Genetic Resources

Genetic Resources (2020) Vol. 1, Issue 2
DOI: 10.46265/genresj.2020.2
www.genresj.org
ISSN: 2708-3764

© Hailu et al

© 2020 ECPGR, CC-BY-NC-SA

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, animal and forest genetic resources involved in monitoring, collecting, maintaining, conserving, characterizing and using genetic resources for food, agriculture and forestry. *Genetic Resources* publishes original research as well as methods, strategies, guidelines, case studies or reviews on a variety of topics of interest on the present and future use of genetic resources. These may include the 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 welcomes contributions from all world regions.

Cover illustration:
Highland sheep flock in Tahtay Maichew district, Tigray, Ethiopia. Hailu *et al* pp. 12-22.

The designations employed, and the presentation of material in the periodical, and in maps which appear herein, do not imply the expression of any opinion whatsoever on the part of ECPGR concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries. Similarly, the views expressed are those of the authors and do not necessarily reflect the views of ECPGR.

GenRes Bridge
Genetic resources for food-secure and forested Europe

This issue has been produced in the frame of GenRes Bridge, a project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817580

Editorial Office:

ECPGR Secretariat
c/o Bioversity International
Via dei Tre Denari 472A
00054 Maccarese (RM), Italy

Submissions to www.genresj.org

Editorial Board:

Managing editor:

Sandra Goritschnig (European Cooperative Programme for Plant Genetic Resources, Italy)

Plant Genetic Resources:

Joana Brehm (University of Birmingham, UK)

Luigi Guarino (The Crop Trust, Germany)

Igor Loskutov (Vavilov Institute of Plant Genetic Resources, Russia)

Lorenzo Maggioni (European Cooperative Programme for Plant Genetic Resources, Italy)

Nigel Maxted (University of Birmingham, UK)

Jade Phillips (University of Birmingham, UK)

Alvaro Toledo (International Treaty on Plant Genetic Resources for Food and Agriculture, FAO, Italy)

Animal Genetic Resources:

Enrico Sturaro (University of Padova, Italy)

Odd Vangen (Norwegian University of Life Sciences, Norway)

Etienne Verrier (AgroParisTech, France)

Genetic Resources

Genetic Resources (2020), 1 (2)
DOI: 10.46265/genresj.2020.1
https://www.genresj.org
ISSN: 2708-3764

TABLE OF CONTENTS

Original Articles

Other Articles

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2020), 1 (2), 1–11
DOI: 10.46265/genresj.BJCV8100
https://www.genresj.org
ISSN: 2708-3764

Assessment of wild tomato accessions for fruit yield, physicochemical and nutritional properties under a rain forest agro-ecology

Dorcas O. Ibitoye[a], **Adesike O. Kolawole** [*,b] **and Roseline T. Feyisola**[c]

[a] *Genetic Resources Unit, National Horticultural Research Institute, PMB 5432, Ibadan, Nigeria*
[b] *Department of Crop Production and Soil Science, Ladoke Akintola University of Technology, PMB 4000, Ogbomoso, Nigeria*
[c] *Department of Plant Science, Olabisi Onabanjo University, Ago-Iwoye, Nigeria*

Abstract: Tomato (*Solanum lycopersicum* L.) is a broadly consumed fruit vegetable globally. It is one of the research mandate vegetable of the National Horticultural Research Institute (NIHORT), Ibadan, Nigeria. The institute's genebank contains diverse collections of tomato accessions and wild relatives, without utilization information for the African continent. With the decline in diversity and potential of cultivars, a robust tomato breeding pipeline with broad genetic base that eliminates redundancy in the development of lines with desired horticultural traits is paramount. This study evaluated the mean performance and variations of thirteen wild tomato accessions obtained from the C.M. Rick Tomato Genetic Resource Center, University of California, Davis, USA, evaluated for agronomic, nutritional and physicochemical traits under a rain forest agro-ecology zone in Nigeria. The accessions were planted and grown in three replications with randomized complete block design. Agronomic traits, physicochemical and nutritional parameters were measured and analyzed. There was significant ($P < 0.001$) variation among accessions for all traits measured. Accession LA0130 was separated from others by cluster analysis and was outstanding for its unique attributes which include: fruit yield parameters, total soluble solids, titratable acidity and lycopene content. The principal component analysis suggests fruit yield related traits, titratable acidity and lycopene contributed most to the variation among the 13 accessions. The results obtained can be used to breed materials adapted to a rain forest agro-ecology. These wild tomato accessions have genes with desirable agronomic, nutritional and physicochemical traits that could be introgressed into breeding lines to improve commercial tomato varieties.

Keywords: agronomic traits, breeding programme, fruit quality, variation, wild relatives

Citation: Ibitoye, D. O., Kolawole, A. O., Feyisola, R. T. (2020). Assessment of wild tomato accessions for fruit yield, physicochemical and nutritional properties under a rain forest agro-ecology. *Genetic Resources* 1 (2), 1–11. doi: 10.46265/genresj.BJCV8100.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Tomato (*Solanum lycopersicum* L. formerly *Lycopersicon esculentum* Mill.) is one of the most famous and broadly consumed vegetable crops throughout the world (Nowicki *et al*, 2013; Ajayi and Hassan, 2019). Nigeria was ranked the largest producer of tomato in West Africa and the 16^{th} largest producer in the World with 4.2 million metric tonnes FAO (2016). These data suggest prospects for Nigeria tomato breeding programmes to enhance production efficiency by improving the quantity and quality of tomato fruit. However, extensive breeding efforts and selection over the years have modified tomato (Blanca *et al*, 2015). The decline in diversity and potential of cultivated germplasm has been reported (Jatoi *et al*, 2008; Chen *et al*, 2009). To enlarge the gene pool of cultivars, breeders now focus on introgression of desirable genes from wild relatives (Singh, 2006) .

Wild tomato species have a rich reservoir of useful genetic traits needed to improve cultivated toma-

*Corresponding author: Adesike O. Kolawole (aokolawole@lautech.edu.ng)

Received: 05.05.2020 Accepted: 09.09.2020 Published online: 21.12.2020

toes and serve as sources of genetic variability (Hanson *et al*, 2007). Miller and Tanksley (1990) reported that the genomes of cultivated tomato contain 5 % of the genetic variation of their wild relatives. These wild tomato species are native to western South America and distributed from central Ecuador, through Peru to northern Chile, and in the Galápagos Islands (Darwin *et al*, 2003). There are 16 wild species of tomato, namely *Solanum habrochaites*, *S. pennellii*, *S. pimpinellifolium*, *S. cheesmaniae*, *S. galapagense*, *S. peruvianum*, *S. corneliomulleri*, *S. chilense*, *S. chmielewskii*, *S. arcanum*, *S. neorickii*, *S. huaylasense*, *S. lycopersicoides*, *S. ochranthum*, *S. jugandifolium*, and *S. sitiens* (Rick and Fobes, 1975; Peralta *et al*, 2008; Zuriaga *et al*, 2009). The closest wild ancestor to cultivated tomato is *S. pimpinellifolium* L. found in the centre of origin of tomato from the northern part of Chile to Colombia. Most of these wild relatives are vulnerable to extinction because of their small population sizes (Bai and Lindhout, 2007; Onyia *et al*, 2019). Therefore, wild tomato accessions stored in genebanks need to be evaluated in time and space in order to identify significant traits and valuable potential. Documented information about the performance of wild tomato accessions will aid future use in tomato breeding programmes.

Currently, the breeding programme at the National Horticultural Research Institute (NIHORT), Ibadan, Nigeria has made diverse collections of landraces, cultivars and wild tomato accessions. Only the landraces and cultivars have been extensively studied for their agronomic performance and resistance to abiotic and biotic stresses (Olaniyi *et al*, 2010; Nnabude *et al*, 2015). However, the wild accessions obtained from the C.M. Rick Tomato Genetic Resource centre, University of California, Davis, USA have not been studied for the identification of useful agronomical traits, nutritional and physicochemical parameters in Nigeria.

In addition to fruit yield and improved agronomical traits, the fruit quality and nutritional parameters are essential breeding objectives from a consumer's point of view and even in the processing industries (Bauchet and Causse, 2012; Bergougnoux, 2014). Fruit quality and physicochemical parameters are cultivar dependent (Riahi *et al*, 2009; Ilahy *et al*, 2011; Domínguez *et al*, 2012). Fertilization, cultural practices and post-harvest storage could influence fruit quality and physicochemical parameters (Rosales *et al*, 2011; Beckles, 2012). Production of quality tomato fruits depends on the climatic conditions, sunlight availability, good agronomic practices and genetic variability among cultivars (Causse *et al*, 2003; Peixoto *et al*, 2018).

In order to identify, select and develop novel tomato lines with desired horticultural traits for a Nigerian breeding programme, it is crucial to evaluate tomato accessions which are wild to the Nigerian germplasm. These could then be subjected to diverse breeding methodologies and agronomic practices (Chitarra and Chitarra, 2005). Knowledge of desirable traits from the evaluated tomato accessions will help to identify those that could be used as parents in a tomato breeding programme, hence, promoting improved nutrition and increased production (Causse *et al*, 2003; Álvaro Toledo and Burlingame, 2006).

This study evaluated mean performance and variation among wild tomato accessions based on agronomic, nutritional and physicochemical traits under rainforest agro-ecology in Nigeria.

Materials and Methods

Germplasm

The accessions evaluated in this study are from the C.M. Rick Tomato Genetic Resource Centre (TGRC), University of California, Davis, USA (Table 1). They are part of the tomato wild relatives core collection of the NIHORT germplasm bank without previous utilization information that could be included in breeding programmes.

Nursery and field operations

Seeds were sown into perforated nursery trays filled with sterilized soil and grown for three weeks in a greenhouse at NIHORT, Ibadan, Oyo State (Rain forest zone; 3° 56'E, 7° 33' N; 168 meters above sea level). The perforated nursery trays were kept moist by regular watering on daily basis at sunrise and sunset with tap water using a watering can. Seedlings were transplanted to the field in paired rows in plots that were 2 m long with spacing of 0.5 m between rows and 0.5 m between plants within a row. Spacing between plots was 1 m. Seedlings were arranged in a randomized complete block design with three replications. N-P-K (15-15-15) fertilizer was applied at the rate of 120 kg/ha three weeks after transplanting. The plants were trellised to prevent lodging and loss of fruits due to diseases and pests. Manual weeding was carried out at two-week intervals. To protect the leaves from defoliating pests, plants were sprayed with the pyrethroid insecticide Cymbush containing cypermethrin at 2, 6, and 9 weeks after transplanting at the rate of 450 ml of active ingredients per 100 liters of water per hectare using a knapsack sprayer. No disease infestation was observed during the experiment and data were collected on five randomly selected plants per plot.

Sample preparation

Uniformly ripe, healthy fruit at the red-ripe stage were harvested (Hanson *et al*, 2004). A total of 10-15 representative fruit were collected from pre-tagged plants (from the first 3 clusters) to minimise intra-plant variability (Borja *et al*, 1998). Tomato samples (100 g) were homogenized in 50 mL of water in a water bath at 4°C and low light (to reduce antioxidant loss) for physicochemical analysis. All analyses were done in triplicate for each sample at the Product Development Laboratory of NIHORT.

Table 1. Description of the 13 species of wild tomato accessions used for the experiment

No.	Accession ID	Species	Origin (Area of collection)
1	LA0103	*Solanum peruvianum* (L.) Miller	Cajamarquilla Lima, Peru
2	LA0130	*Solanum chilense* Dunal	Moquegua, Peru
3	LA0411	*Solanum pimpinellifolium* (L.) Miller.	Pichilingue, Los Rios, Ecuador
4	LA1028	*Solanum chmielewskii* (C.M. Rick et al.)	Casinchichua, Apurimac, Peru
5	LA1041	*Solanum cheesmanii* L. Riley	Santa Cruz, El Cascajo, Galapagos Islands, Ecuador
6	LA1136	*Solanum cheesmanii* L. Riley	Gardner st, Floreana Islet, Galapagos Islands, Ecuador
7	LA1208	*Solanum esculentum* var. *cerasiforme* Dunal	Sierra Nevada, Colombia
8	LA1272	*Solanum pennellii* (Correll) D'Arcy	Pesquera, Lima, Peru
9	LA1293	*Solanum peruvianum* f. *glandulosum* (C.F. Mull)	Matucana, Lima, Peru
10	LA2641	*Solanum parviflorum* (C.M. Rick et al.)	Apurimac, Peru
11	LA4113	*Solanum sitiens* I.M. Johnst.	Estación Ceres, Antofagasta Chile
12	LA4115	*Solanum sitiens* I.M. Johnst.	Quebrada Desde Cerro, Oeste De Paqui, Antofagasta, Chile
13	LA4138	*Solanum pimpinellifolium* (L.) Miller	El Corregidor, La Molina, Lima, Peru

Physicochemical and nutritional analyses

Total soluble solid (°Brix) (g 100 g^{-1}) of the juice was measured using the Eclipse hand-held refractometer [PN# 45-01 (0-15 °Brix)] and the pH of the fruit juice was measured using a benchtop pH meter (Sper scientific benchtop) with the pH meter calibrated with standard buffers pH 4 or 9. For determination of titratable acidity (g 100 g^{-1}) and vitamin C content (mg 100 g^{-1}), 10 mL of juice from 10 fruits was diluted in 100 mL of distilled water and titrated with NaOH (0.1 N) to pH 8.2. For vitamin C, the solution was titrated with iodine (0.1 N) until a colour change was observed (International Plant Genetic Resources Institute, 1996) .

To determine lycopene content (mg 100 g^{-1}), 5 mL of acetone-n-hexane mixture in the ratio 4:6 was added to 0.8 g of tomato pulp for each sample and mixed well. The mix was centrifuged at 5000 rpm for 5 min at 4°C; the supernatant was extracted and absorbance measured with a spectrophotometer (model 6400, Jenway) at 503 nm using the acetone-n-hexane mix as blank (Rosales *et al*, 2006). Lycopene content was calculated using an extinction coefficient ($E^{1\%}$) of 3150.

Agronomic data collection

Data were collected on the following traits: number of leaves at maturity (NLM), plant height at maturity (PH), number of clusters per plant (NCP), number of fruits per cluster (NFC), fruit weight (FW), fruit length (FL), fruit circumference (FC), number of fruits per plant (NFP) and fruit size index (FSI). Fruit yield of tomato was adjusted to t ha^{-1}using the following formula: Fruit yield (t ha^{-1}) = fruit yield per plot (kg) x 10,000 / plot area (m^2) x 1,000.

Statistical analysis

The data was subjected to analysis of variance (ANOVA) using PROC GLM in SAS (SAS Institute, 2010) . Means were separated using Fisher's least significant difference (LSD) test ($P < 0.05$). A rank summation index (RSI) (Mulumba and Mock, 1978) was constructed to create the aggregate trait by ranking accessions with regard to high fruit weight, fruit yield, improved agronomics, nutritional and physicochemical traits. Ranks were summed for each accession to select the top five. Pearson's correlation analysis was done to determine associations among all traits measured with SAS. Hierarchical cluster analysis was performed using SAS PROC CLUSTER based on centroid distance and a dendrogram constructed by PROC TREE in SAS to identify divergent groups. To identify patterns of morphological variation, principal component analysis (PCA) was conducted. Those PCs with Eigen values >1 were selected (Jeffers, 1967). The PCA analysis reduces dimensions of a multivariate data to a few principal axes, generates an Eigen vector for each axis and produces component scores for characters (Sneath and Sokal, 1973).

Results

The ANOVA produced significant mean squares for all agronomic, nutritional and physicochemical traits of the tomato accessions indicating genetic variations for all measured traits (Table 2). The coefficient of variation (CV) used to measure the precision of the experiment indicated the data was reliable (Table 2). Phenotypic variation in the biological growth stages of tomato accessions revealed that LA4113 was tallest and LA2641 shortest (Table 3). The most fruits per cluster were observed for accession LA0103 and least for LA1041. The most fruits per plant were observed for accession LA0411, the least for LA4138. For fruit yield related

Table 2. Mean squares from analysis of variance of agronomic, nutritional andphysicochemical traits of wild tomato accessions

Source	df	NLM	PH (cm)	NCP	NFC	FW (g)	Fruit yield (t/ha)	FL (cm)	FC (cm
Replication	2	752.03	65.89	4.58	31.57**	57571.79**	1.44**	0.19*	0.15
Accessions	12	3254.65***	148.54***	31.89**	28.52***	24458.12*	0.61*	0.37***	0.65***
Error	24	408.16	37.18	10.53	4.80	11010.68	0.28	0.05	0.11
CV		17.77	13.67	33.11	29.22	68.32	68.32	13.20	19.79

Source	df	NFP	FSI	VIT C (mg/100 g)	TSS (^{o}Brix) (g/100 g)	TA (g/100g)	Fruit juice pH	Lycopene (mg/100 g)
Replication	2	1381.87	0.02	2.3	0.0004	0.0003	0.003	0.51
Accessions	12	4605.24***	0.08***	147.40***	0.51***	0.17***	0.08***	106.12***
Error	24	841.79	0.02	2.54	0.03	0.002	0.002	0.31
CV		42.25	12.37	4.94	4.14	4.74	0.95	2.38

*, **, *** Significant at 0.05, 0.01 and 0.001 probability levels, respectively.
CV = Coefficient of variation, NLM = Number of leaves at maturity, PH = Plant height at maturity, NCP = Number of cluster per plant, NFC = Number of fruits per cluster, FW = Fruit weight, FL = Fruit length, FC = Fruit circumference, NFP = Number of fruits per plant, FSI = Fruit size index, VIT C = Vitamin C, TSS = Total soluble solids, TA = Titratable acidity.

traits, tomato accession LA0130 had the heaviest fruit and most fruit yield, LA1293 had the lowest fruit weight and least fruit yield. Accession LA0411 had the highest concentration of vitamin C. LA1028 had the highest levels of total soluble solids while accessions LA4113, LA4138, LA1041 had the lowest levels. Accessions LA1208 and LA4133 had the lowest titratable acidity, LA2641, LA1293 and LA0130 had the lowest fruit juice pH. Lycopene content was highest in tomato accession LA0130 and lowest in LA1208 (Table 3).

Based on a rank summation index (RSI) of 13 accessions, LA0130 was identified as best performing among all tested accessions, with the best fruit yield performance, desirable agronomic, nutritional and physicochemical traits. Tomato accession LA0130 was characterized by moderate plant height, highest number of fruits per cluster, fruit weight and fruit yield (Table 3). Accession LA0130 also had the highest titratable acidity and lycopene content, but moderate fruit juice pH (Table 4).

Pearson's correlation coefficient was calculated to determine associations among traits and showed variation for some trait combinations (Table 5). Fruit yield was significantly positively correlated with fruit weight coupled with a significant negative correlation with number of leaves at maturity. Fruit size index was significantly negatively correlated with fruit circumference and significantly positively correlated with vitamin C content. Total soluble solid was significantly positively correlated with number of leaves at maturity and vitamin C content. Titratable acidity was significantly positively correlated with number of fruit per plants and total soluble solid. Fruit juice pH was significantly negatively correlated with number of fruit per plant, total soluble solid and titratable acidity. Lycopene content was positively significantly correlated with vitamin C content and titratable acidity.

All agronomic traits, nutritional and physicochemical parameters measured which showed significant variations were adopted to construct a hierarchical cluster based on the centroid distances among the 13 tomato accessions as in Figure 1. Cluster analysis differentiated the accessions into 4 distinct groups, where LA0130 differed from the other three groups. Cluster I consisted of five accessions, cluster II had four accessions, and clusters III and cluster IV had three and one accessions, respectively indicating variation among the accessions.

Additionally, contribution of each measured trait to the total variation within the accession was further determined through Principal Component Analysis (PCA) based on correlation matrix of the variables. The Scree plot of the PCA indicated six eigenvalues corresponding to the entire percent variance with eigenvalues >1. PCA1 accounted for about 22 % of variation, PCA2 for 19 %, PCA3 for 15 %, PCA4 for 11 %, PCA5 for 10 % and PCA6 for 6 % (Table 6). The first principal component axis (PCA1) was mainly loaded positively by fruit yield, fruit yield related traits and titratable acidity. In PCA2 traits which had positive contribution were number of leaves at maturity, plant height at maturity, titratable acidity and lycopene. Fruit length, fruit circumference, total soluble solids and lycopene had positive contributions in PCA3. In PCA4 plant height at maturity, number of clusters per plant, fruit weight, fruit yield and fruit juice pH had positive contributions were. In PCA5 number of clusters per plant, number of fruits per cluster, fruit length, fruit circumference and total soluble solids had positive contributions. In PCA6 traits which had positive contributions were fruit length, fruit size index and Vitamin C content.

Discussion

Significant phenotypic variations among the accessions for all agronomic, nutritional and physicochemical traits validate availability of genetic diversity in the collection from the C.M. Rick Tomato Genetics

Table 3. Mean ranking of agronomic traits of wild tomato accessions evaluated under the rainforest agro-ecology zone in Nigeria

S/N	Accession	PH (cm)	NFC	NFP	FW (g)	Fruit yield (t/ha)	NLM	NCP	FL (cm)	FC (cm)	FSI	RSI
1	LA0130	44.14	11.72	95.42	380.00	1.90	99.67	9.72	2.30	2.05	1.15	30
2	LA0411	52.97	8.64	143.78	113.33	0.57	137.67	9.78	1.22	1.17	1.05	51
3	LA1136	45.22	8.50	79.33	236.67	1.18	109.67	8.44	1.33	1.20	1.13	63
4	LA2641	29.64	10.94	106.72	140.00	0.70	96.00	7.28	1.67	1.40	1.19	65
5	LA0103	48.39	10.20	86.23	143.33	0.72	108.00	18.39	1.85	1.90	0.98	67
6	LA1272	35.96	9.39	94.50	156.67	0.78	107.00	7.73	1.60	1.90	0.86	69
7	LA1208	39.74	4.28	28.95	220.00	1.10	49.00	8.55	1.50	2.27	0.67	71
8	LA1028	46.29	5.67	28.83	93.33	0.47	171.67	8.89	2.35	2.79	0.84	79
9	LA1293	50.92	6.34	95.72	40.00	0.20	145.00	10.11	1.38	1.45	0.99	80
10	LA4115	44.96	3.95	27.94	200.00	1.00	138.33	9.39	1.58	1.67	0.95	83
11	LA4138	40.85	3.67	20.72	90.00	0.45	138.33	7.16	1.78	1.37	1.31	83
12	LA1041	44.91	3.33	30.50	130.00	0.65	68.00	7.17	1.71	1.69	1.02	84
13	LA4113	56.00	10.89	53.98	53.33	0.27	110.00	14.78	1.32	1.40	0.95	85
	Minimum	17.92	2.67	11.25	10.00	0.05	34.00	5.83	1.20	1.00	0.66	
	Maximum	62.08	15.83	205.83	500.00	2.50	176.00	29.17	2.41	2.90	1.46	
	Mean of Top 5	44.07	10.00	102.30	202.67	1.01	110.20	10.72	1.67	1.54	1.10	
	Grand mean	44.61	7.50	68.66	153.59	0.77	113.72	9.80	1.66	1.71	1.01	
	Sel. Differential (%)	-1.22	33.32	48.98	31.95	31.95	-3.09	9.41	0.78	-9.83	9.40	
	LSD (5 %)	10.28	3.69	48.89	176.83	0.88	34.05	5.47	0.37	0.57	0.21	

PH = Plant height at maturity, NFC = Number of fruits per cluster, NFP = Number of fruits per plant, FW = Fruit weight, NLM = Number of leaves at maturity NCP = Number of cluster per plant, FL = Fruit length, FC = Fruit circumference, FSI = Fruit size index. RSI = Rank Summation Index. Sel. Differential = Selection differential is estimated as a proportion (%) of mean of all accessions

Table 4. Ranking of the mean performance of nutritional and physicochemical parameters of wild tomato accessions evaluated under the rainforest agro-ecology zone in Nigeria

S/N	Accession	Vit C (mg/100 g)	TSS (^{o}Brix)	TA (g/100 g)	Fruit juice pH	Lycopene (mg/100 g)	RSI
1	LA0130	42.46	4.50	1.38	4.85	32.54	30
2	LA0411	46.28	4.50	1.26	4.9	31.02	51
3	LA1136	29.21	3.75	0.76	5.3	24.42	63
4	LA2641	32.07	3.88	0.98	4.8	15.11	65
5	LA0103	33.63	3.75	0.74	5.15	21.57	67
6	LA1272	21.02	3.75	0.95	5.00	16.13	69
7	LA1208	26.01	4.00	0.63	5.30	13.43	71
8	LA1028	34.56	4.85	0.82	5.05	22.16	79
9	LA1293	25.53	4.50	1.31	4.85	26.16	80
10	LA4115	25.38	3.75	1.02	5.10	27.65	83
11	LA4138	40.21	3.50	0.70	5.05	26.51	83
12	LA1041	31.38	3.50	0.85	5.30	31.24	84
13	LA4113	30.64	3.50	0.69	5.15	18.63	85
	Minimum	20.90	3.25	0.6	4.80	12.64	
	Maximum	46.29	5.20	1.39	5.30	32.81	
	Mean of Top 5	36.73	4.08	1.02	5.00	24.93	
	Grand mean	32.18	3.98	0.93	5.06	23.58	
	Sel. Differential (%)	14.13	2.42	10.01	-1.22	5.73	
	LSD (5 %)	2.95	0.31	0.08	0.09	1.04	

VIT C = Vitamin C, TSS = Total soluble solids, TA = Titratable acidity. RSI = Rank Summation Index. Sel. Differential = Selection differential is estimated as a proportion (%) of mean of all accessions

Table 5. Pearson's correlation coefficient (r) of agronomic, nutritional and physicochemical traits of wild tomato accessions evaluated under a rainforest agro-ecology in Nigeria

	NLM	FW (g)	Fruit yield (t/ha)	FC (cm)	NFP	FSI	VIT C (mg/100 g)	TSS (^{o}Brix)	TA (g/100 g)	pH
FW	-0.32*									
Fruit yield	-0.32*	1.00***								
FC	0.03	0.01	0.01							
NFP	-0.05	0.18	0.18	-0.34*						
FSI	0.18	0.05	0.05	-0.60***	0.18					
VIT C	0.19	0.08	0.08	-0.07	0.22	0.40*				
TSS	0.37*	0.08	0.08	0.30	0.30	-0.10	0.38*			
TA	0.24	0.15	0.15	-0.08	0.50***	0.18	0.32	0.66***		
pH	-0.35*	0.04	0.04	0.03	-0.49***	-0.21	-0.31	-0.51***	-0.76***	
Lycopene	0.27	0.15	0.15	-0.14	0.10	0.35*	0.58***	0.32*	0.57***	-0.14

*, *** Significant at 0.05 and 0.001 probability levels, respectively.
NLM = Number of leaves at maturity, FW = Fruit weight, FC = Fruit circumference, NFP = Number of fruits per plant, FSI = Fruit size index, VIT C = Vitamin C, TSS = Total soluble solids, TA = Titratable acidity, pH = Fruit juice pH.

