significant polymorphism. Markers XCFD168 (rainfed-CAT), XGWM350 (rainfed-APX), XGWM124(a_2) (yield, rainfed-SOD, multiple indices), XGWM136 (irrigated yield, ATI, TOL, SSPI, SSI, PEV), XGWM410(a1) (yield in both environments, irrigated-CAT, multiple indices), and XGWM2(a2) (irrigated yield, rainfed-APX, ATI) exhibited significant trait associations. These markers are strongly recommended for marker-assisted breeding to improve yield and drought tolerance.

Acknowledgments

We gratefully acknowledge the administration of Razi University, Kermanshah, for providing the laboratory and field facilities necessary to conduct this research. We also extend our thanks to the Karaj Seed and Plant Improvement Institute for supplying the plant seeds. It is important to note that while Razi University provided the research infrastructure, it did not provide specific funding for the publication of this article.

Author contributions

Conceptualization: Ezatollah Farshadfar, Kianoosh Cheghamirza; Data curation: Fatemeh Bavandpouri; Methodology: Fatemeh Bavanpouri; Formal analysis: Fatemeh Bavandpouri, Mohsen Farshadfar; Investigation: Ezatollah Farshadfar; Funding acquisition: Ezatollah Farshadfar, Mohsen Farshadfar; Project administration: Fatemeh Bavanpouri, Kianoosh Cheghamirza; Visualization: Fatemeh Bavanpouri, Mohsen Farshadfar; Supervision: Ezatollah Farshadfar, Kianoosh Cheghamirza; Resources: Fatemeh Bavandpouri, Mohsen Farshadfar; Writing – original draft preparation: Fatemeh Bavanpouri; Writing – review and editing: Fatemeh Bavandpouri, Ezatollah Farshadfar. All authors have read and agreed to the published version of the manuscript.

Conflict of interest statement

The authors have declared that no competing interests exist.

Data availability statement

The datasets collected and analyzed for this study are available upon reasonable request.

References

Ahmed, H. G. M. D., Kashif, M., Rashid, M. A. R., Sajjad, M., and Zeng, Y. (2020). Genome wide diversity in bread wheat evaluated by SSR markers. *International Journal of Agriculture & Biology* 24, 263–272. doi: https://doi.org/10.17957/IJAB/15.1433

Ahmed, Sh. F., Ahmed, J. U., Hasan, M., and Mohi-Ud-Din, M. (2023). Assessment of genetic variation among wheat genotypes for drought tolerance utilizing microsatellite markers and morpho-physiological characteristics. *Heliyon* 9, e21629. doi: https://doi.org/10.1016/j.heliyon. 2023. e21629

Ahmed, H. Gh. M. D., Yang, T., Akram, M. I., Iqbal, R., AlGhamdi, A. A., and Al Farraj, D. A. (2024). Molecular marker based analysis of allelic variation in the spring wheat genome. *Genetic Resources and Crop Evolution* 1–17. doi: https://doi.org/10.1007/s10722-024-02274-y

Al-Ashkar, I., Alderfasi, A., Romdhane, W. B., Seleiman, M. F., El-Said, R. A., and Al-Doss, A. (2020). Morphological and genetic diversity within salt tolerance detection in eighteen wheat genotypes. *Plants* 9(287), 1–21. doi: https://doi.org/10.3390/plants9030287

Al-Naggar, A. M. M., El-Shafi, M. A. E. M. A., El-Shal, M. H., and Anany, A. H. (2020). Evaluation of Egyptian wheat landraces (*Triticum aestivum* L.) for drought tolerance, agronomic, grain yield and quality traits. *Plant Archives* 20(1), 3487–3504. url: https://www.researchgate.net/publication/340284945

Altintas, S., Toklu, F., Kafkas, S., Kilian, B., Brandolini, A., and Ozkan, H. (2008). Estimating genetic diversity in durum and bread wheat cultivars from turkey using AFLP and SAMPL markers. *Plant breeding* 127, 9–14. doi: https://doi.org/10.1111/j.1439-0523.2007.01424.x

Amalova, A., Griffiths, S., Abugalieva, S., and Turuspekov, Y. (2024). Genome-wide association study of yield-related traits in a nested association mapping population grown in Kazakhstan. *Agronomy* 14, 1848. doi: https://doi.org/10.3390/agronomy14081848

Amare, A. (2023). Genetic variability, correlation and path coefficient analysis of bread wheat (*Triticum aestivum* L.) genotypes under irrigation in Raya Azebo district, south Tigray. *International Journal of Novel Research in Interdisciplinary Studies* 10(6), 1–6. doi: https://doi.org/10.5281/zenodo.10074538

Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D., and Sorrells M. E. (1993). Optimizing parental selection for genetic linkage maps. *Genome* 36(1), 181–186. doi: https://doi.org/10.1139/g93-024

Bates, L., Waldren, R., and Teare, I. (1973). Rapid determination of free proline for water-stress studies. *Plant and Soil* 39, 205–207. doi: https://doi.org/10.1007/BF00018060

Batool, N., Ilyas, N., Shahzad, A., Hauser, B. A., and Arshad, M. (2018). Quantitative trait loci (QTLs) mapping for salt stress tolerance in wheat at germination stage. *Pakistan Journal of Agricultural Sciences* 55(1), 47–55. doi: https://doi.org/10.21162/PAKJAS/18.5426

Bavandpouri, F., Farshadfar, E., Cheghamirza, K., Farshadfar, M., Bihamta, M. R., Mahdavi, A. M., and Jelodar, N. B. (2025). Identification of molecular markers associated with genomic regions controlling agronomic traits in bread wheat genotypes under different moisture conditions. *Plant molecular biology reporter* 43, 631-651. doi: https://doi.org/10.1007/s11105-024-01494-x

Beauchamp, C., and Fridovich, I. (1971). Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. *Analytical Biochemistry* 44, 276–287. doi: https://doi.org/10.1016/0003-2697(71)90370-8

Bouslama, M., and Schapaugh, W. (1984). Stress tolerance in soybeans. I. evaluation of three screening techniques for heat and drought tolerance. *Crop Science* 24, 933–937. doi: https://doi.org/10.2135/cropsci1984.0011183X002400050026x

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein Dye binding. *Analytical Biochemistry* 72, 248–254. doi: https://doi.org/10.1006/abio.1976.9999

Budak, H., Kanter, M., and Kurtoglu, K. Y. (2013). Drought tolerance in modern and wild wheat. *The Scientific World Journal* 15(2013), 548246. doi: https://doi.org/10.1155/2013/548246

Chance, B., and Maehly A. C. (1995). Assay of catalase and peroxidase. In: Culowic SP, & Kaplan NO, eds. *Methods in enzymology*, Vol 2. New York: Academic Press Inc., 764–765.

Choudhary, R. C., Sharma, N. K., Kumar, R., and Kumar, M. (2016). SSR-based genetic diversity assessment among hexaploid wheat (*Triticum aestivum* L.) cultivars. *Indian Journal of Plant Genetic Resources* 29(2), 137–143. doi: https://doi.org/10.5958/0976-1926.2016.00019.X

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19, 11–15.

Dukamo, B. H., Gedebo, A., Tesfaye, B., and Degu, H. D. (2023). Genetic diversity of Ethiopian durum wheat (*T. turgidum* subsp. Durum) landraces under water stressed and non-stressed conditions. *Heliyon* 9(e18359), 1–17. doi: https://doi.org/10.1016/j.heliyon.2023.e18359

El-demery, S. M. M., Bakry, B. A., Younis, A. E-S. M., Sayed, M. A., and Abdellatif, K. F. (2022). QTL analysis of grain yield-related traits for terminal heat stress tolerance in wheat using SSR markers. *Pakistan Journal of Biological Sciences* 25(6), 516–530. doi: https://doi.org/10.3923/

pjbs.2022.516.530
Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* 5(6), 435–445. doi: https://doi.org/10.1038/nrg1348
El-Rawy, M. A., and Hassan, M. I. (2021). Assessment of genetic diversity in durum and bread wheat genotypes based on drought tolerance and SSR markers. *Plant Breeding and Biotechnology* 9(2), 89–103. doi: https://doi.org/10.9787/PBB.2021.9.2.89
Emre, I., Ozgur, T., Fatma, A. T., and Muzaffer, T. (2011). Determination of tolerance level of some wheat genotypes to post-anthesis drought. *Turkish Journal of Field Crops* 19(1), 59–63.
Farshadfar, E. (2010). New topics in biometric genetics. Islamic Azad University publications, 722p.
Farshadfar, M. (2023). Molecular plant breeding. Payam Noor University press, 424p.
Faysal, A. S. M., Ali, L., Azam, M. G., Sarker, U., Ercisli, S., Golokhvast, K. S., and Marc, R. A. (2022). Genetic variability, character association, and path coefficient analysis in transplant Aman rice genotypes. *Plants* 11, 2952. doi: https://doi.org/10.3390/plants11212952
Feltaous, Y. M. (2019). Genetic diversity among some Egyptian bread wheat cultivars based on morphological characters and SSR markers. *Assiut Journal of Agricultural Sciences* 50(4), 35–50. doi: https://doi.org/10.21608/ajas.2020.70069
Fernandez, G. C. (1992). Effective selection criteria for assessing plant stress tolerance. *Proceedings of the international symposium on adaptation of vegetables and other food crops in temperature and water stress*, AVRDC publication, Tainan, Taiwan, 257–270.
Firouzian, A., Shafeinia, A., Ghaffary, S. M. T., Mohammadi, V., and Sadat, S. (2023). Terminal heat tolerance in bread wheat determined by agronomical traits and SSR markers. *Journal of Plant Growth Regulation* 42(3), 1–12. doi: https://doi.org/10.1007/s00344-022-10680-8
Fischer, R., and Maurer, R. (1978). Drought resistance in spring wheat cultivars. I. grain yield responses. *Crop and Pasture Science* 29, 897–912. doi: https://doi.org/10.1071/AR9780897
Galal, A. A., Safhi, F. A., El-Hity, M. A., Kamara, M. M., El-Din, E. M. G., Rehan, M., Farid, M., Behiry, S. I., El-Soda, M., and Mansour, E. (2023). Molecular genetic diversity of local and exotic durum wheat genotypes and their combining ability for agronomic traits under water deficit and well-watered conditions. *Life* 13(2293), 1–20. doi: https://doi.org/10.3390/life13122293
Giovenali, G., Kuzmanovi´c, L., Capoccioni, A., and Ceoloni, C. (2023). The response of chromosomally engineered durum wheat-*Thinopyrum ponticum* recombinant lines to the application of heat and water-deficit stresses: effects on physiological, biochemical and yield-related traits. *Plants* 12(704), 1–29. doi: https://doi.org/10.3390/plants12040704
Gupta, A. K., Agrawal, M., Yadav, H., Mishra, G., Gupta, R., Singh, A., Katiyar, D., Singh, P., and Srivastava A. (2024). Drought stress and its tolerance mechanism in wheat. *International Journal Environment and Climate Chang* 14(1), 529–544. url: https://doi.org/10.9734/ijecc/2024/v14i13866
Halder, T., Liu, H., Chen, Y., Yan, G., and Siddique, K. H. M. (2023). Chromosome groups 5, 6 and 7 harbor major quantitative trait loci controlling root traits in bread wheat (*Triticum aestivum* L.). *Frontiers in Plant Science* 14, 1092992. doi: https://doi.org/10.3389/fpls.2023.1092992
Haque, M. Sh., Saha, N. R., Islam, M. T., Islam, M. M., Kwon, SJ., Roy, S. K., and Woo, SH. (2020). Screening for drought tolerance in wheat genotypes by morphological and SSR markers. *Journal of Crop Science and Biotechnology* 1-14. doi: https://doi.org/10.1007/s12892-020-00036-7
Hassan M. Z, Hasanuzzaman, M., Uddin, M. J., Alomgir, M. B., Akter, M. S., Mohanto, A. Ch., and Kayess, M. O. (2025). Assessing genetic diversity and population structure of f4:f5 wheat genotypes using morphological and microsatellite markers under heat stress. *Discover Agriculture* 3: 14. doi: https://doi.org/10.1007/s44279-025-00161-3
Heath, R. L., and Packer, L. (1968). Photoperoxidation in isolated chloroplasts: I. kinetics and stoichiometry of fatty acid peroxidation. *Archives of Biochemistry and Biophysics* 125, 189–198. doi: https://doi.org/10.1016/0003-9861(68)90654-1
Heidari, B., Barjoyifard, D., Mazal-Mazraei, T., and Govindan, V. (2024). Assessment of genetic biodiversity and association of micronutrients and agronomic traits using microsatellites and staining methods which accelerates high micronutrients variety selections within different wheat groups. *Scientific Reports* 14, 27419. doi: https://doi.org/10.1038/s41598-024-78964-5
Ilyas, N., Amjid, M. W., Saleem, M. A., Khan, W., Wattoo, F. M., Rana, R. M., Maqsood, R. H., Zahid, A., Shah, G. A., Anwar, A., Ahmad, M. Q., Shaheen, M., Riaz, H., and Ansari, M. J. (2020). Quantitative trait loci (QTL) mapping for physiological and biochemical attributes in a Pasban90/Frontana recombinant inbred lines (RILs) population of wheat (*Triticum aestivum*) under salt stress condition. *Saudi Journal of Biological Sciences* 27(1), 341–351. doi: https://doi.org/10.1016/j.sjbs.2019.10.003
Islam, S., Haque, M. S., Emon, R. M., Islam, M. M., and Begum, S. N. (2012). Molecular characterization of wheat (*Triticum aestivum* L.) genotypes through SSR markers. *Bangladesh Journal of Agricultural Research* 37(3), 389–398. doi: https://doi.org/10.3329/bjar.v37i3.12082
Jabari, M., Golparvar, A., Sorkhilalehloo, B., and Shams, M. (2023). Investigation of genetic diversity of Iranian wild relatives of bread wheat using ISSR and SSR markers. *Journal of Genetic Engineering Biotechnology* 21(73), 1–16. doi: https://doi.org/10.1186/s43141-023-00526-5
Kara, K., Rached-Kanouni, M., Mnasri, S., Khammar, H., and Ben Naceur, M. B. (2020). Genetic variability assessment in bread wheat (*Triticum aestivum*) grown in Algeria using microsatellites SSR markers. *Biodiversitas* 21, 2638–2644. doi: https://doi.org/10.13057/biodiv/d210635
Kaur, V., Singh, S., and Behl, R. K. (2016). Heat and drought tolerance in wheat: integration of physiological and genetic platforms for better performance under stress. *Ekin Journal of Crop Breeding and Genetics* 2(1), 1-14. url: https://www.ekinjournal.com
Kaur, S. J., Talekar, N., Delvadiya, I., Singh, S. K., and Raut, A. (2023). Genetic profiling of bread wheat (*Triticum aestivum* L.): analyzing variation, associations, path analysis, and diversity to revolutionize crop enhancement. *Journal of Food Chemistry and Nanotechnology* 9(S1), S21–S27. doi: https://doi.org/10.17756/jfcn.2023-s1-005
Kearsey, J. M., and Pooni. S. H. (1996). *The genetic analysis of quantitative traits*. 1st ed. Chapman and Hall, London. doi: https://doi.org/10.1007/978-1-4899-4441-2
Khan, A., Ahmad, M., Ahmed, M., Gill, K. S., and Akram, Z. (2021). Association analysis for agronomic traits in wheat under terminal heat stress. *Saudi Journal of Biological Sciences* 28, 7404–7415. doi: https://doi.org/10.1016/j.sjbs.2021.08.050
Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R., and Pandy,

B. K. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics Journal* 2(4), 141–162. url: https://www.pomics.com/Pradeep_2_4_2009_141_162.pdf

Kumar, H., Chugh, V., Kumar, M., Gupta, V., Prasad, S., Kumar, S., Singh, Ch. M., Kumar, R., Singh, B. K., Panwar, G., and Kumar, M. (2023a). Investigating the impact of terminal heat stress on contrasting wheat cultivars: a comprehensive analysis of phenological, physiological, and biochemical traits. *Frontiers in Plant Science* 14(1189005), 1–17. doi: https://doi.org/10.3389/fpls.2023.1189005

Kumar, R., Kumar, S., Azad, C. S., and Baranwal, D. (2023b). Genetic variability, correlation and path coefficient analysis for yield components and grain minerals in wheat (*Triticum aestivum* L.). *The Pharma Innovation Journal* 12(9), 1140–1144. url: https://www.thepharmajournal.com

Kumari, M., Sharma, H., Dadeech, A., and Dashora, A. (2025). Molecular characterization and genetic diversity assessment of bread wheat [*Triticum aestivum* (L.) em. Thell] genotypes using SSR markers. *International Journal of Current Microbiology and Applied Sciences* 14(05), 139–147. doi: https://doi.org/10.20546/ijcmas.2025.1405.014

Li, Y., Tao, F., Hao, Y., Tong, J., Xiao, Y., He, Zh., and Reynolds, M. (2023). Variations in phenological, physiological, plant architectural and yield-related traits, their associations with grain yield and genetic basis. *Annals of Botany* 131, 503–519. doi: https://doi.org/10.1093/aob/mcad003

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., L. Moral, G., Maalouf, F., Nachit, M., Nserallah, N., Ouabbou, H., Rhouma, S., Royo, C., Villegas, D., and Tuberosa, R. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *Journal of Experimental Botany* 62(2), 409–438. doi: https://doi.org/10.1093/jxb/erq287

Ma, J., Zhao, D., Tang, X., Yuan, M., Zhang, D., Xu, M., Duan, Y., Ren, H., Zeng, Q., Wu, J., Han, D., Li, T., and Jiang, L. (2022). Genome-wide association study on root system architecture and identification of candidate genes in wheat (*Triticum aestivum* L.). *International Journal of Molecular Sciences* 23, 777–790. doi: https://doi.org/10.3390/ijms23031843

Mallick, N., Jha, S. K., Agarwal, P., Mall, A., M., N., Kumar, S., Choudhary, M. K., Bansal, S., Saharan, M. S., Sharma, J. B., and Vinod. (2022a). Marker-assisted improvement of bread wheat variety HD2967 for leaf and stripe rust resistance. *Plants* 11, 1152. doi: https://doi.org/10.3390/plants11091152

Mallick, N., Jha, S. K., Agarwal, P., Kumar, S., Mall, A., M., N., Choudhary, M. K., Chandra, A. K., Bansal, S., Saharan, M. S., Sharma, J. B., and Vinod. (2022b). Marker-assisted transfer of leaf and stripe rust resistance from *Triticum turgidum* var. durum cv. Trinakria to wheat variety HD2932. *Frontiers in Genetics* 13, 941287. doi: https://doi.org/10.3389/fgene.2022.941287

Miller, P. A., Williams, J. C., Robinson, H. F., and Comstock, R. I. (1958). Estimates of genotypic and environmental variances and covariance in upland cotton and their implications in selection. *Agronomy Journal* 50(3), 126–131. doi: https://doi.org/10.2134/agronj1958.00021962005000030004x

Mkhabela, S. S., Shimelisa, H., Odindoa, A. O., and Mashilo, J. (2019). Response of selected drought tolerant wheat (*Triticum aestivum* L.) genotypes for agronomic traits and biochemical markers under drought-stressed and non-stressed conditions. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science* 69(8), 674–689. doi: https://doi.org/10.1080/09064710.2019.1641213

Mohammadi, S. A., and Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants: salient statistical tools and considerations. *Crop Science* 43, 1235–1248. doi: https://doi.org/10.2135/cropsci2003.1235

Moosavi, S., Yazdi Samadi, B., Naghavi, M., Zali, A., Dashti, H., and Pourshahbazi, A. (2008). Introduction of new indices to identify relative drought tolerance and resistance in wheat genotypes. *Desert* 12, 165–178. doi: https://doi.org/10.22059/jdesert.2008.27115

Nakano, Y., and Asada, K. (1981). Hydrogen peroxide is scavenged by ascorbate-specific peroxidase in spinach chloroplasts. *Plant and cell physiology* 22, 867–880. doi: https://doi.org/10.1093/oxfordjournals.pcp.a076232

Naroui Rad, M. R., Abdul Kadir, M., Rafii, M. Y., Jaafar, H. Z. E., and Naghavi, M. R. (2012). Bulked segregant analysis for relative water content to detect quantitative trait loci in wheat under drought stress. *Genetics and Molecular Research* 11(4), 3882–3888. doi: http://dx.doi.org/10.4238/2012.November.12.5

Negisho, K., Shibru, S., Matros, A., Pillen, K., Ordon, F., and Wehner, G. (2022). Association mapping of drought tolerance indices in Ethiopian durum wheat (*Triticum turgidum* ssp. Durum). *Frontiers in Plant Science* 13(838088), 1–15. doi: https://doi.org/10.3389/fpls.2022.838088

Nourmand-moaied, F., Rostami. M. A., and Ghannadha, M. R. (2001). A study of morpho-physiological traits of bread wheat (*Triticum aestivum* L.) relationship with grain yield under normal and drought stress conditions. *Iranian Journal of Agricultural Science* 32(4), 785–794. url: https://sid.ir/paper/436197/fa

Oguz, M. C., Aycan, M., Oguz, E., Poyraz, I., and Yildiz, M. (2022). Drought stress tolerance in plants: interplay of molecular, biochemical and physiological responses in important development stages. *Physiologia* 2, 180–197. doi: https://doi.org/10.3390/physiologia2040015

Pour-Aboughadareh, A., Jadidi, O., Shooshtari, L., Poczai, P., and Mehrabi, A. A. (2022). Association analysis for some biochemical traits in wild relatives of wheat under drought stress conditions. *Genes* 13, 1–14. doi: https://doi.org/10.3390/genes13081491

R Core Team (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. url: https://www.R-project.org/.