Table 6. Eigenvalue, proportion of variability and estimated traits of wild tomato accessions contributing to first six principal components

Traits	PC1	PC2	PC3	PC4	PC5	PC6
NLM	-0.08	0.31	0.20	-0.04	0.00	-0.07
PH (cm)	-0.16	0.34	0.06	0.32	-0.07	-0.20
NCP	-0.08	0.18	-0.12	0.48	0.21	0.13
NFC	0.29	0.07	-0.28	0.12	0.24	0.15
FW (g)	0.37	-0.26	0.06	0.27	-0.06	-0.08
Fruit yield (t/ha)	0.37	-0.2	0.06	0.27	-0.06	-0.08
FL (cm)	0.03	-0.05	0.41	0.07	0.30	0.51
FC (cm)	-0.09	-0.13	0.32	0.09	0.54	0.09
NFP	0.32	0.22	-0.21	-0.05	0.06	-0.08
FSI	0.14	0.11	0.03	-0.10	-0.48	0.53
VIT C (mg/100 g)	0.16	0.25	0.22	0.16	-0.16	0.33
TSS (^{o}Brix)	0.19	0.21	0.31	-0.11	0.24	-0.29
TA (g/100 g)	0.30	0.26	0.20	-0.18	0.02	-0.24
Fruit juice pH	-0.24	-0.27	-0.10	0.29	-0.17	-0.08
Lycopene (mg/100 g)	0.13	0.19	0.34	0.12	-0.34	-0.11
Eigenvalue	4.15	3.68	2.82	2.12	1.98	1.08
Proportion (%)	22	19	15	11	10	6
Cumulative (%)	21	41	56	67	78	83

NLM = Number of leaves at maturity, PH = Plant height at maturity, NCP = Number of cluster per plant, NFC = Number of fruits per cluster, FW = Fruit weight, FL = Fruit length, FC = Fruit circumference, NFP = Number of fruits per plant, FSI = Fruit size index, VIT C = Vitamin C, TSS = Total soluble solids, TA = Titratable acidity.

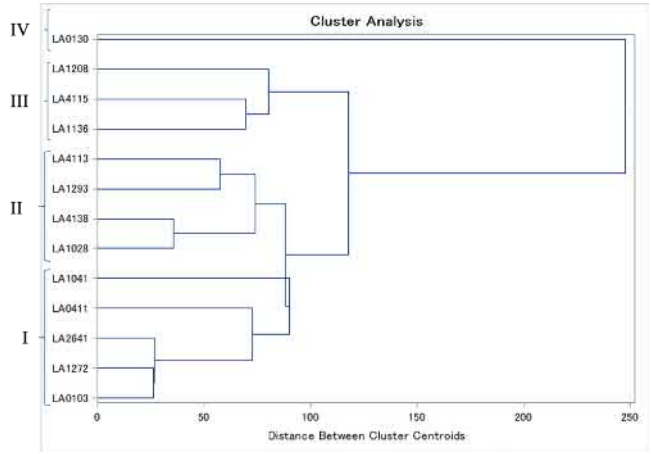

Figure 1. Dendrogram of 13 wild tomato accessions based on agronomic traits, nutritional and physicochemical parameters generated by centroid hierarchical cluster analysis

Resource Center (Chetelat, 2004, 2006). Previous research reported significant variations for agronomic traits for cultivated tomato varieties grown in various environments in Africa (Chernet and Zibelo, 2014; Shiberu, 2016; Regassa *et al*, 2016). This study indicates a wealth of genetic variability for fruit quality traits of wild tomato accessions. Accessions LA0411 and LA2641 had the highest number of fruit per plant which could be ascribed to genetic variation in flower abortion (Kanneh *et al*, 2017). Numbers of fruit per plant from this study were higher than the values reported by Ceballos-Aguirre and Vallejo-Cabrera (2012), but similar with the report of Agong *et al* (2001). The mean performances for fruits per cluster and fruit weight in our study were higher than the results presented by Ceballos-Aguirre and Vallejo-Cabrera (2012) who worked on wild tomato accessions from the Tomato Genetics Resources Center (TGRC), University of California-Davis. Disparities in the results from this study may be due to difference in the accessions evaluated, number of days before transplanting, agronomic practices used, and the environment.

Important quality traits that determine flavor, shelf life and market-related attributes of tomato are total soluble solids, fruit juice pH, titratable acidity, lycopene and Vitamin C content. The quality of tomato fruit for industrial processing and paste production depends on a high value of total soluble solids. Total soluble solids recorded in this present study ranged from 3 to 5 ^{o}Brix which is comparable to the minimum value of total soluble solids (4.5 ^{o}Brix) reported by Campos *et al* (2006) but considered low for industrial tomatoes. Previous studies have reported a range from 4 to 6 ^{o}Brix for total soluble solids of tomato fruits (Alcántar *et al*, 1999; Cramer *et al*, 2001; Pascale *et al*, 2001). High total soluble solid increases tomato paste efficiency and must be between 5.0 and 6.5 % in industrial tomatoes (Teka, 2013). The range of 4.80 – 5.30 for tomato fruit juice pH reported in this study is considerably high. Tomato fruit juice pH values can vary from 4.25 to 4.78 and fruits with high pH values may not be recommended for fresh tomato consumption or industrial processing (Paulson and Stevens, 1974; Anthon *et al*, 2011; Rajae *et al*, 2018). A pH below 4.50 is desirable because it reduces proliferation of microorganisms and indicates quality (Mohammed *et al*, 1999; Tigist *et al*, 2013). However, the pH of ripe tomatoes may exceed 4.50 because a higher pH value is associated with flavor (Stevens, 1972). Titratable acidity in this study was higher than previously reported (George *et al*, 2004; Tigist *et al*, 2013; Rajae *et al*, 2018). Tomatoes are considered the main source of lycopene compounds and a major source of carotenoids in the human diet (Willcox *et al*, 2003). Lycopene imparts the red color to tomato and affects quality. The range for lycopene content reported in the literature is between 0.58-6.50 mg 100 g^{-1} (Rickman *et al*, 2007; Saha *et al*, 2010), which is lower than reported in this study. The Vitamin C concentrations reported in this study for all accessions were higher than reported by Aoun *et al* (2013), but consistent with the range reported by Franke *et al* (2004) and Saha *et al* (2010). Our results show that wild tomato accessions contain significant antioxidants and may be useful for nutritional improvement in tomato breeding programmes (Tigchelaar, 1986) . All fruit quality and nutritional traits measured in this study reveal the value of the wild tomato accession as a source of useful alleles and their utilization as interesting donor parents in cultivar development.

Selection of the top outstanding five accessions with RSI may be useful as donor parent through intra and interspecific hybridization (Ghani *et al*, 2020) and may result in a significant increase in tomato fruit weight and fruit yield. This gain in fruit weight and yield could also be associated with improvement in number of fruits per cluster, number of fruits per plant, titratable acidity, lycopene content and fruit juice pH. To improve breeding efficiency and selection indices in crop improvement, knowledge about correlation among traits is essential (Nzuve *et al*, 2014). Results from Pearson's correlation coefficients indicate that as tomato fruit yield increases, number of leaves at maturity decreases significantly. Selection based on fruit weight and reduction in number of leaves at maturity could lead to tomato fruit yield improvement. There were significant negative correlations between fruit juice pH and titratable acidity. This implies that increased fruit juice pH was accompanied by a decrease in titratable acidity and acid concentrations and is associated with maturity (Teka, 2013). Significant positive correlations between total soluble solids and titratable acidity in this study corroborate findings of Aoun *et al* (2013), and also indicated that plants with high sugar content have more free organic acids than plants with low sugar content (Saliba-Colombani *et al*, 2001; Georgelis, 2002; Getinet *et al*, 2008). With positive correlations, genes controlling these traits could be linked to, or be under control of, pleiotropic effects (Boćanski *et al*, 2009). Positive and/or negative desirable relationships among

some agronomic, nutritional and physicochemical traits indicate that desirable genes in these wild accessions could be exploited in further breeding activities for cultivar improvement (Sujiprihati *et al*, 2003).

Furthermore, the wild tomato accessions were arranged in 4 clusters; with cluster IV appearing as the most phenotypically diverse. The best performing accession LA0130 in cluster IV has the highest number of fruits per cluster; highest fruit weight, fruit yield, vitamin C concentration, lycopene content, and moderate total soluble solids, titratable acidity and fruit juice pH. This accession might harbor novel traits that are lacking in cultivated tomato and may be used as potential parent in tomato breeding to develop high yielding cultivars with desirable nutritional and physicochemical traits. The eigenvalue from PCA indicates importance of each principal component axis and its contribution to variability in traits of the tomato accessions. Fruit size index and vitamin C concentration play a role in explaining the variation but are less important than the first four factors.

Conclusion

This study identifies variability among the 13 wild tomato accessions evaluated. Accession LA0130 was outstanding for its unique attributes which included high number of fruits per cluster, fruit weight, fruit yield, total soluble solids, titratable acidity and lycopene content amongst others. Thus, this wild tomato accession may be considered promising to broaden the genetic variability for tomato improvement programmes. Consequently, this accession may be incorporated into the tomato breeding programme in the national institutes and could be used in hybridization for developing lines with desirable horticultural traits. Documentation of the agronomic, nutritional and physicochemical performance of the evaluated wild tomato accessions is informative for their utilization in breeding programmes. These results are useful for breeders working on the development and improvement of tomato, as desirable traits from these wild tomatoes can be transferred into the commercial tomato varieties suitable for the growth conditions in the rainforest agro-ecology zone of Nigeria and to boost production and diversity.

Conflicts of Interest:

The authors declare no conflict of interest.

Acknowledgements

The authors appreciate Mrs. Bola Bello for providing technical support during planting, field management, data collection and tomato harvesting.

Author contributions

D.O.I. designed the experiment. D.O.I., A.O.K. and R.T.F. executed the experiment. A.O.K. performed the data analysis. A.O.K. and D.O.I. wrote the manuscript. All the authors contributed to writing the article, read and approved its submission.

References

Agong, S. G., Schittenhelm, S., and Friedt, W. (2001). Genotypic variation of Kenyan tomato (Lycopersicon esculentum L.) germplasm. *Journal of Food Technology in Africa* 6(1), 13–17. url: https://dx.doi.org/10.4314/jfta.v6i1.19277. doi: 10.4314/jfta.v6i1.19277

Ajayi, A. M. and Hassan, G. F. (2019). Response of selected tomato (Solanum lycopersicum L.) cultivars to on-field biotic stress. *J. Agric. Crop Res.* 7(3), 38–46. doi: https://doi.org/10.33495/jacr_v7i3.19.110

Alcántar, G. G., Villarreal, M. R., and Aguilar, A. S. (1999). Tomato growth (Lycopersion Esculentum Mill.), and nutrient utilization in response to varying fertigation programs. *Acta Horticulturae* 481, 385–392. url: https://dx.doi.org/10.17660/actahortic.1999.481.45. doi: 10.17660/actahortic.1999.481.45

Álvaro Toledo and Burlingame, B. (2006). Biodiversity and nutrition: A common path toward global food security and sustainable development. *Journal of Food Composition and Analysis* 19(6-7), 477–483. url: https://dx.doi.org/10.1016/j.jfca.2006.05.001. doi: 10.1016/j.jfca.2006.05.001

Anthon, G. E., LeStrange, M., and Barrett, D. M. (2011). Changes in pH, acids, sugars and other quality parameters during extended vine holding of ripe processing tomatoes. *Journal of the Science of Food and Agriculture* 91(7), 1175–1181. url: https://dx.doi.org/10.1002/jsfa.4312. doi: 10.1002/jsfa.4312

Aoun, A. B., Lechiheb, B., Benyahya, L., and Ferchichi, A. (2013). Evaluation of fruit quality traits of traditional varieties of tomato (Solanum lycopersicum) grown in Tunisia. *Afr. J. Food Sci.* 7(10), 350–354. doi: https://doi.org/10.5897/AJFS2013.1067

Bai, Y. and Lindhout, P. (2007). Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future? *Annals of Botany* 100(5), 1085–1094. url: https://dx.doi.org/10.1093/aob/mcm150. doi: 10.1093/aob/mcm150

Bauchet, G. and Causse, M. (2012). Genetic diversity in tomato (Solanum lycopersicum) and its wild relatives. In *Genetic diversity in plants,* ed. Çalıs 133-162.

Beckles, D. M. (2012). Factors affecting the postharvest soluble solids and sugar content of tomato (Solanum lycopersicum L.) fruit. *Postharvest Biology and Technology* 63(1), 129–140. url: https://dx.doi.org/10.1016/j.postharvbio.2011.05.016. doi: 10.1016/j.postharvbio.2011.05.016

Bergougnoux, V. (2014). The history of tomato: From domestication to biopharming. *Biotechnology Advances* 32(1), 170–189. url: https://dx.doi.org/10.1016/j.biotechadv.2013.11.003. doi: 10.1016/j.biotechadv.2013.11.003

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., van der

Knaap, E., and Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16(1), 257–257. url: https://dx.doi.org/10.1186/s12864-015-1444-1. doi: 10.1186/s12864-015-1444-1

Boćanski, J., Srećkov, Z., and Nastasić, A. (2009). Genetic and phenotypic relationship between grain yield and components of grain yield of maize (Zea mays L.). *Genetika* 41(2), 145–154. url: https://dx.doi.org/10.2298/gensr0902145b. doi: 10.2298/gensr0902145b

Borja, A., Angosto, T., Capel, J., Abad, J., Anatasio, G., and Lozano, R. (1998). Intra and inter individual variations in components of tomato flavor for fresh consumption and wild lines. In Proceedings of XI Conference on Selection and Improvement of Horticultural Plants, 49-56.

Campos, C. A. B., Fernandes, P. D., Gheyi, H. R., Blanco, F. F., Goncalves, C. B., and Campos, S. A. F. (2006). Yield and fruit quality of industrial tomato under saline irrigation. *Sci. Agr* 63, 146–152.

Causse, M., Buret, M., Robini, K., and Verschave, P. (2003). Inheritance of Nutritional and Sensory Quality Traits in Fresh Market Tomato and Relation to Consumer Preferences. *Journal of Food Science* 68(7), 2342–2350. url: https://dx.doi.org/10.1111/j.1365-2621.2003.tb05770.x. doi: 10.1111/j.1365-2621.2003.tb05770.x

Ceballos-Aguirre, N. and Vallejo-Cabrera, F. A. (2012). Evaluating the fruit production and quality of cherry tomato (Solanum lycopersicum var. cerasiforme). *Rev. Fac. Nac. Agron. Medellín* 65(2), 6593–6604. url: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0304-28472012000200004&lng=en&nrm=iso.

Chen, J., Wang, H., Shen, H., Chai, M., Li, J., Qi, M., and Yang, W. (2009). Genetic variation in tomato populations from four breeding programs revealed by single nucleotide polymorphism and simple sequence repeat markers. *Scientia Horticulturae* 122(1), 6–16. url: https://dx.doi.org/10.1016/j.scienta.2009.03.025. doi: 10.1016/j.scienta.2009.03.025

Chernet, S. and Zibelo, H. (2014). Evaluation of Tomato Varieties for Fruit Yield and Yield Components in Western Lowland of Tigray, Northern Ethiopia. *International Journal of Agricultural Research* 9(5), 259–264. url: https://dx.doi.org/10.3923/ijar.2014.259.264. doi: 10.3923/ijar.2014.259.264

Chetelat, R. T. (2004). Revised list of wild species stocks. url: https://tgrc.ucdavis.edu/Wild%20species%20stock%20list-2013-v2.pdf.

Chetelat, R. T. (2006). Revised list of miscellaneous stocks. url: https://tgrc.ucdavis.edu/Misc-stocks%20list%202015.pdf.

Chitarra, M. I. F. and Chitarra, A. B. (2005). Pós-colheita de frutas e hortaliças: fisiologia e manuseio. (Post-harvest of fruits and vegetables: physiology and handling) 785-785, (in Portuguese).

Cramer, M. D., Oberholzer, J. A., and Combrink, N. J. (2001). The effect of supplementation of root zone dissolved inorganic carbon on fruit yield and quality of tomatoes (cv 'Daniella') grown with salinity. *Scientia Horticulturae* 89(4), 269–289. url: https://dx.doi.org/10.1016/s0304-4238(00)00243-0. doi: 10.1016/s0304-4238(00)00243-0

Darwin, S. C., Knapp, S., and Peralta, I. E. (2003). Taxonomy of tomatoes in the Galápagos Islands: Native and introduced species ofSolanum-sectionLycopersicon(Solanaceae). *Systematics and Biodiversity* 1(1), 29–53. url: https://dx.doi.org/10.1017/s1477200003001026. doi: 10.1017/s1477200003001026

Domínguez, I., Ferreres, F., del Riquelme, F. P., Font, R., and Gil, M. I. (2012). Influence of preharvest application of fungicides on the postharvest quality of tomato (Solanum lycopersicum L.). *Postharvest Biology and Technology* 72, 1–10. url: https://dx.doi.org/10.1016/j.postharvbio.2012.04.010. doi: 10.1016/j.postharvbio.2012.04.010

FAO (2016). FAOSTAT. url: http://faostat.fao.org. accessed date: 2020-09-12

Franke, A. A., Custer, L. J., Arakaki, C., and Murphy, S. P. (2004). Vitamin C and flavonoid levels of fruits and vegetables consumed in Hawaii. *Journal of Food Composition and Analysis* 17(1), 1–35. url: https://dx.doi.org/10.1016/s0889-1575(03)00066-8. doi: 10.1016/s0889-1575(03)00066-8

George, B., Kaur, C., Khurdiya, D. S., and Kapoor, H. C. (2004). Antioxidants in tomato (Lycopersium esculentum) as a function of genotype. *Food Chemistry* 84(1), 45–51. url: https://dx.doi.org/10.1016/s0308-8146(03)00165-1. doi: 10.1016/s0308-8146(03)00165-1

Georgelis, N. (2002). High fruit sugar characterization, inheritance and linkage of molecular markers in tomato (Gainesville).

Getinet, H., Seyoum, T., and Woldetsadik, K. (2008). The effect of cultivar, maturity stage and storage environment on quality of tomatoes. *Journal of Food Engineering* 87(4), 467–478. url: https://dx.doi.org/10.1016/j.jfoodeng.2007.12.031. doi: 10.1016/j.jfoodeng.2007.12.031

Ghani, M. A., Abbas, M. M., Amjad, M., Ziaf, K., Ali, B., Shaheen, T., Awan, F. S., and Khan, A. N. (2020). Production and characterisation of tomato derived from interspecific hybridisation between cultivated tomato and its wild relatives. *The Journal of Horticultural Science and Biotechnology* 95(4), 506–520. url: https://dx.doi.org/10.1080/14620316.2019.1689182. doi: 10.1080/14620316.2019.1689182

Hanson, P. M., Sitathani, K., Sadashiva, A. T., yu Yang, R., Graham, E., and Ledesma, D. (2007). Performance of Solanum habrochaites LA1777 introgression line hybrids for marketable tomato fruit yield in Asia. *Euphytica* 158(1-2), 167–178. url: https://dx.doi.org/

10.1007/s10681-007-9440-4. doi: 10.1007/s10681-007-9440-4

Hanson, P. M., yu Yang, R., Wu, J., tzu Chen, J., Ledesma, D., Tsou, S. C., and Lee, T.-C. (2004). Variation for Antioxidant Activity and Antioxidants in Tomato. *Journal of the American Society for Horticultural Science* 129(5), 704–711. url: https://dx.doi.org/10.21273/jashs.129.5.0704. doi: 10.21273/jashs.129.5.0704

Ilahy, R., Hdider, C., Lenucci, M. S., Tlili, I., and Dalessandro, G. (2011). Antioxidant activity and bioactive compound changes during fruit ripening of high-lycopene tomato cultivars. *Journal of Food Composition and Analysis* 24(4-5), 588–595. url: https://dx.doi.org/10.1016/j.jfca.2010.11.003. doi: 10.1016/j.jfca.2010.11.003

International Plant Genetic Resources Institute (1996). Descriptors for tomato (Lycopersicon spp.) 1-44. url: https://www.bioversityinternational.org/e-library/publications/detail/descriptors-for-tomato-lycopersicon-spp/.

Jatoi, S. A., Fujimura, T., Yamanaka, S., Watanabe, J., and Watanabe, K. N. (2008). Potential loss of unique genetic diversity in tomato landraces by genetic colonization of modern cultivars at a non-center of origin. *Plant Breeding* 127(2), 189–196. url: https://dx.doi.org/10.1111/j.1439-0523.2007.01446.x. doi: 10.1111/j.1439-0523.2007.01446.x

Jeffers, J. N. R. (1967). Two Case Studies in the Application of Principal Component Analysis. *Applied Statistics* 16(3), 225–225. url: https://dx.doi.org/10.2307/2985919. doi: 10.2307/2985919

Kanneh, S. M., Quee, D. D., Ngegba, P. M., and Musa, P. D. (2017). Evaluation of Tomato (Solanum lycopersicum L.) Genotypes for Horticultural Characteristics on the Upland in Southern Sierra Leone. *Journal of Agricultural Science* 9(6), 213–213. url: https://dx.doi.org/10.5539/jas.v9n6p213. doi: 10.5539/jas.v9n6p213

Miller, J. C. and Tanksley, S. D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon. *Theoretical and Applied Genetics* 80(4), 437–448. url: https://dx.doi.org/10.1007/bf00226743. doi: 10.1007/bf00226743

Mohammed, M., Wilson, L. A., and Gomes, P. I. (1999). Postharvest sensory and physiochemical attributes of processing and nonprocessing tomato cultivars. *Journal of Food Quality* 22(2), 167–182. url: https://dx.doi.org/10.1111/j.1745-4557.1999.tb00549.x. doi: 10.1111/j.1745-4557.1999.tb00549.x

Mulumba, N. N. and Mock, J. J. (1978). Improvement of yield potential in the Eto Blanco maize (Zea mays L.) population by breeding for plant traits. *Egypt. J. Genet. Cytol* 7, 40–51.

Nnabude, P. C., Nweke, I. A., and Nsoanya, L. N. (2015). Response of three varieties of tomatoes (Lycopersicon esculentum) to liquid organic fertilizer (alfa life) and inorganic fertilizer (NPK 20:10:10) and for soil improvements. *European Journal of Physical and Agricultural Sciences* 3, 28–37.

Nowicki, M., Kozik, E. U., and Foolad, M. R. (2013). Late blight of tomato. In *Translational Genomics for Crop Breeding,* ed. Varshney, R. K. and Tuberosa, R. (John Wiley and Sons Ltd), 241-265.

Nzuve, F., Githiri, S., Mukunya, D. M., and Gethi, J. (2014). Genetic Variability and Correlation Studies of Grain Yield and Related Agronomic Traits in Maize. *Journal of Agricultural Science* 6(9), 166–166. url: https://dx.doi.org/10.5539/jas.v6n9p166. doi: 10.5539/jas.v6n9p166

Olaniyi, J. O., Akanbi, W. B., Adejumo, T. A., and Akande, O. G. (2010). Growth, fruit yield and nutritional quality of tomato varieties. *Afr. J. Food Sci* 4, 398–402.

Onyia, V. N., Chukwudi, U. P., Ogwudu, V. C., Atugwu, A. I., Eze1, S. C., Ene, C. O., and Umeh, S. (2019). Evaluation of Tomato Genotypes Growth, Yield, and Shelf Life Enhancement in Nigeria. *J. Agr. Sci .Tech* 21, 143–152. url: https://jast.modares.ac.ir/article-23-16444-en.pdf.

Pascale, S. D., Maggio, A., Fogliano, V., Ambrosino, P., and Ritieni, A. (2001). Irrigation with saline water improves carotenoids content and antioxidant activity of tomato. *The Journal of Horticultural Science and Biotechnology* 76(4), 447–453. url: https://dx.doi.org/10.1080/14620316.2001.11511392. doi: 10.1080/14620316.2001.11511392

Paulson, K. N. and Stevens, M. A. (1974). Relationships among titratable acidity, pH and buffer composition of tomato fruits. *Journal of Food Science* 39(2), 354–357. url: https://dx.doi.org/10.1111/j.1365-2621.1974.tb02893.x. doi: 10.1111/j.1365-2621.1974.tb02893.x

Peixoto, J. V. M., Garcia, L. G. C., Nascimento, A. R., Moraes, E. R., Ferreira, T. A. P., Fernandes, M. R., and Pereira, V. D. A. (2018). Postharvest evaluation of tomato genotypes with dual purpose. *Food Sci. Technol.* 38, 255–262. doi: http://dx.doi.org/10.1590/1678-457x.00217

Peralta, I. E., Spooner, D. M., Knapp, S., and Anderson, C. (2008). Taxonomy of wild tomatoes and their relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). *Syst. Bot. Monogr* 84(2), 1–186.

Rajae, A., Addi, M., Mingeot, D., Elamrani, A., Serghini, H. C., Mihamou, A., and Abid, M. (2018). Morphological, molecular, and physic-chemical characterization of traditional Moroccan tomato (Solanum Lycopersicum L.) genotypes. *J. Biotech Res* 9, 58–69. url: http://www.btsjournals.com/assets/2018v9p58-69.pdf.