Rabieyan, E., Bihamta, M. R., Esmaeilzadeh Moghaddam, M., Alipour, H., Mohammadi, V., Azizyan, K., and Javid, S. (2023). Analysis of genetic diversity and genome-wide association study for drought tolerance related traits in Iranian bread wheat. *BMC Plant Biology* 23(431), 1–27. doi: https://doi.org/10.1186/s12870-023-04416-3

Ramachandra Reddy, A., Chaitanya, K. V., Jutur, P. P., and Sumithra, K. (2004). Differential antioxidative responses to water stress among five mulberry (*Morus alba* L.) cultivars. *Environmental and experimental botany* 52(1), 33–42. doi: https://doi.org/10.1016/j.envexpbot.2004.01.002

Rashid, U., Yasmin, H., Hassan, M. N., Naz, R., Nosheen, A., Sajjad, M., Ilyas, N., Keyani, R., Jabeen, Z., Mumtaz, S., Alyemeni, M. N., and Ahmad, P. (2022). Drought-tolerant *Bacillus megaterium* isolated from semi-arid conditions induces systemic tolerance of wheat under drought conditions. *Plant Cell Reports* 41(3), 549–569. doi: https://doi.org/10.1007/s00299-020-02640-x

Reddy, S. S., Saini, D. K., Singh, G. M., Sharma, S., Mishra, V. K., and Joshi, A. K. (2023). Genome-wide association mapping of genomic regions associated with drought stress

tolerance at seedling and reproductive stages in bread wheat. *Frontiers in Plant Science* 14(1166439), 1–17. doi: https://doi.org/10.3389/fpls.2023.1166439

Rosewarne, G. M., Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Lan, C. X., and He, Z. H. (2013). Quantitative trait loci of stripe rust resistance in wheat. *Theoretical Applied Genetics* 126(10), 2427–2449. doi: https://doi.org/10.1007/s00122-013-2159-9

Rossielli, A., and Hamblin, A. (1981). Theorical aspects of selection for stress and non-stress environment. *Crop Science* 21, 1441–1446. doi: https://doi.org/10.2135/cropsci1981.0011183X002100060033x

Ruan, C. J., Li, H., and Mopper, S. (2009). Characterization and identification of ISSR markers associated with resistance to dried-shrink disease in Sea Buckthorn. *Molecular Breeding* 24(3), 255–268. doi: https://doi.org/10.1007/s11032-009-9288-5

Saed-Moucheshi, A., Razi, H., Dadkhodaie, A., Ghodsi, M., and Dastfal, M. (2019). Association of biochemical traits with grain yield in triticale genotypes under normal irrigation and drought stress conditions. *Australian Journal of Crop Science* 13, 272–281. doi: https://doi.org/10.21475/ajcs.19.13.02.p1403

Saeed, I., Chen, X., Bachir, D. G., Chen, L., and Hu, Y-G. (2017). Association mapping for photosynthesis and yield traits under two moisture conditions and their drought indices in winter bread wheat (*Triticum aestivum* L.) using SSR markers. *Australian Journal of Crop Science* 11(03), 248–257. doi: https://doi.org/10.21475/ajcs.17.11.03pne252

Sallam, A., Alqudah, A. M., Dawood, M., Baenziger, P. S., and Borner, A. (2019). Drought stress tolerance in wheat and barley: advances in physiology, breeding and genetics research. *International Journal of Molecular Sciences* 20, 1–36. doi: https://doi.org/10.3390/ijms20133137

Sallam, M., Ghazy, A., Al-Doss, A., and Al-Ashkar, I. (2024a). Combining genetic and phenotypic analyses for detecting bread wheat genotypes of drought tolerance through multivariate analysis techniques. *Life* 14, 183. doi: https://doi.org/10.3390/life14020183

Sallam, M., Al-Ashkar, I., Al-Doss, A., Al-Gaadi, K. A., Zeyada, A. M., and Ghazy, A. (2024b). Assessing heat stress tolerance of wheat genotypes through integrated molecular and physio-biochemical analyses. *Agronomy* 14, 1999. doi: https://doi.org/10.3390/agronomy14091999

Sinha, A. K. (1972). Colorimetric assay of catalase. *Analytical Biochemistry* 47, 389–394. doi: https://doi.org/10.1016/0003-2697(72)90132-7

Shah, A. A., Bhat, R. A., Bhat, B. A., and Mondal, S. K. (2019). Genetic evaluation of winter wheat genotypes under rainfed conditions. *International Journal of Chemical Studies* 7(1), 1064–1071. url: https://www.chemijournal.com/archives/2019/vol7issue1/PartS/6-4-853-700.pdf

Sowadan, O., Xu, S., Li, Y., Muleke, E. M., Sitoe, H. M., Dang, X., Jiang, J., Dong, H., and Hong, D. (2024). Genome-wide association analysis unravels new quantitative trait loci (QTLs) for eight lodging resistance constituent traits in rice (*Oryza sativa* L.). *Genes* 15(105), 1–18. doi: https://doi.org/10.3390/genes15010105

Sunil kumar, V. P., Krishna, H., Devate, N. B., Manjunath, K. K., Chauhan, D., Singh, S., Sinha, N., Singh, J. B., Prakasha, T. L., Pal, D., Sivasamy, M., Jain, N., Singh, G. P., and Singh, P. K. (2023). Marker-assisted selection for transfer of QTLs to a promising line for drought tolerance in wheat (*Triticum aestivum* L.). *Frontiers in Plant Science* 14(1147200), 1–13. doi: https://doi.org/10.3389/fpls.2023.1147200

Vaillancourt, A., Nkongolo, K., Michael, P., and Mehes, M. (2008). Identification, characterization, and chromosome locations of rye and wheat specific ISSR and SCAR markers useful for breeding purposes. *Euphytica* 159(3), 297–306. doi: https://doi.org/10.1007/s10681-007-9492-5

Yusuf, Z., Mohammed, W., Zeleke, H., Hussein, Sh., and Arno, H. (2021). Coheritability and genetic advances of agromorphological and oil quality traits in groundnut (*Arachis hypogaea* L.) genotypes from Ethiopia. *International Journal of Agronomy* 5148772(5). doi: https://doi.org/10.1155/2021/5148772

Zhao, J., Sun, L., Gao, H., Hu, M., Mu, L., Cheng, X., Wang, J., Zhao, Y., Li, Q., Wang, P., Li, H., and Zhang, Y. (2023). Genome-wide association study of yield-related traits in common wheat (*Triticum aestivum* L.) under normal and drought treatment conditions. *Frontiers in Plant Science* 13(1098560), 1–20. doi: https://doi.org/10.3389/fpls.2022.1098560

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 194–204
DOI: 10.46265/genresj.KWVH7038
https://www.genresj.org
ISSN: 2708-3764

Seed germinability of *Stylosanthes* spp. (Fabaceae) accessions under water stress conditions

Vitor Oliveira dos Santos*, Aritana Alves da Silva, Robson de Jesus Santos, Uasley Caldas de Oliveira, Marilza Neves do Nascimento, Claudineia Regina Pelacani

Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, Brazil

Abstract: This study aimed to evaluate and compare the germinative performance of seeds from accessions of *Stylosanthes* spp. held in the Forage Germplasm Bank of the State University of Feira de Santana (BGF-UEFS), Brazil, under water stress conditions during the initial phases of germination. Germination tests were conducted using seeds from six accessions subjected to different osmotic potentials (0.0MPa – distilled water, -0.2, -0.4, -0.6 and -0.8MPa) prepared with polyethylene glycol 6000 (PEG 6000). The experimental design was completely randomized, using 25 seeds per replicate for each treatment. The seeds were evaluated over a period of 5 days under water stress conditions, followed by an additional 5-day recovery period in distilled water for the seeds remaining from the -0.8MPa treatments. The following variables were measured: germination percentage (G%), mean germination time (MGT), germination speed index (GSI), and germination recovery (GR). The results indicated an interaction between factors affecting the germinative behaviour of the *Stylosanthes* spp. accessions for all variables. The genotypes showed significant reductions in G%, with accessions BGF 12-014, BGF 10-018 and BGF 10-029 exhibiting the best performance under the most severe osmotic potential (-0.8MPa). MGT and GSI were also significantly affected by increased water stress. Accessions BGF 12-014, BGF 10-018 and BGF 10-029 were the most promising based on their germinative performance under water stress conditions simulated with PEG 6000 during the early germination phase.

Keywords: Forage legume, polyethylene glycol 6000, osmotic potential, germination, water deficit.

Citation: Oliveira dos Santos, V., Alves da Silva, A., de Jesus Santos, R., Caldas de Oliveira, U., Neves do Nascimento, M., and Pelacani, C. R. (2025) “Seed germinability of *Stylosanthes* spp. (Fabaceae) accessions under water stress conditions”, *Genetic Resources*, 6(12), pp. 194–204. doi: 10.46265/genresj.KWVH7038.

Introduction

The Brazilian Semi-Arid region (SAB) comprises 1,262 municipalities, most of which are located in the northeastern part of the country (IBGE, 2023). This territory is notably characterized by a hot, dry climate and highly irregular rainfall throughout the year, which directly impacts the region’s socioeconomic development (Simões *et al*, 2022), as water scarcity significantly influences local agricultural and livestock activity.

With regard to livestock farming in the SAB, the region is home to approximately 65% and 90% of the country’s sheep and goat herds, respectively, and about 14.8% of the national cattle herds (IBGE, 2018). Most of this livestock production is carried out under extensive systems that rely on low-yielding forages, especially during periods of reduced rainfall (Souza *et al*, 2020). Moreover, the climatic conditions of the SAB limit water availability for the development of forage species, and animal feeding is often partially compromised (Gusha *et al*, 2015). During the driest periods, producers are forced to seek alternative feed sources because pasture forage production becomes insufficient to meet the animals’ nutritional requirements, resulting in increased costs.

In addition, few studies focused on improving forage plants adapted to the SAB have been conducted to date. Consequently, the search for alternatives to reduce livestock production costs in the region remains limited to a few studies. In light of these challenges, it is essential to explore plant

* Corresponding author: Vitor Oliveira dos Santos (vitor.agro.uefs@gmail.com)

Received: 10.09.2025 | Accepted: 22.10.2025 | Published online: 10.12.2025

genetic resources that are tolerant to the climatic conditions of the SAB to mitigate the impacts of irregular rainfall in the region. Among the forage species that demonstrate tolerance to water stress are elephant grass (*Cenchrus purpureus* (Schumach.) Morrone.), lead tree (*Leucaena leucocephala* (Lam.) de Wit), mexican lilac (*Gliricidia sepium* Jacq.), birdwood grass (*Cenchrus ciliaris* L.), and species of the genus *Stylosanthes* Sw.

The genus *Stylosanthes* Sw. (Fabaceae Lindl.) is widely distributed across the Americas, with Brazil harbouring the greatest diversity of species – 38 in total, 17 of which are endemic to the country (Flora e Funga do Brasil, 2020). Additionally, several species within the genus are considered plant genetic resources because of their suitability for animal feed (Canzi *et al*, 2021) owing to their high forage potential, biomass production and protein content. Furthermore, their adaptation to acidic soils and tolerance to water scarcity (Gonzalez *et al*, 2000; Liu *et al*, 2019; Habermann *et al*, 2021) make them particularly suitable for cultivation in regions with edaphoclimatic conditions, such as the SAB.

Notably, the semi-arid region is considered one of the centres of diversity for the *Stylosanthes* genus, as evidenced by expeditions carried out in the semi-arid mesoregions of Bahia between 2007 and 2019 (Santos Júnior *et al*, 2022). Genotypes collected during these expeditions are conserved in the Forage Germplasm Bank of the State University of Feira de Santana (BGF-UEFS), located in the municipality of Feira de Santana, Bahia, with approximately 370 accessions catalogued from these regions (Santos Júnior *et al*, 2022; Silva *et al*, 2024). However, these materials still lack comprehensive studies on their genotypic, morphoagronomic and physiological traits, as well as their performance under different environmental conditions, especially under water stress, which is one of the main abiotic factors limiting productivity.

Water stress affects numerous physiological processes in plants, including the reduction in transpiration rates and degradation of photosynthetic pigments, ultimately impacting photosynthetic efficiency (Lawlor and Cornic, 2002; Hussain *et al*, 2018). Nevertheless, Hussain *et al* (2018) highlighted that those plants in water-limited environments exhibit a range of adaptive responses that confer tolerance to water stress.

During germination, the seed rehydrates its tissues through water absorption, triggering the embryo's metabolic processes and resuming its growth (Bradford, 1990; Bewley and Black, 1994). In this context, water availability directly influences the success of this process, as germination may be inhibited or delayed under water deficits (Bradford, 1990; Roberts, 1973; Bewley and Black, 1994). Consequently, when seeds are exposed to such environmental conditions in productive areas, germination may be interrupted or postponed, leading to field emergence irregularities, inefficient land use, and reduced forage production, negatively impacting the feed supply for livestock.

Thus, assessing seed germinability under water deficit or limiting moisture conditions provides valuable insights for identifying genotypes tolerant to semi-arid regions. This is especially relevant because germination is one of the most critical stages in the plant life cycle. Polyethylene glycol 6000 (PEG 6000) has been widely used to simulate water restriction during seed germination (Braccini *et al*, 1996), as its high molecular weight prevents it from penetrating the plant cell membrane, thereby avoiding toxicity (Marcos Filho, 2002).

In this context, understanding the germinative behaviour of genotypes subjected to water-limited conditions is essential for selecting materials suitable for cultivation in arid environments. Therefore, this study aimed to evaluate and compare the germinability of seeds from *Stylosanthes* spp. accessions held in BGF-UEFS under water stress during germination.

Materials and methods

The experiment was conducted in the Seed Germination Laboratory (LAGER) at the State University of Feira de Santana (UEFS), Brazil. The genetic materials used in this study were conserved in the Forage Germplasm Bank of UEFS (BGF-UEFS) and were obtained from accessions propagated between 2014 and 2020, with progenitors collected from the semi-arid region of Bahia (Table 1) and stored following the methodology of Gómez-Campo (2006), kept in properly labelled kraft paper envelopes, sealed with galvanized staples, and placed in airtight containers with silica gel as a moisture indicator, at room temperature (approximately 25.1ºC); with seed moisture content ranging approximately between 4.5% and 8%. Moreover, the accessions were selected based on seed viability above 80%.

Table 1. Passport data of *Stylosanthes* spp. accessions stored in the Forage Germplasm Bank of the State University of Feira de Santana (BGF-UEFS)

Accession	Species	Year of seed regeneration	Municipality	Coordinates
BGF 10-016	*S. scabra*	2020	Queimadas	10°54’40”S/39°12’17”W
BGF 12-014	*S. humilis*	2020	Canarana	11º48’597”S/41º42’066”W
BGF 014-P137-2	*S. scabra*	2020	Seabra	12º27’311”S/42º11’452”W
BGF 10-018	*S. scabra*	2014	Candeal	11°49’49,8”S/ 39°07’08,5”W
BGF 10-034	*S. scabra*	2014	Feira de Santana	12°09’719”S/38°57’696”W
BGF 10-029	*S. viscosa*	2014	Canudos	09°54’29,9”S/39°03’17,2”W

To overcome seed coat dormancy, mechanical scarification was performed using sandpaper (sandpaper no. 150) (Silva *et al*, 2024). Additionally, to reduce contamination during the experiment, the seeds were disinfected in a 0.5% sodium hypochlorite solution for 10 min and then rinsed with distilled water.

Seeds from the accessions were subjected to treatment with PEG 6000 solutions at different osmotic potentials (0.0MPa for pure distilled water, -0.2MPa, -0.4MPa, -0.6MPa and -0.8MPa), simulating water stress, as proposed by Villela *et al* (1991), with adjustments made at a temperature of 30°C. Germination tests were carried out in 60mm glass Petri dishes lined with two sheets of sterilized germination paper (germitest type) moistened with 2ml of the respective pre-prepared PEG 6000 solutions. Every two days, the Petri dishes and germination papers were replaced, and the osmotic solutions were replenished to maintain the target osmotic potentials for each treatment.

The tests were conducted in biochemical oxygen demand (B.O.D.) germination chambers, under constant temperature (30°C) and in the dark. Germination was defined as the emergence of the radicle (≥ 2mm), with daily counts recorded throughout the 5-day evaluation period. The choice of 5 days was based on preliminary tests (unpublished data), which showed that after this period, the germination of the remaining seeds was not significant.