Regassa, D., Tigre, W., and Shiferaw, A. (2016). Tomato (Lycopersicon esculentum Mill.) varieties evaluation in Borana zone, Yabello district, Southern Ethiopia. *Journal of Plant Breeding and Crop Science* 8(10), 206–210. doi: https://doi.org/10.5897/JPBCS2015.0543

Riahi, A., Hdider, C., Sanaa, M., Tarchoun, N., Kheder, M. B., and Guezal, I. (2009). Effect of conventional

and organic production systems on the yield and quality of field tomato cultivars grown in Tunisia. *Journal of the Science of Food and Agriculture* 89(13), 2275–2282. url: https://dx.doi.org/10.1002/jsfa.3720. doi: 10.1002/jsfa.3720

Rick, C. M. and Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club* 102(6), 376–376. url: https://dx.doi.org/10.2307/2484764. doi: 10.2307/2484764

Rickman, J. C., Barrett, D. M., and Bruhn, C. M. (2007). Nutritional comparison of fresh, frozen and canned fruits and vegetables. Part 1. Vitamins C and B and phenolic compounds. *Journal of the Science of Food and Agriculture* 87(6), 930–944. url: https://dx.doi.org/10.1002/jsfa.2825. doi: 10.1002/jsfa.2825

Rosales, M. A., Cervilla, L. M., Sánchez-Rodríguez, E., del Mar Rubio-Wilhelmi, M., Blasco, B., Ríos, J. J., Soriano, T., Castilla, N., Romero, L., and Ruiz, J. M. (2011). The effect of environmental conditions on nutritional quality of cherry tomato fruits: evaluation of two experimental Mediterranean greenhouses. *Journal of the Science of Food and Agriculture* 91(1), 152–162. url: https://dx.doi.org/10.1002/jsfa.4166. doi: 10.1002/jsfa.4166

Rosales, M. A., Ruiz, J. M., Hernández, J., Soriano, T., Castilla, N., and Romero, L. (2006). Antioxidant content and ascorbate metabolism in cherry tomato exocarp in relation to temperature and solar radiation. *Journal of the Science of Food and Agriculture* 86(10), 1545–1551. url: https://dx.doi.org/10.1002/jsfa.2546. doi: 10.1002/jsfa.2546

Saha, S., Hedau, N. K., Mahajan, V., Singh, G., Gupta, H. S., and Gahalain, A. (2010). Textural, nutritional and functional attributes in tomato genotypes for breeding better quality varieties. *Journal of the Science of Food and Agriculture* 90(2), 239–244. url: https://dx.doi.org/10.1002/jsfa.3802. doi: 10.1002/jsfa.3802

Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., and Buret, M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics* 102(2-3), 259–272. url: https://dx.doi.org/10.1007/s001220051643. doi: 10.1007/s001220051643

SAS Institute (2010). Statistical Analysis Software (SAS) Release 9.3. url: https://support.sas.com/documentation/93/.

Shiberu, T. (2016). Evaluation of improved tomato varieties (Lycopersicon esculentum Mill.) performance against major insect pests under open field and glasshouse conditions. *International Journal of Research Studies in Agricultural Sciences* 2(3), 1–7. doi: 10.20431/2454-6224.0203001

Singh, R. J. (2006). Genetic Resources, Chromosome Engineering, and Crop Improvement: Vegetable Crops, ed. Singh, R. J., volume 3 (CRC Press), 1-352.

Sneath, P. H. A. and Sokal, R. R. (1973). Numerical taxonomy. The principles and practice of numerical classification (San Francisco, CA: W.H Freeman), 1-537.

Stevens, M. A. (1972). Relationships between components contributing to quality variation among tomato lines. *J. Am. Soc. Hortic. Sci* 97, 70–73.

Sujiprihati, S., Saleh, G. B., and Ali, E. S. (2003). Heritability, performance and correlation studies on single cross hybrids of tropical maize. *Asian J. Plant Sci* 2(1), 51–57. doi: 10.3923/ajps.2003.51.57

Teka, T. A. (2013). Analysis of the effect of maturity stage on the postharvest biochemical quality characteristics of tomato (Lycopersicon esculentum Mill.) fruit. *Int. Res J Pharm. App Sci.* 3(5), 180–186.

Tigchelaar, E. C. (1986). Tomato Breeding, ed. Basset, J., M., Ed., and Intercept, B. V. C. (Andover) 135-171.

Tigist, M., Workneh, T. S., and Woldetsadik, K. (2013). Effects of variety on the quality of tomato stored under ambient conditions. *Journal of Food Science and Technology* 50(3), 477–486. doi: 10.1007/s13197-011-0378-0

Willcox, J. K., Catignani, G. L., and Lazarus, S. (2003). Tomatoes and Cardiovascular Health. *Critical Reviews in Food Science and Nutrition* 43(1), 1–18. url: https://dx.doi.org/10.1080/10408690390826437. doi: 10.1080/10408690390826437

Zuriaga, E., Blanca, J., and Nuez, F. (2009). Classification and phylogenetic relationships in Solanum section Lycopersicon based on AFLP and two nuclear gene sequences. *Genetic Resources and Crop Evolution* 56(5), 663–678. url: https://dx.doi.org/10.1007/s10722-008-9392-0. doi: 10.1007/s10722-008-9392-0

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2020), 1 (2), 12–22
DOI: 10.46265/genresj.SHBD3744
https://www.genresj.org
ISSN: 2708-3764

Phenotypic characterization of sheep populations in Tahtay Maichew district, Northern Ethiopia

Abebe Hailu[a], **Amine Mustefa** *,[a], **Tesfalem Aseged**[a], **Abraham Assefa**[a], **Seble Sinkie**[a] **and Semere Tsewene**[b]

[a] *Ethiopian Biodiversity Institute, P.O.Box 30726, Addis Ababa, Ethiopia*
[b] *Ethiopian Biodiversity Institute, Mekelle Biodiversity Center, P.O.Box 30726, Mekelle, Ethiopia*

Abstract: Eighteen quantitative measurements and fourteen qualitative characteristics taken from 306 adult sheep (57 rams and 249 ewes) were used to phenotypically characterize sheep populations of Tahtay Maichew district, Ethiopia. Most traits showed significant variation by agro-ecological zone, sex and age groups with higher values generally recorded for rams as compared to ewes. Middle age group animals displayed highest values for several traits, reflecting the optimal production age. Agro-ecological zone affected ewes more than rams. The highland sheep had shortest height at withers, widest shoulder points and longest hair, indicative of adaptation to their environment. Qualitative characteristics of the studied sheep populations such as tail shape, plain coat color pattern, unpigmented skin, hairy fiber and the absence of horn, toggle, ruff and beard suggest that they constitute a previously undescribed sheep breed. Tan coat color differentiated high and midland sheep from lowland sheep where white and brown colors were dominant. Canon bone length, height at withers and tail length were the three most important morphometric variables used in discriminating the sheep populations. On average 66% of the animals could be classified into their respective agro-ecological zone. Our data suggest that highland sheep populations are distantly related to lowland sheep, while midland sheep are more closely related to lowland sheep. It can be concluded that breeding programs specific to each agro-ecological zone need to be designed for sustainable utilization and conservation of the studied sheep populations. Furthermore, molecular based studies might allow further characterization of Ethiopian sheep breeds.

Keywords: Indigenous breed, Sheep genetic resources, Morphological characterization, Qualitative traits, Tigray region, Ethiopia

Citation: Hailu, A., Mustefa, A., Aseged, T., Assefa, A., Sinkie, S., Tsewene, S. (2020). **Phenotypic characterization of sheep populations in Tahtay Maichew district, Northern Ethiopia**. *Genetic Resources* 1 (2), 12–22. doi: 10.46265/genresj.SHBD3744.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

The Ethiopian livestock sector, which is mainly dominated by indigenous animal genetic resources, contributes significantly to the economy and food security of the country (Central Statistical Agency, 2018), providing livelihood for 37-87% of the country's population (Central Statistical Agency, 2005). This sector contributes 15 — 17% and 35 — 49% of the total and agricultural Gross Domestic Product, respectively (Michael *et al*, 2016). Within the livestock sector, small ruminants, especially sheep, provide a sustainable option for smallholder low input-output production systems. Indigenous sheep genetic resources play a major role in developing countries like Ethiopia, as are better adapted to environments which are harsh, marginal and degraded, have low body weight and excellent grazing skills (Misra and Singh, 2002; Degen, 2007). The indigenous sheep genetic resources account for 99.81% of the total sheep population in Ethiopia (Central Statistical Agency, 2018).

Conducting phenotypic characterization is a prerequisite for sustainable utilization, conservation and improvement of a breed through designing appropriate sheep breeding programs (FAO, 2012). This will fur-

*Corresponding author: Amine Mustefa (aminemustefa32@gmail.com)

Received: 04.06.2020 Accepted: 20.10.2020 Published online: 21.12.2020

ther maximize sustainable food security while minimizing pressure on the environment. Ethiopia, one of the major gateways for domestic sheep to Africa (Devendra and Mcleroy, 1982), is believed to have the largest livestock population in Africa with 31.3 million sheep (Central Statistical Agency, 2018) categorized into 14 traditionally recognized and phenotypically distinct sheep populations (9 breeds within 6 breed groups) (Gizaw *et al*, 2008; EBI, 2016). The 9 sheep breeds of Ethiopia are Simien, Short fat tailed, Washera, Gumz, Horro, Arsi, Bonga, Afar and Black Head Somali (BHS) (Gizaw *et al*, 2008). In literature, sheep populations of the current study area were generally classified as Sekota traditional population and further categorized under the Short fat tailed breed (Gizaw *et al*, 2008). However, samples were not taken from nearby areas of the current study areas.

Sekota sheep populations were characterized as short fat tail turned-up at end and fused with main part. The population is medium-sized, predominantly brown or white coat color, few blacks with brown belly. The white animals have finer hair or wooly udder-coat, semi-pendulous or rudimentary ears in Wag Himra and Tigray while predominantly rudimentary in Tekeze valley. Sekota sheep population were reared by Agew, Tigray and Amhara communities (Gizaw *et al*, 2008).

However, due to the country's high ecological and production system variations, some of the breeds were re-characterized in more recent studies, including Simien sheep (Melaku *et al*, 2019), Short fat tailed sheep (Hayelom *et al*, 2014; Bimerow *et al*, 2011; Getachew *et al*, 2009), Washera sheep (Mengistie *et al*, 2010), Arsi sheep (Worku, 2018); Afar sheep (Getachew *et al*, 2009). In addition to these studies, some work was done in the Tigray region on Abergelle sheep (Tajebe *et al*, 2011) and Tigray Highland sheep (Gebreyowhens and Tesfay, 2016).

Despite the efforts made to characterize the Ethiopian indigenous sheep genetic resources as mentioned above, they have not yet been exhaustive in covering all regions of the country in general and the Tigray region in particular. They also focused on only a few specifically well-known sheep populations. As a consequence, a high sheep diversity remains unstudied, along with the associated diversified ecology, production systems and ethnic groups. Therefore, there is an urgent need for continued characterization and identification to understand the relationships within and among breeds. Thus, the present study was initiated to cover these gaps and phenotypically describe the indigenous sheep populations of Tahtay Maichew district, Central Tigray zone, Ethiopia.

Materials and methods

Study Areas

The study was carried out in the Tahtay Maichew district, which is located in the central zone of Tigray National Regional State (Figure 1). The district covers a total area of 18,618 km^2 with estimated livestock population size of 247,907, consisting 75,707 cattle, 55,517 goats, 110 mules, 6,716 donkeys, 25,195 sheep and 84,102 poultry (Atsbeha *et al*, 2015). The studied areas were categorized as highland, midland and lowland based on the climatic factors in Table 1.

Table 1. Climatic factors and sheep population size of the threeagro-ecological zones in the Tahtay Maichew district of Tigray region, Ethiopia.

Variables	Highland	Midland	Lowland
Altitude in meters	> 2500	1500 – 2500	< 1500
Temperature in °C (mean)	9.9	19.9	30.3
Annual rainfall in mm (mean)	600 – 700	500 – 600	400 – 500
Sheep population size	11,816	8,903	4,476

Site selection and data collection

Available background information on the existence of unstudied sheep populations adapted to different agro-ecological zones was captured through short pilot survey and focus group discussions with livestock experts and keepers. Two sites (kebeles) were sampled randomly from each agro-ecological zone (see Table 1). Quantitative and qualitative data were recorded from a total of 306 adult sheep (57 rams and 249 ewes) based on data collection procedures outlined in FAO guidelines (FAO, 2012). Studied animals were carefully handled by trained personnel. Quantitative measurements were taken early in the morning of the day before feeding and watering when the animals were calm and standing in an upright position on flat ground.

Eighteen quantitative measurements were collected: body length (cm), body weight (kg), heart girth (cm), height at withers (cm), chest depth (cm), shoulder point width (cm), subs height (cm) rump length (cm), rump width (cm), tail length (cm), tail width (cm), head length (cm), head width (cm), shin circumference (cm), horn length (cm), hair length (cm), canon bone length (cm), ear length (cm), testis circumference (cm). Fourteen qualitative characteristics were also collected: coat color pattern, coat color, skin color, fiber type, ear orientation, head profile, back profile, rump profile, tail type, tail shape and presence of toggle, horn, beard and ruff were recorded.

Statistical Analysis

Data were entered and managed using Microsoft Excel© worksheet. Detection of outliers and testing the normality of the quantitative measurements data was performed using the UNIVARIATE procedure of Statistical Analysis Software (SAS) 9.0 (SAS Institute, 2002). Analysis of data on quantitative measurements and qualitative characteristics was carried out using the General Linear Model (GLM) procedure and the frequency (FREQ) procedure of SAS 9.0 software,

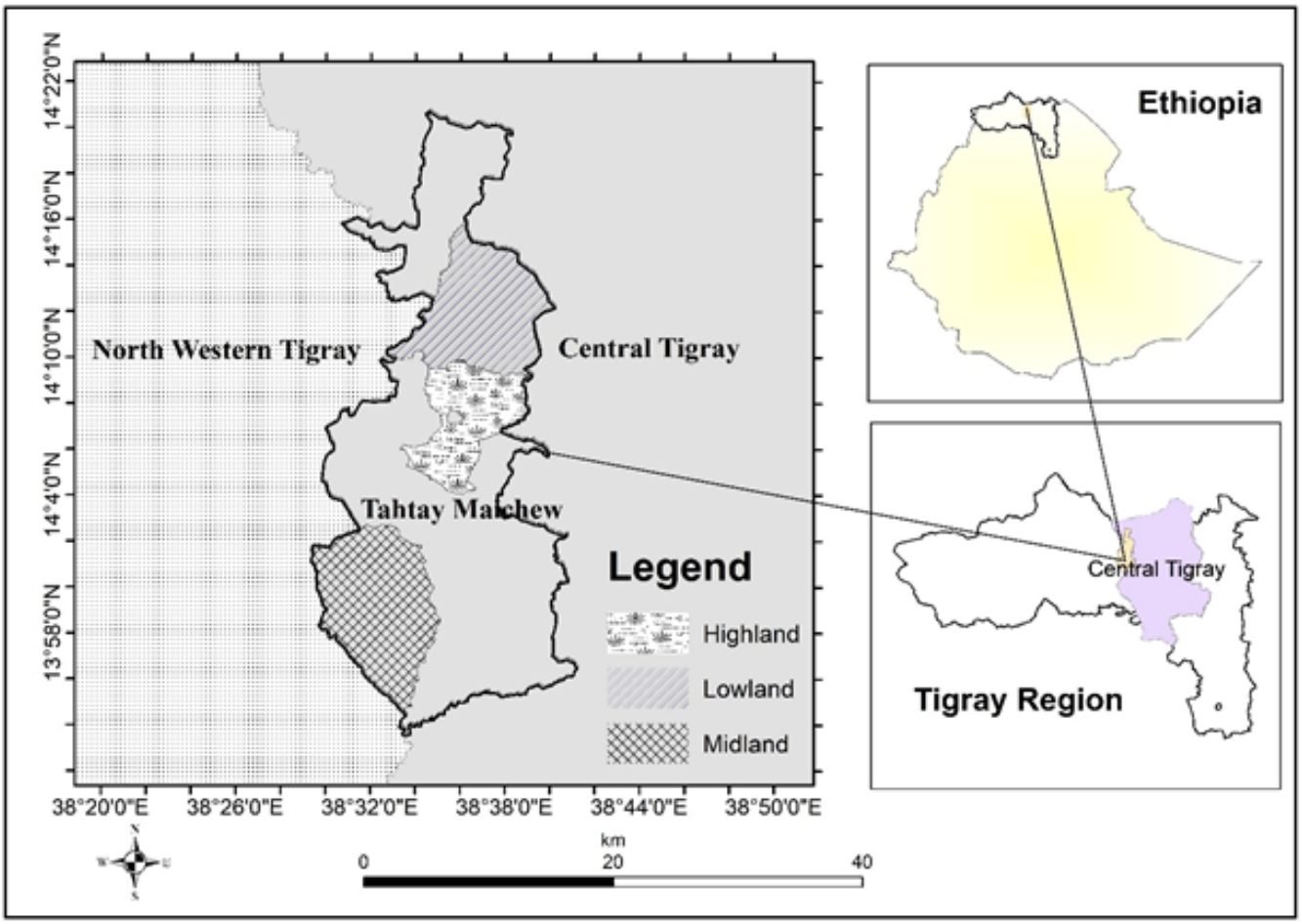

Figure 1. Map of the studied areas.

respectively. Linear measurements Least Square Means (LSM) were separated using the adjusted Tukey-Kramer test (SAS Institute, 2002). Quantitative and qualitative data were analyzed using the following model:$Y_{ijk} = \mu + A_i + B_j + C_k + e_{ijk}$ where Y_{ijk} is an observation, μ is the overall mean, A_i is the fixed effect of environment, B_j is the fixed effect of the sex, C_k is the fixed effect of age group and e_{ijk} is the random error attributed to the nth observation. Environment, sex and age group were fitted as class variables throughout the analysis, while sex effect was removed from the class variables when the analysis was done separately for each sex.

Multivariate analysis was performed on quantitative measurements. Stepwise discriminant function analysis (STEPDISC) with forward selection procedure was used to find out the quantitative variables that better discriminate populations from different environment. Percentage assignment of observations to environment and probabilities of misclassifications were evaluated by discriminant function analysis (DISCRIM). Canonical discriminant function analysis (CANDISC) was also performed to find out linear combination of quantitative variables that provide maximal separations between environments. The scored canonical variables were used to plot pairs of canonical variables to get visual interpretation of environmental differences. Pairwise squared Mahalanobis distances between environments were computed as: $D^2\,(i|j) = (x_i - x_j)^{'} cov^{-1}\ (x_i - x_j)$. Where $D^2\,(i|j)$ is the distances between environments zones i and j, cov^{-1} is the inverse of the covariance matrix of measured variables, x_i and x_j are the means of variables in the i^{th} and j^{th} populations.

Results

Quantitative measurements

Level of significance (P-values) outputs of the class variables for both the overall analysis and separately for each sex are presented in Table 2. Overall, most of the studied traits were significantly affected by agro-ecological zone, age and sexual differences. Effect of agro-ecological zone on some quantitative traits (heart girth, height at withers, shoulder point width, rump length, tail length, head length, hair length, canon bone length, and ear length) was more significant on ewes than rams.

The overall mean with the respective standard error and deviation, and the effect of agro-ecological zone, sex and age on the quantitative measurements are presented in Tables 3, 4 and 5. The highland sheep population had the shortest height at withers, widest shoulder points and longest hair, while midland sheep population

Table 2. Level of significance for the overall analysis and separately for both sexes

Traits	Overall			Rams		Ewes	
	Agro-eco zone	Age	Sex	Agro-eco zone	Age	Agro-eco zone	Age
BL (cm)	0.2688	<0.0001	0.3394	0.3930	0.0028	0.3873	0.0004
BW (kg)	0.5309	<0.0001	0.9515	0.3384	0.0061	0.6135	<0.0001
HG (cm)	0.0330	<0.0001	0.6983	0.0772	0.0002	0.0399	<0.0001
HAW (cm)	0.0005	0.0250	0.0617	0.0562	0.0009	0.0075	0.5451
CD (cm)	0.0486	0.0013	0.0267	0.1216	0.0120	0.2068	0.0397
SPW (cm)	0.0012	<0.0001	0.0017	0.8511	0.0008	<0.0001	0.0078
SH (cm)	0.5553	0.5924	0.2194	0.1241	0.9660	0.8881	0.7606
RL (cm)	<0.0001	0.2364	0.2377	0.8307	0.8157	<0.0001	0.1604
RW (cm)	0.7008	0.0028	0.4727	0.2884	0.0001	0.3402	0.0504
TL (cm)	0.0003	0.9584	0.0040	0.1598	0.2868	0.0007	0.2635
TW (cm)	0.3219	0.3538	<0.0001	0.5269	0.0064	0.2726	0.8968
HL (cm)	<0.0001	0.0002	0.0410	0.2124	0.0213	<0.0001	0.0062
HW (cm)	0.0283	0.0092	<0.0001	0.2161	0.9169	0.0597	0.0122
SC (cm)	0.4564	0.0003	<0.0001	0.5994	0.0001	0.6909	0.2794
HRL (cm)	<0.0001	0.0076	0.0007	0.1643	0.3147	<0.0001	0.0123
CBL (cm)	<0.0001	0.0377	0.2666	0.6187	0.0349	<0.0001	0.1417
EL (cm)	<0.0001	0.0365	0.6856	0.5261	0.4147	0.0002	0.0153
TC (cm)	-	-	-	0.0253	0.0114	-	-

BL = body length, BW = body weight, HG = heart girth, HAW = height at withers, CD = chest depth, SPW = shoulder point width, SH = subs height, RL =rump length, RW = rump width, TL = tail length, TW = tail width, HL =head length, HW = head width, SC = shin circumference, HRL = hair length, CBL = canon bone length, EL = ear length, TC = testis circumference.

possessed the shortest tail and ear. Almost half of the measured traits were affected by sex of the animals showing higher values for males.

Most of the overall agro-ecological zone differences were due to the differences within the ewes. However, testis circumference, the only trait among the rams which is affected by agro-ecological zone, increases significantly as we shift from highland to lowland.

The majority of the quantitative measurements were significantly affected by the age of the animals (Figures 2, 3 and 4). Accordingly, values of some traits (body weight, chest depth, shoulder point width, rump width, testis circumference, body length, heart girth, and height at withers) gradually increased towards the optimum age of three years and then decreased towards the oldest age (5 years; Figures 2 and 3). However, this was not true in some traits (head length, head width, shin circumference, hair length, canon bone length, and ear length; Figure 4). On the other hand, age did not affect subs height, rump length, rump width, tail length, and tail width.

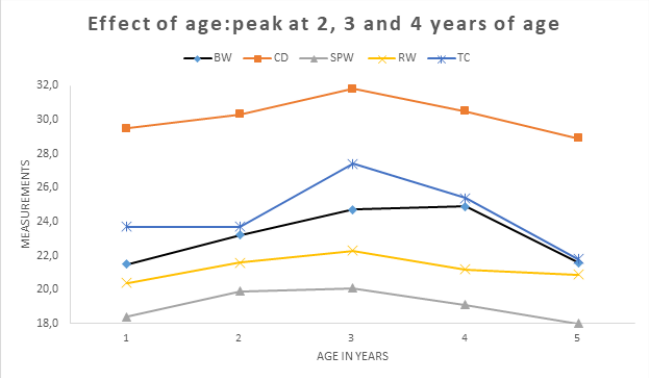

Figure 2. Effect of age on body weight(BW), chest depth (CD), shoulder point width (SPW), rump width (RW), and testis circumference (TC).

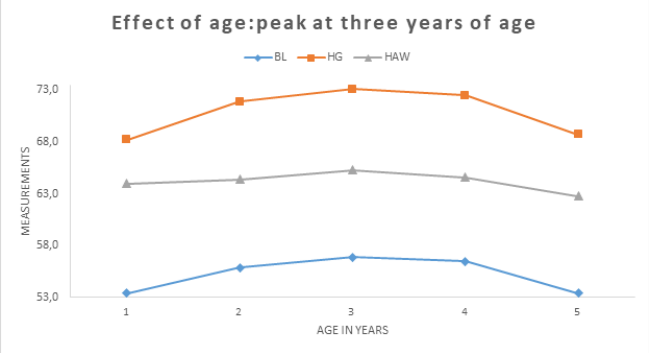

Figure 3. Effect of age on body length (BL), heart girth (HG), and height at withers (HAW)

Qualitative characteristics

The outputs of the chi-squared tests, if the qualitative characteristics of the sheep populations from the three agro-ecological zones differ, are presented in Table 5. Accordingly, ear orientation, back profile, head profile, rump profile and coat color of the three agro-ecological zones were significantly different (Table 5). On the other hand, the coat color pattern, fiber type, skin color, tail type, tail shape and presence of toggle, horn, beard

Table 3. Overall mean ($\overline{X}$), standard error (SE), standard deviation (SD) and pairwise mean comparison (least square means and standard errors) for the effect of agro-ecological zone and sex. Means within a column bearing different superscripts are significantly different; a is given to the highest value.