On the fifth day of evaluation, seeds that did not show radicle protrusion in -0,8MPa were removed from the PEG solution, rinsed to eliminate residual PEG 6000, and transferred to new Petri dishes with fresh germitest paper moistened with 2ml of distilled water. This set of remaining seeds was returned to the B.O.D. chamber for an additional 5-day period, referred to as the 'germination recovery' phase. Germination during this phase was recorded, and the results were expressed as the percentage of recovered seeds, based on a total of 25 sown seeds per replicate.

The germination percentage (G%) was calculated according to Laboriau and Valadares (1976): G% = (N/A) × 100, where: G% = germination percentage, N = total number of germinated seeds, A = total number of seeds tested. Mean germination time (MGT; days) was calculated following Labouriau (1993): MGT = $(\Sigma n_i \cdot t_i)/\Sigma ni$, where: n_i = number of seeds germinated at time interval t_i in days. Germination speed index (GSI; seeds·day^{-1}) was calculated according to Maguire (1962): GSI = G1/N1 + G2/N2 + ... + Gn/Nn, where: G1, G2, ..., Gn = number of seeds germinated on each respective day, N1, N2, ..., Nn = number of days from the start of the test to each respective counting day.

The experiment followed a completely randomized design with four replicates, each consisting of one Petri dish containing 25 seeds. The data were tested for the assumptions of residual normality using the Shapiro–Wilk test and for homogeneity of variances using Bartlett's test. For G%, because the assumptions required for analysis of variance (ANOVA) were not initially met, the data were transformed using the arcsine square root function: arcsin√(x/100).

Once the assumptions were satisfied, ANOVA was conducted. Upon observing a significant interaction between the factors (P-value < 0.05), regression curves were constructed using the best-fit model (based on R^2 values). To compare accession performance within each osmotic potential, the means were grouped using the Scott–Knott test (Scott & Knott, 1974).

All analyses and plots were conducted using R statistical software (version 2024.12.0+467) (R Core Team, 2024).

Results

The analysis of variance (Table 2) revealed a significant interaction between accessions and osmotic potentials for all variables analyzed: germination (G%), germination speed index (GSI) and mean germination time (MGT) indicating that the combination of osmotic potentials and genotypes had a differential effect on seed germination behaviour.

Table 2. Summary of the analysis of variance (ANOVA) for the variables evaluated during germination of six accessions of the genus *Stylosanthes* spp. subjected to different water potentials of PEG 6000. DF, degrees of freedom; G%, germination; GSI, germination speed index; MGT, mean germination time; CV, coefficient of variation; **, significant (p-value < 0.01) and *, (p-value < 0.05) by the F-test, respectively.

Variation source	DF	G%	GSI	MGT
Accession	5	12.41**	17.13*	4.62**
Osmotic potential	4	80.24**	499.60**	85.32**
Accession X osmotic potential	20	4.38**	6.23**	5.07**
Residual	90			
CV(%)		13.19	12.19	21.58

Figure 1 shows the influence of osmotic potential on the germination of seeds from different *Stylosanthes* spp. accessions based on regression analysis. Differences among accessions were evident, as indicated by the distinct slopes and intercepts of the regression curves. Additionally, the R^2 values associated with each regression model indicated a good fit, ranging from 88.51% to 97.61%, demonstrating a strong correlation between osmotic potential and germination under the tested conditions. At osmotic potentials of 0.0, -0.2 and -0.4MPa, the accessions exhibited similar performance, with little variation in the germination percentage. However, at lower osmotic potentials (-0.6 and -0.8MPa), more pronounced differences in the performance of the tested genetic materials were observed than at higher potentials, highlighting that the genotypes exhibit divergent performances, especially under more severe water stress conditions.

Accessions BGF 10-029 (Figure 1f), BGF 12-014 (Figure 1b), and BGF 10-018 (Figure 1d) exhibited less steep regression curves than the other materials, indicating lower sensitivity to water stress. Moreover, under the most negative osmotic potential, these accessions maintained relatively high mean germination percentages, ranging from 63% to 66%. In contrast, accession BGF 014-P137-2 (Figure 1c) showed intermediate performance under the same conditions, with a germination rate of 38%. Accessions BGF 10-034 (Figure 1e) and BGF 10-016 (Figure 1a) showed the poorest performance under the imposed water stress conditions, with germination percentages at the most negative osmotic potential reaching only 19% and 6%, respectively, reflecting substantial reductions compared to the control treatment (0.0MPa), and indicating that under low water availability, these genetic materials are notably disadvantaged in terms of germination occurrence.

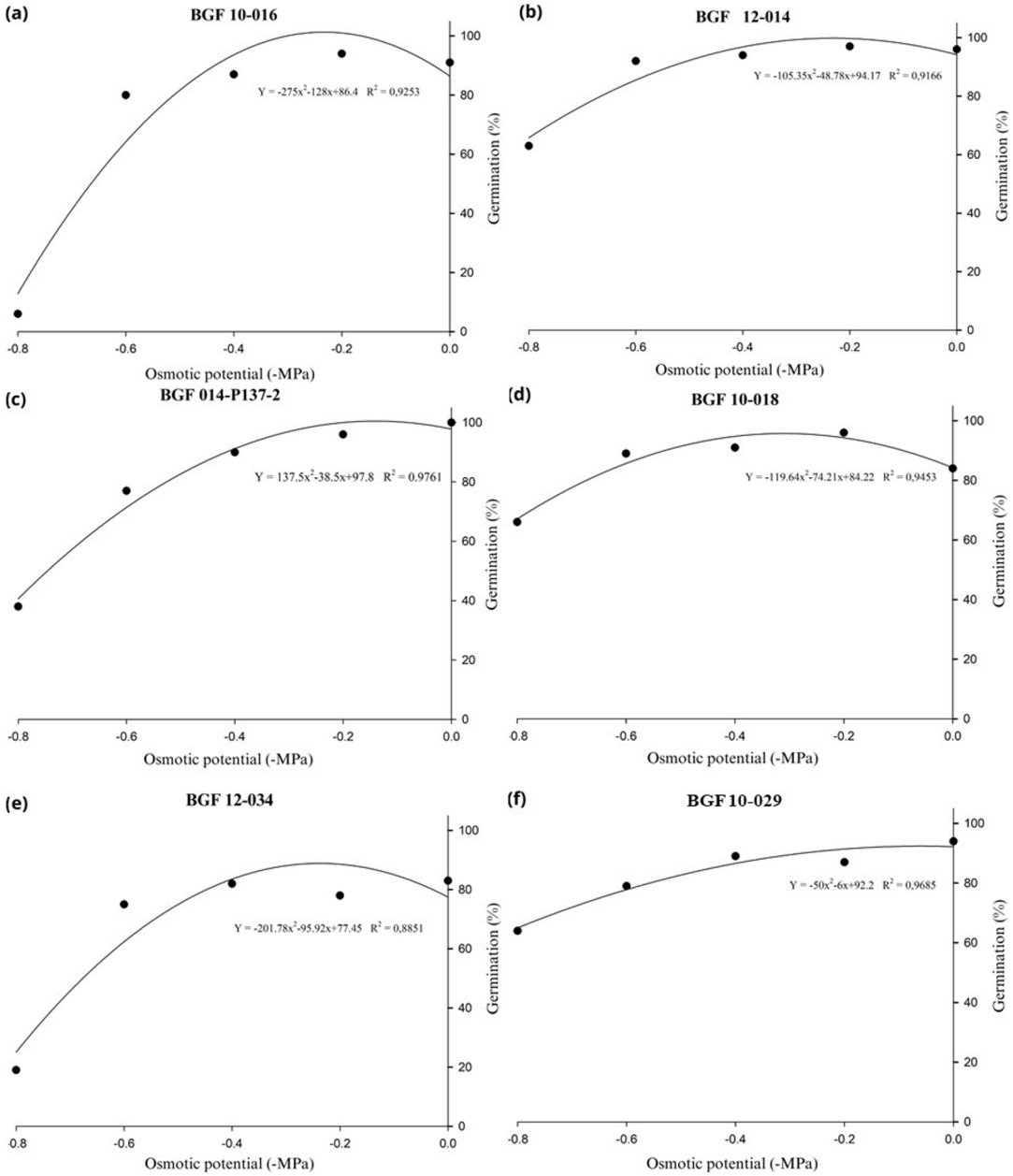


Figure 1. Regression plots showing the effect of different concentrations of PEG 6000 on the osmotic potential and seed germination of *Stylosanthes* spp. accessions. PEG 6000 concentrations were used to simulate water stress, ranging from 0.0 to -0.8MPa. Each point represents the mean of four replications, and the lines represent the fit of the data using a quadratic regression model.

The breakdown of accession performance at each osmotic potential is presented in Table 3. According to the Scott–Knott clustering test, under control conditions, the materials were divided into two groups, with accessions BGF 10-016, BGF 12-014, BGF 014-P137-2 and BGF 10-029 exhibiting higher germination rates. Additionally, at -0.2MPa, all accessions were grouped together, except for BGF 10-036, whose mean value was significantly lower than that of the other genotypes. At osmotic potentials of -0.4 and -0.6MPa, no significant differences were detected among the accessions, as a single group was formed based on the genotype.

Furthermore, at the most negative osmotic potential, germination percentages among accessions once again showed significant differences, as this level of the factor yielded the greatest variation in the results. The accessions that demonstrated the best performance under these conditions were BGF 12-014, BGF 10-018 and BGF 10-029, with germination percentages of 63.0%, 66.0% and 64.0%, respectively. At the same osmotic potential, all other accessions were considered statistically different from the aforementioned group and from each other, as BGF 014-P137-2, BGF 10-034 and BGF 10-016 were each assigned to separate groups.

Table 3. Germinability of seeds from *Stylosanthes* spp. accessions subjected to different concentrations of PEG 6000. Means followed by the same letter in the columns do not differ from each other according to the Scott–Knott test at 5%.

	Germination percentage (G%)				
	Osmotic Potential (-MPa)				
Accession	**0.0**	**-0.2**	**-0.4**	**-0.6**	**-0.8**
BGF 10-016	91.0 a	94.0 a	87.0 a	80.0 a	6.0 d
BGF 12-014	96.0 a	97.0 a	94.0 a	92.0 a	63.0 a
BGF 014-P137-2	100.0 a	96.0 a	90.0 a	77.0 a	38.0 b
BGF 10-018	84.0 b	96.0 a	91.0 a	89.0 a	66.0 a
BGF 10-034	83.0 b	78.0 b	82.0 a	75.0 a	19.0 c
BGF 10-029	94.0 a	87.0 a	89.0 a	79.0 a	64.0 a

The regression graphs showing the effect of PEG 6000 on MGT of the tested materials are presented in Figure 2. Accession BGF 10-016 did not fit either the linear or quadratic regression models, preventing the construction of a curve to represent the data. Most of the remaining accessions were best described by a quadratic regression model, with determination coefficients (R^2) ranging from 93.75% to 99.16%, indicating a strong fit to observed trends. For accessions BGF 014-P137-2 (Figure 2b), BGF 10-034 (Figure 2d), and BGF 10-029 (Figure 2e), significant increases in MGT began to occur at -0.4MPa, indicating that this osmotic potential was already sufficient to delay the germination process in these genotypes. In contrast, for accessions BGF 12-014 (Figure 2a) and BGF 10-018 (Figure 2c), this sharp increase was only observed starting at -0.6MPa, indicating that these materials are less sensitive to osmotic stress compared to the others, as changes in MGT were only triggered under lower water availability conditions.

Additionally, the Scott–Knott clustering test for MGT (Table 4) showed that at osmotic potentials of 0.0, -0.2 and -0.6MPa, the accessions did not differ significantly from one another, indicating similar performances under these environmental conditions. However, at -0.4MPa, accessions BGF 12-014, BGF 10-018 and BGF 10-029 displayed superior MGT values, differing significantly from the other genetic materials, which exhibited higher MGTs, highlighting the variation in sensitivity to water stress among the accessions. At the most negative potential (-0.8MPa), only BGF 10-016 showed a significantly different MGT compared to the others, with a lower mean germination time than the remaining accessions.

Table 4. Mean germination time (MGT) of seeds from *Stylosanthes* spp. accessions subjected to different concentrations of PEG 6000. Means followed by the same letter in the columns do not differ from each other according to the Scott-Knott test at 5%.

	Mean germination time (MGT) (days)				
	Osmotic Potential (-MPa)				
Accession	**0.0**	**-0.2**	**-0.4**	**-0.6**	**-0.8**
BGF 10-016	1.10 a	1.10 a	1.72 a	3.00 a	1.00 a
BGF 12-014	1.17 a	1.45 a	1.55 a	2.32 a	3.30 b
BGF 014-P137- 2	1.05 a	1.37 a	2.25 b	2.40 a	3.20 b
BGF 10-018	1.12 a	1.15 a	1.25 a	2.55 a	3.10 b
BGF 10-034	1.22 a	1.35 a	2.02 b	2.72 a	3.42 b
BGF 10-029	1.07 a	1.02 a	1.67 a	2.40 a	3.00 b

The relationship between osmotic potential and GSI was modelled using quadratic regression equations (Figure 3), with R^2 values ranging from 0.9110 to 0.9991, indicating a strong dependence of GSI on the osmotic conditions imposed by PEG 6000 treatments. The accessions showed distinct GSI responses as the osmotic potential decreased. Under the control treatment (0MPa), only accession BGF 10-034 exhibited a GSI below 20, while the other accessions showed similar values, approximately at the same level. The fitted equations suggest that the accessions respond differently to water availability, with greater or lesser reductions in GSI as water stress increases across the evaluated materials.

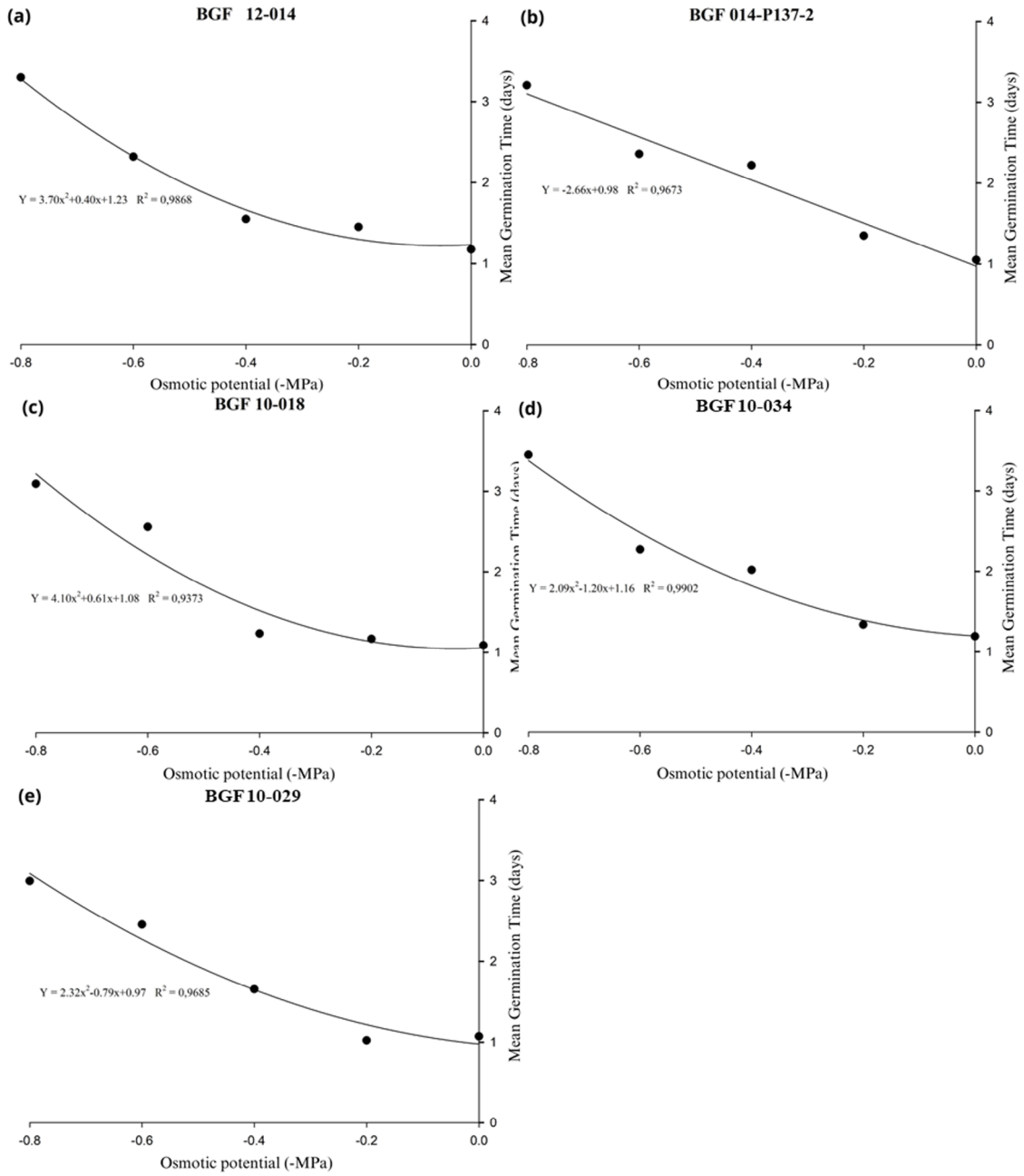


Figure 2. Regression plots showing the effect of different concentrations of PEG 6000 on the mean germination time (MGT) of *Stylosanthes* spp. accessions. PEG 6000 concentrations were used to simulate water stress, ranging from 0.0 to -0.8MPa. Each point represents the mean of four replications, and the lines represent the fit of the data using a quadratic or linear regression mode. Accession BGF 10-016 is not included in the figure because it did not fit either the linear or quadratic regression models, preventing the construction of a representative curve.