Traits	Overall		Agro-ecological zone			Sex	
	$\overline{X}$±SE	SD	Highland	Midland	Lowland	Rams	Ewes
N	306		102	102	102	57	249
BL (cm)	55.2±0.2	4.0	54.7±0.5	55.4±0.4	55.5±0.4	55.6±0.6	54.9±0.3
BW (kg)	23.5±0.2	3.4	22.9±0.4	23.4±0.3	23.3±0.4	23.2±0.5	23.2±0.2
HG (cm)	71.0±0.3	4.9	71.8±0.5^{a}	70.1±0.5^{b}	70.7±0.5^{ab}	71.0±0.7	70.7±0.3
HAW (cm)	63.9±0.2	3.2	63.2±0.4^{b}	64.5±0.3^{a}	64.9±0.4^{a}	64.8±0.5	63.7±0.2
CD (cm)	29.8±0.2	3.1	30.0±0.4	30.0±0.3	30.8±0.3	30.9±0.5	29.6±0.2
SPW (cm)	18.7±0.1	2.3	19.7±0.3^{a}	18.6±0.3^{b}	19.0±0.3^{b}	19.8±0.3	18.4±0.2
SH (cm)	32.9±0.1	2.5	32.8±0.3	33.1±0.3	33.1±0.3	33.3±0.4	32.7±0.2
RL (cm)	14.9±0.1	2.2	15.6±0.3^{a}	15.2±0.2^{a}	14.1±0.2^{b}	15.2±0.3	14.7±0.2
RW (cm)	21.3±0.1	2.3	21.2±0.3	21.4±0.2	21.2±0.3	21.1±0.4	21.4±0.2
TL (cm)	16.8±0.2	3.1	17.5±0.4^{a}	16.3±0.3^{b}	18.0±0.3^{a}	18.1±0.5	16.4±0.2
TW (cm)	16.8±0.2	3.0	18.0±0.3	17.5±0.3	17.5±0.3	19.0±0.5	16.3±0.2
HL (cm)	13.9±0.1	1.5	14.4±0.2^{a}	14.0±0.2^{a}	13.4±0.2^{b}	14.2±0.2	13.6±0.1
HW (cm)	10.2±0.1	1.3	10.4±0.1^{b}	10.8±0.1^{a}	10.4±0.1^{b}	11.1±0.2	9.9±0.1
SC (cm)	6.8±0.1	0.8	7.0±0.1	7.1±0.1	6.9±0.1	7.3±0.1	6.6±0.1
HRL (cm)	4.5±0.1	1.3	5.2±0.1^{a}	4.6±0.1^{b}	4.5±0.1^{b}	5.2±0.2	4.4±0.1
CBL (cm)	12.3±0.1	1.2	12.7±0.1^{a}	12.5±0.1^{a}	11.8±0.1^{b}	12.5±0.2	12.2±0.1
EL (cm)	7.6±0.2	3.0	7.8±0.4^{a}	6.4±0.3^{b}	8.4±0.4^{a}	7.7±0.5	7.4±0.2

N = number of observations, BL = body length, BW = body weight, HG = heart girth, HAW = height at withers, CD = chest depth, SPW = shoulder point width, SH = subs height, RL =rump length, RW = rump width, TL = tail length, TW = tail width, HL =head length, HW = head width, SC = shin circumference, HRL = hair length, CBL = canon bone length, EL = ear length.

Table 4. Pairwise mean comparison (least square means and standard errors) for the effect of agro-ecological zonewithin each sex. Means within a column bearing different superscripts are significantly different; a is given to the highest value.

Traits	Rams			Ewes		
	Highland	Midland	Lowland	Highland	Midland	Lowland
N	14	26	17	88	76	85
BL (cm)	56.3±1.3	56.1±0.9	57.6±1.0	54.7±0.5	55.5±0.5	55.3±0.5
BW (kg)	23.2±1.1	23.7±0.7	24.8±0.8	23.3±0.4	23.7±0.4	23.3±0.4
HG (cm)	73.2±1.5	70.6±1.0	73.3±1.2	72.2±0.6^{a}	70.6±0.6^{b}	70.6±0.6^{b}
HAW (cm)	65.0±1.1	65.7±0.7	67.6±0.8	63.1±0.4^{b}	64.4±0.4^{a}	64.4±0.4^{a}
CD (cm)	30.2±1.2	30.8±0.8	32.5±0.9	29.7±0.3	29.4±0.4	30.2±0.4
SPW (cm)	20.0±0.9	19.9±0.6	20.3±0.7	19.6±0.2^{a}	18.1±0.3^{b}	18.5±0.2^{b}
SH (cm)	32.0±1.0	34.0±0.6	33.3±0.7	32.7±0.3	32.7±0.3	32.8±0.3
RL (cm)	15.3±0.8	14.8±0.5	14.9±0.6	15.5±0.3^{a}	15.2±0.3^{a}	13.7±0.3^{b}
RW (cm)	22.6±0.8	21.4±0.5	22.1±0.6	21.6±0.3	21.9±0.3	21.4±0.3
TL (cm)	19.2±1.2	17.2±0.8	18.7±0.9	16.9±0.3^{a}	15.7±0.4^{b}	17.4±0.3^{a}
TW (cm)	19.9±1.1	18.7±0.7	19.3±0.8	16.9±0.3	16.4±0.4	16.3±0.3
HL (cm)	15.2±0.6	14.6±0.4	14.1±0.4	14.2±0.2^{a}	13.8±0.2^{a}	13.2±0.2^{b}
HW (cm)	10.4±0.6	11.4±0.4	11.2±0.5	9.8±0.1	10.2±0.1	9.7±0.1
SC (cm)	7.4±0.3	7.7±0.2	7.5±0.3	6.7±0.1	6.7±0.1	6.6±0.1
HRL (cm)	4.9±0.5	5.4±0.3	4.5±0.4	5.0±0.1^{a}	4.1±0.2^{b}	4.2±0.1^{b}
CBL (cm)	12.6±0.4	12.9±0.2	12.6±0.3	12.8±0.1^{a}	12.4±0.1^{a}	11.7±0.1^{b}
EL (cm)	7.6±1.1	7.2±0.7	8.3±0.8	7.5±0.4^{a}	6.1±0.4^{b}	8.2±0.4^{a}
TC (cm)	23.2±1.1^{b}	24.0±0.7^{ab}	26.1±0.8^{a}	-	-	-

N = number of observations, BL = body length, BW = body weight, HG = heart girth, HAW = height at withers, CD = chest depth, SPW = shoulder point width, SH = subs height, RL = rump length, RW = rump width, TL = tail length, TW = tail width, HL = head length, HW = head width, SC = shin circumference, HRL = hair length, CBL = canon bone length, EL = ear length, TC = testis circumference.

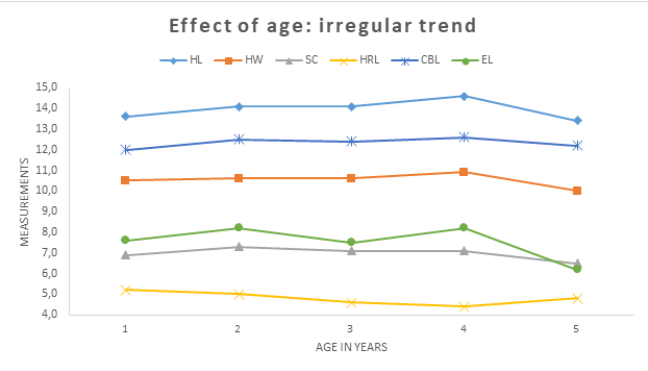

Figure 4. Effect of age on head length(HL), head width (HW), shin circumference (SC), hair length (HRL), canon bone length (CBL), and ear length (EL)

and ruff were not significantly different among the studied agro-ecological zones. Accordingly, the sheep populations can be characterized as hair type sheep with plain coat color pattern. Additionally, all of the studied sheep populations possess cylindrical thin tail with turned up at end, and straight head profile. The results also revealed that almost none of the sheep sampled have pigmented skin, horns, toggle, ruff and beard. Tan coat color was dominantly observed in the high and midland agro-ecological zones while white and brown colors were dominant in the lowland sheep (Figures 5, 6 and 7).

Multivariate analysis for discrimination of sheep populations

According to stepwise discriminant function analysis, canon bone length, height at withers and tail length were the three most important morphometric variables used in discriminating the sheep populations from different agro-ecological zones (Table 6). Chest depth, subs height, body length, and tail width were found not to be useful variables due to their lowest discriminatory power (Table 6).

The probabilities of all main multivariate tests over the canonical structures were significant (Table 7).

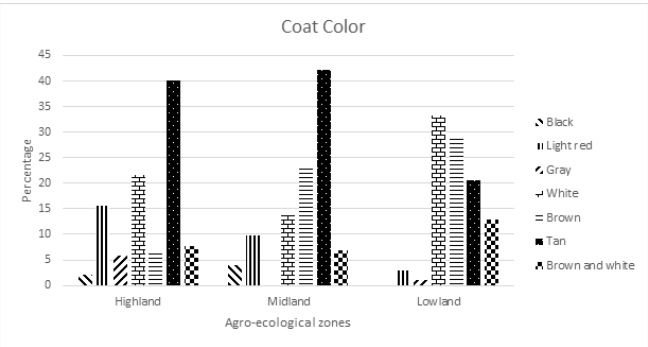

Figure 5. Coat color of sheep populations from different agro-ecological zones; significant ($p<0.0001$) effect of agro-ecological zones was observed over the coat color of the studied sheep populations.

Table 5. Percentages and their respective chi-squared probabilities of some qualitative characteristics of the sheep populations from different agro-ecological zones. χ^2values in bold are statistically significant.

Qualitative traits	Agro-ecological zones			χ^2
	Highland	Midland	Lowland	
Coat color pattern				0.1445
Plain	77.5	85.3	82.4	
Patchy	17.6	12.7	17.6	
Spotted	4.9	2.0	NR	
Fiber type				0.8211
Hairy	85.3	86.3	88.2	
Wooly	14.7	13.7	11.8	
Ear orientation				**0.0127**
Erect	1.0	1.0	4.9	
Semi-pendulous	48.0	34.3	37.2	
Pendulous	25.5	16.7	21.6	
Carried horizontal	25.5	48.0	36.3	
Head profile				**0.0100**
Straight	71.6	65.7	85.3	
Concave	12.7	11.8	8.8	
Convex	15.7	22.5	5.9	
Back profile				**0.0002**
Straight	70.5	40.2	62.8	
Slopes up towards the rump	25.5	56.9	34.3	
Slopes down from withers	2.0	2.9	2.9	
Curved	2.0	0	0	
Rump profile				**<0.0001**
Flat	27.5	59.8	50.0	
Sloping	72.5	40.2	48.0	
Roofy	0	0	2.0	
Toggle				0.0987
Present	3.9	11.8	10.8	
Absent	96.1	88.2	89.2	

Canonical correlation coefficients of the quantitative variables and class means outputs from the two canonical structures are shown in Table 8. The first canonical structure (Can 1) explains the majority (69%) of the variability with eigenvalue of 0.48. The first canonical correlation (57%) was the greatest multiple correlation with the classes that was achieved by using the linear combination of the quantitative variables. The results revealed that Can 1 separates the sheep populations (class means) from different agro-ecological zones.

Results of a discriminant function analysis (Table 9) shows the classification of data into a known agro-ecological zone. Accordingly, an average of 66% of the sampled animals were classified into their respective

Table 6. Summary of the stepwise discriminant function analysis; ascending order of traits used in discriminating the sheep populations from different agro-ecological zones.

Step	Variables entered	Partial R-squared	F value	Pr > F	Wilk's Lambda	Pr < Lambda
1	Canon bone length	0.1237	21.39	<0.0001	0.8763	<0.0001
2	Height at withers	0.0564	9.02	0.0002	0.8269	<0.0001
3	Tail length	0.0557	8.87	0.0002	0.7809	<0.0001
4	Rump length	0.0554	8.79	0.0002	0.7376	<0.0001
5	Head width	0.0529	8.36	0.0003	0.6986	<0.0001
6	Hair length	0.0460	7.18	0.0009	0.6664	<0.0001
7	Head length	0.0525	8.23	0.0003	0.6315	<0.0001
8	Shoulder point width	0.0308	4.70	0.0098	0.6120	<0.0001
9	Shin circumference	0.0258	3.91	0.0210	0.5962	<0.0001
10	Heart girth	0.0172	2.58	0.0777	0.5860	<0.0001
11	Body weight	0.0268	4.03	0.0188	0.5703	<0.0001
12	Rump width	0.0147	2.18	0.1154	0.5619	<0.0001
-	Chest depth	0.0073	1.07	0.3432	-	-
-	Subs height	0.0013	0.18	0.8314	-	-
-	Body length	0.0010	0.15	0.8595	-	-
-	Tail width	0.0008	0.12	0.8876	-	-

agro-ecological zone. The overall error rate was 34%, while higher error rates were obtained from the classification of midland sheep populations.

Pairwise squared distances between agro-ecological zones are shown in Table 10. All distances were significant. Highland sheep populations are distantly related to the lowland sheep. On the other hand, midland sheep relates more towards lowland sheep than highland sheep.

Table 7. Multivariate statistics and F approximates

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.5559	6.14	32	576	<0.0001
Pillai's Trace	0.5015	6.04	32	578	<0.0001
Hotelling-Lawley Trace	0.6957	6.24	32	510.82	<0.0001
Roy's Greatest Root	0.4811	8.69	16	289	<0.0001

Table 8. Eigen values, canonical correlations and class means

	Can 1	Can 2
Multivariate Statistics		
Canonical correlation	**0.5699**	0.4203
Eigenvalue	**0.4811**	0.2146
Proportion	**0.6916**	0.3084
Class (agro-ecological zones)		
Highland	**0.9139**	-0.2290
Lowland	**-0.7539**	-0.4141
Midland	**-0.1600**	0.6430

Discussion

Along with the most observable qualitative traits, quantitative measurements produce reliable information in characterization and differentiation of sheep populations. In our current study, more traits showed significant differences among ewes than among rams in different agro-ecological zones. This might be due to either the larger sample size taken for ewe populations or due to the similarity of rams over the studied agro-ecological zones, which could be attributed to common markets

Table 9. Number and percent of observations classified into agro-ecological zones.

From	Highland	Lowland	Midland	Total
Highland	70 (68.6)	20 (19.6)	12 (11.8)	102 (100)
Lowland	12 (11.8)	71 (69.6)	19 (18.6)	102 (100)
Midland	19 (18.6)	23 (22.6)	60 (58.8)	102 (100)
Total	101 (33.0)	114 (37.3)	91 (29.7)	306 (100)
Error rate	0.31	0.30	0.41	0.34
Priors	0.33	0.33	0.33	

Table 10. Squared Mahalanobis distance between agro-ecological zones; output of the multivariate analysis calculated using the quantitative measurements

From	Highland	Lowland	Midland
Highland	0		
Lowland	2.82***	0	
Midland	1.91***	1.47***	0

*** shows the significance of the distance calculations at p<0.0001.

Figure 6. Lowland sheep flock in Tahtay Maichew district.

where farmers select and purchase rams for sire purposes.

Most of the studied traits were affected by agro-ecological zone differences, which might be due to the differences in physiological adaptation mechanism of sheep types to different environments, management, availability of different feed and nutrition and/or the variations being caused by genetic factors. For example, the majority of lowland sheep possess higher values for height at withers, testis circumference and short hair which might help them to adapt to a hot environment. These results are in line with the results of Getachew *et al* (2009) and Gizaw *et al* (2008) for Menz and Afar sheep who reported that such measurements of the lowland Afar sheep were higher than the highland Menz sheep populations.

Most of the body measurements were higher for rams than ewes, which might be attributed to enhanced muscle mass and skeletal development in males due to testosterone hormone secretions (Baneh and Hafezian, 2009). These results follow Rensch's rule where the males of a particular species are usually heavier than the females (Rensch, 1950). Size differences may be ascribed to the differences in the endocrine system of the two sexes; estrogen hormone was shown to have a limited effect on growth in females (Baneh and Hafezian, 2009). These results are in agreement with Mustefa *et al* (2019) and Getachew *et al* (2009) who reported that males were higher than the females in most growth traits in goats and sheep respectively.

In agreement with the results of the current study, most scholars report differences in traits between the sexes with rams being dominant over ewes (Rensch, 1950; Baneh and Hafezian, 2009; Mustefa *et al*, 2019; Getachew *et al*, 2009). However, differences due to sex was not observed in Tigray Highland sheep populations (Hayelom *et al*, 2014). In contrast to this, Hayelom *et al* (2014) report the dominance of ewes over the rams on most of the linear body measurements of Elle sheep populations. In a different study, no significant differences were observed among body weight of Simien sheep ewes and rams (Melaku *et al*, 2019).

The overall mean body weight (23.50 kg) presented in the current study was higher than those reported by Tajebe *et al* (2011) for Abergelle sheep (21.25 kg), and by Gebreyowhens and Tesfay (2016) for Tigray Highland sheep (22.10 kg), while the values were lower than those reported by Hayelom *et al* (2014) for Tigray Highland sheep (27.52 kg), and Melaku *et al* (2019) for Simien sheep (24.90 kg) which show their difference from the sheep populations of the neighboring areas. Although the body weight of rams and ewes presented in this study were higher than those reported for

Figure 7. Highland sheep flock in Tahtay Maichew district.

Abergelle and Tigray highland sheep (Tajebe *et al*, 2011; Gebreyowhens and Tesfay, 2016), other traits such as body length, height at withers, heart girth and tail length were comparable between these sheep populations. On the other hand, higher values for all morphological measurements were reported by Edea *et al* (2010) for the country's most known sheep breeds (Bonga and Horro sheep).

The results also revealed that linear body measurements among the studied sheep population differ with age. Three-year-old sheep showed the highest values for most of the measurements, reflecting the optimum growth age. These results are in contrast with results of Getachew *et al* (2009) for Menz and Afar sheep and Melaku *et al* (2019) for Simien sheep who reported that the body weight of the sheep continued to increase with age.

In addition to the quantitative measurements, the qualitative characteristics of a population also allow to easily differentiate genetic resources. Among the obvious qualitative characteristics which differentiate the current sheep populations from the previously characterized sheep populations are the complete absence of beard, horn, ruff, toggle and pigmented skin. Similarly, variations in coat color were also observed among the different agro-ecological zones. Accordingly, the majority of the highland and midland sheep populations from the current study possess tan coat color which makes them unique among the other Tigray highland sheep populations. Sheep populations sampled from the lowland area display dominantly white and light colors which is in agreement with the report of Getachew *et al* (2009) for the lowland Afar sheep. The majority of the sheep populations from the current study possessed hairy fiber type coats, which was in contrast to the results of Hayelom *et al* (2014), who report course wool for Tigray Highland sheep. However, huge variations were not observed among the other qualitative characteristics of the sheep types from the current study and earlier studies of Gebreyowhens and Tesfay (2016) and Hayelom *et al* (2014) for Tigray Highland sheep.

Discriminant function analysis allowed the classification of an average 66% of the studied animals into their respective environments zone. Lowest classification of individuals into their respective agro-ecological zone was observed in midland sheep populations, indicating a lack of uniqueness within them. All of the pairwise comparisons between populations from different agro-ecological zones were found to be highly significant with the largest difference observed between the highland and lowland sheep populations. These results reflect the large altitudinal differences between the two agro-ecological zones. The shortest distance calculated

between the lowland and midland sheep populations show their relative similarities as compared with the highland sheep. These differences among different agro-ecological zones show the presence of potential genetic resource variations which can be useful for maintaining diversity and further selection-based genetic improvement programs.

In conclusion, using a combination of quantitative and qualitative characteristics we were able to discriminate the sheep populations from three agro-ecological zones in the Tahtay Maichew district, Tigray region of Ethiopia and to group them into two distinct populations (the highland and the lowland sheep). Sheep populations from the midland agro-ecological zone were considered to be part of the lowland group. Therefore, it is better to consider the highland and lowland sheep as different traditional populations until molecular characterization results provide further evidence for population differentiation. Additionally, the molecular characterization studies will show the within population genetic diversity and level of inbreeding which can be used for selecting appropriate genetic improvement plans (selection or crossbreeding). According to the reports of Gizaw *et al* (2008), the sheep genetic resources of most parts of the Tigray region were generally referred as Sekota sheep population under the short fat tailed breed. However, results from our study indicate that there are several sheep populations that cannot be categorized under the Sekota traditional sheep population. Therefore, it is advisable to include these genetic resources for further molecular studies to understand the genetic diversity within and among populations.

Acknowledgements

The authors are highly indebted to the Ethiopian Biodiversity Institute (EBI) for covering all the budget needs of the work. Our special appreciation also goes to the smallholder farmers /breeders for providing their animal to this work for free. We also take this opportunity to thank the animal science experts and development agents in the district for their endless help during the data collection. A special word also goes to our friend and work partner Mr. Tadesse Hunduma for mapping the study area.

Author contributions

All authors contributed to the study conception and design. Material preparation, and data collection were performed by AH, TA, SS and ST. Data analysis and writing the first draft of the manuscript was performed by AM and AH. AA commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflict of interest statement

The authors declare no conflicts of interest.

References

Atsbeha, G., Demissew, S., Woldu, Z., and Edwards, S. (2015). Floristic composition of herbaceous flowering plant species in Lalay and Tahtay Michew Districts, Central Zone of Tigray, Ethiopia. *African Journal of Ecology and Ecosystems* 2(6), 159–169.

Baneh, H. and Hafezian, S. H. (2009). Effect of environmental factor on growth traits in Ghezel sheep. *African Journal of Biotechnology* 8, 2903–2907. url: https://www.ajol.info/index.php/ajb/article/view/60943.

Bimerow, T., Yitayew, A., Taye, M., and Mekuriaw, S. (2011). Morphological Characteristics of Farta Sheep in Amhara Region, Ethiopia. *Online Journal of Animal and Feed Research* 1(6), 299–305.

Central Statistical Agency (2005). Ethiopian agricultural sample survey 2012/2013 (2005 E.C.). In Report on livestock and livestock characteristics., volume II, Central Statistical Agency, Addis Ababa, Ethiopia.

Central Statistical Agency (2018). Ethiopian agricultural sample survey 2017/18 (2010 E.C.). In Report on livestock and livestock characteristics, volume II, Central Statistical Agency, Addis Ababa, Ethiopia.

Degen, A. A. (2007). Sheep and goat milk in pastoral societies. *Small Ruminant Research* 68(1-2), 7–19. url: https://dx.doi.org/10.1016/j.smallrumres.2006.09.020. doi: 10.1016/j.smallrumres.2006.09.020

Devendra, C. and Mcleroy, G. B. (1982). Goat and sheep production in the tropics Intermediate tropical agriculture series, (London: Longman), 271p.

EBI (2016). Ethiopian National Strategy and Plan of Action for conservation and utilization of Animal Genetic Resources (Addis Ababa, Ethiopia: Ethiopian Biodiversity Institute), 114-114. url: https://www.ebi.gov.et/wp-content/uploads/2013/01/Final%20ENSAP_final_submitted%20(1).pdf.

Edea, Z., Haile, A., Tibbo, M., Sharma, A. K., Assefa, D., Sölkner, J., and Wurzinger, M. (2010). Morphological characterization of Bonga and Horro indigenous sheep breeds under smallholder conditions in Ethiopia. *Ethiopian Journal of Animal Production* 9(1), 117–133.

FAO (2012). Phenotypic characterization of Animal Genetic Resources. url: www.fao.org/docrep/015/i2686e/i2686e00.pdf.

Gebreyowhens, W. and Tesfay, Y. (2016). Morphological Characterization of Indigenous Highland Sheep Population of Tigray, Northern Ethiopia. *Journal of Natural Sciences Research* 6(1), 96–104. url: https://www.iiste.org/Journals/index.php/JNSR/article/view/28281/0.

Getachew, T., Haile, A., Tibbo, M., Sharma, A. K., Kifle, A., Terefe, E., Wurzinger, M., and Sölkner, J. (2009). Morphological characters and body weight of Menz and Afar sheep within their production system. *Ethiopian Journal of Animal Production* 9(1), 99–115.

Gizaw, S., Komen, H., Hanotte, O., and Arendonk, J. A. M. V. (2008). Indigenous sheep resources of Ethiopia: types, production systems and farmers preferences.

Animal Genetic Resources Information 43, 25–39. url: https://dx.doi.org/10.1017/s1014233900002704. doi: 10.1017/s1014233900002704

Hayelom, M., Abegaz, S., and Mekasha, Y. (2014). Within Breed Phenotypic Diversity of Sokota/Tigray Sheep in Three Selected Zones of Tigray, Northern Ethiopia. *Journal of Biology Agriculture and Healthcare* 4(17), 148–157. url: https://www.iiste.org/Journals/index.php/JBAH/article/view/15182.

Melaku, S., Kidane, A., Abegaz, S., Tarekegn, A., and Tesfa, A. (2019). Phenotypic characterization of Simien sheep in Simien Mountain Region. *Ethiopia. International Journal of Agriculture and Biosciences* 8(4), 178–185. url: http://www.ijagbio.com/volume-8-no-4-2019/.

Mengistie, T., Girma, A., Solomon, G., Sisay, L., Abebe, M., and Markos, T. (2010). Traditional management systems and linear body measurements of Washera sheep in the western highlands of the Amhara National Regional State, Ethiopia. *Livestock Research for Rural Development* 22(169). url: http://www.lrrd.org/lrrd22/9/taye22169.htm.

Michael, A., Kefelegn, K., and Yoseph, M. (2016). Phenotypic Characterization of Indigenous Sheep Types in Northern Ethiopia. *Journal of Natural Sciences Research* 6(15), 16–27. url: https://iiste.org/Journals/index.php/JNSR/article/view/32220/33107.

Misra, A. K. and Singh, K. (2002). Effect of water deprivation on dry matter intake, nutrient utilization and metabolic water production in goats under semi-arid zones of India. *Small Ruminant Research* 46, 159–165. doi: https://doi.org/10.1016/S0921-4488(02)00187-6

Mustefa, A., Gizaw, S., Banerjee, S., Abebe, A., Taye, M., Areaya, A., and Besufekad, S. (2019). Growth performance of Boer goats and their F1 and F2 crosses and backcrosses with Central Highland goats in Ethiopia. *Livestock Research for Rural Development* 31(6), 89. url: http://www.lrrd.org/lrrd31/6/amine31089.html.

Rensch, B. (1950). Die Abhängigkeit der relativen Sexualdifferenz von der Körpergröße. *Bonner Zoologische Beitrage* 1, 58–69.

SAS Institute (2002). Statistical Analysis Software (SAS) url: https://www.sas.com/en_us/software/stat.html.

Tajebe, S., Gangwar, S. K., and Kebede, K. (2011). Performance and Physical Body Measurement of Abergele Sheep Breed in Traditional Management System of Tigray Regional State, Northern Ethiopia. *International Journal of Science and Nature* 2(2), 225–230. url: https://hdl.handle.net/10568/70955.

Worku, A. (2018). Morphometrical Characterization and Traditional Breeding Objectives of Native Sheep Reared in East Shoa and West Arsi Zone of Oromia Regional State. M.Sc. thesis, Hawassa University, Hawassa, Ethiopia.