Accession BGF 10-018 (Figure 3d) maintained a high GSI up to -0.4MPa, with a sharp decline only observed at -0.6MPa, standing out as the material with the least variation in GSI as the osmotic potential decreased. Accessions BGF 10-016 (Figure 3a), BGF 10-034 (Figure 3f) and BGF 10-029 (Figure 3e) maintained high GSI values at the two least concentrated PEG levels (0.0 and -0.2MPa) but exhibited significant reductions at -0.4MPa and again at -0.8MPa. Similarly, accessions BGF 12-014 (Figure 3b) and BGF 014-P137-2 (Figure 3c) also showed marked decreases in GSI as osmotic potential declined; however, an osmotic potential of only -0.2MPa was already sufficient to significantly affect the GSI of these genotypes, as this condition caused noticeable differences compared to the control, highlighting

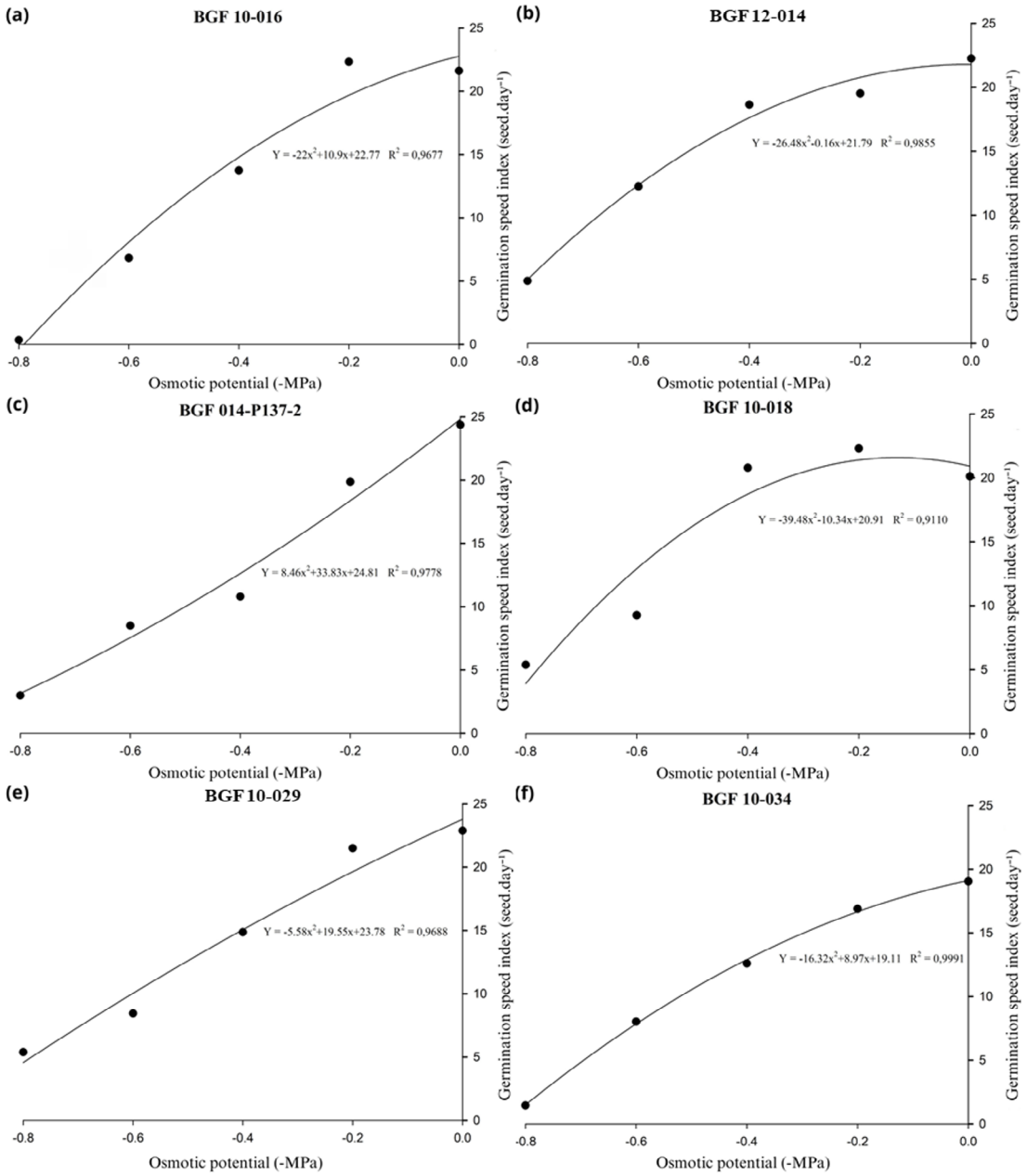


Figure 3. Regression plots showing the effect of different osmotic potentials on the germination seed index (GSI) of *Stylosanthes* spp. accessions. PEG 6000 concentrations were used to simulate water stress, ranging from 0.0 to -0.8MPa. Each point represents the mean of four replications, and the lines represent the fit of the data using a quadratic regression mode.

that even a slight reduction in water availability is enough to exert a significant influence on this variable.

According to the Scott–Knott test (Table 5), variability in GSI among the accessions was evident at each osmotic potential. Under the control condition (0.0MPa), accessions BGF 10-034 and BGF 10-018 had the lowest GSI values, whereas the other accessions had higher GSI means. Moreover, variations in genotype responses were observed across the remaining osmotic potentials. Under the most severe water stress (-0.8MPa), accessions BGF 12-014, BGF 10-018 and BGF 10-029 showed the highest GSI values, standing out under this condition of limited water availability.

Table 5. Germination speed index of seeds from *Stylosanthes* spp. accessions subjected to different PEG 6000 concentrations. Means followed by the same letter in the columns do not differ from each other according to the Scott-Knott test at 5%.

	Germination speed index (GSI) (seeds.days-1) Osmotic Potential (-MPa)				
Accession	**0.0**	**-0.2**	**-0.4**	**-0.6**	**-0.8**
BGF 10-016	21.62 a	22.33 a	13.75 b	6.83 b	0.35 b
BGF 12-014	22.25 a	19.52 b	18.64 a	12.25 a	4.87 a
BGF 014-P137-2	24.37 a	19.87 b	10.80 c	8.50 b	2.99 b
BGF 10-018	20.12 b	22.30 a	20.79 a	9.26 b	5.39 a
BGF 10-034	19.05 b	16.89 c	12.60 c	8.03 b	1.45 b
BGF 10-029	22.89 a	21.50 a	14.87 b	8.45 b	5.39 a

The analysis of variance (Table 6) for the stage referred to as 'germination recovery' – which aimed to assess the proportional germination of the remaining seeds from the accessions previously subjected to the most severe water stress treatment (-0.8MPa), now placed in distilled water (0.0MPa) – revealed significant differences among the accessions.

Table 6. Summary of analysis of variance (ANOVA) for germination recovery of *Stylosanthes* spp. accessions after exposure to water stress at a potential of -0.8MPa. DF, degrees of freedom; REC, germination recovery (%); CV, coefficient of variation; **, significant (p-value < 0.05).

Variation source	**DF**	**REC**
Accession	5	4.78**
Residual	18	
CV (%)		30.44

The Scott–Knott mean grouping test revealed the formation of two distinct groups for germination recovery. Accessions BGF 10-016 and BGF 014-P137-2 exhibited the highest recovery percentages of 89.6% and 88.7%, respectively. All other accessions showed significantly lower recovery rates than these two, highlighting that water stress during the initial germination phase had a significant impact, leading to greater loss of seed viability in these genotypes (Figure 4).

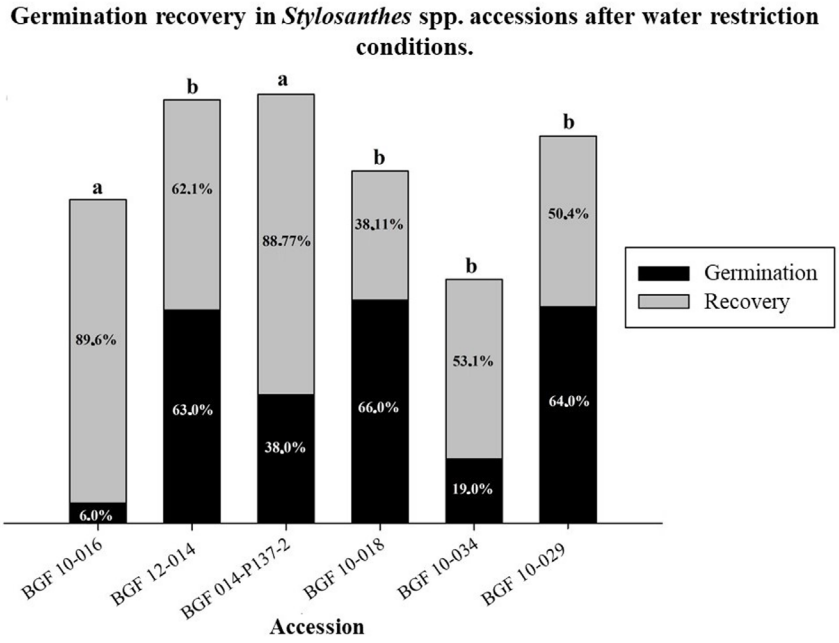


Figure 4. Graph of germination recovery in distilled water of *Stylosanthes* spp. accessions after water stress at a potential of -0.8MPa. Different letters indicate significant differences in recovery rates according to the Scott–Knott test (5%).

Discussion

The germination percentages at -0.8MPa (Figure 1) confirmed that this concentration of PEG 6000 reduced the number of germinated seeds. Yamashita *et al* (2018) evaluated the effect of water stress induced by PEG 6000 at different osmotic potentials (0.0; -0.2; -0.4; -0.6; -0.8; and -1MPa) on the germinability of the species *Stylosanthes capitata* Vogel, and found that the more negative the solution potential, the greater the reduction in seed germination percentage – results that are consistent with those observed in the present study.

Moreover, Braccini *et al* (1996) explained the reduction effect caused by the use of PEG 6000, who considered the high molecular weight of the substance, its high viscosity, and low O_2 diffusion rate as factors that directly impair oxygen availability to seeds during germination. Additionally, Oliveira *et al* (2017) reiterated that the high viscosity of PEG 6000 is also directly related to the seed's ability to absorb water, because as the osmotic potential of the solution decreases, the availability of water falls below the level required for seeds to resume metabolic activities and, consequently, for the embryonic axis to grow.

Tolerance to water stress is an important characteristic to consider when recommending genotypes capable of withstanding different osmotic potentials and water scarcity, especially in ecologically challenging areas with low water availability and saline characteristics (Rego *et al*, 2011). Moreover, plants adapted to such conditions not only survive and establish successfully but also complete their reproductive cycle, ensuring population persistence in harsh environments (Baskin & Baskin, 2014; Nicotra *et al*, 2010). The BGF-UEFS accessions presented both intra- and interspecific genetic variability, and combined with the fact that they were collected from different locations in the semi-arid region of Bahia (Oliveira and Queiróz, 2016), this may explain the differing germination performances of the *Stylosanthes* spp. genotypes under water stress induced by PEG 6000, as observed in this study. The genetic materials BGF 12-014, BGF 10-018, and BGF 10-029 were superior to the other genotypes.

The regression analysis for MGT of the accessions (Figure 2) suggested that as PEG 6000 concentration increased, seeds required more time to emit the radicle. Duarte *et al* (2018), when observing the effect of water stress on the germination of white angico (*Anadenanthera colubrina* var. cebil (Vell.) Brenan), also found results similar to those of this study, with an increased MGT under lower water availability. This can be explained by the fact that changes in water potential can affect the hydraulic properties of the seed coat, and from this perspective, the lower the potential, the lower the water diffusibility (Antunes *et al*, 2011). This phenomenon delays water absorption by the seed and, consequently, germination (Bradford 1990; Ávila *et al*, 2007).

The divergence in MGT results among the accessions may be attributed to genotypic variation, as they originated from different localities and may have different seed physiological qualities. The heterogeneity among them leads to differences in the average MGT (Kolchinski *et al*, 2005). From an ecological perspective, delayed germination under low water potential can act as a selective filter, favouring genotypes capable of maintaining viability and responding only under favourable moisture conditions (Bradford, 2002). In this regard, accession BGF 10-016 exhibited the most favourable germination pattern under the highest water stress level. However, the low number of germinated seeds in this accession suggests caution in interpreting this result, as the apparent performance may reflect a small number of highly vigorous seeds rather than a consistent tolerance across the seed lot. Moreover, Santos *et al* (2016), investigating water stress simulated by PEG 6000 in seeds of *Caatinga* species, Catingueira (*Poincianella pyramidalis* (Tul.) L. P. Queiroz) and White angico (*Anadenanthera colubrina* (Vell.) Brenan), observed a higher MGT compared to all *Stylosanthes* spp. accessions tested in this study at -0.8MPa, being 4.85 and 3.48 days, respectively.

As pointed out by Bewley and Black (1994), when the available moisture is below the necessary level, enzymatic activity is significantly reduced, hindering seed hydration and compromising the ability to metabolize internal reserves. In this sense, the authors emphasized that this phenomenon results in slower and less efficient germination, as water stress limits the biochemical reactions essential for resuming seed metabolism. Therefore, the results of this study are consistent with those reported in the literature, as higher PEG 6000 concentrations led to a reduction in the GSI of the *Stylosanthes* spp. accessions tested.

Simioni *et al* (2011) found similar results in sorghum seeds (*Sorghum bicolor* (L.) Moench) under water deficit conditions, with GSI decreasing from -0.4MPa. The same was observed by Oliveira *et al* (2017), who analyzed the germination behaviour of cotton genotypes (*Gossypium hirsutum* L.). Azerêdo *et al* (2016) also found that water deficit simulated with PEG 6000 at -0.2MPa was sufficient to reduce the GSI of white angico seeds, a species also used in animal feed.

PEG 6000, owing to its high molecular weight, cannot penetrate seed structures and thus only limits water availability (Marcos Filho, 2002). Accordingly, Marcos Filho (2002) also noted that, since it does not exert toxic effects on the seeds, those that tolerate this temporary interruption in water supply during germination can survive and germinate when moisture becomes favourable. This explains the recovery of germination in some genotypes evaluated in this study. The differences observed among them may be attributed to genotypic variation, as some genotypes may not tolerate partial hydration. In such cases, metabolic activity may be initiated and then halted, where the continued lack of water becomes crucial for maintaining seed viability and successful germination.

This study highlights the variability among BGF-UEFS genotypes in terms of their performance under water stress conditions.

The results enhance our understanding of germination dynamics under water-limited conditions and reveal significant variability among *Stylosanthes* spp. accessions. This variation is particularly important for identifying promising genotypes with potential use as forage resources in semi-arid environments, where water scarcity poses a major challenge to seedling establishment and pasture productivity. While these findings provide valuable insights into early responses to water stress, further research is needed to assess the agronomic potential of these accessions under field conditions. Field trials or common-garden experiments represent a logical next step to evaluate seedling establishment, early

plant development, and overall performance in more variable and realistic environments. Such approaches are essential for selecting resilient accessions suitable for forage production systems facing increasingly unpredictable climatic conditions.

Conclusion

The *Stylosanthes* spp. accessions from BGF-UEFS exhibited different germination performances under water stress conditions for the variables G%, MGT and GSI.

The evaluated germination parameters of the Stylosanthes spp. accessions were negatively affected by the reduction in osmotic potential.

Accessions BGF 012-014, BGF 10-018 and BGF 10-029 were the most promising due to their performance in the evaluated germination parameters under water stress, simulated by PEG 6000, during the initial germination phase.

Author contributions

VOS was responsible for conceptualization and methodology. VOS, AAS, and RJS handled data collection, while VOS and RJS also performed formal analysis. VOS and UCO wrote the original draft. MNN and CRP focused on methodology and visualization. VOS, UCO, MNN, and CRP performed revision and editing. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge FAPESB (Foundation for Research Support of the State of Bahia) for financial support during the execution of this work.

Conflict of interest statement

The authors report no known financial or personal conflicts of interest that could have affected the research or findings presented in this article.

References

Antunes, C. G. C., Pelacani, C. R., Ribeiro, R. C., Souza C. L. M., Castro R. D. (2011). Germinação de sementes de *Caesalpinia pyramidalis* Tul. (catingueira) submetidas a deficiência hídrica. *Revista Árvore*. 35, 1007-1015. doi: https://doi.org/10.1590/S0100-67622011000600006

Ávila, M. R., Braccini, A. L., Scapim, C. A., Flagiari, J. R., Santos J. L. (2007). Influência do estresse hídrico simulado com manitol na germinação de sementes e crescimento de plântulas de canola. *Revista Brasileira de Sementes*. 29, 98-106. doi: https://doi.org/10.1590/S0101-31222007000100014

Azerêdo, G. A., Paula, R. C., Valeri, S. V. (2016). Germinação de sementes de *Piptadenia moniliformis* Benth. sob estresse hídrico. *Ciência Florestal*. 26,193-202. doi: https://doi.org/10.5902/1980509821112

Bewley, J. D., Black, M. (1994) Seeds: physiology of development and germination (New York: Plenum), 445p.

Braccini, A. L., Ruiz, H. A., Braccini, M. C.L., Reis, M. S. (1996). Germinação e vigor de sementes de soja sob estresse hídrico induzido por soluções de cloreto de sódio, manitol e polietileno glicol. *Revista Brasileira de Sementes*. 18, 10-16.

Bradford, K. J. (1990). Influence of water stress on seed germination rate in *Cucumis sativus*. *Plant Physiology*, 80, 351–357. doi: https://doi.org/10.1104/pp.94.2.840

Bradford, K. J. (2002). Applications of hydrothermal time to quantifying and modeling seed germination and dormancy. *Weed Science* 50, 248–260. doi: https://doi.org/10.1614/0043-1745(2002)050[0248:AOHTTQ]2.0.CO;2

Baskin, C., Baskin, J.M. (2014). Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination. *Academic Press*, 8, 150-162.

Canzi, G. M., Gai, V. F., Souza, G. B. P., Effting, P. B. (2021). Adubação nitrogenada sobre a produtividade do Stylosanthes Campo Grande em solo argiloso. *Revista Cultivando o Saber*. 14, 86-94.

Duarte, M. M., Kratz, D., Carvalho, R. L. L., Nogueira, A. C. (2018). Influência do estresse hídrico na germinação de sementes e formação de plântulas de angico branco. *Advances in Forestry Science*. 5, 375-379. doi: https://doi.org/10.34062/afs.v5i3.5521

Flora e Funga do Brasil. *Stylosanthes*. Jardim Botânico do Rio de Janeiro. Available at: https://floradobrasil.jbrj.gov.br/FB115.

Gómez-Campo, C. (2006) Long-term seed preservation: updated standards are urgent (Spain: Universidad Politécnica de Madrid, 168p.

Gonzalez, L. M., Lopez, R. C., Fonseca, I., Ramirez, R. (2000). Growth, stomatal frequency, yield and accumulation of ions in nine species of grassland legumes grown under saline conditions. *Pastos y Forrajes*. 23, 299-308. doi: https://doi.org/10.5555/20013029138

Gusha, J., Halimanti, T. E., Kantsandes, S., Zvinorova, P. I. (2015). The effect of *Opuntia ficus-indica* and forage legumes-based diets on goat productivity in the smallholder sector in Zimbabwe. *Small Ruminant Research*. 125, 21-25. doi: https://doi.org/10.1016/j.smallrumres.2015.02.018

Habermann, E., Oliveira, E. A. D., Delvecchio, G., Belisário, R., Barreto, R. F., Viciado, D. O., Rossingnoli, N. O., Costa, K. A. P., Prado, R. M., Gonzalez-Meler, M., Martinez, C. A. (2021). How does leaf physiological acclimation impact forage production and quality of a warmed managed pasture of *Stylosanthes capitata* under different conditions of soil water availability? *Science of the Total Environment*. 759, 143505. doi: https://doi.org/10.1016/j.scitotenv.2020.143505

Hussain, H. A., Hussain, S., Khaliq, A., Ashraf, U., Anjum, S. A., Men, S., Wang, L. (2018) Chilling and Drought Stresses in Crop Plants: Implications, Cross Talk, and Potential Management Opportunities. *Frontiers in Plant Science*. 9, 1-21. doi: https://doi.org/10.3389/fpls.2018.00393/full

IBGE – Instituto Brasileiro de Geografia e Estatística (2018) Novo Censo Agropecuário mostra crescimento de efetivo de caprinos e ovinos no Nordeste. Avaliable at: https://www.embrapa.br/cim-inteligencia-e-mercado-de-caprinos-e-ovinos/busca897de-noticias//noticia/36365362/novo-censo-agropecuario-mostra-crescimento-de-efetivo-de-caprinos-e-ovinos-no-nordeste.