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2020), 1 (2), 23–28
DOI: 10.46265/genresj.LLTJ1737
https://www.genresj.org
ISSN: 2708-3764

Isolation and Characterization of Fifteen Microsatellite Loci for the Use in Breeding of *Gmelina arborea* Roxb. (Lamiaceae)

Oscar J Rocha [*,a], **Dawa Méndez-Álvarez** [b], **Fabiana Rojas Parajeles** [b] **and Olman Murillo Gamboa** [b]

[a] *Department of Biological Sciences, Kent State University, Ohio, 44242, Kent, USA*
[b] *School of Forest Engineering, Instituto Tecnológico de Costa Rica, 159-7050, Cartago, Costa Rica*

Abstract: *Gmelina arborea* (melina) is a valuable tree species throughout tropical areas, and there are extensive commercial plantations of this species in Southeast Asia, West Africa, and Latin America. As part of a research program for the genetic improvement and management of *G. arborea* at Instituto Tecnólogico de Costa Rica, we developed, validated, and optimized fifteen microsatellite loci. We used 23 clones belonging to five different companies currently using clonal selection to manage their commercial plantations. Our results showed that all fifteen loci were polymorphic and together had 75 alleles (2-7 alleles/locus). We also found that eleven loci showed lower heterozygosity than expected under Hardy-Weinberg equilibrium (HWE). We calculated the genetic similarity among all clone pairs using the number of shared alleles to examine the potential of these loci for clone discrimination. Overall, pairwise similarity among clones ranged from 0.36 to 0.83, and our findings also showed that clones from the same commercial plantation tended to be more similar to each other than to clones from other plantations. These microsatellite loci will contribute toward the characterization of the genetic diversity, the identification of elite clone lines for timber production, and breeding and adequate management of commercial plantations of *G. arborea*.

Keywords: breeding, clonal forestry, gene markers, melina, population genetics

Citation: Rocha, O. J., Méndez-Álvarez, D., Rojas Parajeles, F., Murillo Gamboa, O. (2020). Isolation and Characterization of Fifteen Microsatellite Loci for the Use in Breeding of *Gmelina arborea* Roxb. (Lamiaceae). *Genetic Resources* 1 (2), 23–28. doi: 10.46265/genresj.LLTJ1737.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Gmelina arborea (melina) is a valuable timber species that grows throughout tropical areas. Native to Southeast Asia and India, it is an important commercial timber species in tropical regions worldwide, particularly in Southeast Asia, West Africa, and South America, where G. arborea is grown in large plantations. Its low-density wood is durable and yields reasonable quantities of relatively uniform, stable, and light color pulp (Dvorak, 2004; Wee *et al*, 2012). The Panel of Experts on Forest Genetic Resources of the Food and Agriculture Organization (FAO) describes *G. arborea* as an important tree species with high potential and utility (Lauridsen and Kjaer, 2002).

G. arborea is the second most planted timber species in Costa Rica because of its rapid growth rate, easy establishment, high productivity, a wide range of tolerance to site conditions and excellent regrowth capacity (Rosero *et al*, 2011; Ávila Arias *et al*, 2015a,b; Vergara *et al*, 2017). This species was first introduced into Costa Rica in 1966 for pulp production by the local paper company and to serve as a seed source for the establishment of plantations in the Jari Project in Brazil in Eastern Amazonia (Rounda, 1988). This initial introduction consisted of seeds from twenty independent origins, i.e., provenances from different regions throughout its native range in Asia

*Corresponding author: Oscar J Rocha (orocha@kent.edu)

Received: 26.03.2020 | Accepted: 03.11.2020 | Published online: 28.12.2020

(India, Pakistan, and Bangladesh), and commercial plantations in Africa (Nigeria and Cameroon) and British Honduras (now Belize). These provenances were planted separately in >100 ha blocks, to provide a broad base for genetic improvement. More than 20 years later, seeds from this plantation were collected from healthy trees with desirable phenotypes, initiating its spread throughout the region.

Breeding efforts of *G. arborea* in Costa Rica started in the early '90s, leading to the development of highly productive genetic stock for timber production at a regional scale (Ávila Arias *et al*, 2014, 2015a,b). The most successful melina breeding programs in the region use clonal propagation to establish their commercial plantation, as this strategy provides a reliable stock of propagules that are easy to produce and plant and results in fast-growing trees and high productivity. Moreover, researchers have used variables such as trade volume and quality of wood and other indicators of each clone line's performance for the selection of the genetic stock to be planted in sites with different soil characteristics, flooding, and land use-history (Ávila Arias *et al*, 2015a,b).

Here, we describe fifteen microsatellite loci developed to support ongoing breeding programs of melina in Costa Rica using a small number of clones selected for their rapid growth and high productivity. These markers will be used for clone identification and potentially for marker-assisted breeding of *G. arborea*.

Materials and Methods

Development of Microsatellite Markers

The microsatellite markers were developed using the magnetic bead protocol described by Cullings (1992) and Li *et al* (1997) and modified by Glenn and Schable (2005). Genomic DNA from a sample of five *G. arborea* trees was digested using *HaeIII/PshA1* restriction enzymes (Invitrogen; Carlsbad, CA). Two linkers were added to the digested genomic DNA (M28 5'CTCTTGCTTGAATTCGGACTA 3' and M29 5'pTAGTCCGAATTCAAGCAAGAGCACA 3') and M28 was used as a primer for subsequent polymerase chain reactions (PCR). Finally, the digested genomic DNA was amplified in multiple PCR reactions and their product concentrated to gain enough DNA for the following bead hybridization process.

Two arbitrary repeat motifs (CA20 and AG17) were selected as probes for the bead hybridization reactions based upon Cardle *et al* (2000). The short tandem repeat (STR) probes from Integrated DNA Technologies (Coralville, IA, USA) had a biotin label on the 5' end. The STR probes were added to a bead hybridization reaction to select for DNA fragments that contained the repeat motif of the probe. This bead hybridization process aimed to allow the fragments containing repeats to anneal to the biotin-labeled probes. After the hybridization, the selected fragments were isolated from the rest of the genomic DNA using streptavidin-coated magnetic beads, which bind to the biotin-labeled probes. These fragments were then eluted and re-amplified using the M28 primer in additional PCR reactions. The bead hybridization and PCR pre-amplification processes were repeated one more time to enrich for genomic DNA containing the selected repeats.

After completing the bead hybridization and selection process, the repeat sequences enriched DNA was ligated into a pGEM-T vector from Promega (Madison, WI, USA) to begin the sequencing phase of this protocol. We cloned the vectors into electrocompetent *Escherichia coli* cells. We later plated transformed E. coli cells onto selective media containing 0.1 mg/mL ampicillin, 0.05 mg/mL X-Gal, and 1mM IPTG. All positive clones were sequenced on an ABI PRISM 377 DNA Sequencer using universal M13 forward (F) and reverse (R) primers (Schuelke, 2000). The sequencing reactions were standard 20 ml reactions using the ABI PRISM BigDye Terminator sequencing kits (Applied Biosystems, Foster City, CA, USA) and 3.2 pmol of PCR product for the template. Primers for each of the fifteen microsatellite loci were designed from sequences containing multiple copies of the repeated motif and with sufficiently long flanking regions on the 5' and the 3' end of the repeated region pairs using Primer 3.0 software (Rozen and Skaletsky, 2000).

Microsatellite Loci Characterization

All primer pairs were tested for amplification and polymorphism using DNA obtained from 23 promising genotypes (clones) of *G. arborea* belonging to five different privately operated clonal breeding programs. Two ramets from each clone were gathered from a clonal collection maintained in a greenhouse at the Instituto Tecnológico de Costa Rica to validate all alleles by genotyping them separately. As described above (Doyle and Doyle, 1987; Lodhi *et al*, 1994), total genomic DNA was extracted at the Forest Molecular Genetic Laboratory, in the Forest Innovation Research Center (CIF) at the Instituto Tecnológico de Costa Rica, Cartago, Costa Rica. Copies of these clones are maintained in the mini clonal garden facility and could be made accessible upon request.

PCR Amplification and Fragment Analysis

Polymerase chain reactions were performed in a final volume of 15 μl, containing approximately 50 ng of genomic DNA, 10 mM Tris buffer, pH 8.0, 10 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 μM of each primer, and 1 U of Taq polymerase (Fermentas®) using an Eppendorf® Mastercycler EP thermal cycler. The PCR program used included an initial step of 2 min of denaturation at 94 °C, 30 cycles of 15 s at 94 °C, 15 s at 55 °C and 30 s at 72 °C, and a final extension cycle of 1 min at 72 °C. To genotype each individual, we conducted electrophoresis for fragment separation using a QIAxcel Advanced fragment analyzer from QIAGEN® at Centro de Investigación en Biología Celular y Molecular (CIBCM) at Universidad de Costa Rica. Once

all of the data scorings were complete, random samples were re-amplified and re-ran to assess reproducibility and confirm scoring and allele sizes.

Genetic analysis

GenAlex 6.3 (Peakall and Smouse, 2006) was used to calculate common indicators of genetic diversity, including the number of alleles (N_a) per locus and the expected (H_e) and observed heterozygosity (H_o). GenAlex was also used to calculate deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium. Genotype errors due to stutter bands, allele dropout, and null alleles were estimated using the MICRO-CHECKER software (van Oosterhout *et al*, 2004).

To examine the potential of these loci for discrimination among the 23 clones, the multilocus genotype of each clone was determined using the presence and absence of alleles to estimate genetic similarities for all pairwise comparisons among clones. Genetic similarity among each pair of clones was calculated based on the number of alleles common among the clones according to the following equation proposed by Dice (1945), where

$$GSxy = 2a/(2a + b + c),$$

where a is the number of alleles common to clones x and y, b the number of alleles present only in clone x, and c the number of alleles present only in clone y. A cluster analysis based on sequential, agglomerative, hierarchical, and nested clustering methods (SAHN, UPGMA; NTSYS-pc-p package; (Rohlf, 1993) was conducted to describe the relationship between the clones.

Results

Microsatellite loci

Table 1lists the loci names, corresponding accession numbers in Genbank, repeated motifs, forward and reverse primer sequences, the size range of PCR products, and annealing temperatures for each of the fifteen microsatellite loci isolated for *Gmelina arborea*. All loci were polymorphic, with the number of alleles per locus ranging from 2 to 7 (Table 2). We found 75 different alleles across all loci (Supplemental Table 1), with an average of 5.00 ± 0.41 alleles per locus. Average observed and expected heterozygosities were also high ($H_o = 0.504$ and $H_e = 0.645$, Table 2). Moreover, our findings did not show evidence of scoring error due to stuttering or significant allele dropout for any of the fifteen polymorphic loci.

Our analyses revealed significant deviations from Hardy-Weinberg proportion in most loci (Table 2). We observed heterozygote deficiencies in eleven loci and an excess of heterozygotes in one locus (Meldi-12; Table 2). However, given the small sample size used to validate these loci and the high number of alleles found in most of them, it is reasonable to expect that they will not be in Hardy-Weinberg equilibrium (HWE). Moreover, the clones used to validate these microsatellite loci do not represent a sample of a natural population of *G. arborea*, but a collection of promising genotypes selected by the timber industry. We also caution that two loci; namely, Meldi11 and Meldi11.2, which include different tandem repeats, were derived from the same sequence.

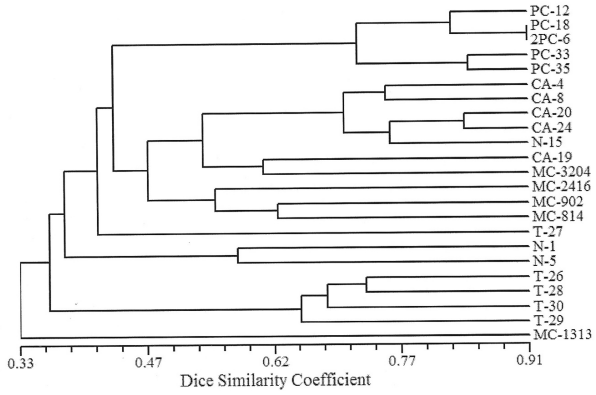

Figure 1. Dendrogram of the 23 clones of *Gmelina arborea* based on Dice similarity coefficient (Dice, 1945). Letters preceding the clone identification number indicate the breeding programs from which each clone was obtained. Clones from two programs (PC and CA) tended to cluster together while one clone from each of the other three programs (MC, N, and T) clustered with clones from the other programs.

Our analysis using the software MICRO-CHECKER did not reveal evidence for genotype errors due to stutter bands or allele dropout. Our analysis suggested the presence of null alleles in nine loci (Table 2), but such findings might result from a deviation from Hardy-Weinberg proportions. MICRO-CHECKER uses deviations of Hardy-Weinberg proportions to identify loci likely to have null alleles. We need to reiterate that our sample did not represent a natural population of *G. arborea*. For that reason, deviations from Hardy-Weinberg proportions are likely to occur in multiple loci.

Our results also showed that all 23 clones exhibited a unique combination of alleles (Supplemental Table 1), resulting in genetic similarities (Dice) ranging from 0.36 to 0.83 (Figure 1). Overall, most of the clones clustered according to their origin or breeding program. All clones from programs PC and CA clustered together while some clones from programs MC, N, and T grouped with clones from the other programs.

Discussion

We described fifteen polymorphic microsatellite loci for the fast-growing timber tree *Gmelina arborea*. These new microsatellite loci proved to be very informative, accurate, and with a reliable discrimination power for assessing genotype identity. The process of allele validation provides confidence for utilizing this set of microsatellite loci for multiple purposes. Overall, we found high levels of allelic diversity, suggesting a broad genetic base in the original material from which these

Table 1. Development and optimization of microsatellite markers for *Gmelina arborea*. Locus name, accession number, repeat motifs, oligonucleotide primer sequences, PCR product size range, and annealing temperatures for the fifteen microsatellite loci isolated for *G. arborea*.

Locus name	Accession no.	Repeat motif	Primer sequence	Range of PCR products (bp)	Annealing Temp (°C)
Meldi1	MH249250	CT_{12}	R: 5'-CCCCACTTTTGATGCTCTCT-3'	113-131	56
			F: 5'-TAGTTGGTGAAATGAAAATTCGC-3'		
Meldi2	MH230082	$TC_9.GA_{15}.GA_9$	R: 5'-TAGTCGCATAAGGAAAAGTTAG-3'	299-317	55
			F: 5'-CCAATTCTGTGCTGCTAAGG-3'		
Meldi4	MH249251	TC_{17}	R: 5'-ACTCCAGACCTTTCTCGCTC-3'	214-237	58
			F: 5'-CCTTCTTCAAAATCCGATCTTTC-3'		
Meldi5	MH249252	GA_{10}	R: 5'-CAATTTCCTTCAGTTATTCCCAC-3'	283-307	50.8
			F: 5'-CAAAATAGATTAAGCCCACATC-3'		
Meldi6	MH249253	$TC_{12}TC_9.TC_{10}$	R: 5'-CCCATATCACCTGCTCTTAG-3'	292-320	56
			F: 5'-GGATGTCACGATGTCGGATG-3'		
Meldi7	MH249255 MH249254	$CT_{22}...CT_{19}$	R: 5'-CCTAACAATCATGGAAAAGAG-3'	352-376	54
			F: 5'-CACTCAGCGTGGCAAAGAG-3'		
Meldi10	MH249256	CT_{19}	R: 5'-CAATTTCTCCCTAGCAAGTTC-3'	174-210	56
			F: 5'-AGTACATCTACTCAAAGGAGGT-3'		
Meldi11	MH249257	GA_{12}	R: 5'-GCTTGTGGTTTGGGTGACGA-3'	168-182	57
			F: 5'-AGTGATGCTCCTCATTCAGGA-3'		
Meldi11.2	MH249257	TC_9	R: 5'-TTCATAATGAGCAAAGGGGAC-3'	140-160	56
			F: 5'-GGCCACTTCAGTATCCACAC-3'		
Meldi12	MH249258	TC_{15}	R: 5'-CCTAGATTAGCCATTTATATTG-3'	206-214	56
			F: 5'-GGTGGAGCAGAAGAAGAGAGA-3'		
Meldi16	MH249260	TG_{23}	R: 5'-ACCAGGACACCCTCGACAC-3'	213-225	55.2
			F: 5'-GTTGAGGAGCAGTTTATAAGAC-3'		
Meldi17	MH249261	$TG_{10}.GA_{17}$	R: 5'-GGTTATTATAATCACTTCTGCC-3'	153-167	55
			F: 5'-CCAATGTAATAGGACAAAACTC-3'		
Meldi20	MH249263	$GT_8.GA_{17}$	R: 5'-GGGAAAATCAGATACACTGCC-3'	205-231	57
			F: 5'-ATCAACCAAGGGGACAAGCC-3'		
Meldi21	MH249264	$GA_{14}..TG_n$	R: 5'-GCATGCAACAGAGAAAAGAAG-3'	231-262	55
			F: 5'-GATAATCGTACTAAAGTTGAGG-3'		
Meldi24	MH249267	$CA_6..GA_{17}$	R: 5'-CCTTTCCCTAAGTCCACACA-3'	211-219	53
			F: 5'-TTTTGTCCAAATGAGGCTCC-3'		

23 clones were selected. We expected to encounter high genetic diversity among the clones used in this study because they represent a sample taken from collections of *G. arborea* selected by growers because of their performance. Moreover, the plantations where these clones were selected have different soil types, precipitation regimes, and topography.

We found that all clones from two clonal breeding programs clustered together in the dendrogram (PC and AC; Figure 1). However, this is not true for clones from all breeding programs, as clones from the same program may not group in the same cluster. For example, clone T-27 did not cluster together with the other four genotypes in the same program (T-26, T-28, T-29, and T-30). Similarly, clones N-15 and MC-1313 did not group with the other trees from their program. However, clones from the same breeding program tended to group, suggesting that the process of selecting promising clones, based on what breeders considered desirable phenotypes, varies among breeding programs. Furthermore, this finding also implies that promising clone lines could perform well in a given environment. Therefore, it suggests that the degree of similarity of allelic composition among clones may indicate similarities in their ability to respond to environmental conditions.

Ávila Arias *et al* (2014) conducted a field trial using different clone lines planted in two locations in south-western Costa Rica. They found significant differences in diameter at breast height (DBH), commercial height, commercial volume of the trunk over bark, trunk quality, and the volume and quality of the wood among clone lines two years after planting. Their analysis also showed significant genotype by environment interaction in clonal performance, as some accessions grew well in their site of origin but not in other locations. Murillo-Gamboa *et al* (2016) reported differences in the tolerance to melina's wilt, a critical disease in Costa Rica,

Table 2. Genetic analysis of microsatellite loci on 23 *Gmelina arborea* clones. Observed number of alleles per locus, observed and expected heterozygosity, significance value for test for the deviation from Hardy-Weinberg proportions (χ^2 test), and test for null alleles are listed for each of the fifteen polymorphic loci developed for *G. arborea*.

	Number	Heterozygosity		χ^2 test	Test for Null Alleles
Locus name	of alleles	Observed	Expected	P-value	P-value
Meldi1	3	0.435	0.446	ns	†
Meldi2	5	0.348	0.712	***	***
Meldi4	6	0.652	0.671	ns	†
Meldi5	6	0.652	0.750	**	ns
Meldi6	7	0.217	0.774	***	***
Meldi7	7	0.550	0.839	***	**
Meldi10	6	0.435	0.722	***	**
Meldi11	3	0.227	0.404	***	†
Meldi11.2	6	0.500	0.792	*	**
Meldi12	2	0.609	0.423	*	†
Meldi16	3	0.565	0.542	*	†
Meldi17	4	0.348	0.661	***	***
Meldi20	6	0.522	0.705	ns	*
Meldi21	5	0.435	0.746	***	**
Meldi24	6	0.435	0.775	***	***

ns = not significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, †A single allele contributed to more than 50% of the observation in this locus. Binomial analysis could not be performed.

among clone lines used in the field trial conducted by Ávila Arias *et al* (2015b,a). These findings indicate that clone selection is biased toward genotypes performing well in particular environments, thus suggesting that genetic markers could play a role in identifying promising genotypes.

In summary, the fifteen polymorphic microsatellite markers we described here have great potential use for the breeding of *G. arborea*, including genotyping the breeding collections, as well as keeping the identity and assessing the purity in clonal gardens. In this respect, there are eleven additional loci available to expand the multilocus genotype of each clonal line (Liao *et al*, 2010) to increase the possibilities for genetic analysis and marker-assisted selection of *G. arborea*.

Acknowledgements

This work was supported by the Fundación CR-USA, the Department of Biological Sciences of Kent State University and Instituto Tecnológico de Costa Rica.

Supplemental data

Supplemental Table 1: Alleles observed in each of the 23 clones of *Gmelina arborea*. Summary of the presence (1) or absence (0) of the 75 alleles found in the fifteen polymorphic microsatellite loci in each clone. This information was used to estimate genetic similarities for all pairwise comparisons among clones

Author contributions

O.J.R. and O.M-B. contributed to the conception and design of the submitted manuscript. O.J.R. isolated and developed the fifteen microsatellite loci, D.M-A and F.R-P. validated and optimized each of the fifteen polymorphic loci. O.J.R. drafted the manuscript, and all authors contributed the drafting, revision, and final approval of the submitted manuscript.

Conflict of interest statement

The Authors declare no conflict of interest.

References

Ávila Arias, C., Murillo-Cruz, R., Murillo-Gamboa, O., and Sandoval-Sandoval, C. (2014). Desarrollo juvenil de clones de *Gmelina arborea* Roxb. de dos procedencias, en sitios planos del Pacífico Sur de Costa Rica. *Revista Forestal Mesoamericana Kurú* 12(28), 23–23. (in Spanish). doi: https://doi.org/10.18845/rfmk.v12i28.2097

Ávila Arias, C., Murillo-Cruz, R., and Murillo-Gamboa, O. (2015a). Selección de clones superiores de dos conjuntos genéticos de *Gmelina arborea* en el Pacífico Sur de Costa Rica. *Revista de Ciencias Ambientales* 49(1), 17–35. (in Spanish). doi: https://doi.org/10.15359/rca.49-1.2.

Ávila Arias, C., Murillo-Cruz, R., Murillo-Gamboa, O., and Sandoval-Sandoval, C. (2015b). Interacción genotipo sitio para dos conjuntos clonales de *Gmelina arborea* Roxb., en sitios planos del Pacífico Sur de Costa Rica. *Revista Forestal Mesoamericana Kurú* 12(29), 02–14. url: https://revistas.tec.ac.cr/index.php/kuru/article/view/2250/2033.

Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., and Waugh, R. (2000). Computational and experimental characterization of physically clustered

simple sequence repeats in plants. *Genetics* 156, 847–854.

Cullings, K. W. (1992). Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology* 1(4), 233–240. doi: https://doi.org/10.1111/j.1365-294x.1992.tb00182.x

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology* 26(3), 297–302. doi: https://doi.org/10.2307/1932409

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. url: https://webpages.uncc.edu/~jweller2/pages/BINF8350f2011/BINF8350 Readings/Doyle plantDNAextractCTAB 1987.pdf.

Dvorak, W. S. (2004). World view of *Gmelina arborea*: opportunities and challenges. *New Forests* 28(2/3), 111–126. doi: https://doi.org/10.1023/b:nefo.0000040940.32574.22.

Glenn, T. C. and Schable, N. A. (2005). Isolating microsatellite DNA loci. In *Molecular Evolution: Producing the Biochemical Data,* ed. Zimmer, E. A. and Roalson, E. H. (Academic Press), volume 395 of *Methods in Enzymology*, 202-222.

Lauridsen, E. B. and Kjaer, E. D. (2002). Provenance research in *Gmelina arborea* Linn., Roxb. A summary of results from three decades of research and a discussion of how to use them. *The International Forestry Review* 4(1), 20–29. url: https://www.jstor.org/stable/43740942.

Lega Rounda, F. F. (1988). Estudios de la forma de *Gmelina arborea* (Roxb). Análisis de las plantaciones de Manila, Siquirres. MSc Thesis, CATIE; UCR, Turrialba, Costa Rica (in Spanish). url: http://hdl.handle.net/11554/4991

Li, S.-H., Huang, Y.-J., and Brown, J. L. (1997). Isolation of tetranucleotide microsatellites from the Mexican jay *Aphelocoma ultramarina*. *Molecular Ecology* 6(5), 499–501. doi: https://doi.org/10.1046/j.1365-294x. 1997.00215.x

Liao, M., Li, C., Qi, J., and Hong, Y. (2010). Isolation and characterization of 11 polymorphic microsatellite loci for *Gmelina arborea*. *Molecular Ecology Resources* 10, 1106–1108. doi: https://doi.org/10.1111/j.1755-0998.2010.02916.x

Lodhi, M. A., Ye, G.-N., Weeden, N. F., and Reisch, B. I. (1994). A simple and efficient method for DNA extraction from grapevine cultivars and *Vitis* species. *Plant Molecular Biology Reporter* 12(1), 6–13. doi: https://doi.org/10.1007/bf02668658

Murillo-Gamboa, O., Salas-Rodríguez, A., Murillo-Cruz, R., and Ávila Arias, C. (2016). Tasa de avance de la pudrición del tronco en melina *Gmelina arborea* Roxb. y posibilidades de manejo. *Revista Forestal Mesoamericana Kurú* 13(32), 40–40. (in Spanish). doi: https://doi.org/10.18845/rfmk.v0i0.2551

Peakall, R. and Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1), 288–295. doi: https://doi.org/10.1111/j.1471-8286.2005.01155.x

Rohlf, F. J. (1993). NTSYS-pc : numerical taxonomy and multivariate analysis system.

Rosero, C., Argout, X., Ruiz, M., and Teran, W. (2011). A drought stress transcriptome profiling as the first genomic resource for white teak - Gamhar - (*Gmelina arborea* Roxb) and related species. *BMC Proceedings* 5 (Suppl 7), P178–P178. doi: https://doi.org/10.1186/1753-6561-5-s7-p178

Rozen, S. and Skaletsky, H. J. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols,* ed. Misener, S. and Kravetz, S., (Totowa, NJ: Humana Press), volume 132 of *Methods in Molecular Biology™*, 365-386.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18(2), 233–234. doi: https://doi.org/10.1038/72708

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3), 535–538. doi: https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vergara, A., Mercado, T., Jarma, A., and Gatti, K. (2017). Requerimiento Hídrico de *Gmelina arborea* en Etapa de Vivero Bajo Condiciones Controladas (Water Requirement of *Gmelina arborea* on Nursery Stage Under Controlled Conditions). *Floresta e Ambiente* 24, e20150129. (in Spanish). doi: http://doi.org/10.1590/2179-8087.012915

Wee, A. K. S., Li, C., Dvorak, W. S., and Hong, Y. (2012). Genetic diversity in natural populations of *Gmelina arborea*: implications for breeding and conservation. *New Forests* 43(4), 411–428. doi: https://doi.org/10.1007/s11056-011-9288-2

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2020), 1 (2), 29–41
DOI: 10.46265/genresj.BNHB8715
https://www.genresj.org
ISSN: 2708-3764

Robust microsatellite markers for hybrid analysis between domesticated pigs and wild boar

Donovan Anderson [*,a], **Yuki Negishi** [a], **Rio Toma** [a], **Junco Nagata** [b], **Hidetoshi Tamate** [c] **and Shingo Kaneko** [a,d]

[a] *Symbiotic Systems Science and Technology, Fukushima University, Fukushima City, Fukushima, Japan*
[b] *Forestry and Forest Products Research Institute, Ibaraki, Tsukuba, Japan*
[c] *Department of Biology, Yamagata University, Yamagata City, Yamagata, Japan*
[d] *Institute of Environmental Radioactivity, Fukushima University, Fukushima City, Fukushima, Japan*

Abstract: Hybridization between wild boar (*Sus scrofa*) and their domestic relative, pigs, is a global issue and gene flow between these populations has been known to negatively impact biodiversity with increased aggression, litter sizes, and growth. However, establishing a cost-effective analysis for long-term monitoring of possible gene flow of wild pigs into wild boar populations is challenging due to common alleles at multiple loci and often it is difficult to distinguish boar specific lineages. Therefore, there is a need to select loci with lineage specific alleles for hybrid detection. To determine these loci, we calculated allele frequencies and polymorphism measurements from successfully amplified microsatellite loci with DNA extracted from domestic pigs and wild boar populations from the period prior to, and after, the evacuations and disasters in Fukushima, Japan, in 2011, which resulted in an uncontrolled release of domestic pigs. Thirty-two microsatellite loci showed pig putative alleles suggesting these selected loci can be useful genetic markers. Seventeen loci successfully distinguished pig and wild boar hybridization in Fukushima populations. Identified loci from this study provide a cost-efficient tool for genetic analysis and will provide a wealth of information on how an uncontrolled release of domestic livestock from natural or anthropogenic disasters may impact their wild relatives.

Keywords: microsatellite, hybridization, alleles, pigs, polymorphism

Citation: Anderson, D., Negishi, Y., Toma, R., Nagata, J., Tamate, H., Kaneko, S. (2020). Robust microsatellite markers for hybrid analysis between domesticated pigs and wild boar. *Genetic Resources* 1 (2), 29–41. doi: 10.46265/genresj.BNHB8715.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Hybridization between wild species and their domesticated relatives has been detected in multiple environments across the globe (Pierpaoli *et al*, 2003; Godinho *et al*, 2011; Goedbloed *et al*, 2013a). Invasive species and hybridized individuals compete with native populations, and cause negative impacts to biodiversity (Rhymer and Simberloff, 1996; Randi, 2008; Harrison and Larson, 2014). Invasive pigs are known to successfully disperse in wild environments and cause considerable impact on the gene pool of native wild boar populations (Vernesi *et al*, 2003; Koutsogiannouli *et al*, 2010; Goedbloed *et al*, 2013b). Multiple countries have implemented management programs to reduce wild boar population expansion (Waithman *et al*, 1999; Scandura *et al*, 2008; Saito *et al*, 2011), but hybrid individuals may have increased litter sizes, aggression, and growth rates (Goedbloed *et al*, 2013b; Dzialuk *et al*, 2018). Areas of suspected hybridization between invasive pigs and wild boar populations should be continuously monitored to understand the extent of introgression of pig genes in the wild boar gene pool.

Microsatellite marker analysis is a well-established monitoring tool to evaluate possible introgression of invasive species and hybridization detection (Nijman *et al*, 2003; Randi, 2008; Uemura *et al*, 2018). The selection of reliable microsatellite markers by optimizing amplification protocols prior to monitoring a target population is of great importance because it has consequences for subsequent genotyping (Hoffman and Amos, 2005; Kolodziej *et al*, 2012). However,

*Corresponding author: Donovan Anderson (s1871003@ipc.fukushima-u.ac.jp)

Received: 22.06.2020 | Accepted: 30.11.2020 | Published online: 29.12.2020

genotyping hybrid individuals (e.g. crossed pig and boar) can be challenging due to common or shared alleles at multiple loci (Larson *et al*, 2005; Grossi *et al*, 2006; Choi *et al*, 2014) and it is necessary to select suitable loci with lineage specific alleles for hybridization detection. Validating appropriate microsatellite markers with possible hybridized wild boar and the pigs involved in the hybridization will provide the necessary genetic composition data to develop a cost-efficient monitoring tool to evaluate the introgression of pig genes to the wild boar gene pool. Such cost-efficient analyses have provided monitoring opportunities to estimate abundancy of hybrids (Qi *et al*, 2010; Matsumoto *et al*, 2019), population characteristics (Goedbloed *et al*, 2013a; Sharma *et al*, 2013) and local genetic structures (Tadano *et al*, 2016; Touma *et al*, 2020) in animals.

Genetic diversity and ancestry of wild boar have been well studied, including areas of South East Asia, and information from hybridization occurrences with domestic pigs is of increasing interest due to possible genetic alterations (Choi *et al*, 2014; Todesco *et al*, 2016). Wild boar populations inhabiting Fukushima prefecture, in Japan, are suggested to be threatened by hybridization following the uncontrolled release of domesticated pigs after mandated evacuations due to the Fukushima nuclear disasters in 2011 (Okuda *et al*, 2018; Anderson *et al*, 2019). Additionally, hybridization in this area has not altered the morphological characteristics of wild boar (Anderson *et al*, 2019) and possible hybrids can only be detected using DNA. Thus, estimating appropriate genotypes of wild boar from the period prior to 2011, after 2011, and from domestic pigs in this area, with microsatellite markers will provide an important source of information for better understanding hybridization effects with native species following such events. Adequate selection of microsatellite markers from this area will establish a cost-efficient tool to easily distinguish if a wild boar population has been impacted by hybridization.

In this study, we selected robust microsatellite markers used in European and Asian pig studies (Rohrer *et al*, 1994; Krause *et al*, 2002; Karlskov-Mortensen *et al*, 2007; FAO, 2011) that could differentiate wild boar or pig alleles. Our goal for this study was two-fold: First, we aimed to select useful microsatellite markers for hybrid analysis between domesticated pig and wild boar populations; and second, to use these loci to perform a preliminary check of the introgression of pig alleles into wild boar populations in Fukushima prefecture following the disasters in 2011.

Materials and Methods

Analysed samples and DNA extraction

Thirty-one muscle tissue samples were collected from three populations (hereafter referred to as Pop1, Pop2 and Pop3) and were selected based on mitochondrial DNA (mtDNA) haplotype and year sampled. Sample haplotype and date were prioritized for optimal determination of reliable microsatellite screening of hybridization between wild boar and domesticated pigs after the Fukushima disasters in 2011. Pop1 samples were from 10 unrelated domestic pigs (*Sus scrofa domesticus*) that were collected from a local pig slaughterhouse or local markets within Fukushima prefecture in 2016-2017. Pop2 samples were from 13 wild boar (*Sus scrofa*) muscle samples that were collected in 2003-2004, prior to the evacuations and Fukushima disasters, from a wild boar population in northern Ibaraki prefecture, south of Fukushima prefecture. The mtDNA analysis has shown that this population is the same haplotype (D42172) and has extremely high genetic similarity to the wild boar population in eastern Fukushima prefecture (Nagata *et al*, 2006). Pop3 samples were collected in 2015-2016, after the Fukushima disasters, from eight suggested hybridized wild boar that had a typical mtDNA haplotype of pig (suggested pig ancestor in maternal lineage; MK801664, see Anderson *et al* (2019)). All animals were legally culled by licensed hunters, and this entire study was approved by Fukushima University's Institutional Animal Care and Use Committee. All experiments were performed in accordance with relevant guidelines and regulations. All samples were stored individually at −20 °C in 99.5% ethanol until extraction. Total genomic DNA was extracted from muscle tissue using the Gentra Puregene Tissue Kit (QIAGEN), according to manufacturer's instructions.

Microsatellite loci genotyping

A total of 52 unlinked microsatellite loci were selected from previously developed phage libraries (Rohrer *et al*, 1994; Krause *et al*, 2002; Karlskov-Mortensen *et al*, 2007) and recommended microsatellite markers from the Food and Agriculture Organisation of the United Nations database (FAO, 2011) and screened for amplification success on all 31 samples.

PCR amplification was performed in 5 μL reactions using the QIAGEN Multiplex PCR Kit (QIAGEN) and a protocol for fluorescent dye-label (Blacket *et al*, 2012). Each sample reaction contained 10 to 20 ng of genomic template DNA, 2.5 μL of Multiplex PCR Master Mix, 0.1 μM of forward primer, 0.2 μM of reverse primer, and 0.1 μM of fluorescently labeled primer. Amplification conditions consisted of 95 °C for 15 minutes followed by 33 cycles of denaturation at 94 °C for 30 seconds, annealing at 57 °C for 1.5 minutes, and extension at 72 °C for 1 minute and an extension at 60 °C for 30 minutes. All thermal cycling conditions used in T100 thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Product sizes were determined using an ABI PRISM 3130 Genetic Analyzer and GeneMapper software (Applied Biosystems, Foster City, CA, USA).

Characterization of microsatellite markers

Successful markers were identified after our initial screening by clear peak patterns following amplifications. Number of alleles (N_A), observed heterozygosity

Table 1. Characteristics of 32 microsatellite markers selected. T_a = annealing temperature

Locus	Primer sequence (5' → 3') Forward/Reverse	Repeat motif	Range of alleles (bp)	T_a (°C)	Fluorescent label	Genebank accession No.
Sw632	TGGGTTGAAAGATTTCCCAA	$(AC)_{21}$	115-138	55	FAM	AF225099
	GGAGTCAGTACTTTGGCTTGA					
S0090	CCAAGACTGCCTTGTAGGTGAATA	$(AC)_{24}$	227-253	55	FAM	M95002
	GCTATCAAGTATTGTACCATTAGG					
Sw24	CTTTGGGTGGAGTGTGTGC	$(GT)_{13}$	99-135	55	VIC	AF235245
	ATCCAAATGCTGCAAGCG					
Swr1941	AGAAAGCAATTTGATTTGCATAATC	$(TG)_{20}$	215-255	55	VIC	AF253904
	ACAAGGACCTACTGTATAGCACAGG					
Sw857	TGAGAGGTCAGTTACAGAAGACC	$(CA)_{22}$	165-187	55	NED	AF225105
	GATCCTCCTCCAAATCCCAT					
S0228	GGCATAGGCTGGCAGCAACA	$(AC)_{17}$	93-112	55	PET	L29195
	AGCCCACCTCATCTTATCTACACT					
Sw2008	CAGGCCAGAGTAGCGTGC	$(GT)_{25}$	148-170	55	FAM	AF253773
	CAGTCCTCCCAAAAATAACATG					
Sw240	AGAAATTAGTGCCTCAAATTGG	$(TG)_{17}$	164-186	55	VIC	AF235246
	AAACCATTAAGTCCCTAGCAAA					
S0097	GACCTATCTAATGTCATTATAGT	$(AC)_{28}$	135-155	55	NED	M95020
	TTCCTCCTAGAGTTGACAAACTT					
UMNp147	GCCTTCGTTACATGGCATTC	$(GT)_{23}$	151-167	58	PET	AF511119
	TCTCTGTGAGGTCATGGTGG					
UMNp239	CTTACAAAACCACCACCATCG	$(AC)_{18}$	96-112	60	FAM	AF511146
	TCAATATCAACATTGCGTGTTG					
UMNp296	CAGGGAACTCTCTTCAATATCC	$(TG)_{13}$	151-181	58	NED	AF511184
	ACATTTGATTTCCAAAGTTGTG					
UMNp298	GCTATAAGAACCGCCTCATTG	$(GT)_{22}$	157-169	58	NED	AF511185
	TGTGTGCTGCTGAAGCATG					
UMNp351	TCAGTGTCACCCCTCATCAC	$(AC)_{15}$	143-169	58	FAM	AF511222
	TCTCCTTGACCTTCTAAGCACC					
UMNp358	AAGTCATTTCACACCTCTGTGC	$(CA)_{22}$	160-176	58	VIC	AF511230
	CGTTGCAGTTACTATTCCAAGC					
UMNp362	GATGTGTAGCTGATTTGCAATG	$(AC)_{21}$	125-135	60	PET	AF511231
	GACAAGAATCTGAAAAGGAGCG					
UMNp381	CCGATTAGACCCCTAGTCTGG	$(AC)_{22}$	169-185	60	NED	AF511244

Continued on next page

Table 1 continued

Locus	Primer sequence (5' → 3') Forward/Reverse	Repeat motif	Range of alleles (bp)	T_a (°C)	Fluorescent label	Genebank accession No.
	CAGATTAGCGTTCCCTGTTTG					
UMNp405	CAGAGTTCACCTCTCCCTTTAC	$(AC)_{21}$	148-162	62	VIC	AF511255
	TCCTTGCTGAGTCCCAGG					
UMNp442	ATCCAAGCTGCTGAAGTTGG	$(TG)_{12}$	122-124	60	NED	AF511283
	AAACATTTCCACAAGAAAATGG					
UMNp453	TCATTCTCTATCTCAAGATGCATG	$(AC)_{17}$	122-140	58	PET	AF511291
	CTGAGGTACCTTTGCCTAGAGG					
UMNp480	AGTGATTTCTGCCCAGGATG	$(TG)_{21}$	143-155	58	VIC	AF511308
	CCTAGGAATTTCCCTCTGCC					
UMNp485	CCTCAGGCTCAGCTCTGC	$(TG)_{17}$	187-213	57	PET	AF511313
	GTTGTCCGTGAGTCCCTAGC					
UMNp489	AAGCACCATAGGAGAAGACTGG	$(AC)_{12}$	115-141	60	PET	AF511317
	CTCGGAAGCAAGTAAGTGGG					
UMNp494	CTGCCTGATTGGCACATTAG	$(AC)_{23}$	114-142	60	FAM	AF511320
	GGTAATGGGAAAGCCTAGCC					
UMNp500	TGAGGCTATCACCTGCAGTG	$(AG)_{24}$	229-251	60	FAM	AF511324
	GACTGAACCCTTAACAGATGGG					
UMNp502	TGGCAAACGTTGCTTTAGG	$(GT)_{22}$	164-172	60	VIC	AF511325
	TAGGGAAATATCTGAAATCTAAAATG					
UMNp509	AAACTACATCCATTCTCTTGGG	$(GT)_{21}$	138-164	60	FAM	AF511328
	GTTGTGCCAGTTACACTTCTGC					
UMNp511	GATCACTGTGTGAGTGCATGC	$(GT)_{14}$	107-117	60	VIC	AF511329
	AACAGAGTTCCATTTTGCGG					
UMNp539	CAACGTTGCTGTGGCTGTAG	$(CA)_{32}$	171-181	60	NED	AF511346
	TTCTGGTTTATGGTTCCCATG					
UMNp548	TCCAAGTTAGACTGCCTGCC	$(CA)_{14}$	172-180	60	NED	AF511353
	ACTGCTGCTTATTTCTCAAGGG					
UMNp610	CTTTGGCTCAATCTCATTCATG	$(AC)_{33}$	168-178	60	VIC	AF511389
	TGGGCTTTTGAAAATTTAAATG					
UMNp640	TATGCCATGTGCGTGGTC	$(AC)_{13}$	123-145	60	FAM	AF511399
	ACAAACTGCACCACAGAATAGC					

(H_O), and expected heterozygosity (H_E) were calculated using GenAIEX version 6.5 (Peakall and Smouse, 2012) for successfully amplified loci. Calculation of inbreeding coefficients (F_{IS}) and test of deviation from Hardy–Weinberg equilibrium (HWE) for polymorphic loci were tested using FSTAT version 2.9.3 (Goudet, 1995). Allele frequency in each locus for Pop1, Pop2, and Pop3 was calculated using GenAlEx version 6.41 (Peakall and Smouse, 2006). Genetic differentiation among Pop1, Pop2, and Pop3 (i.e. differentiation among pig, wild boar, and hybrids) was evaluated using AMOVA, calculating pairwise codominant genotypic distances (Smouse and Peakall, 1999), and performing principal coordinates analysis (PCoA) using GenAlEx version 6.41 (Peakall and Smouse, 2006).

Results

Of the initial 52 microsatellite loci selected, 32 loci were successfully amplified with all wild boar and pig samples. Marker information is provided in Table 1. Twenty loci were eliminated based on low amplification success or unclear peak patterns in wild boar DNA samples collected from Pop2 and Pop3. Polymorphism measurements for the 32 amplified microsatellite loci in Pop1, Pop2, and Pop3 are summarized in Table 2. For Pop1, H_O and H_E per locus ranged from 0.10 to 1.00 (mean, 0.64) and from 0.10 to 0.82 (mean, 0.65), respectively. The range of F_{IS} was -0.46 to 0.47 (mean, 0). For Pop2, H_O and H_E per locus ranged from 0.00 to 0.77 (mean, 0.36) and from 0.00 to 0.73 (mean, 0.39), respectively. The range of F_{IS} was -0.28 to 0.85 (mean, 0.07). For Pop3, the H_O and H_E per locus ranged from 0.00 to 0.88 (mean, 0.45) and from 0.00 to 0.76 (mean, 0.44), respectively. The range of F_{IS} was -0.62 to 1.00 (mean, -0.01). All 32 loci showed no evidence of significant deviation from HWE ($P > 0.05$).

In total, 231 putative alleles were identified that ranged from 1 to 8 per locus (mean, 4), as outlined in Table 2 (No. alleles). The mean number of alleles was 5.1, 3.2, and 3.3 for Pop1, Pop2, and Pop3, respectively. Of the 231 alleles, 52 (23%) were putative to the wild boar populations and 95 (41%) were putative to domestic pigs (Table 3). Additionally, 68 (30%) alleles were shared by pigs and one of the wild boar populations or by all three populations. Among the 68 shared alleles, 21 were shared by pigs and the wild boar population from the period after the Fukushima disasters in 2011 (bolded alleles in Table 4), indicating introgression of pig genes into the wild boar gene pool. The allele frequencies of amplified microsatellite markers, including those that distinguished these shared alleles between Pop1 and Pop3, are provided in Table 4. AMOVA suggested strong genetic differentiation between the three populations ($F_{ST} = 0.318$, $p < 0.001$). Genetic differentiation is also well-supported by clear divisions among the three populations with PCoA (Figure 1). Pop1 is uniquely distinguished along the first axis and Pop2/Pop3 are distinguished along the second axis. Furthermore, codominant genotypic distances describe 35.40% and 7.4% of the variation with the first and second axes, respectively. Taken together, our data strongly indicate genetic differentiation of pigs, wild boar from the period prior to, and after, the disasters in Fukushima in 2011.

Discussion

All 32 loci selected from previous studies (Rohrer *et al*, 1994; Krause *et al*, 2002; Karlskov-Mortensen *et al*, 2007; FAO, 2011) showed pig putative alleles suggesting they can be useful for wild boar and pig hybrid analysis in Fukushima and elsewhere. The presence of pig-specific alleles at certain loci depends on the genetic composition of the target population and the pig population involved in the hybridization. In this study, pig samples were from slaughterhouses and farms nearby the evacuated area to improve the likelihood of detecting newly introgressive pig alleles in the hybridized wild boar. One of the pigs sampled had the mtDNA haplotype that was the same as hybrid wild boar in Fukushima (Anderson *et al*, 2019), and the sampled pig population in this study had high genetic variation (mean $N_A = 5.1$). Therefore, we were confident in the representation of the pig genetic composition involved in the hybridization for this study and were able to distinguish an appropriate set of markers for hybrid analysis. Our selective use of markers with low frequencies of common alleles in source pig individuals and target wild boar populations is highly suggested for cost-effective analysis.

The highest number of alleles was observed in the pig population, which was expected because of human mediated translocations of domestic populations with high genetic diversity (Scandura *et al*, 2008; Yang *et al*, 2017). Additionally, if the 95 pig putative alleles were excluded, then 27% of the detected alleles were shared by all three populations in this study (Table 3). The high percentage of shared alleles between pig and wild boar verifies the challenge of identifying appropriate markers for hybridization analysis.

Seventeen of the 32 microsatellite loci distinguished hybridization between pigs and wild boar in this study and these can be used as robust markers, specifically for wild boar populations in Fukushima. These seventeen markers detected at least one of the 21 alleles that were only shared between Pop1 and Pop3 (bolded alleles in Table 4). Alleles only shared by Pop1 and Pop3, and not detected in Pop2, would suggest that the alleles were introgressive through mixing of pigs and wild boar during the period after the 2011 evacuations and Fukushima disasters. The higher percent of total shared alleles between Pop1 and Pop3 (9%), compared to Pop1 and Pop2 (4%), indicates that there is likely more genetic mixing between Pop1 and Pop3 (Table 3), which would also support the hypothesis of hybridization occurring after disasters in 2011.

Table 2. Polymorphism measurements of microsatellite loci of each sampled population. n = No. samples, N_A = No. alleles, H_O = observed heterozygosity, H_E = heterozygosity, F_{IS} = breeding coefficient

Locus	Pop1 (n = 10)				Pop2 (n = 13)				Pop3 (n = 8)			
	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}
Sw632	5	0.80	0.79	-0.01	2	0.38	0.50	0.23	3	0.50	0.40	-0.25
S0090	6	0.90	0.77	-0.17	2	0.08	0.07	-0.04	3	0.43	0.36	-0.20
Sw24	6	0.56	0.77	0.27	4	0.31	0.39	0.21	6	0.88	0.76	-0.15
Swr1941	5	0.60	0.72	0.16	2	0.15	0.26	0.41	2	0.14	0.13	-0.08
Sw857	6	1.00	0.75	-0.33	3	0.31	0.27	-0.14	3	0.38	0.32	-0.17
S0228	5	0.40	0.59	0.32	5	0.77	0.69	-0.11	4	0.38	0.66	0.44
Sw2008	5	0.90	0.70	-0.29	4	0.31	0.38	0.19	3	0.63	0.48	-0.31
Sw240	6	0.70	0.77	0.08	2	0.31	0.43	0.28	3	0.50	0.55	0.10
S0097	6	0.60	0.79	0.24	4	0.38	0.48	0.19	4	0.63	0.63	0.01
UMNp147	7	0.80	0.81	0.01	4	0.08	0.51	0.85	3	0.13	0.23	0.45
UMNp239	4	0.50	0.62	0.19	4	0.62	0.56	-0.10	4	0.75	0.68	-0.10
UMNp296	7	0.80	0.81	0.01	3	0.69	0.61	-0.14	5	0.75	0.66	-0.14
UMNp298	6	0.80	0.76	-0.05	3	0.23	0.21	-0.10	2	0.13	0.49	0.75
UMNp351	5	0.40	0.76	0.47	5	0.62	0.49	-0.26	3	0.50	0.57	0.12
UMNp358	7	0.60	0.81	0.25	4	0.31	0.33	0.08	2	0.50	0.38	-0.33
UMNp362	2	0.30	0.38	0.20	1	0.00	0.00	N/A	1	0.00	0.00	N/A
UMNp381	2	0.10	0.10	-0.05	5	0.62	0.73	0.15	3	0.13	0.23	0.45
UMNp405	3	0.40	0.34	-0.19	1	0.00	0.00	N/A	2	0.13	0.12	-0.07
UMNp442	3	0.30	0.27	-0.13	1	0.00	0.00	N/A	2	0.00	0.22	1.00
UMNp453	6	0.80	0.73	-0.10	2	0.08	0.07	-0.04	1	0.00	0.00	N/A
UMNp480	5	0.80	0.60	-0.34	2	0.31	0.36	0.13	2	0.38	0.30	-0.23
UMNp485	6	0.80	0.61	-0.32	3	0.38	0.52	0.26	3	0.88	0.54	-0.62
UMNp489	4	0.56	0.52	-0.07	3	0.54	0.48	-0.13	5	0.86	0.65	-0.31
UMNp494	4	0.50	0.59	0.15	4	0.62	0.48	-0.28	4	0.75	0.67	-0.12
UMNp500	5	0.70	0.76	0.07	4	0.46	0.63	0.27	6	0.88	0.64	-0.37
UMNp502	4	0.80	0.67	-0.20	5	0.31	0.34	0.09	5	0.50	0.50	0.00
UMNp509	5	0.70	0.67	-0.05	4	0.38	0.49	0.21	5	0.63	0.50	-0.25
UMNp511	4	0.90	0.62	-0.46	4	0.77	0.67	-0.16	4	0.38	0.62	0.39
UMNp539	8	0.70	0.82	0.14	3	0.46	0.41	-0.12	4	0.63	0.55	-0.13
UMNp548	3	0.50	0.51	0.01	2	0.15	0.36	0.57	2	0.13	0.12	-0.07
UMNp610	6	0.60	0.72	0.17	3	0.15	0.14	-0.06	4	0.38	0.41	0.09
UMNp640	7	0.70	0.70	-0.01	3	0.69	0.54	-0.27	4	0.75	0.65	-0.16

Table 3. Number of putative and shared alleles by population with putative allele origin. Percentage indicates proportion of alleles related to total alleles detected in this study.