IBGE – Instituto Brasileiro de Geografia e Estatística (2023) Brasil em síntese. Avaliable at: https://brasilemsintese.ibge.gov.br/territorio.html.

Kolchinski, E. M., Schuch, L. O. B., Peske, S. T. (2005). Vigor de sementes e competição intraespecífica em soja. *Ciência Rural*. 35, 1248-1256. doi: https://doi.org/10.1590/S0103-84782005000600004

Laboriau, L. G.; Valadares, M. B. (1976). On the germination of seeds of *Calotropis procera*. *Anais da Academia Brasileira*

de Ciências, 48, 174–186.
Labouriau, L. G. (1993). A germinação das sementes (Washington: OEA), 174 p.
Lawlor D. W., Cornic G. (2002). Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants. *Plant, Cell & Environment*, 25(275-294). doi: https://doi.org/10.1046/j.0016-8025.2001.00814.x
Liu, P., Huang, R., Hu, X., Jia, Y., Li, J., Luo, J., Liu, Q., Luo, L., Liu, G., Chen, Z. (2019). Physiological responses and proteomic changes reveal insights into *Stylosanthes* response to manganese toxicity. *BMC Plant Biology*, 19, 202–223. doi: https://doi.org/10.1186/s12870-019-1822-y
Maguire, J. D. (1962). Speed of germination-aid in selection and evaluation for seedling emergence and vigor. *Crop Science*, 2,176-177. doi: https://doi.org/10.2135/cropsci1962.0011183X000200020033x
Marcos Filho, J. (2002). Fisiologia das Sementes de Plantas Cultivadas. (Brazil: Fealq), 237p.
McKeon, G.M. (1985). Pasture seed dynamics in a dry monsoonal climate, II The effect of water availability, light and temperature on germination speed and seedling survival of *Stylosanthes humilis* and *Digitaria ciliaris*. *Australian Journal of Ecology*, 10, 149-163. doi: https://doi.org/10.1111/j.1442-9993.1985.tb00876.x
Michel B.E., Kaufmann, M.R. (1973). The osmotic potential of polyethylene glycol 6000. *Plant Physiology*, 5, 914–916. doi: https://doi.org/10.1104/pp.51.5.914
Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M. D., Richards, C. L., Valladares, F., & van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, 15, 684–692. doi: https://doi.org/10.1016/j.tplants.2010.09.008
Oliveira, H., Nascimento, R., Leão, A. R., Cardoso, J. A. F., Guimarães, R. F. B. (2017). Germinação de sementes e estabelecimento de plântulas de algodão submetidas a diferentes concentrações de NaCl e PEG 6000. *Revista Espacios*. 38, 13. doi: https://doi.org/10.5555/19621604893
Oliveira, R. S., Queiroz, M. A. (2016). GENETIC DIVERSITY IN ACCESSIONS OF *Stylosanthes* spp. USING MORPHOAGRONOMIC DESCRIPTORS. *Revista Caatinga*. 29, 101-112.
R Core Team (2024). R: A Language and Environment for Statistical Computing. Avaliable at: https://www.R-project.org/
Rego, S. S., Ferreira, M. M., Nogueira, A. C., Grossi, F., Sousa, R. K., Brondani, G. E., Araujo, M. A., Silva A. L. L (2011). Estresse hídrico e salino na germinação de sementes de *Anadenanthera colubrina* (Velloso) Brenan. *Journal of Biotechnology and Biodiversity*. 2, 37- 42.
Roberts, E. H. (1973). Predicting the storage life of seeds. *Seed Science and Technology*. 1, 499-514. Santos Júnior, S. R. A, Pelacani, C. R., Santos, V. O., Silva, A. A., Fernandes, S. M., Gissi, D. S., Oliveira R. S. (2022). Banco de Germoplasma de Forrageiras da Universidade Estadual de Feira de Santana (BGF-UEFS). *Revista RG News*. 8, 5-15.
Santos, C. A., Silva, N. V., Walter, L. S., Silva, E. C. A., Nogueira, R. J. M. C. (2016). Germinação de duas espécies da caatinga sob déficit hídrico e salinidade. *Pesquisa Florestal Brasileira*, 36, 219-224. doi: https://doi.org/10.4336/2016.pfb.36.87.1017
Silva, A. A., Pelacani, C. R., Grilo, J. S. T. F., Pereira, L. S., Oliveira, R. S. (2024). Analysis of dormancy and physiological quality of *Stylosanthes* spp. Stored in FGB-UEFS. *Revista Ceres*. 71, 1-7. doi: https://doi.org/10.1590/0034-737X2024710012
Simioni, T. A., Gomes, F. J., Teixeira, U. H. G., Fernandes, G. A., Botini, L. A., Mousquer, C. J., Costa, J. W. R., Hoffman, A. (2014). Potencialidade da consorciação de gramíneas e leguminosas forrageiras em pastagens tropicais. *PubVet*. 8, 1551-1697. doi: https://doi.org/10.22256/pubvet.v8n13.1742
Simões, W. L., Oliveira, A. R., Guimarães, M. J. M., Silva, J. S., Oliveira, C. R. S., Voltolini, T. V., Barbosa, K. V. S. (2022). Arranjo populacional do sorgo forrageiro irrigado para um cultivo eficiente no Semiárido brasileiro. *Brazilian Journal of Development*. 8, 16305-16320. doi: https://doi.org/10.34117/bjdv8n3-053
Scott, A. J., Knott, M. A. A. (1974) A cluster analysis method for grouping means in the analysis of variance. *Biometrics*. 30, 507-512.
Souza, R., Hartzell, S., Feng, X., Dantas, A. C., Souza, E. S., Menezes, R. S. C., Porporato, A. (2020). Optimal management of cattle grazing in a seasonally dry tropical forest ecosystem under rainfall fluctuations. *Journal of Hydrology*. 588, 125102.
Villela, F. A., Doni Filho, L., Sequeira, E. L. (1991). Tabela de Potencial Osmótico em Função da Concentração de Polietilenoglicol 6.000 e da Temperatura. *Pesquisa Agropecuária Brasileira*. 26, 1957-1968.
Yamashita, O. M., Petry, R. D. S., Baio, F. H. R., Roque, C. G., Camillo, M. A., Carvalho, R. D. (2012). *Stylosanthes capitata* germination response to environmental and chemical stimulus. *Gaia Scientia*, 12, 161-169.
Zambão, J., Bittencourt, H. H., Bonome, L. T. S., Trezzi, M. M., Fernandes, A. C. P.P. (2020). Water restriction, salinity and depth influence the germination and emergence of sourgrass. *Planta Daninha*. 38, 1-9. doi: https://doi.org/10.1590/S0100-83582020380100057

Genetic Resources

SHORT COMMUNICATION

Genetic Resources (2025), 6 (12), 205–210
DOI: 10.46265/genresj.SRQL5271
https://www.genresj.org
ISSN: 2708-3764

A small-scale assessment of the availability of EURISCO accessions

Erik Wijnker*, Theo van Hintum

Centre for Genetic Resources, The Netherlands, Wageningen University and Research, P.O. Box 16, 6700AA Wageningen, The Netherlands

Abstract: A critical assessment of plant genetic resource (PGR) availability in Europe reveals a significant gap between documented accessions and those that are practically obtainable for researchers and breeders. While the EURISCO database lists over two million accessions, a study of 100 random accessions found nearly 60% to be unavailable, challenging the assumption that a large number of documented accessions equals usability. Material from 52% of the approached genebanks could not be obtained within five months. The primary barrier was the inability to contact genebank staff, which points to a systemic issue in which PGR access is not always prioritized. Insufficient material for distribution and geopolitical issues were further causes of low availability. This indicates a threat to the effective utilization of European PGR and highlights an urgent need for genebanks to improve communication and operational capacity to ensure these vital resources become accessible for crop development and food security.

Keywords: Availability, EURISCO, PGR, plant genetic resources, SMTA

Citation: Wijnker, E. and van Hintum, T. (2025) "A small-scale assessment of the availability of EURISCO accessions", *Genetic Resources*, 6(12), pp. 205–210. doi: 10.46265/genresj.SRQL5271

Introduction

The International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) (FAO, 2009) defines plant genetic resources (PGR) for food and agriculture as: "any genetic material of plant origin of actual or potential value for food and agriculture." PGR underpin global food security by enabling crop improvement to face the challenges of future food production.

In Europe alone, hundreds of genebanks have been established to conserve PGR *ex situ*. Their holdings are listed in aggregated databases such as Genesys PGR (Genesys, 2025) and FAO's World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture (WIEWS) (FAO, 2025a). According to policy documents addressing the *ex situ* conservation of plant genetic resources (PGR), such as the influential *Third Report on the State of the World's Plant Genetic Resources for Food and Agriculture*, over four million accessions are documented in Genesys with over two million accessions conserved in Europe (FAO, 2025b). Global Crop Conservation Strategies also use the aggregated databases as a source of information to determine the conservation status of crops and set priorities for future activities (Dulloo and Khoury, 2023). At first glance, such numbers suggest that a wealth of PGR is available and conserved. However, this assumption may be misleading as conserved material is not always accessible for use.

Because PGR are tangible assets, it is of crucial importance that stakeholders can access and use them. The availability of material in genebanks, however, is an issue. Personal communications with the user community, but also rare

* Corresponding author: Erik Wijnker (erik.wijnker@wur.nl)

Received: 03.09.2025 | Accepted: 27.10.2025 | Published online: 12.12.2025

studies like Bjørnstad *et al* (2013), make it apparent that PGR access is limited. Bjørnstad *et al* (2013) requested PGR from genebanks globally that are contracting parties to the ITPGRFA and observed that 'facilitated access' is not that straightforward. They received material from only 44 countries out of the 121 approached. If PGR are to be of "value for food and agriculture" as noted in the definition of PGR by FAO, they must be available.

Defining PGR availability is not easy, as it has many dimensions. One prerequisite is that genebanks physically have material for distribution. *Ex situ* collections require a targeted infrastructure to ensure long-term safekeeping of material, which depends on stable genebank funding. Accessions are to be held under good conditions and be subjected to good genebank practice: proper documentation, regular viability monitoring, proper regeneration practices and the creation of safety backups. The FAO genebank standards (FAO, 2014; 2022) provide excellent guidance on such practices.

Equally important, however, is the ability of genebanks to deliver material to users. This requires efficient distribution systems, compliance with phytosanitary and legal requirements, and ensuring clear conditions of use. For example, material that is distributed *for research purposes only* – and therefore not eligible for commercialization – may be considered 'available' but has little, if any, value to commercial breeding and food production. In practice, availability means that users have good access to accession information, are able to order and obtain accessions quickly (i.e. within weeks) under permissive and uniform conditions. The Standard Material Transfer Agreement (SMTA) is a well-established, standardized contract used for plant genetic resources shared under the provisions of the ITPGRFA.

The intention of this research was to provide an impression of current PGR availability in Europe for policymakers, researchers and the genebank community. To better understand the actual availability of PGR material in Europe, a small-scale study was conducted in which plant genetic resources were requested to assess their availability. PGR were selected from EURISCO, a large and complete European PGR database that is often used by stakeholders (Kotni *et al*, 2023).

This study was not designed to evaluate the functioning of individual genebanks. The names of individual genebanks and requested accessions have therefore purposely been omitted from this report. This study should also not be interpreted as a critique of EURISCO. EURISCO itself notes that the presence of data does not guarantee that the material will be supplied: "The presence of data listed in EURISCO does not provide any warranty that the respective collection holders will be able to provide any plant material to interested parties" (ECPGR, 2025).

Materials and methods

Selection of the material to be requested

The general approach involved directly requesting a random set of PGR accessions and systematically observing the requesting process, the receipt of material and the conditions governing its use. Material was selected from EURISCO, the European Search Catalogue for Plant Genetic Resources that is maintained by the European Cooperative Programme for Plant Genetic Resources (ECPGR) (Kotni *et al*, 2023), and serves as an important data source for both Genesys and WIEWS. Because of its completeness and frequent use by stakeholders, the EURISCO database provided a perfect starting point for this research. Availability was defined pragmatically as the ability to order material and receive it within five months. Material that was not obtainable within this timeframe was considered unavailable.

Given the high costs associated with maintaining and regenerating genebank materials, the number of requested accessions was deliberately kept low. A total of 100 accessions was randomly selected, representing 0.005% of records in EURISCO. Anticipating that no more than half of the requested accessions would be successfully obtained, the study was expected to use material from approximately 50 accessions or fewer. While the limited sample size constrains the statistical reliability of the availability estimate, it was deemed sufficient to provide a general impression of the availability of PGR accessions currently conserved by European genebanks.

On 14 January 2025, a complete EURISCO dataset in CSV format was downloaded from the EURISCO website (ECPGR, 2025), comprising 2,101,833 accession records. To refine the dataset and focus on plant genetic resources (PGR) that are conserved *ex situ*, two filtering steps were applied. First, the 682,541 records from the Nottingham Arabidopsis Stock Centre (EURISCO descriptor INSTCODE = "GBR140") were excluded. Additionally, 5,697 records identified as *in situ* conserved material (EURISCO descriptor STORAGE = 60) were removed. Following these steps, 1,413,596 accessions remained.

To ensure proportional representation of genebanks while selecting 100 accessions at random, the 418 contributing institutes were sorted according to the number of accessions they recorded in EURISCO. One accession was randomly selected from the smallest institutes contributing to the cumulative first 1% of the total records, another from the next 1%, and so forth, up to the point where an individual institute contributed 1% or more of the total accessions. For those larger institutes, the number of selected accessions was determined proportionally to their contribution. For instance, the largest institute contributed 14.2% of the accessions in EURISCO; therefore, 14 accessions were randomly selected from its records. This selection methodology ultimately resulted in 100 accessions being drawn from 52 institutes (of which 38 contributed only one accession). With 52 institutes (spread over 23 countries), this selection includes 12.4% of all genebanks that contribute to EURISCO. It comprises 52 different species of plants, of which barley (12 accessions), bread wheat (7), common bean (6) and maize (6) were most prevalent.

Requesting the material

Guidelines for requesting seeds were established in advance, including the use of a standardized text for all requests through email. This standard text did not disclose the purpose of our requests, but when a statement of purpose was explicitly required (occasionally during online requests or later communication, the following text was provided: "In the framework of a methodological research project we are exploring the content of European genebanks in EURISCO and

access to these valuable resources." The depletion of valuable seed stocks was minimized by requesting small quantities (< 25 seeds) whenever the requesting process allowed specification of seed amounts. Requests for vegetatively propagated material were withdrawn once the holding genebank indicated that the requested accession could be provided. The requesting procedure consisted of two rounds of requests: an initial request in which all genebanks were approached, and a follow-up in which contact was sought with genebanks in which the initial request did not result in a response.

The initial request was made on 28 February 2025. An online search was conducted for the 52 selected genebanks to identify contact details or online ordering facilities. Material was preferentially requested via online ordering. When no clear online order form or selection system was available, requests were submitted via email or online contact/feedback forms using standardized texts. Requests were directed primarily to the email address indicated for ordering; if unavailable, contact was sought with the genebank manager, the genebank, or the associated institute, in that order. In the absence of a clear email address, the WIEWS listing corresponding to FAO of the institute (FAO, 2025a) was used. In five cases, emails were returned as "undeliverable," prompting a search for alternative addresses. For one accession, a renewed search identified an online ordering facility, and in two cases, the initial email request led to referral to the online order site. The request cycle concluded once all institutions had been contacted via online order, email or online forms.

Approximately four weeks later (between 21 and 25 March 2025), follow-up emails were sent to the 23 genebanks that had not responded, or to inquire further about the status of our request. These reminders were directed to the previously used email addresses or submitted via online forms. Where possible, the institutes' email addresses listed in the FAO WIEWS database were included in cc if no contact had been established during the initial request. When the initial request had been submitted through an online form, a separate email was sent to the WIEWS-listed address.

The received seed material was stored in the drying chamber of the Centre for Genetic Resources, the Netherlands (CGN), and seed amounts were determined. Two vegetatively propagated accessions were received, of which the number of viable cuttings was recorded.

Results

Requesting the material

For two institutions (with one selected accession each), online searches for the genebanks yielded no leads, and WIEWS contact details were incomplete (no website or email). In a third case (one accession), the genebank indicated on its website that it was unable to accommodate new requests due to resource constraints. Consequently, material could be requested from 49 genebanks, covering 97 accessions.

When possible, material was requested through online ordering. Of the 52 selected genebanks (conserving 100 selected accessions), 17 (maintaining 55 accessions) allowed online searches of their holdings, including 4 genebanks (conserving a total of 9 accessions) using the GRIN-Global interface. Not all institutes permitting online searches offered online ordering. Seven initial requests were placed online, with an additional 5 following further inquiries, resulting in 12 institutions (23%) providing 31 accessions via online ordering.

Online ordering was not always straightforward, often requiring prior registration, navigating counterintuitive interfaces, or addressing technical errors that necessitated email contact. Some failures to order had other causes: one accession could not be located via the online interface and was later confirmed as no longer held, while two accessions listed as 'unavailable' on a GRIN-Global interface were requested via a contact form but received no response. Overall, experiences with online ordering varied from straightforward to frustrating.

Of the 49 institutions (97 accessions) that could be approached, 7 (14%) were initially contacted through online ordering. The remaining 42 were contacted through email (39 genebanks) or online forms (3 genebanks), making email the principal mode of communication for requesting genebank material. Email communication presented its very specific challenges: replies to requests were often sent from alternative addresses, causing changes in subject lines that complicated tracking of individual requests. Telephone requests were not attempted.

The initial 42 messages resulted in a reply in 18 cases (43%) but were not answered in most cases (57%): either no response was received (19 cases; 45%) or emails were returned as 'undeliverable' (5 cases; 12%). Repeated efforts, including renewed searches, alternative addresses and reminder emails, allowed us to establish contact with 33 institutions (including the one indicating unavailability of all material in the collection on its website), while 19 genebanks (37%) remained unreachable, representing 40 accessions.

Effectiveness of our requests highly depended on live and actively monitored addresses. During the request process, WIEWS-listed addresses associated with the FAO code for the genebank used in EURISCO were often necessary, as no other contacts were available online. Of the 52 selected institutions, 6 had no email addresses listed in WIEWS, 6 addresses resulted in undeliverable emails, and 9 did not respond. Thus, at least 21 of 52 institutes (40%, representing 29 accessions) could not be reached via WIEWS-listed addresses. Some were nevertheless contactable through other online addresses. No systematic assessment of WIEWS address functionality was conducted; therefore, it cannot be assumed that the remaining 60% are reliably reachable.