Population(s)	Putative allele origin	Number of alleles (% of total alleles)
Pop1 (pig)	Pig	95 (41%)
Pop2 (wild boar)	Wild boar	18 (8%)
Pop3 (hybrid boar)	Pig and/or wild boar	16 (7%)
Shared Pop1 and Pop2	Pig and/or wild boar	10 (4%)
Shared Pop1 and Pop3	Pig	21 (9%)
Shared Pop2 and Pop3	Wild boar	34 (15%)
Shared Pop1, Pop2, and Pop3	Pig and/or wild boar	37 (16%)
	Total alleles	231

Table 4. Allele frequencies for selected microsatellite markers. Pop1 n=10; Pop2 n=13; Pop3 n=8. Allele^{P} indicates putative pig alleles. Allele^{B} indicates putative wild boar alleles. Bolded alleles indicate pig alleles putatively introgressed into the wild boar population.

Locus	Allele	Frequency			Locus	Allele	Frequency		
		Pop1	Pop2	Pop3			Pop1	Pop2	Pop3
Sw2008	102^{B}	0	0.15	0.69	*UMNp362*	124^{P}	0.25	0	0
	106^{B}	0	0.77	0.19		126^{P}	0.75	0	0
	108	0.05	0.04	0		132^{B}	0	1	1
	110^{P}	0.25	0	0	*UMNp381*	**165^{P}**	0.05	0	**0.06**
	112	0.45	0.04	0		167^{B}	0	0.35	0
	114^{P}	0.1	0	**0.13**		173	0.95	0.27	0.88
	116^{P}	0.15	0	0		175^{B}	0	0.08	0.06
Sw24	121	0	0	0.06		179^{B}	0	0.27	0
	123^{P}	0.28	0	0		181^{B}	0	0.04	0
	125	0.06	0.77	0.13	*UMNp405*	140	0	0	0.06
	127^{B}	0	0.08	0.06		144	0.15	1	0.94
	129	0.33	0.08	0.38		148^{P}	0.05	0	0
	131	0.17	0.08	0.25		156^{P}	0.8	0	0
	133^{P}	0.11	0	0	*UMNp442*	118	0	0	0.13
	135^{P}	0.06	0	0		120	0.1	1	0.88
	139	0	0	0.13		122^{P}	0.85	0	0
Sw240	107^{P}	0.1	0	0		124^{P}	0.05	0	0
	111^{P}	0.4	0	0	*UMNp453*	122	0.05	0.96	1
	113^{P}	0.15	0	0		130^{P}	0.05	0	0
	119^{B}	0	0.69	0.5		132^{P}	0.1	0	0
	121^{P}	0.1	0	0		134^{P}	0.1	0	0
	123^{P}	0.1	0	0		136	0.3	0.04	0
	125	0.15	0.31	0.44		138^{P}	0.4	0	0
	127	0	0	0.06	*UMNp480*	136^{P}	0.15	0	0
Sw632	160	0	0.5	0.75		138^{B}	0	0.77	0.81
	172^{P}	0.25	0	0		144^{P}	0.6	0	0
	174	0.2	0.5	0.19		146	0.1	0.23	0.19
	180^{P}	0.25	0	0		148^{P}	0.1	0	0
	182^{P}	0.15	0	**0.06**		152^{P}	0.05	0	0
	184^{P}	0.15	0	0	*UMNp485*	185^{B}	0	0.58	0.38
Sw857	156^{P}	0.05	0	0		193^{B}	0	0.38	0.56
	164^{P}	0.15	0	0		195^{B}	0	0.04	0
	166^{P}	0.35	0	0		203^{P}	0.1	0	0
	168	0.05	0.12	0.06		207^{P}	0.6	0	0
	170	0.3	0.85	0.81		209^{P}	0.1	0	0
	172	0.1	0.04	0.13		**211^{P}**	0.05	0	**0.06**
						219^{P}	0.05	0	0
						225^{P}	0.1	0	0

Continued on next page

Table 4 continued

Locus	Allele	Frequency			Locus	Allele	Frequency		
		Pop1	Pop2	Pop3			Pop1	Pop2	Pop3
Swr1941	224^{B}	0	0.85	0.93	*UMNp489*	**116^{P}**	0.11	0	**0.07**
	228	0.3	0.15	0.07		**126^{P}**	0.67	0	**0.07**
	232^{P}	0.4	0	0		128	0.11	0.65	0.5
	234^{P}	0.15	0	0		130	0.11	0.04	0
	236^{P}	0.1	0	0		134	0	0	0.07
	238^{P}	0.05	0	0		140^{B}	0	0.31	0.29
S0090	252^{P}	0.05	0	0	*UMNp494*	108^{P}	0.5	0	0
	254	0.35	0.96	0.79		114	0	0	0.06
	256	0.15	0.04	0.07		126^{B}	0	0.15	0.31
	258^{P}	0.2	0	**0.14**		128^{B}	0	0.12	0.19
	260^{P}	0.2	0	0		130	0.4	0.69	0.44
	262^{P}	0.05	0	0		132^{P}	0.05	0	0
S0097	230^{P}	0.1	0	0		134^{P}	0.05	0	0
	232^{P}	0.2	0	0		141^{B}	0	0.04	0
	238	0	0	0.13	*UMNp500*	219^{P}	0.3	0	0
	240^{B}	0	0.08	0		221^{P}	0.15	0	0
	244^{B}	0	0.04	0		223	0.05	0.12	0.06
	250^{P}	0.3	0	0		**225^{P}**	0.3	0	**0.13**
	252^{B}	0	0.69	0.5		227^{B}	0	0.54	0.56
	254^{P}	0.1	0	0		229	0	0	0.13
	256^{P}	0.25	0	**0.06**		237^{B}	0	0.12	0.06
	258^{P}	0.05	0	0		239^{B}	0	0.23	0.06
	260^{B}	0	0.19	0.31		245^{P}	0.2	0	0
S0228	239^{P}	0.6	0	0	*UMNp502*	156	0.2	0.81	0.69
	241^{P}	0.05	0	0		158	0.05	0.04	0.06
	243^{B}	0	0.42	0		160	0.45	0.04	0.06
	245	0.05	0.04	0		162^{P}	0.3	0	0
	247^{B}	0	0.23	0.19		164	0	0	0.13
	251	0	0	0.19		166^{B}	0	0.08	0.06
	255^{B}	0	0.27	0.13		168^{B}	0	0.04	0
	257	0.2	0.04	0.5	*UMNp509*	133	0	0	0.06
	259^{P}	0.1	0	0		141^{B}	0	0.08	0.13
UMNp147	141^{P}	0.3	0	0		143^{B}	0	0.12	0.69
	147^{P}	0.1	0	0		145^{B}	0	0.69	0.06
	148^{P}	0.05	0	0		147^{P}	0.05	0	0
	149^{P}	0.2	0	0		151	0.4	0.12	0
	153^{B}	0	0.04	0.06		**153^{P}**	0.4	0	**0.06**
	157^{P}	0.2	0	0		155^{P}	0.1	0	0
	159	0.05	0.23	0.88		157^{P}	0.05	0	0
	163	0	0	0.06					
	165^{B}	0	0.65	0					
	167	0.1	0.08	0					

Continued on next page

Table 4 continued

Locus	Allele	Frequency			Locus	Allele	Frequency		
		Pop1	Pop2	Pop3			Pop1	Pop2	Pop3
UMNp239	90^{B}	0	0.54	0.31	*UMNp511*	98	0	0	0.13
	92^{P}	0.35	0	0		102^{B}	0	0.46	0.56
	94^{P}	0.5	0	**0.13**		**106^{P}**	0.35	0	**0.13**
	96^{P}	0.1	0	0		108	0.1	0.12	0.19
	98^{P}	0.05	0	**0.13**		110	0.5	0.31	0
	104^{B}	0	0.38	0.44		112	0.05	0.12	0
	106^{B}	0	0.04	0	*UMNp539*	159^{P}	0.05	0	0
	108^{B}	0	0.04	0		161^{P}	0.05	0	0
UMNp296	147^{B}	0	0.15	0.06		165^{P}	0.25	0	0
	149^{P}	0.1	0	**0.06**		167	0.05	0.23	0.19
	151^{P}	0.15	0	0		169^{P}	0.2	0	0
	155^{P}	0.2	0	**0.06**		171	0.05	0.04	0
	159^{P}	0.3	0	0		173	0.25	0.73	0.63
	161^{P}	0.05	0	0		175	0	0	0.13
	167^{P}	0.15	0	0		**177^{P}**	0.1	0	**0.06**
	171	0.05	0.35	0.44	*UMNp548*	**168^{P}**	0.65	0	**0.06**
	177^{B}	0	0.5	0.38		170	0.25	0.23	0
UMNp298	153	0.25	0.88	0.56		176^{B}	0	0.77	0.94
	155	0.2	0.08	0.44		178^{P}	0.1	0	0
	167	0.35	0.04	0	*UMNp610*	**162^{P}**	0.45	0	**0.06**
	169^{P}	0.05	0	0		164	0.2	0.92	0.75
	185^{P}	0.1	0	0		166^{B}	0	0.04	0
	191^{P}	0.05	0	0		170^{B}	0	0.04	0.06
UMNp351	130^{P}	0.1	0	0		**174^{P}**	0.15	0	**0.13**
	136^{P}	0.3	0	0		178^{P}	0.05	0	0
	140	0.3	0.15	0.56		180^{P}	0.05	0	0
	142	0.2	0.08	0		186^{P}	0.1	0	0
	144^{P}	0.1	0	**0.13**	*UMNp640*	113^{P}	0.05	0	0
	156^{B}	0	0.04	0		117	0.45	0.23	0.38
	162^{B}	0	0.69	0.31		119	0	0	0.06
	166^{B}	0	0.04	0		121^{P}	0.3	0	0
UMNp358	154	0.05	0.81	0.75		127	0.05	0.62	0.44
	158	0.25	0.08	0.25		**129^{P}**	0.05	0	**0.13**
	160^{P}	0.2	0	0		135^{P}	0.05	0	0
	164^{P}	0.05	0	0		137^{P}	0.05	0	0
	166^{B}	0	0.08	0		139^{B}	0	0.15	0
	168^{B}	0	0.04	0					
	169^{P}	0.05	0	0					
	170^{P}	0.15	0	0					
	172^{P}	0.25	0	0					

The loci identified in this study (Table 4) provide a unique tool to contribute to determining a timeline of hybridization for these populations. Similar frequencies of pig alleles in other wild boar populations may suggest early stages of hybridization, as our data indicates recent occurrence of hybridization in Fukushima prefecture, following the release of domestic pigs into the wild boar populations in 2011 (Okuda *et al*, 2018). Additionally, the identified loci can contribute to determining if the introgressive alleles are being retained or lost due to natural causes (e.g. backcross) in the hybridized wild boar population using introgressive allele frequencies over time.

Studies have been published to determine wild boar and pig hybridization hotspots, recent occurrences, and genetic impacts using variable genetic markers, such as mtDNA sequence (Ishiguro *et al*, 2002; McCann *et al*, 2014), RAD-seq analysis (Goddard and Hayes, 2007; Iacolina *et al*, 2018) or microsatellite markers (Murakami *et al*, 2014). However, Next Generation Sequencing (NGS), such as RAD-seq, have disadvantages including that a large amount of high quality DNA is required, and the amount of data to be analyzed becomes demanding. Therefore, general genetic markers, such as microsatellite markers, are still useful for analysis of degraded DNA extracted from feces in the field and old specimens of bones (Kierepka *et al*, 2016). Selected robust markers from our study will show their advantages in future hybrid analysis and are cost-effective for immediate or continuous monitoring for small sample sizes or DNA analysis of degraded samples. Also, comparing NGS and microsatellite marker data from a common population in future studies will not only give more in-depth information about that target population, but will more clearly show the advantages and disadvantages of each marker.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

Thanks to Dr Hiroko Ishiniwa and Dr Kei Okuda for support with sampling. Additionally, we are thankful to all prefectural hunters for their support in obtaining samples. Funding for this study was partially provided by the Nippon Life Insurance Foundation.

Author contributions

DA and SK contributed to the writing, drafting, and execution of the manuscript. DA, YN and SK, contributed to data analysis and interpertation. All authors contributed to study design, revision, and approval of the submitted manuscript. All authors declare that the submitted work is their own and that copyright has not been breached in seeking its publication. Additionally, the submitted work has not been previously published and is not being considered elsewhere.

Conflict of interest statement

The authors declare no conflict of interest.

References

Anderson, D., Toma, R., Negishi, Y., Okuda, K., Ishiniwa, H., Hinton, T. G., Nanba, K., Tamate, H. B., and Kaneko, S. (2019). Mating of escaped domestic pigs with wild boar and possibility of their offspring migration after the Fukushima Daiichi Nuclear Power Plant accident. *Scientific Reports* 9, 11537. doi: https://doi.org/10.1038/s41598-019-47982-z

Blacket, M. J., Robin, C., Good, R. T., Lee, S. F., and Miller, A. D. (2012). Universal primers for fluorescent labelling of PCR fragments-an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources* 12(3), 456–463. doi: https://dx.doi.org/10.1111/j.1755-0998.2011.03104.x

Choi, S. K., Lee, J. E., Kim, Y. J., Min, M. S., Voloshina, I., Myslenkov, A., Oh, J. G., Kim, T. H., Markov, N., Seryodkin, I., Ishiguro, N., Yu, L., Zhang, Y. P., Lee, H., and Kim, K. S. (2014). Genetic structure of wild boar (*Sus scrofa*) populations from East Asia based on microsatellite loci analyses. *BMC Genetics* 15, 85. doi: https://doi.org/10.1186/1471-2156-15-85

Dzialuk, A., Zastempowska, E., Skórzewski, R., Twarużek, M., and Grajewski, J. (2018). High domestic pig contribution to the local gene pool of free-living European wild boar: a case study in Poland. *Mammal Research* 63(1), 65–71. doi: https://dx.doi.org/10.1007/s13364-017-0331-3

FAO (2011). Molecular genetic characterization of animal genetic resources. *FAO Animal Production and Health Guidelines* 9. url: http://www.fao.org/3/i2413e/i2413e00.pdf.

Goddard, M. E. and Hayes, B. J. (2007). Genomic selection. *Journal of Animal Breeding and Genetics* 124(6), 323–330. doi: https://dx.doi.org/10.1111/j.1439-0388.2007.00702.x

Godinho, R., Llaneza, L., Blanco, J. C., Lopes, S., Álvares, F., García, E. J., Palacios, V., Cortés, Y., Talegón, J., and Ferrand, N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology* 20(24), 5154–5166. doi: https://dx.doi.org/10.1111/j.1365-294x.2011.05345.x

Goedbloed, D. J., Megens, H. J., Hooft, P. V., Herrero-Medrano, J. M., Lutz, W., Alexandri, P., Crooijmans, R. P. M. A., Groenen, M., Wieren, S. E. V., Ydenberg, R. C., and Prins, H. H. T. (2013a). Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. *Molecular Ecology* 22(3), 856–866. doi: 10.1111/j.1365-294x.2012.05670.x

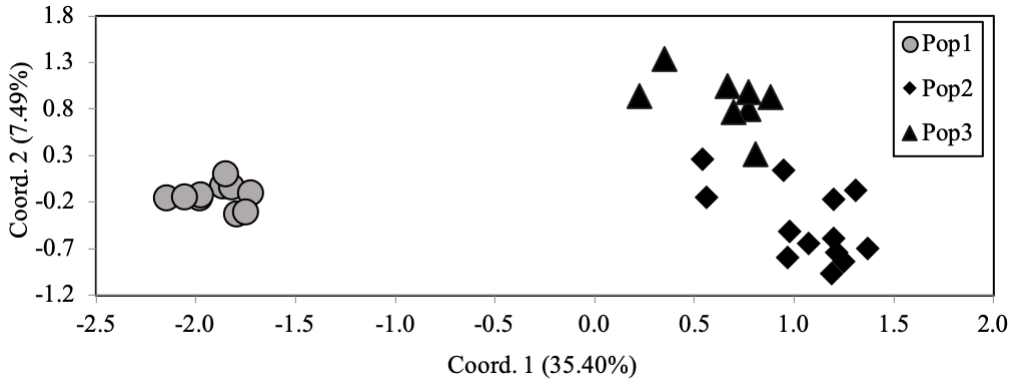

Figure 1. Principal Coordinates Analysis Plot of Pop1 (grey circle), Pop2 (black diamond), and Pop3 (black triangle) based on codominant genotypic distances. Axis 1 explains 35.4% of the variance and axis 2 explains 7.49% of the variance.

Goedbloed, D. J., van Hooft, P., Megens, H.-J., Langenbeck, K., Lutz, W., Crooijmans, R. P., van Wieren, S. E., Ydenberg, R. C., and Prins, H. H. (2013b). Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar. *BMC Genetics* 14(1), 43–43. doi: https://dx.doi.org/10.1186/1471-2156-14-43

Goudet, J. (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity* 86(6), 485–486. doi: https://dx.doi.org/10.1093/oxfordjournals.jhered.a111627

Grossi, S. F., Lui, J. F., Garcia, J. E., and Meirelles, F. V. (2006). Genetic diversity in wild (*Sus scrofa scrofa*) and domestic (*Sus scrofa domestica*) pigs and their hybrids based on polymorphism of a fragment of the D-loop region in the mitochondrial DNA. *Genetics and Molecular Research* 5(4), 564–568.

Harrison, R. G. and Larson, E. L. (2014). Hybridization, Introgression, and the Nature of Species Boundaries. *Journal of Heredity* 105(S1), 795–809. doi: https://dx.doi.org/10.1093/jhered/esu033

Hoffman, J. I. and Amos, W. (2005). Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* 14(2), 599–612. doi: https://dx.doi.org/10.1111/j.1365-294x.2004.02419.x

Iacolina, L., Pertoldi, C., Amills, M., Kusza, S., Megens, H.-J., Bâlteanu, V. A., Bakan, J., Cubric-Curik, V., Oja, R., Saarma, U., Scandura, M., Šprem, N., and Stronen, A. V. (2018). Hotspots of recent hybridization between pigs and wild boars in Europe. *Scientific Reports* 8(1), 1–10. doi: https://dx.doi.org/10.1038/s41598-018-35865-8

Ishiguro, N., Naya, Y., Horiuchi, M., and Shinagawa, M. (2002). A Genetic Method to Distinguish Crossbred Inobuta from Japanese Wild Boars. *Zoological Science* 19(11), 1313–1319. doi: https://dx.doi.org/10.2108/zsj.19.1313

Karlskov-Mortensen, P., Hu, Z. L., Gorodkin, J., Reecy, J. M., and Fredholm, M. (2007). Identification of 10 882 porcine microsatellite sequences and virtual mapping of 4528 of these sequences. *Animal Genetics* 38(4), 401–405. doi: https://dx.doi.org/10.1111/j.1365-2052.2007.01609.x

Kierepka, E. M., Unger, S. D., Keiter, D. A., Beasley, J. C., Rhodes, O. E., Cunningham, F. L., and Piaggio, A. J. (2016). Identification of robust microsatellite markers for wild pig fecal DNA. *The Journal of Wildlife Management* 80(6), 1120–1128. doi: https://dx.doi.org/10.1002/jwmg.21102

Kolodziej, K., Theissinger, K., Brün, J., Schulz, H. K., and Schulz, R. (2012). Determination of the minimum number of microsatellite markers for individual genotyping in wild boar (*Sus scrofa*) using a test with close relatives. *European Journal of Wildlife Research* 58(3), 621–628. doi: https://dx.doi.org/10.1007/s10344-011-0588-9

Koutsogiannouli, E. A., Moutou, K. A., Sarafidou, T., Stamatis, C., and Mamuris, Z. (2010). Detection of hybrids between wild boars (*Sus scrofa scrofa*) and domestic pigs (*Sus scrofa* f. *domestica*) in Greece, using the PCR-RFLP method on melanocortin-1 receptor (MC1R) mutations. *Mammalian Biology* 75(1), 69–73. doi: https://dx.doi.org/10.1016/j.mambio.2008.08.001

Krause, E., Morrison, L., Reed, K. M., and Alexander, L. J. (2002). Radiation hybrid mapping of 273 previously unreported porcine microsatellites. *Animal Genetics* 33(6), 477–485. doi: https://dx.doi.org/10.1046/j.1365-2052.2002.00938_9.x

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, F., Andersson, L., and Cooper, A. (2005). Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307(5715), 1618–1621. doi: 10.1126/science.1106927

Matsumoto, Y., Takagi, T., Koda, R., Tanave, A., Yamashiro, A., and Tamate, H. B. (2019). Evaluation of introgressive hybridization among Cervidae in Japan's Kinki District via two novel genetic markers developed from public NGS data. *Ecology and*

Evolution 9(10), 5605–5616. doi: https://dx.doi.org/10.1002/ece3.5131

McCann, B. E., Malek, M. J., Newman, R. A., Schmit, B. S., Swafford, S. R., Sweitzer, R. A., and Simmons, R. B. (2014). Mitochondrial diversity supports multiple origins for invasive pigs. *The Journal of Wildlife Management* 78(2), 202–213. doi: https://dx.doi.org/10.1002/jwmg.651

Murakami, K., Yoshikawa, S., Konishi, S., Ueno, Y., Watanabe, S., and Mizoguchi, Y. (2014). Evaluation of genetic introgression from domesticated pigs into the Ryukyu wild boar population on Iriomote Island in Japan. *Animal Genetics* 45(4), 517–523. doi: https://dx.doi.org/10.1111/age.12157

Nagata, J., Maruyama, T., Asada, M., Ochiai, K., Yamazaki, K., Yamada, F., Kawaji, N., and Yasuda, M. (2006). Genetic characteristics of the wild boars in Tochigi prefecture and neighboring prefectures. *Wildlife in Tochigi Pref* 32, 58–62.

Nijman, I. J., Otsen, M., Verkaar, E. L. C., de Ruijter, C., Hanekamp, E., Ochieng, J. W., Shamshad, S., Rege, J. E. O., Hanotte, O., Barwegen, M. W., Sulawati, T., and Lenstra, J. A. (2003). Hybridization of banteng (Bos javanicus) and zebu (*Bos indicus*) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. *Heredity* 90(1), 10–16. doi: https://dx.doi.org/10.1038/sj.hdy.6800174

Okuda, K., Toma, R., Negishi, Y., Hinton, T. G., Smyser, T. J., Tamate, H. B., and Kaneko, S. (2018). Did domestic pigs that escaped after the Fukushima Daiichi nuclear power plant accident cause genetic contamination of the wild boar population? *Japanese Journal of Conservation Ecology* 23(1), 137–144. doi: https://doi.org/10.18960/hozen.23.1_137

Peakall, R. and Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1), 288–295. doi: https://dx.doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R. and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics* 28(19), 2537–2539. doi: https://dx.doi.org/10.1093/bioinformatics/bts460

Pierpaoli, M., Birò, Z. S., Herrmann, M., Hupe, K., Fernandes, M., Ragni, B., Szemethy, L., and Randi, E. (2003). Genetic distinction of wildcat (Felis silvestris) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology* 12(10), 2585–2598. doi: 10.1046/j.1365-294x.2003.01939.x

Qi, X. B., Jianlin, H., Wang, G., Rege, J. E. O., and Hanotte, O. (2010). Assessment of cattle genetic introgression into domestic yak populations using mitochondrial and microsatellite DNA markers. *Animal Genetics* 41(3), 242–252. doi: https://dx.doi.org/10.1111/j.1365-2052.2009.01989.x

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology* 17(1), 285–293. doi: https://doi.org/10.1111/j.1365-294X.2007.03417.x

Rhymer, J. M. and Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* 27(1), 83–109. doi: https://dx.doi.org/10.1146/annurev.ecolsys.27.1.83

Rohrer, G. A., Alexander, L. J., Keele, J. W., Smith, T. P., and Beattie, C. W. (1994). A microsatellite linkage map of the porcine genome. *Genetics* 136(1), 231–245.

Saito, M., Momose, H., and Mihira, T. (2011). Both environmental factors and countermeasures affect wild boar damage to rice paddies in Boso Peninsula, Japan. *Crop Protection* 30(8), 1048–1054. doi: https://dx.doi.org/10.1016/j.cropro.2011.02.017

Scandura, M., Iacolina, L., Crestanello, B., Pecchioli, E., di Benedetto, M. F., Russo, V., Davoli, R., Apollonio, M., and Bertorelle, G. (2008). Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? *Molecular Ecology* 17(7), 1745–1762. doi: https://dx.doi.org/10.1111/j.1365-294x.2008.03703.x

Sharma, S., Dutta, T., Maldonado, J. E., Wood, T. C., Panwar, H. S., and Seidensticker, J. (2013). Spatial genetic analysis reveals high connectivity of tiger (*Panthera tigris*) populations in the Satpura-Maikal landscape of Central India. *Ecology and Evolution* 3(1), 48–60. doi: https://dx.doi.org/10.1002/ece3.432

Smouse, P. E. and Peakall, R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82(5), 561–573. doi: https://dx.doi.org/10.1038/sj.hdy.6885180

Tadano, R., Nagai, A., and Moribe, J. (2016). Local-scale genetic structure in the Japanese wild boar (*Sus scrofa leucomystax*): insights from autosomal microsatellites. *Conservation Genetics* 17(5), 1125–1135. doi: https://dx.doi.org/10.1007/s10592-016-0848-z

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., and Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications* 9(7), 892–908. doi: https://dx.doi.org/10.1111/eva.12367

Touma, S., Arakawa, A., and Oikawa, T. (2020). Evaluation of the genetic structure of indigenous Okinawa Agu pigs using microsatellite markers. *Asian-Australasian Journal of Animal Sciences* 33(2), 212–218. doi: https://dx.doi.org/10.5713/ajas.19.0034

Uemura, Y., Yoshimi, S., and Hata, H. (2018). Hybridization between two bitterling fish species in their sympatric range and a river where one species is native and the other is introduced. *PLoS One* 13(9), e0203423. doi: https://doi.org/10.1371/journal.pone.0203423

Vernesi, C., Crestanello, B., Pecchioli, E., Tartari, D., Caramelli, D., Hauffe, H., and Bertorelle, G. (2003).