Four requests (four accessions) were withdrawn by requestors during the requesting process. In two cases, requests were terminated after verification that material was available. This concerned a vegetatively propagated *Rhododendron* accession and an accession for which the holding institute indicated that the requested seeds were in short supply, but that a small amount could nevertheless be distributed. Two other requests (two accessions) were terminated without clarity on the availability status. These both concerned accessions that required special permissions to be granted (either by government or by breeders) prior to distribution: one CITES-protected species and one commercial grape variety. Both these accessions would – if permission were sought – be subject to MTAs rather than standard SMTAs and be available for research only. These accessions were classified as 'possibly available.'

Table 1. Availability status of requested material. Accessions have been grouped according to availability status, in which availability was defined as the ability to receive material within a 5-month period. *, Four genebanks could not provide all accessions requested and appear in both cells marked with an asterisk; **, The genebank total indicates the number of genebanks included in this study.

Availability	Subgroup	No. of genebanks	No. of accessions	Notes
Available		18 + 4*	38	Material was either received (33 accessions), the request was withdrawn by requestors to save seed stock of prevent vegetative propagation (2 accessions) or material was lost in the mail (3 accessions)
Possibly available	Request withdrawn	2	2	Material may be obtainable if special permission is requested and granted (availability for research only; MTA)
	Wrong accession was received	1	1	
Not available	Genebank confirmed non-availability	7 + 4*	18	Material not available at genebank
	Long delivery time	1	1	Ordering time exceeds the research limit of 5 months
	Genebank did not respond	19	40	Material could not be requested due to no response
Total		**52****	**100**	

Availability of accessions

On 31 July 2025, five months after the initial request, the request period for this research was stopped. Starting from an initial list of 52 genebanks (100 accessions), it was possible to assess the availability of selected accessions at 33 genebanks (63%; 60 accessions). Contact was made with 32 institutes, and for one genebank (one accession), information for availability was derived from the website that indicated distribution was not possible. The remainder of 19 genebanks (37%; 40 accessions) could not be contacted through the channels used.

Data on accession availability are presented in Table 1, in which material is listed as 'available' if it could be obtained within a 5-month period, given the used communication/order methods. Of the 100 selected accessions, 38 were classified as 'available'. Of these, 33 accessions were physically received from 19 genebanks (37%). Five accessions were not physically received, but availability was inferred from communication with genebanks: two accession requests were withdrawn (because of low stocks or vegetative material) and three accessions were lost in transit (sent but not received). A further three accessions were considered 'possibly available': one delivered as the wrong accession and two requiring additional, special permissions. The remaining 59 accessions were considered 'unavailable': 1 accession still in process, 18 accessions explicitly unavailable according to genebank feedback, and 40 accessions held by institutes that could not be contacted (Table 1).

For 60 accessions (33 genebanks), we obtained information on their availability (other genebanks could not be contacted or did not respond). Of those, 18 accessions (30%) were unavailable. Reasons for their unavailability were provided for 14 accessions, with a 'lack of material' being the most prevalent reason (8 accessions, 3 genebanks). Legal requirements prevented the distribution of two accessions, international armed conflicts prevented distribution of two accessions (two genebanks), one genebank indicated the inability to distribute any material on their website, and one accession was no longer maintained in the collection (Table 2). Availability via online ordering did not appear to correlate with seed availability. Of 31 accessions requested online (12 genebanks), 8 accessions (26%) were unavailable, similar to the overall unavailability rate of 30% for genebanks that provided feedback on holdings.

Table 2. Reasons for non-availability of requested material as provided by genebanks.

Reason for not sending	No. of genebanks	No. of accessions
Genebank cannot distribute	1	1
International agreements prevent distribution	1	1
Lack of material/too few seeds	3	8
Material is not disease tested	1	1
No longer in collection	1	1
War prevents sending	2	2
No reason provided	2	4
Total	11	18

.

Distribution details, payments and conditions for use

By recording the moment of parcel reception, the approximate delivery time for the requested PGR could be inferred (Figure 1). The majority of seed shipments (14 out of 17 packages) were received between two and eight weeks after initiating the request. The average time from first contact to accession receipt was six weeks. All received material arrived undamaged. Seed quantities per accession varied widely, ranging from 11 seeds (three accessions) to 381 seeds (one accession), with a median of 50 seeds. Clonal material (cuttings) was received from two institutes (two accessions) and was alive and viable upon arrival.

In three cases, payments were required as part of the request process. Two genebanks (two stations of the same institute) charged handling fees (€10 per request plus €2 per accession), and one genebank charged handling fees (€17) plus phytosanitary testing costs (€177) for two accessions.

The majority of received accessions (obtained from 19 genebanks) required a material transfer agreement (MTA). For 16 genebanks (84%; 30 accessions), a standard MTA (SMTA) was used, either signed manually with electronic exchange of documents (eight cases) or via click-wrap/easy-SMTA (eight cases). Two genebanks (providing one accession each) provided accessions without any user restrictions, and one genebank distributed under a different MTA. All 30 accessions received under SMTA were obtained from genebanks located in countries that ratified the ITPGRFA (FAO, 2009). The one accession obtained under a different MTA was received from a genebank located in a country that did not ratify the ITPGRFA.

Discussion

Access to PGR is a fundamental prerequisite for their effective utilization, and constitutes a central theme in international policy discourse, particularly in relation to access and benefit-sharing (ABS) mechanisms under the ITPGRFA. Despite the prominence of this issue, the practical conditions governing physical access to PGR remain insufficiently documented. Empirical investigations of this domain are scarce, with Bjørnstad *et al* (2013) being among the few to demonstrate that presumed availability of PGR is not always substantiated in practice. In light of this, a systematic assessment of PGR availability within Europe is both timely and imperative.

Although locating accessions via the EURISCO catalogue is relatively straightforward, the process of requesting material from the respective holding institutions proved considerably more challenging. For 22 genebanks (42%), we did not find a website dedicated to genebank activities (like describing collections and featuring genebank contact details). When websites did exist, the procedures for ordering material were frequently ambiguous. As a result, email emerged as the primary mode of communication. The request protocol was standardized and limited to two contact attempts per genebank, which was deemed a reasonable threshold for effort. While alternative strategies, such as translating emails into the local language, telephone outreach or leveraging personal networks, might have improved response rates, they were not employed in this study.

Of the 52 approached genebanks, 22 genebanks (42%) sent (at least part of) the requested material, which might have increased to 25 genebanks (48%) with additional effort (see Table 1). Likewise, it is estimated that the current availability of PGR accessions ranges between 38% and 41%, with the upper bound including accessions that may be obtainable through additional effort (see Table 1). Because of the small sample size, we suggest that this number be interpreted with caution. The average time required to receive seed material was approximately 6–7 weeks, a duration considered acceptable for most research and breeding purposes. Nevertheless, the study reveals that material from 27 genebanks (52%) could not be obtained, leaving 59% of accessions presently inaccessible. This highlights a significant and alarming proportion of inaccessible material in European genebanks.

The principal barrier to accessing PGR appears to be the difficulty in establishing communication with genebank personnel. The initial round of requests elicited responses from only half of the 52 genebanks contacted, which increased to two-thirds (33 genebanks) following a second round. Notably, one-third of the genebanks (19 out of 52; 37%) remained entirely unresponsive. The difficulty of contacting

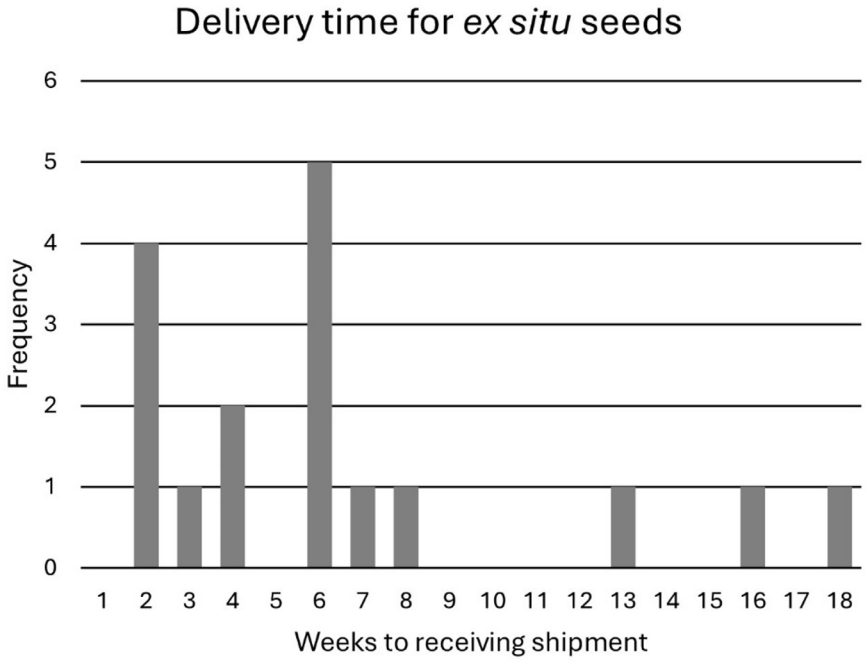


Figure 1. Time required for receiving ordered *ex situ* accessions from 17 genebanks. The histogram shows the time (in weeks) between the moment a genebank was contacted and the receipt of the parcel.

genebanks was also due to the unclarity and unreliability of contact information on websites or in databases, like WIEWS. While the reasons for non-responsiveness of genebanks are speculative, the data suggest that not all genebanks prioritize the facilitation of PGR access. These findings align with those of Bjørnstad *et al* (2013), who reported a similarly high rate of non-responsiveness among contracting parties in a global study on facilitated access.

Although EURISCO serves as a valuable repository for documenting PGR holdings, our findings indicate that actual access to these resources remains a significant challenge. Extrapolating the estimated availability rate of 38–41% to the broader EURISCO database (excluding *Arabidopsis* accessions) suggests that only approximately 481,000 to 594,000 accessions may be readily obtainable. This figure stands in stark contrast to the more than two million accessions currently documented, highlighting a substantial gap between nominal documentation and actual accessibility.

The sample size for our research was deliberately kept small: 0.005% of EURISCO accessions and 12.4% of contributing genebanks (including all those that contribute > 1% of all EURISCO accessions). Our estimate for availability is therefore to be interpreted with caution. Likewise, the small dataset limits the ability to statistically discern possible causes underlying seed availability. For example, requests to ten genebanks (19%) were directed to genebanks in four countries that did not ratify the ITPGRFA. Of the 25 requested accessions, 92% were classified as non-available. But assessing any effect of ratification of the ITPGRFA on availability would be confounded by the fact that nine of these requests were made to institutions in countries involved in armed conflicts during our research period (Israel, Russia and Ukraine). In this case, feedback provides more insight. Of the four genebanks from these countries that replied to our emails, ITPGRFA ratification was not mentioned as prohibiting seed distribution. Two of the genebanks did explicitly mention that the ongoing war prevented distribution (Table 2), indicating that PGR availability can be affected by geopolitical conflicts.

Europe hosts over 400 genebanks and collection holders in 43 countries that are listed in EURISCO, yet it remains unclear whether all institutions are actively conserving and distributing their holdings. This raises critical questions regarding institutional capacity and commitment to resource sharing. The observed challenges in distribution point to a disconnect between the documented inventory of accessions and their actual availability to users.

Many factors affect the availability of accessions from genebanks, ranging from expired email addresses to war. Even though this research could not address the underlying causes, elucidating these will be of the greatest interest to both genebanks and their stakeholders. For many genebanks, having material available for users touches directly on their raison d'être and may affect societal support for their work. For other stakeholders, a clear grasp of the underlying causes may help take measures to improve and promote PGR availability.

In conclusion, the findings of this study underscore the urgent need to address systemic barriers to PGR access in Europe. Failure to do so risks eroding valuable genetic resources, decreasing societal support for genebanks and impeding the capacity of researchers and breeders to develop resilient crop varieties essential for food security.

Author contributions

Both authors contributed equally to the manuscript

Acknowledgements

We would like to thank the many enthusiastic genebank staff we interacted with during the course of our research. This work was carried out in the framework of the Programme Genetic Resources (WOT-03 and KB-34) funded by the Dutch Ministry of Agriculture, Fisheries, Food Security and Nature.

Conflict of interest statement

The authors have no relevant financial or non-financial interests to disclose.

Data availability

No data accompany this manuscript, but details on sampling procedures and collected data may be requested by contacting the authors.

References

Bjørnstad A., Tekle S., Goransson M. (2013). "Facilitated access" to plant genetic resources: does it work? *Genet. Resour. and Crop Evol.* 60, 1959-1965. doi: https://doi.org/10.1007/s10722-013-0029-6

Dulloo E., Khoury C. K. (2023). Towards mainstreaming global crop conservation strategies (Germany, Bonn: Global Crop Diversity Trust). doi: https://doi.org/10.5281/zenodo.7610356

ECPGR (2025). European search catalogue for plant genetic resources (EURISCO). http://eurisco.ecpgr.org

FAO (2009). International treaty on plant genetic resources for food and agriculture (ITPGRFA) (Italy, Rome; Food and Agriculture Organization of the United Nations). https://www.fao.org/plant-treaty/overview/text-treaty/en

FAO (2014). Genebank Standards for Plant Genetic Resources for Food and Agriculture. Rev. ed. (Italy, Rome: Food and Agriculture Organization of the United Nations). https://www.fao.org/3/i3704e/i3704e.pdf

FAO (2022). Practical guide for the application of the Genebank Standards for Plant Genetic Resources for Food and Agriculture: Conservation of orthodox seeds in seed genebanks (Italy, Rome: FAO Commission on Genetic Resources for Food and Agriculture). doi: https://doi.org/10.4060/cc0021en

FAO (2025a). World information and early warning system on plant genetic resources for food and agriculture (WIEWS). https://www.fao.org/wiews/data/ex-situ-sdg-251/search/en

FAO (2025b). The Third Report on the State of the World's Plant Genetic Resources for Food and Agriculture (Italy, Rome: FAO Commission on Genetic Resources for Food and Agriculture), 374 p. doi: https://doi.org/10.4060/cd4711en

Genesys (2025). Genesys plant genetic resources portal (Genesys PGR). https://www.genesys-pgr.org/

Kotni P., van Hintum T. J. L., Maggioni L., Oppermann M., Weise S. (2023). EURISCO update 2023: the European search catalogue for plant genetic resources, a pillar for documentation of genebank material. *Nucleic Acid Res.* gkac852. doi: https://doi.org/10.1093/nar/gkac852

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2025), 6 (12), 211–220
DOI: 10.46265/genresj.PEGR2598
https://www.genresj.org
ISSN: 2708-3764

What farmers value matters for the management of local breeds: A case study of the Pyrenean goat breed

Anne Lauvie[a,*], Marie-Odile Nozières-Petit[a], Fanny Thuault[b,c], Nathalie Couix[d,e]

[a] *SELMET, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France*
[b] *Association La chèvre de race Pyrénéenne, Foix, France*
[c] *Present affiliation: AREDEAR Occitanie, Foix, France*
[d] *AGIR, Université de Toulouse, INRAE, Castanet-Tolosan, France*
[e] *Present affiliation: LESSEM, INRAE, Grenoble, France*

Abstract: Local breeds are often kept in farming systems where their locally adapted traits are an asset. The agroecological transition movement has brought renewed interest in these breeds with their locally adapted traits. Better understanding the factors underpinning trait preferences of local-breed goat farmers can inform efforts to manage genetic diversity and animal selection. This article examines what goat farmers value in their animals to gain a better understanding of their decisions and expectations. Based on evidence from 20 interviews with farmers using the local Pyrenean goat breed, we show that although breeders attach importance to various animal performance factors, hardiness consistently stands out as a key value. We also show that they attach importance to other factors, such as animal diversity, temperament and appearance (breed standard and aesthetics). Breeders relate to their animals in a way that mobilizes senses, feeling, experiential or emotional dimensions more than just a set of hard traits, and when breeders voice what they value and how that value manifests in practice, their narratives translate the importance of balance and trade-offs.

Keywords: Local breeds, goats, France, hardiness, values

Citation: Lauvie, A., Nozières-Petit, M.-O., Thuault, F. and Couix, N. (2025) "What farmers value matters for the management of local breeds: A case study of the Pyrenean goat breed", *Genetic Resources*, 6(12), pp. 211–220. doi: 10.46265/genresj.PEGR2598.

Introduction

Livestock selection must address new challenges around animal genetics to move the agroecological transition forward. The genetics community highlights the need to design breed selection programmes with specific objectives, such as resilience, relevant for livestock in transition toward agroecology. Additionally, they should consider the reduced control over environmental conditions in agroecological farming systems (Phocas *et al*, 2016). There is consequently growing interest in efforts to improve breed hardiness and integrate diverse objectives that combine a wider set of traits (for example, production-related traits with traits tied to reproduction, survival, health and welfare). Animal genetics is experiencing a fast-paced change in methodologies, with the development of genomics soon to be followed by high-throughput genotyping and phenotyping workflows (Boichard *et al*, 2015).

Local breeds, and especially those with smaller populations, have been relatively sidelined from selection programmes that have focused on dominant breeds. However, local-breed animals often thrived in farming systems such as agropastoralism, where their locally adapted traits and hardiness proved an asset (Moulin and Perucho, 2023). The agroecological transition movement has brought renewed interest in these breeds with their locally adapted traits (Dumont *et al*, 2013).

For any goat farm, deciding which breed or breeds to select as a starting point and deciding which individual animals to use to produce the next generation are pivotal decisions that shape and scaffold a coherent production system (Vissac, 2002).

* Corresponding author: Anne Lauvie (anne.lauvie@inrae.fr)

Received: 12.03.2025 | Accepted: 03.09.2025 | Published online: 12.12.2025

Better understanding the factors underpinning these trait-preference decisions of local-breed farmers can valuably inform the work of geneticists who deliver the tools and strategies driving genetic selection programmes. What the farmers want and expect to see in their herds – in other words, what they value in each animal – is a key factor driving these decisions. However, beyond the market value of their products that fundamentally drives revenues, local breeds – more than other breeds – are attributed a variety of other values that reflect each breeder's relationship with individual animals from that breed and with the breed itself (Couix *et al*, 2023). We posit that examining these values can bring key insights to inform the design of genetic resources management programmes geared to address the challenges facing livestock farming today. This inquiry is especially relevant in the case of local and rare breeds investigated here, as they have so far remained sidelined from the big selection programmes but are now attracting renewed interest to help agroecological transition move forward.

Here we addressed these challenges by studying farmers' objects of value around the local and rare Pyrenean goat breed from south-western France.

The Pyrenean goat is a local breed that farmers have progressively abandoned. A Pyrenean goat conservation organization was created in 2004, and Pyrenean goat numbers today amount to just 5,000 animals kept by around 200 farmers, mostly in farms across the French Pyrenees (Pyrenean goat breed association, 2024a). In 2023, most of the farmers (207) were situated in the two regions concerned by the Pyrenean Mountain range, Occitanie and Nouvelle-Aquitaine regions. Twenty-two flocks were located in other French Regions (Pyrenean goat breed association, 2024b). About a third of these farmers raise their flock for milk and cheese. The remaining two-thirds produce kid meat, often alongside other farm products or additional professional activities ('pluriactivity farmers'). The Pyrenean goat exhibits strong phenotypic variability, especially in terms of colour. A breed standard describing the defining traits was adopted in 2008 (Pyrenean goat breed association, 2024a).

Adopting the same kind of pragmatist stance as Couix *et al* (2023), we considered value formation as a series of what John Dewey (1939) theorized as valuation processes, where the formation of values spans both the immediate valuing appraisal and the higher-level evaluation (Bidet *et al*, 2011) and considers the dynamic dimension of values as they change over the course of time and in the course of action. This led us to examine both the way farmers appraise and qualify the animals they tend, and the related livestock farming practices. Consequently, the aim of this article was to study what goat farmers value in their breed and in their animals.

First, we present the approach adopted, which is based on analyzing evidence captured through interviews. We then show that the farmers voiced a diversity of objects of value that accommodate diverse biological and even non-biological traits beyond pure animal performance factors – and thus reveal a diversity of values. We then discuss how the diverse values attached to local-breed animals, driven by the farming system in which they are raised, can help inform strategies for addressing the challenges in managing farm-animal genetic resources.

Materials and methods

Theoretical background

We based our work on 20 semi-structured interviews led in 2021 with Pyrenean goat farmers. The objective of this analysis was to sift through their discourses to identify how the values they attach to their individual animals and the animal breed are formed (Lauvie *et al*, 2022). Here we mobilized John Dewey's theory of valuation (Dewey, 1939), which posits that values are shaped by projected consequences of actions (the means) led – in this case, by the goat farmers – to bring about desirable outcomes (the ends). We therefore focused our attention on what farmers said they wanted and expected to see in their flock animals, in other words, what they valued in each animal. However, we went beyond simply surfacing the diversity of sources of value that emerged from farmers' discourses, to focus on what they told us about how this diversity of values manifests into practice, especially when they connected their expectations of the animals to practices they employ to get a flock that suits them. We also attended to what they said about the responses they saw in the animals. John Dewey (1939) argued that values are not immutable but always potentially open to revision as they can be made to change in the course of interactions between individuals and their environment. Consequently, we also prepared to focus on the way farmers revised their objects of value in response to interactions with their animals.

Data collection

In order to capture 'what farmers value' in their work with a local breed, we collected data using a process designed to capture the diversity of farming situations found in our case study. We selected the sample of farmers to cover various criteria. These included the diversity of geographical configurations, by choosing farmers in different locations, and both dairy systems and suckling systems. The sample included mainly Pyrenean goat breed association members, as well as non-members and former members. Additionally, it included breed-association board or committee members, and farmers with varying lengths of experience with Pyrenean goats. The interview guide helped to collect focused data on the history of the goat farm, the reasons that prompted the farmer to choose the Pyrenean goat and flock management system – particularly reproduction and culling/replacement – and the farmer's relationship with the Pyrenean goat breed association.

A student (senior internship), under the supervision of the authors, performed 20 interviews: 11 with suckling-system farmers and 9 with dairy-system farmers. The vast majority of the farmers surveyed (i.e. 16 out of 20) were members of the Pyrenean goat breed association. Six of the interviews were attended by two people from the farm (goatherd, partner or associate) who conferred and spoke together. One of the interviews led into further exchange with a second (former) farmer (who worked on another farm) who joined the first interviewee. We counted these two exchanges as two separate interviews. All the interviews were recorded and transcribed, except one which was the object of a very detailed report. An interdisciplinary and transdisciplinary team supervised the student and later conducted the analysis presented in this paper (two researchers in animal sciences with a systemic

approach, one researcher in organization sciences and the person in charge of the facilitation and extension for the Pyreanean goat breed association).

Data analysis

We explored the qualitative interviews by coding the data using NVivo v11 software to develop themes. Coding produced the six themes that feature in the node map in Figure 1, and each of these six themes contained subthemes. Here, we focused our analysis more specifically on the core theme tagged 'objects of value in the animal' and its node-mapped subthemes (see Figure 1).

We did not set out to work on speech as an object, but like Macé, we considered spoken discourse as a space that operates connections and attachments and therefore warrants attentive analysis (Macé, 2021). It is for this reason that here, unlike in most livestock science papers, we organized the presentation of our results around verbatim content. The original versions of those verbatim (in French) are available as Supplemental Material 1.

Results

Diverse considerations attached to animal performance factors

Farmers have varied expectations regarding milk yield

The farmers who saw milk yield as an object of value in the animals referred to both milk quantity and milk quality. Certain farmers mentioned milk protein yields or milk fat yields, whereas others referred to a more global, end-in-view criterion, i.e. cheese yield, which in some cases they connected to specific milk parameters and, therefore, to frames of evaluation involving other, non-yield-related criteria. One farmer, for example, talked about the acidity of their flock's milk, while another – likely inspired by a previous study in the same breed – cited casein polymorphism in their flock's milk, and several farmers spoke about the taste of the milk as a dimension of milk quality. Still connected to milk production but going beyond quantitative and qualitative traits of milk output, some farmers stated that they attached importance to having goats that were easy to milk, which they sometimes connected to udder conformation, but sometimes not.

Looking at the process of evaluation for these objects of value, certain farmers talked about their own criteria, especially on milk quality, while others also mentioned the standardized national milk test records system called 'contrôle laitier' that some saw as a valuable tool for community-wide Pyrenean goat breed conservation and management efforts. On a farm, they highlighted that the milk records tool provided benefits for the wider breed community as well as for their own personal management.

Lastly, certain interviewees explained the link between the performance of their animals and their flock management decisions:

> *"This year, well, increasingly, but especially this year, maybe because we give alfalfa now, as we have an alfalfa plot, but it's now 30–40 crottins [type of cheese] a day. With no weight differential. That's a huge yield right there."*

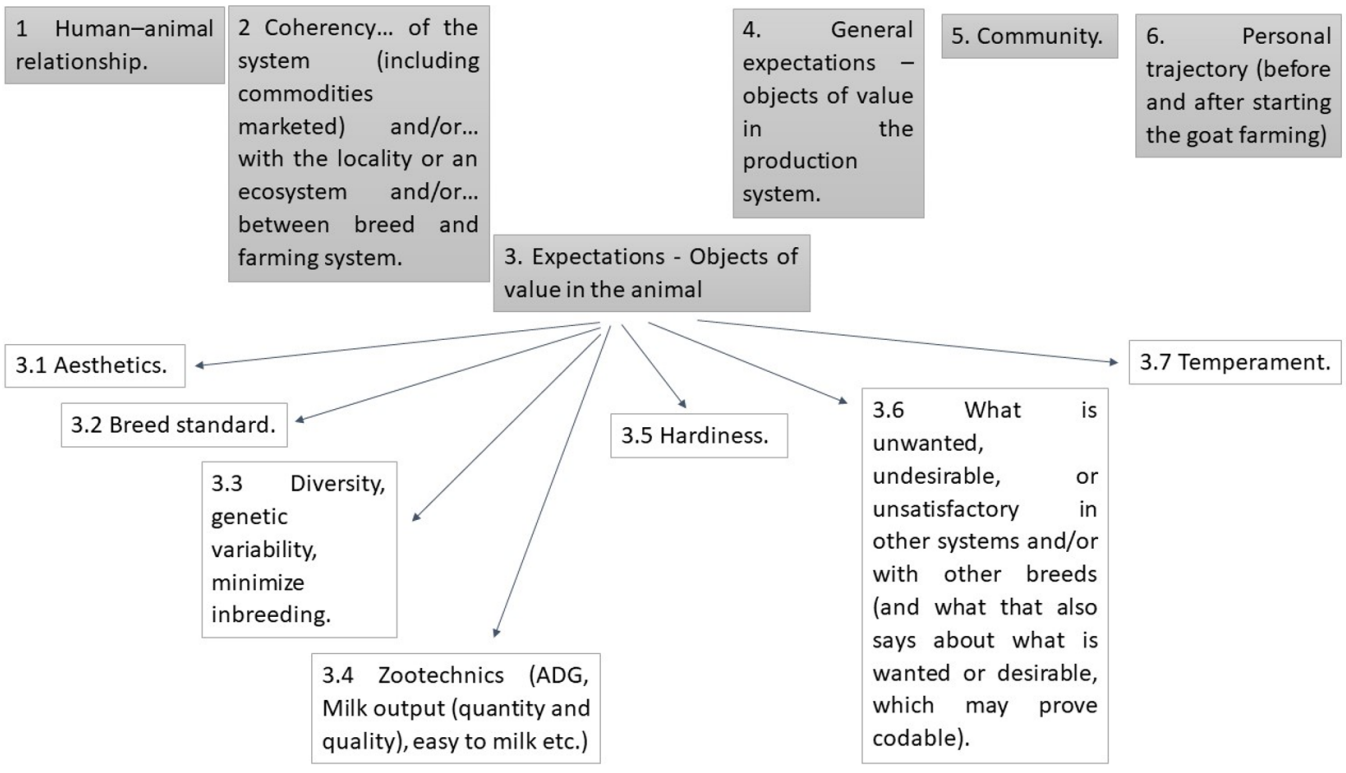


Figure 1. Node map featuring the six core themes identified and the subthemes identified specifically for theme no. 3.

Meat production: it's not just average daily gain that counts

Farmers also emphasized the importance of factors related to kid meat production. These factors include size, weight, growth and conformation, which one farmer summed up as the need to have kids "that grow out well". These size and conformation factors that work towards zootechnic performances were sometimes more or less explicitly verbalized as connected to the breed standard, which we discuss in further detail below. Some farmers attached importance to more specific criteria, such as birth weight, minimum expectations on liveweight at a given age milestone, average daily gain (ADG), or carcass weight.

As seen above for milk yield, certain interviewees explained the link between the performance of their animals and their flock management decisions:

> *"Because if you care for your does, then the kid will automatically birth big, so you weigh it. It might weigh 5 kg at birth or 3 depending on how you've cared for the doe".*

Coherency between conformation and dairying: a challenge for a breed claimed as dual-purpose?

The Pyrenean goat farmers studied here have the option of running a dairy system or a suckling system, as the breed is considered dual-purpose. However, certain farmers mentioned that, in the past, the breed had to navigate periods of tension between the two production uses that proved challenging to manage as a community:

> *"In fact, the buck committee[1] was originally set up about 10 or 15 years ago when everyone in the breed association realized that divisions were emerging between the dairy farmers who only wanted to select for milk, and the suckling systems farmers who only wanted to select for type".*

Several farmers advanced a linkage here between kid health and/or growth and milk quality of the nursing does. For instance, one farmer, who farms their flock for milk and cheese, said:

> *"Kid growth […] is important [to consider] because if the mother is giving milk packed with protein, then I reckon the kid is also going to grow out a lot faster. So that's another way [suckling system farmers] can look at it, even if they're not dairying".*

Another farmer formulated the same idea in a different way:

> *"But a dairy breed and a meat breed are just the same thing. Because goats that produce the most meat also give the most milk. That's just common sense."*

[1] Group of farmers within the breed conservation association that sets and selects breeding bucks

Prolificacy: a flock trait framed by the system

The last category of animal-factor performances that farmers discussed is prolificacy of the flock, which is a trait one farmer actively worked on:

> *"For does farmed 100% free-range and browsing a 'nutrient-poor' environment […] prolificacy isn't even two. So, for me, if there was something to improve my production, it would be about that, some improvement in prolificacy".*

However, high prolificacy is not always a preference, and farmers are conscious that it comes with limits:

> *"We have a crazy-high prolificacy rate, which has downsides too, because we end up with triplets. We've had triplets making up 30% of the herd, and that's not great. […] it screws up the lactation cycles, the animals are exhausted, and then you have triplet kids weighing in at 4kg."*

Hardiness is valued as a constant

A notion that accommodates a diversity of traits

Hardiness was a term widely used by the farmers interviewed, and for many, it was one of the features that led them to choose the breed, and sometimes even the first factor they mentioned. One farmer put it in these terms:

> *"Ah! Hardiness! Well, that's what we like about the Pyrenean goat, its hardiness. It can eat woody brush, and never really get sick. Hardiness means goats that range and graze what they need, never getting sick, kept outside all the time – you just arrange to bring them in for kidding, and that's pretty much it, that's what we want."*

When expanding on what they mean by hardiness, the farmers talked about a wide range of abilities: ability to range and use all vegetation resources available on the farm, ability to stay outside even in harsh weather, ability to mother their kids, ability to take care of their own needs, ability to tolerate parasitism problems and health problems in general, and so on. Certain abilities, such as ranging widely and using all vegetation resources available on the farm, were more widely recognized among breeders. Depending on farmers, various abilities were emphasized in different ways.

A notion that pulls together animal–environment–herd management factors

Several farmers made a connection between their vision of hardiness and their animal management practices or even their livestock system performance goals, with one farmer asserting that:

> *"It's that opportunity to exploit that hardiness, with prospects for saving on costs, that ultimately pushed me to look for the friction point that would mean it ends up making economic sense".*

Some farmers talked about the benefit of finding a compromise between animal-factor performances and hardiness, as illustrated here:

> *"Obviously, you need your animals [...] to give you enough milk to get by, to make a livelihood, put food on the table. That's your bottom line – the rest is just about finding the right compromise, the right balance. But for us it's also super important to have well-adapted animals that will happily stay outside without any kind of problems".*

Certain interviewees also underlined that hardiness drives animal-factor performances. For example, one farmer made the connection between the ability to use specific resources and production performances:

> *"It would be a hardy goat, that has learnt very young to range for all the food around in its environment, like woods, brush and wildland, and yet still manages to give you a goodish ADG."*

An important feature of this notion of hardiness is that it is tightly connected to how the animals are managed, such as free-ranging for example, or in resource-squeezed systems, and some farmers also make the connection to local range habitat, such as this farmer who explained:

> *"The breed is adapted to our work conditions and the local topography, even if today […] it's been raining non-stop for a month now, and there's no other breed that could withstand the wet like she did, that's for sure. Plus, it's a goat that uses little if any bought feed, so it's pretty cheap to keep."*

Hardiness as all-round self-sufficient animals

Finally, in addition to underpinning a range of abilities and traits linked to farm-system conditions, the notion of hardiness was also articulated in more global approaches that refer to what might be described as self-sufficiency, i.e. the fact that the animals demand less human care and attention, and fend for themselves:

> *"Hardiness also means needing as little care from us as possible."*

One of the farmers surveyed expanded on this idea by bringing the domestication factor into play:

> *"That's what it is: livestock that can... Listen, I'd be exaggerating if I thought my goats could happily live without me, I don't think they could, because there's been some domestication. They are only there to collaborate with me. But my goats can stay outside without getting infested with parasites".*

Note that this self-sufficiency can also be identified as a specific trait in a given animal, in which case it makes that individual animal particularly valued. In one interview, when the interviewer and farmer were standing facing the flock and the farmer was asked which of the goats had the most merit, the farmer answered:

> *"I think it's the white one with the black neck. Her name is [name of the goat]. In milk, hair, type, hardiness, easy, stood in the middle of the herd, never needing special care. She just blends in, so independent that... Well, for me, that's a good balance, a good compromise."*

This self-sufficiency also overlaps into a degree of stubbornness, resistance or refusal to yield to human command – a point touched upon by a farmer:

> *"It's that part of hardiness that has its downsides. Just try and separate the kids at birth, and see how far you get! Nowhere! There's just no way!".*

On the same farm, the following testimony explained why they aim for and prize self-sufficiency:

> *"Like us, an animal will adapt to indoor comfort, but we just cannot keep them penned up indoors around the clock. It's just impossible. They would turn violent. They're too accustomed to being outside… […] We favour that. Because they need to know how to go out and forage, find their own food. Look, here's your proof: we give them a yoghurt pot of barley, it's there on offer, and they don't even look at it. Zero cereal feed".*

Note that in this case, the fact that the goats resist the farmers' command is considered a valued trait.

Other objects of value beyond zootechnical traits

Diversity valued in and of itself

While one farmer spoke of how milk yield in his flock was widely heterogeneous and how they wanted to bring about more homogeneity by removing the less milk-driven stock, most of the farmers who talked about intra-breed diversity approached the issue from the angle of working to conserve this genetic diversity:

> *"Otherwise, you are also impoverishing the gene pool. Because if we keep pushing the population through a funnel, one day we'll end up with clones"*

or again:

> *"Because I really liked those animals, plus I'm naturally for conservation and against losing stuff, like agricultural diversity and even biological diversity in general. [...] But there's also the fact, when you're working on a local breed, you have to be really careful not to bottleneck the population. Which means you just can't let any family die off. It's our job to improve every goat – even poor ones. That's part of what the conservation effort is about – taking a poor goat and improving it. Sometimes it drags on, and sometimes nothing works. There are bad families, hopelessly bad families. [….] That said, for me, there's that overriding criterion of vital genetic variability, as you really can't let any family disappear, because a breed that counts just 4,000 animals is a desperately vulnerable breed."*

The farmer in that last example, like others interviewed, voiced the fact that they are ready to push less hard on animal performance criteria in order to conserve the breed's genetic diversity. One farmer said they used to inbreed heavily, but got to a point where they needed to "look around elsewhere".

Importance attached to the animal's temperament

The animal's behaviour and temperament proved to be equally important objects of value, and the farmers often frame these traits in terms of the type of human–animal relationship they potentiate, whether positive or negative. This is seen in the words of one farmer, who said:

> *"[At the] close of the season, we set away three goats, one of which, [name of a goat], is at practically 2L. We just can't get along at milking. She's a right bitch! […]. Outside of milking, she's great, but at milking, […] she just straight refuses."*

Another puts it in these terms:

> *"As a rule, the one I like best is the one that communicates the most, the one you share most with, the one you call by first name."*

These temperament factors can also be gain importance depending on the types of activities carried out on the farm. One farmer, for instance, highlighted how the goats play a doubly important role because they host visitors on his farm.

The appearance factor: between breed standard and aesthetic preferences

The farmers' accounts featured a substantial number of references to the animal's appearance, either in terms of the breed standard or 'type', i.e. the defining physical traits of the breed, or in terms of the aesthetics factor. Note that several farmers employed the adjective 'beautiful' for both type-related and aesthetics-related traits:

> *"This year we've had several newborns that really are beautiful. When I say beautiful, I don't just mean beautiful as in pretty to look at. I mean, they also show good conformation, real potential... In terms of breed standard and such."*

Farmers who explicitly mentioned the breed standard sometimes voiced the fact that it was co-constructed by collective agreement, and one interviewee distinguished a component based on personal criteria, on top of the other foundational criteria that had been set by the breed community:

> *"That's personal criteria, not Pyrenean goat breed criteria."*

Note too that several farmers outlined how these breed-standard criteria or breed-type criteria are in flux and workable, and that they aim to improve that dimension in their own flocks.

Several farmers prized the aesthetics factor of the breed, as shown in this exchange:

> *"And I just fell in love with the breed. Instantly. The beautiful coat, the lovely long hair, the body form. Plus the fact that it's threatened with extinction – that was added motivation."*

Traits that farmers attach aesthetic importance to or say are a part of what makes the animals beautiful include not just the striking horn set and coat colouring (sometimes with a preference for a palette of colours) but also, for example, the ears, the hair, and the shape of the head or the legs. Some statements also referenced an aesthetics of the goat roaming free outside. One farmer said:

> *"They are prettier when they are ranging free, that's where they are beautiful."*

And another said:

> *"Outside they 're just magnificent".*

It is noteworthy that the role of the senses, feelings, the experiential and/or emotional dimensions emerge strongly in these accounts.

What breeders value: how valuation processes involve more than just hard traits

The experiential and emotional dimensions in livestock farming practice

The previous paragraph, highlighting the aesthetics factor in goat farming, clearly reveals a prominent role of senses or feelings to the goat farmers' decisions. The responses to questions addressing what farmers want and expect to see in their animals also expressed more holistic considerations tied to all-around satisfaction with goat farming. In other words, the farmers' objects of value accommodate more than just a hard set of diverse biological traits, as illustrated in this exchange:

> *"But what I expect from the goats is that everyone is happy – myself and the goats – and that we all get a decent living because, well, because they give good milk. Because I try to make sure that they get the best life possible, and give us milk, and then we sell the cheese and we are happy with that."*

Both the aesthetics factor and the considerations around overall satisfaction, as expressed in the interviews, carry an emotional dimension, and this emotional dimension will translate into how an animal population is managed. It is worth noting that the emotions expressed formed a subtheme of the 'human-animal relationship' theme. Although this is not the focus of this paper, it highlights how emotions are connected to attachment to animals. In an illustrative example, one farmer said:

> *"It's the city; I don't like it there and I miss my goats, so I don't like being away from them for too long."*

And one added:

"In time, you start to love these animals. There's really something special about goat farming."

Farmer narratives on balance and trade-offs

Finally, listening to the farmer's accounts of the broad diversity of features they value, a challenge emerged: how to integrate these diverse features, or in other words, how to find balance or manage trade-offs. The farmers' narratives are effectively punctuated with stories that talk about balance and trade-offs. We identified several ways in which this is expressed that employ a number of terms, such as "package", "rounded", or "balance":

"You ask what I'm looking for in particular? It's a combination, a package"

Or

"When I bought her, she was an instant favourite. She was a goat that had an angular, bone-led frame with a big head, which matched exactly what I was looking for in the Pyrenean. Very very long hair. She only had one udder, but she had been bucks dam for several years in a row. She had so much, she gave so much, gave a lot of milk too. So for me, she counted among the goats that I'd call 'rounded'. She was a real favourite. Gentle, easy temperament – and she even lived to 16 years old."

"You learn pretty quick that a goat that's giving 3–4L of milk puts everything into milk and forgets to invest in immunity. So you have to find the right balance."

How efforts to find balance and manage trade-offs manifest into practice

This effort to find balance is explained through farmers' statements on their selection decisions within the Pyrenean goat breed population, or for renewal of their own herd.

Note that any trade-offs made will primarily depend on the perception of the focal traits and on whether selection can serve as a lever to achieving satisfactory goals. For example, one interviewee said they lend more importance to certain selection criteria on the grounds that other criteria, although important, are considered innate (i.e. inherently part of the breed's make-up, and therefore not requiring active selection):

"Quite honestly, me, personally, I put type first: good legs, a big muzzle, good ears, hair. For me, type comes first. I know, it doesn't sound logical, but... it's because I think any Pyrenean goat has good yields already, it's innate".

In terms of ways to construct a compromise between traits to select for, the process can translate, for example, as (1) expressing a set of several traits to be co-selected together as a set, or (2) prioritizing criteria by first selecting for one and then selecting for others. The following farmer's narrative illustrates, for instance, approach (1):

"Always looking for ways to continue selecting for both type and milk, without putting one ahead of the other. I really think you can manage to get both. You have to work on it, but you can get both."

This other farmer's narrative is more illustrative of approach (2):

"I'm in a performance recording programme, so obviously I look at growth. Growth in my herd first. Then I'll adjust depending on maternal prolificacy, maternal hardiness."

Discussion

A diversity of farmers' objects of value in animals

The variety of farmers' objects of value lead to a reconsideration of how animal diversity is characterized. Multiple objects of value were often observed grouped together, which means that farmers' concerns revolve around accommodating balances and trade-offs. This diversity also shows that while typical zootechnical categories, such as performances and abilities, may work for a number of elements that farmers wish to see, these categories fail to encompass the full range of elements valued. This insight prompted us to further explore the notion of 'breed attribute', as proposed in our previous research, to account for zootechnical characteristics of local-breed animals alongside characteristics connected to other dimensions, such as their importance for sustaining local cultures and communities (Nozières-Petit & Lauvie, 2018; Lauvie *et al*, 2023). Although these attributes effectively connect to diverse functionalities provisioned by local-breed animals, here we found that this purely functionality-driven vision is too narrow, as our results showed that experiential or emotional dimensions may also come into play. The objects of value expressed by farmers are more than just 'zootechnical', reflecting a diverse set of potential reasons for choosing one animal or breed over another. Farming livestock as a livelihood is also associated with other dimensions that carry meaning to farmers. From this perspective, the features of discourse that translate the relational dimension and the emotions involved warrant attentive investigation. There is a compelling rationale for developing research at the intersection of zootechnics and sociology or zootechnics and philosophy that looks at relational bonds between humans and farm animals (see Porcher, 2001; Despret and Meuret, 2016). In a report on the results of her thesis research, Porcher (2001) argues that affectivity is an integral part of working with livestock and that this affectivity component has been underconsidered in research on animal welfare (which was her original starting point). The object of related research has thus shifted towards the study of work, i.e. the study of how animals and humans work together. To refocus research on the management of farm-animal populations, a promising approach is to better account for human–animal relationships in the analysis of livestock farming practices, starting by looking at how breeders talk about and describe these relationships and the way these relationships interact with specific practices and contribute to the satisfaction breeders get out of their work.

It is worth noting the example of a farmer who considered the occasional stubbornness of animals in his herd – when

they refused to obey his command – as a positive trait. Although this only emerged from one interview, it offers an insightful starting point for defining the perimeters (in terms of objects addressed), strategies and paradigms engaged by disciplines dealing with the management of genetic resources. This is because this kind of discourse challenges the traditional strategy. Rather than projecting methodical control, it embraces the idea of 'working with' (*composer avec* in French, borrowing from the title of Despret and Meuret (2016) cited above), of 'following/guiding nature' rather than trying to control it (Morin, 1980). More broadly, attending to what it means for humans to work with living beings would help address a question that is pivotal to the principles of agroecology (Hubert, 2020).

From a methodological perspective, one way to further pursue this research would be to produce the most accurate description possible of farmers' practices, particularly in terms of decisions around which animals to retain in the herd, but also, more broadly, the relationships that form between farmers and animals at key points in livestock management. This methodological approach could usefully combine interviews tailored to capture these points with direct observation of practices. Equally important would be a diachronic approach enabling specific attention to the way farmers effectuate changes in their practices (integrating their own analysis of the outcomes of their practices), in order to introduce a dynamic perspective where values are not seen as immutable. It is at this juncture that we address the formation of values as theorized by John Dewey, and as stated by Bidet *et al* (2011), who conceptualize valuations purely as behaviours that are situationally observable; consequently, a valuation cannot be reduced to a representation.

Pulling together everything that farmers value: selection as a lever in multidimensional and relational approaches

As stated in the introduction, one of the key motivations for focusing on what farmers value in their herds was to inform fresh thought on managing genetic resources in farm-animal populations. However, our findings resonate strongly with the questions being raised in selection programmes for farm animals. One challenge identified for genetics is diversifying the objectives of selection programmes. This involves defining breeding objectives that balance production traits with functional abilities. Such diversification is crucial for advancing the agroecological transition (Phocas *et al*, 2016). This trend finds confirmation in the diversity of traits that the farmers surveyed considered important. At the same time, they emphasize the effort to find balance and trade-offs, which aligns with models that consider herd management and genetic selection together while examining trade-offs among multiple performances and abilities within a herd (see Douhart, 2013). To complement these modelling approaches, characterizing how farmers translate this effort to find the right balance and trade-offs into practice could provide insight into the different ways these choices come together. Furthermore, the strong links between animal traits and flock management, and between animal traits and flock environments, that emerged here in the farmers' narratives are also in line with recent work by geneticists to better characterize livestock breeding environments and account for genotype–environment interactions (Phocas *et al*, 2016; Hazard *et al*, 2017). This observation argues for bridging disciplines to bring scientific communities together – the genetics community and the livestock farming system community, and even reaching out to ecologists and social scientists. Mobilizing farmers' field knowledge of their animals and their local farm habitats – and possibly even the ways they describe how these two factors, i.e. animals and habitats, intersect and interplay – also appears to be a promising direction for developing this research front.

Finally, diversity has also been identified as a pivotal issue and a key resource for making genetic selection better adapted to addressing agroecology challenges (Phocas *et al*, 2016). However, the local-breed farmers studied here considered and valued such a diversity of dimensions that they cannot viably be reduced simply to a gene pool; therefore, it is important to consider the genetic dimension in conjunction with other dimensions of value embodied by local-breed animals. Ducos *et al* (2021) stress that a contribution of animal genetics to the diversity principle of agroecology requires going beyond approaches based on the search for an optimal animal with calibrated performances for standardized and controlled breeding environments, and adopting a more systemic view.

The convergence and complementarity between our findings and some of the central challenges for animal genetics highlight that the plurality of dimensions and relationships emerging from this work demonstrate how overly reductionist approaches fail to capture the full complexity of questions concerning the suitability of livestock animals to farming practices, environments and systems. We arrived at the idea that progress could be made by developing a relational approach that repositions genetics as one of several levers intersecting with livestock animals, farming practices and environments in the broadest sense, i.e. in their ecological dimensions, technical dimensions, social dimensions, and beyond. This would require even greater interdisciplinarity and transdisciplinarity, so that we move beyond simply devising selection programmes toward producing knowledge about a nexus of suitability between animals, habitats/environments, and the objectives breeders attach to their livestock practice. In this new approach, the lever of genetic selection would be reframed to articulate with other actionable levers by crossing boundaries to mobilize disciplines that develop systemic approaches, but also social sciences and the field knowledge of farmers and the wider stakeholder community involved in managing farm-animal populations. The type of approach envisioned here would aim to produce knowledge that is both more global (i.e. holistically including a diversity of knowledge sets and dimensions), situated (to address agroecology challenges) and operable (i.e. borrowing from practices that operate at animal, herd, and breed population scales).

A further avenue for extending this work would be to attach greater importance to the observable valuation processes at work at the breed-society scale. Indeed, the methodology developed in this article rests upon interviews and mainly focuses on individual values and valuation processes. The composition of our work team, built through a transdisciplinary process, included the facilitator of the association, who had good knowledge of the collective dynamics and discussions. This allowed us to check a certain consistency of the analysis. However, to better

tackle the collective dimension of valuation, complementary methodologies could be used, such as focus groups, to investigate how farmers talk about what they value in their animals when they discuss this theme together. Participant observation of moments when collective decisions are made at the scale of the breed association could also be a relevant methodology. It would allow a better understanding of how those values are discussed (through which objects and themes, which kind of actions), and help identify around which values consensus is built (which play a role in holding the collective together). We take the stance that closer attention to situations involving collective management of animal populations that have been sidelined from major selection programmes is a potentially fertile approach to help identify configurations, relationships, questions and practices already at work, which could usefully inform processes of rethinking better ways to mobilize genetic resources – a challenge that is now high on the agenda in the field of genetic selection for both animal and plant species.

Conclusion

This study aimed to describe the full diversity of what farmers value in their local-breed animals. This entry point enabled us to characterize a wide-ranging diversity of their objects of value spanning zootechnical performances, adaptive capacities, behavioural factors, self-sufficiency (requiring little human intervention), appearance and aesthetics, genetic diversity, and more. We also demonstrated the overarching importance of trade-offs and efforts to find balance among different traits. Managing genetics to meet these expectations involves more dimensions than simply focusing on improving individual traits; it requires an integrated approach to livestock farming. This approach is situated, or, in other words, it takes into account the specificities of each situation, and it connects people with tools, practices and animals.

Supplemental data

Supplemental Material 1. Verbatim accounts from the interviews in French, and their translation into English, as reported in the article.

Author contributions

Conceptualization/design of the research questions: AL, NC, MONP, FT; Conceptualization/theoretical framework: NC, AL; Methodology: AL, NC, MONP, FT; Formal analysis: AL, NC, MONP, FT; Supervision of the investigation (interviews conducted by a student): AL, MONP, FT, NC; Writing – first version of the original draft: AL; Writing – Completed original draft: AL, NC, MONP, FT; Writing – review and editing: AL, NC, MONP, FT.

Acknowledgements

The authors thank senior-year intern Blaise Dupuis for conducting the interviews that provided the material for this research. We also thank all the farmers who graciously set aside their time for these interviews. Metaform Langues translated the French draft of this paper into English. We thank the reviewers for their useful comments. This work was supported by the joint technology unit 'UMT PASTO' ['Resources and transformations for pastoral livestock farming in Mediterranean regions'].

Conflict of interest statement

The authors declare no conflict of interest.

Ethics statement

Farmers gave permission to conduct the interviews for the purposes of this research, and were fully informed about its purposes.

References

Bidet, A., Quéré, L., Truc, G. (2011). Ce à quoi nous tenons. Dewey et la formation des valeurs, In Dewey, J. *La formation des Valeurs*, traduit de l'anglais et présenté par Alexandra Bidet, Louis Quéré, Gérôme Truc. (Les empêcheurs de penser en rond/ La découverte) Paris, 5-64.

Boichard, D., Ducrocq, V., Fritz, S. (2015). Sustainable dairy cattle selection in the genomic era. *J. of Animal Breeding and Genetics* 132 (2),135-143. doi: https://doi.org/10.1111/jbg.12150

Couix, N., Lauvie, A., Verrier, E. (2023). Biodiversité domestique animale : à quoi tenons-nous ? In Lauvie, A., Audiot, A., Verrier, E. (eds) *La biodiversité domestique. Vers de nouveaux liens entre élevage, territoires et société.* (QUAE), 111-127.

Despret, V., Meuret, M. (2016). Composer avec les moutons, lorsque des brebis apprennent à leurs bergers à leur apprendre. (Cardère, hors les drailles) 154p.

Dewey, J. (1939). Theory of Valuation. Chicago University Press (French translation in Bidet *et al*, 2011)

Douhard, F. (2013). Towards resilient livestock systems : a resource allocation approach to combine selection and management within the herd environment. PhD Thesis AgroParisTech.

Ducos, A., Douhard, F., Savietto, D., Sautier, M., Fillon, V., Gunia, M., Rupp, R., Moreno, C., Mignon-Grasteau, S., Gilbert, H., Fortun-Lamothe, L., Lamothe, L. (2021). Contributions de la génétique animale à la transition agroécologique des systèmes d'élevage. *INRAE Prod. Anim.* 34(2), 79–96 doi: https://doi.org/10.20870/productions-animales.2021.34.2.4773

Dumont, B., Fortun-Lamothe, L., Jouven, M., Thomas, M., Tichit, M. (2013). Prospects from agroecology and industrial ecology for animal production in the 21st century. *Animal* 7 (6), 1028-1043.

Hazard, D., Larroque, H., González García, E., Francois, D., Hassoun, P., Bouvier, F., Parisot, S., Clement, V., Piacère, A., Masselin-Silvin, S., Buisson, D., Loywick, V., Palhiere, I., Tortereau, F., Lagriffoul, G. (2017). Caractérisation des environnements de production et de nouveaux phénotypes pour améliorer la sélection et l'adaptation des ovins et caprins dans des environnements variés. In FAO/CIHEAM Network for Research and Development in Sheep and Goats. *Joint Seminar of the Subnetworks (Nutrition and Production systems) and Innovation for Sustainability in Sheep and Goats* (iSAGE), Vitoria-Gasteiz, Spain, 2017-10-03 2017.

Hubert, B. (2020). Agriculture et alimentation. Les modèles de production questionnés : l'impératif du changement agroécologique. *Raison présente* 213 (1), 85-96. doi: https://doi.org/10.3917/rpre.213.0085

Lauvie, A., Thuault, F., Nozieres-Petit, M.O., Dupuis, B., Couix, N. (2022). Valuation of local breeds: methodological questions following interviews with Pyrenees goat farmers, In *73rd Annual Meeting of the European Federation of Animal Science*, EAAP, Sep 2022, Porto, Portugal.

Lauvie, A., Alexandre, G., Angeon, V., Couix, N., Fontaine, O., Gaillard, C., Meuret, M., Mougenot, C., Moulin, C.-H., Naves, M., Nozières-Petit, M.-O., Paoli, J.C., Perucho, L., Sorba, J.-M., Tillard, E., Verrier, E. (2023). Is the ecosystem services concept relevant to capture the multiple benefits from farming systems using livestock biodiversity? A framework proposal. *Genet. Res.* 4 (8), 15-28. doi: https://doi.org/10.46265/genresj.MRBT4299

Macé, M. (2021). Parole et pollution. *AOC Media Analyse Opinion Critique* (29.01.21). url: https://aoc.media/opinion/2021/01/28/parole-et-pollution/

Morin, E. (1980). L'écologie généralisé *Oikos*. In *La méthode 2. La vie de la vie* (Ed. du Seuil), 15-97.

Moulin, C.H., Perucho, L. (2023). Biodiversité domestique et résilience des systèmes d'élevage. In Lauvie A., Audiot A., Verrier E. (eds) *La biodiversité domestique. Vers de nouveaux liens entre élevage*, territoires et société. (QUAE), 131-145

Nozières-Petit, M.-O., Lauvie, A. (2018). Diversité des contributions des systèmes d'élevage de races locales. Les points de vue des éleveurs de trois races ovines méditerranéennes. *Cah. Agric.* 27 (6), 65003

Phocas, F., Belloc, C., Bidanel, J., Delaby, L., Dourmad, J.Y., Dumont, B., Ezanno, P., Fortun-Lamothe, L., Foucras, G., Frappat, B., González-García, E., Hazard, D., Larzul, C., Lubac, S., Mignon-Grasteau, S., Moreno, C.R., Tixier-Boichard, M., Brochard, M. (2016). Review: Towards the agroecological management of ruminants, pigs and poultry through the development of sustainable breeding programmes: I-selection goals and criteria. *Animal* 10 (11), 1749-1759. doi: https://doi.org/10.1017/S1751731116000926

Porcher, J. (2001). L'élevage, un partage de sens entre hommes et animaux : intersubjectivité des relations entre éleveurs et animaux dans le travail. *Ruralia* ([En ligne], 09, 2001 accessed 15 February 2024. URL: https://journals.openedition.org/ruralia/278

Pyrenean goat breed association (2024a). Craba e caulet, la chèvre et le chou, bulletin de liaison de l'Association la Chèvre de race pyrénéenne, 41, 5 p.

Pyrenean goat breed association (2024b). Website of the Pyrenean goat breed association, *l'association La chèvre de race pyrénéenne*. https://www.chevredespyrenees.org/ (accessed 13 February 2024).

Vissac, B. (2002). Les vaches de la République: saisons et raisons d'un chercheur citoyen, INRA Editions.