The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*): A microsatellite analysis. *Molecular Ecology* 12(3), 585–595. doi: https://dx.doi.org/10.1046/j.1365-294x.2003.01763.x

Waithman, J. D., Sweitzer, R. A., Vuren, D. V., Drew, J. D., Brinkhaus, A. J., Gardner, I. A., and Boyce, W. M. (1999). Range Expansion, Population Sizes, and Management of Wild Pigs in California. *The Journal of Wildlife Management* 63(1), 298–298. doi: https://dx.doi.org/10.2307/3802513

Yang, B., Cui, L., Perez-Enciso, M., Traspov, A., Crooijmans, R. P. M. A., Zinovieva, N., Schook, L. B., Archibald, A., Gatphayak, K., Knorr, C., Triantafyllidis, A., Alexandri, P., Semiadi, G., Hanotte, O., Dias, D., Dovč, P., Uimari, P., Iacolina, L., Scandura, M., Groenen, M. A. M., Huang, L., and Megens, H.-J. (2017). Genome-wide SNP data unveils the globalization of domesticated pigs. *Genetics Selection Evolution* 49(1), 71–71. doi: https://dx.doi.org/10.1186/s12711-017-0345-y

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2020), 1 (2), 42–50
DOI: 10.46265/genresj.LPJS9396
https://www.genresj.org
ISSN: 2708-3764

Growth performance and immune response to Newcastle disease in four gene pools of indigenous chicken in Rwanda

Richard Habimana [*,a,b], **Kiplangat Ngeno** [b], **Anselme Shyaka** [a], **Martin Ntawubizi** [a], **Janvier Mahoro** [a], **Claire d'Andre Hirwa** [c], **Angelique Ingabire** [c], **Lucy Kiptui** [d], **Isidore Mapendo Gafarasi** [c] **and Tobias Okeno Otieno** [b]

[a] *College of Agriculture, Animal Science and Veterinary Medicine, University of Rwanda, P.O Box 57, Nyagatare, Rwanda*
[b] *Department of Animal Science, Animal Breeding and Genomics Group, Egerton University, P.O. Box 536, Egerton, 20115, Kenya*
[c] *Rwanda Agricultural and Animal Resources Development Board, P. O Box 5016, Kigali, Rwanda*
[d] *Department of Biological Sciences, Egerton University, P.O. Box 536, 20115, Egerton, Kenya*

Abstract: In Rwanda, the most prevalent livestock farming activity carried out to improve livelihoods in rural areas is poultry farming. The most common bird species raised for egg and meat production is indigenous chicken (IC). Despite its economic significance, the improvement programs in Rwanda have given IC little attention. There is, therefore, little documentation on their production and disease resistance performance. This study aimed at evaluating the growth performance and immunologic response of IC against Newcastle disease vaccine in four gene pools in Rwanda. One hundred eighty-nine (189) IC from four gene pools were kept on-station using a completely randomized design. Thereafter, the IC were vaccinated with a live commercial Newcastle disease virus. Data on body weight were collected weekly for 20 weeks whereas serum samples were collected after 35 days. Antibody responses were determined using indirect ELISA. Analysis of variance was performed using the SAS Generalised Linear Model procedure. Growth patterns were modelled using the logistic function. Results revealed a highly significant difference ($P<0.001$) for growth performance and antibody response to Newcastle disease vaccine among the four gene pools. Gene pool A was the heaviest (1.6 kg) in the period of 20 weeks and gene pool C expressed the highest immune response (8,161antibody titres) for Newcastle disease vaccine ($P<0.001$). The results indicated that the selection of gene pool A in breeding could yield IC with good growth performance trait whilst gene pool C could be selected for its higher immunity against Newcastle disease virus.

Keywords: Body weight, indigenous chicken, antibody, Newcastle disease

Citation: Habimana, R., Ngeno, K., Shyaka, A., Ntawubizi, M., Mahoro, J., Hirwa, C. d., Ingabire, A., Kiptui, L., Gafarasi, I. M., Otieno, T. O. (2020). Growth performance and immune response to Newcastle disease in four gene pools of indigenous chicken in Rwanda. *Genetic Resources* 1 (2), 42–50. doi: 10.46265/genresj.LPJS9396.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Indigenous chicken (IC) contribute considerably to the rural farmers' livelihood by providing excellent animal protein (Gebreselassie *et al*, 2015). Though these birds are predominant in most Rwanda households, their production and genetic potential have not been fully exploited (Padhi, 2016). Low genetic potential and high occurrence of diseases like Newcastle disease (ND) are among the main features affecting IC productivity in the tropics (Lwelamira and Katule, 2005). Despite their low egg productivity and growth rate (Mbuza *et al*, 2016), they have a prodigious potential to improve rural livelihoods (Ministry of Agriculture and Animal Resources (MINAGRI), 2012).

Growth identified as an increase in weight, length and height with age (Aggrey, 2002; Yang *et al*, 2006)

*Corresponding author: Richard Habimana (hrichard86@yahoo.fr)

Received: 03.04.2020 | Accepted: 12.12.2020 | Published online: 29.12.2020

is very significant in any IC production systems. An improvement in this trait is expected to have a countless positive impact on farmers livelihoods (Magothe *et al*, 2010). Several mathematical functions are used to describe growth and provide biologically interpretable parameters. In order to estimate the growth of chicken, growth models coupled with fitted longitudinal experimental data are used. The most growth models frequently used are nonlinear functions (Mignon-Grasteau and Beaumont, 2000; Hyánková *et al*, 2001; Karkach, 2006; Ndri *et al*, 2018). The nonlinear models have been utilized because of their several advantages, which include describing weight per age (Hyánková *et al*, 2001), and estimating the association between body weight and feed requirement (Fiengül and Kiraz, 2005). The growth curve parameters determined by the nonlinear models are used to decide the appropriate age whereat to select chicken and to design management approaches for breeding (Fiengül and Kiraz, 2005).

Growth performance is influenced by genes and management practises, especially nutrition and health (Mwalusanya *et al*, 2002; Gondwe and Wollny, 2005; Dana *et al*, 2010). Newcastle disease (ND) is among the most severe problems encumbering IC growth performance in Rwanda (Mazimpaka *et al*, 2018). Newcastle disease, which affects many bird's species globally is brought about by single-strand, non-segmented, negative-sense RNA virus also called Avian paramyxovirus 1 (APMV-1) (Cattoli *et al*, 2011). This disease is an issue of great concern to the farmers because its outbreak can cause a high economic loss of up to 100% flock mortality. Newcastle disease has no effective cure. The spread of this virus is, however, contained by appropriate management strategies such as vaccination (Alexander, 2001). The possibility of eradicating this avian disease currently is completely questionable. This is because pathogen virulence, which renders the vaccine ineffective increases with time. Consequently, requiring new vaccines to be developed (Witter, 1998). Although new vaccines lessen ND epidemics, frequent research to develop the vaccine makes the strategy expensive. Breeding using indirect selection for the development of immunological response trait is, therefore, the best long-term strategy for mitigating diseases (Gavora *et al*, 1983); in this case, immune response trait will be used to develop IC breeds with a high-level of biosecurity (Padhi, 2016).

Although body weight and disease tolerance are the highly preferred traits by poultry farmers (Banerjee, 2012), in Rwanda there is no established structure for breeding programmes, nor for genetic improvement, nor for conservation (Mahoro *et al*, 2018). In an effort to improve genetic potential through within breed selection of IC in Rwanda (Shapiro *et al*, 2017), preliminary studies have been conducted (Mbuza *et al*, 2016; Mahoro *et al*, 2017; Hirwa *et al*, 2019; Habimana *et al*, 2020). Research on the morphological characterisation of IC ecotypes showed a very huge variation among IC in Rwanda (Hirwa *et al*, 2019). Subsequently, a study on genetic diversity and population structure unveiled the presence of four different gene pools among IC in Rwanda (Habimana *et al*, 2020). There is, however, no information on the diversity in growth performance and immune responses to ND among those four gene pools. Availability of such information would be necessary for designing breeding strategies for increased IC productivity in Rwanda. Increasing IC performance would contribute towards improved food security and secure livelihoods. This study aimed at evaluating the growth performance and ND immunity of IC among four gene pools in Rwanda.

Materials and methods

Ethical approval

This study was conducted in line with the principles of the Declaration of Helsinki. Approval was granted by the Research Screening and Ethical Clearance Committee of the College of Agriculture, Animal Sciences and Veterinary Medicine, University of Rwanda (Ref: 031/19/DRI September 2, 2019).

Study sites

The study was carried out on-station at the University of Rwanda-Nyagatare campus in Nyagatare district, Rwanda. The coordinates of the area of study are 1° 18' 0.00"S, 30° 19' 30.00"E (Latitude: -1.3000; Longitude: 30.3250). The location of study is characterized by two main seasons: one long dry season that varies between 3 and 5 months with an annual average temperature varying between 25.3°C and 27.7°C. The monthly distribution of the rains varies from one year to another. Annual rainfalls are very low (827mm/year) and unpredictable to satisfy agriculture and livestock needs.

Study population

Indigenous chickens were selected from four gene pools previously identified (Habimana *et al*, 2018). These gene pools are distributed across the five agro-ecological zones in Rwanda, which are Central North and North West, South West, South West and Central South and East. The aforementioned gene pools were designated as letters A, B, C and D, respectively.

Sampling, data collection and management of experimental chicken

Seven hundred and twenty (720) fertile eggs were collected from four IC gene pools previously identified (Habimana *et al*, 2018) and only 707 eggs were viable to be incubated. Eggs were concurrently incubated but separated based on gene pool within the incubator (LDG-1200). After 21 days, 189-day old chicks were obtained from hatch (Table 1). Each chick was weighed, wing tagged with an identification number, and allocated into a house with four equal compartments (5m x 3m for each one). Brooding of chicks was

done from hatching to week six (Ngeno *et al*, 2010; Sayed *et al*, 2016), in deep litter brooders containing infrared electric bulbs (230V, 250W). Initial brooding temperature, at the level of the chicks, was 36 to 35°C and reduced by approximately 2 to 3°C per week until 21°C was reached. The IC gene pools were reared in different brooders. All birds were kept under the same environmental conditions and management practices. Feeds were provided ad libitum on a starter diet (20% crude protein and 12.9 MJ ME /kg diet) for the first six weeks and a growing diet (16.8% crude protein and 12.5 MJ ME /kg diet) from the 7^{th} to 20^{th} week. Clean water was provided daily ad libitum. Disinfection of the brooding and rearing pens was done procedurally. The IC were then immunized with two commercial ND virus live vaccines, which were AVI ND HB1 in drinking water at 2 days of age and AVI ND Lasota (Intervet International B.V., Boxmeer, Netherlands) by eye drop at 28 days of age according to manufacturer's instructions. At 7 days after the second immunization (35 days of age) as per the protocol outlined by Al-Garib *et al* (2003), blood samples were collected without anticoagulant for separation of sera to detect the titre of antibody against ND vaccine. The Ab responses to ND were detected using indirect ELISA (ID Screen® Newcastle Nucleoprotein Indirect Version 2, IDvet, Grabels, FRANCE). The IDSoft™ data analysis programme was used to compute the Ab titres.

Data analysis

Descriptive statistics were generated using the SAS Generalised Linear Models procedure (SAS Institute, 2011). Body weights and Ab response were subjected to analysis of variance (ANOVA) with gene pools as the main variable in a completely randomized design and fixed effects model (Equation (1)). For the analysis of body weight at hatch, egg weight was fitted in the model as a covariate, whereas in other ages; body weight at hatch was also fitted as a covariate. Tukey's test was used for means separation between gene pools.

$$Y_{ijk} = \mu + G_i + + S_j + (GS)_{ij} + e_{ijk} \quad (1)$$

where: Y_{ijk}: record of k^{th} chicken at particular age from i^{th} gene pool with j^{th} sex, μ: overall mean; G_i: fixed effect of i^{th} gene pool; S_j: fixed effect of j^{th} sex; $(GS)_{ij}$: interaction between gene pool and sex; e_{ijk}: random error.

The analysis of longitudinal growth data was performed using non-linear regression. The following logistic regression model (Equation (2)) using PROC NLIN of SAS (SAS Institute, 2011) was fitted to get growth curve parameters (a, b and k).

$$y_t = a(1 - b\, e^{-kt})^{-1} \quad (2)$$

where: y_t: live weight at age t, a: asymptotic or mature weight, b: scaling parameter, k: maturity index and t: age in weeks.

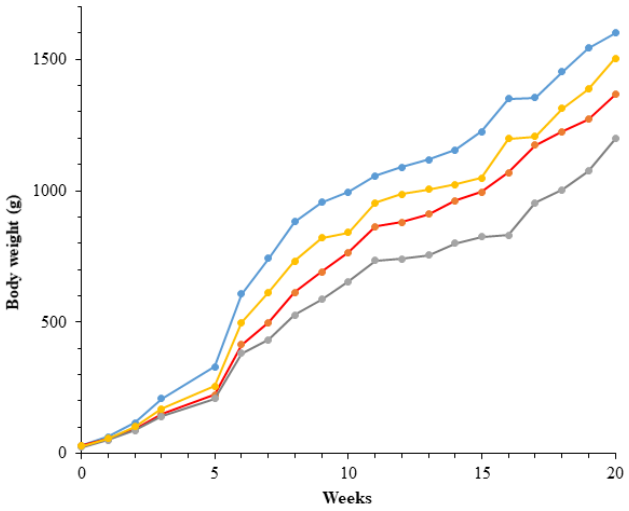

Figure 1. Body weight within and between indigenous chicken gene pools in Rwanda. Colors refer to gene pools: blue = A; red = B; grey = C; yellow = D.

Results

Growth performance

Body weights (BW) from hatch to 20 weeks of age among the four gene pools of IC in Rwanda are presented in Figure 1. The mean BW were significantly different ($p < 0.001$) across gene pools from week 0 to week 20. At hatch, gene pool A and B had the highest body weight (29.63±3.78 and 29.78±3.69g, respectively) and the lowest BW (24.29±3.39g) was recorded in gene pool C. At week one, Gene pool D and A exhibited higher BW (58.24±10.70 and 62.94±11.59g, respectively) followed by gene pool B (55.95±7.79g). At week 2, gene pool A showed the highest BW (116.32±27.73g) and gene pool C exhibited the lowest BW (87.51±24.06g). At week three, gene pool A showed the highest BW (208.19±56.16g) followed by gene pool D (170.69±43.43g) and the lowest BW was recorded in gene pool C (140.40±37.49g). During the remaining rearing period, the trend observed in week 3 was maintained.

The least-square means of the growth parameters (a, b and k), coefficient of determination (R^2) are given in Table 2. The fitted parameters for the logistic model revealed a significant gene pool effect ($P < 0.001$) in the asymptotic weights (a), scaling parameter (b) and the maturity index (k). The lowest for these parameters were found for gene pool C.

Antibody titres for Newcastle disease vaccine

Antibody titres to Newcastle disease were statistically different ($p < 0.001$) across the gene pools. Gene pool C was found to be significantly higher in Ab titre for the ND virus vaccine (8,161). Lower Ab titre to ND vaccine was recorded in gene pool A (4,879) (Table 3).

Table 1. Sampling and management of experimental chicken

Gene pools	Collected fertile eggs	Incubated eggs	Hatched eggs	Chicks in brooder
A	180	179	72	72
B	180	177	34	34
C	180	175	36	36
D	180	176	47	47
Total	720	707	189	189

Table 2. Least-square means (±SE) of growth parameters of indigenous chicken gene pools in Rwanda. Parameters: a, asymptomatic size body weight (g); b, scaling parameter; k, maturity index; R^2, coefficient of determination. ^{ab}Least means in a row with one or more letter superscripts in common are not significantly different (P >0.05).

Parameters	Gene pools				Overall	P-value
	A	B	C	D		
a	1446.57±62.42^{a}	1286.31±58.88^{ab}	1086.38±66.64^{b}	1350.13±69.16^{a}	1309.46±59.82	0.000
b	15.41±5.19^{a}	16.33±4.13^{a}	13.65±3.93^{b}	15.15±4.87^{a}	15.02±4.41	
k	0.35±0.05^{a}	0.30±0.04^{a}	0.29±0.04^{b}	0.32±0.05^{a}	0.321±0.04	
R^2	0.96	0.97	0.96	0.96	0.96	

Discussion

This experiment sought to evaluate diversity in growth performance and Ab response to ND among four different gene pools of IC in Rwanda.

Growth performance

This experiment compared the growth performance of four gene pools of IC in Rwanda. The results of this study showed the existence of a large variation in growth between the four IC gene pools in Rwanda. Growth is a quantitative trait that is influenced by genes and production environment such as health and nutrition (Mwalusanya *et al*, 2002; Gondwe and Wollny, 2005; Dana *et al*, 2010). Since the IC were kept under the same environmental conditions in this study, the witnessed differences between the IC gene pools could result in genetic diversity among these IC populations. In addition to that, gene pool A and gene pool C are morphologically different (Hirwa *et al*, 2019). These results affirmed the previous study undertaken on IC genetic diversity and population structure using microsatellite markers in Rwanda (Habimana *et al*, 2020). The difference in growth performance of four gene pools is due to the difference in feed conversion efficiency (FCE), with heavier gene pool having higher FCE and perhaps due to smaller maintenance requirements (Tadelle *et al*, 2003). Feed conversion efficiency is normally influenced by the climate, feed properties, gut microbiota and genetic predispositions (Reyer *et al*, 2015). In this situation where the IC were kept under similar environment, the difference in FCE is likely due to their genetic constitution. Indigenous chickens in the tropics have major genes which affect their performance (Ibe, 1993). There are typically a big number of genes involved in growth which leads to differences in BW. These genes can be activated or deactivated at different stage depending on the development of the genetic background of the IC (Rouzic *et al*, 2008). Certain genes regulating BW in gene pool A are not found in gene pool C. This occurrence is attributed to mutations and recombination, which are the most important factors responsible for the genetic variances between breeds and populations of chicken (Weigend and Romanov, 2001). The significant differences observed between gene pools in this study for BW measured were confirmed also in Ethiopia (Tadelle *et al*, 2003; Dessie and Ogle, 2001), Tanzania (Msoffe *et al*, 2001), Zimbabwe (McAinsh *et al*, 2004), Botswana (McAinsh *et al*, 2004), Kenya (Ngeno *et al*, 2010, 2013, 2014) and Bangladesh (Bhuiyan *et al*, 2011). These differences among IC gene pools enable selection of parent lines for practical IC breeding (Ngeno *et al*, 2010), which results in their genetic improvement (Li *et al*, 2002).

The growth rate among gene pools varied for the first three weeks. Thus, their ranking was not consistent. This implies that a gene pool by age interaction occurred. For this reason, a constant criterion cannot be used in genetic selection for BW between the gene pools. The genetic relationship between immature BW and

Table 3. Antibody titres (mean, maximum, minimum, geometric mean and coefficient of variation) for Newcastle disease vaccine. GMT: geometric mean titre, CV: Coefficient of variation. Antibody titre distribution is significantly different across gene pools (P<0.05) based on analysis of variance. ^{abc}Means in a row with one different letter superscripts are significantly different (P < 0.05).

Parameters	Gene pools				P-Value
	A (n=56)	B (n=33)	C (n=35)	D (n=46)	
Mean	4,879^{a}	5,089^{a}	8,161^{b}	6,070^{c}	
Minimum	1,105	2,019	2,119	1,615	
Maximum	8,225	7,784	11,758	9,788	0.000
GMT	4,428^{a}	4,748^{a}	7,742^{b}	5,599^{c}	
CV	39%	35%	28%	37%	

age when selection decisions are made and desired matured weight may not be the same for different gene pools. There was an increase in the BW measurements during the birds' maturity period, demonstrating a direct positive relationship between BW and age. This was confirmed by the observed differences in growth curve parameters among the gene pools. This result agrees with previous reports from Nigeria (Ojedapo, 2013), Ethiopia (Tadelle *et al*, 2003) and Kenya (Ngeno *et al*, 2010, 2013).

Growth parameters were also determined to obtain a better description of the growth performance of Rwandese IC gene pools. The overall coefficient of determination (R^2= 0.964) was high in this study, demonstrating the sufficiency of the model in fitting and describing growth data effectively. The objective of a curve fitting was to describe body weight increase in respect to age using equations with few parameters (Ricklefs, 1985). Curve parameters provide information on growth characteristics (Ndri *et al*, 2018) and can be used for pre-selection (Eleroğlu *et al*, 2014). Growth curve patterns have been performed to increase the performance of domestic poultry. Growth parameters are also heritable (Mignon-Grasteau and Beaumont, 2000). Selection approaches have been used to change the shapes of growth curves, leading to quick growth in poultry (Hyánková *et al*, 2001; Maruyama *et al*, 2001). In addition, a detailed understanding of growth characteristics will contribute to design feeding schemes to minimise production costs while maximising production efficiency of IC.

Immune Response

The findings showed that all IC had good immunity ranging from 4,879 to 8,161 (> cut-off: 993), but the chickens in gene pool C had a comparatively higher immunoglobulin titre against ND than the others ($p<0.05$) (Table 3). The high Ab found in the gene pool C may have resulted from their genetic make-up. Genetic diversity in Ab responses of the chicken could be due to the major histocompatibility complex (MHC) (Lwelamira *et al*, 2008), immunoglobulin allotypes, and to genes not associated with either MHC or allotypes (Gavora *et al*, 1983). This result complements Palladino *et al* (1977) study, which showed a quantitative variance between Ab responses. The Ab response to the similar virus varies between IC populations (Pitcovski *et al*, 2001) and selection for an Ab response may increase disease resistance in chicken (Gross *et al*, 1980). The levels of Ab in appropriate response to a pathogen is known as immune competence; the higher the levels, the greater the immune competence (Star *et al*, 2007). A standard immune response to a pathogen results in a disease resistance. This aptitude plays a key role in controlling disease occurrences (Lamont *et al*, 2009). The reduction of disease outbreaks contributes to increased animal production and profitability (Wakchaure and Ganguly, 2016). This result on significant differences between gene pools for Ab responses to ND virus concurs with previous reports, (Hassan *et al*, 2004; Chang *et al*, 2011; Taha *et al*, 2012)

A noticeable trend observed in this research was that gene pool A, the heaviest had a low ND Ab titre, while the slowest growing gene pool C had the highest ND Ab titre, reinforcing the negative genetic correlation between BW and Ab response (Martin *et al*, 1990; Parmentier *et al*, 1998; Siegel *et al*, 2009). The negative correlation between BW and Ab response is caused by pleiotropic effects between genes associated with immunoresponsiveness (Martin *et al*, 1990). The negative relationship between Ab production and BW occur as a result of resources allocation to immune responses rather than growth (Hassan *et al*, 2004; Dunnington and Siegel, 1996). Increasing Ab response consumes large quantities of energy that the IC could have used in increasing the BW (Rauw, 2012). An animal, as an entire organism, has various physiological functions which may be competitive for its inadequate resources during its lifetime (Lamont *et al*, 2009). Protein secretion in antibodies and protein deposition in muscle compete for the restricted accessible nutrient resources (Hassan *et al*, 2004). To achieve a delicate equilibrium of optimal productivity and immunoresponsiveness, knowledge of the nutrient requirements of an activated immune system with respect to individual genes regulating protein synthesis in lymphocytes and muscle tissue is needed (Lamont *et al*, 2009).

The existence of diversity in body weight and immune response to ND among gene pools of IC populations in this study is attributed to the geographical isolation and eras of artificial and natural selection. Earlier studies in Rwanda (Habimana *et al*, 2018), Senegal (Missohou *et al*, 1998); Ethiopia (Halima *et al*, 2007) and results by Hauser *et al* (1995) reported that ecological zones lead to the establishment of population traits with probable variations in genetic effect. This also confirms the genotype by environment interaction (GxE) revealed in previous studies (Mathur and Horst, 1994; James, 2009; Lwelamira, 2012). The genetic variances observed in these different gene pools could be utilised in crossbreeding programmes targeting IC with better growth performance and higher immunoglobulins against the ND virus. Consequently, gene pools A and C could be considered.

Conclusion

This study revealed that gene pool A was the heaviest and the gene pool C expressed the highest immune response for Newcastle disease vaccine. This research work provides the basis for choosing a suitable IC gene pool for breeding programme. From this result, it was evident that the selection of gene pool A in breeding would yield IC with good growth performance trait whilst gene pool C could be selected for its higher immunity against Newcastle disease. For further studies, there is a need to contemplate diversity in

egg production and other important poultry's diseases among IC gene pools in Rwanda. Besides, with regard to immune response, immunoglobulins Y (IgY), M (IgM) and possibly A (IgA) should be examined separately across IC gene pools in Rwanda.

Acknowledgements

This material is based upon work supported by the United States Agency for International Development (USAID), as part of the Feed the Future initiative, under the CGIAR Fund, award number BFS-G-11-00002, and the predecessor fund the Food Security and Crisis Mitigation II grant, award number EEM-G-00-04-00013.

Supplemental data

Supplemental File 1: Abstract in French

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by RH, KN, AS, MN, JM, CAH, AI, LK, FG and TOO. The first draft of the manuscript was written by RH and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

Aggrey, S. E. (2002). Comparison of three nonlinear and spline regression models for describing chicken growth curves. *Poultry Science* 81(12), 1782–1788. doi: https://doi.org/10.1093/ps/81.12.1782

Al-Garib, S. O., Gielkens, A. L. J., Gruys, E., and Kochi, G. (2003). Review of Newcastle disease virus with particular references to immunity and vaccination. *World's Poultry Science Journal* 59(2), 185–200. doi: https://doi.org/10.1079/wps20030011

Alexander, D. J. (2001). Newcastle disease. *British Poultry Science* 42(1), 5–22. doi: https://doi.org/10.1080/713655022

Banerjee, S. (2012). Morphological characterization of indigenous chickens of Sikkim and West Bengal, India. *Animal Genetic Resources / Ressources Génétiques Animales / Recursos Genéticos Animales* 51, 57–71. doi: https://doi.org/10.1017/S2078633612000343

Bhuiyan, A. K. F. H., Bhuiyan, M. S. A., and Deb, G. K. (2011). Indigenous chicken genetic resources in Bangladesh: current status and future outlook. *Animal Genetic Resources Information* 36, 73–84. doi: https://doi.org/10.1017/S1014233900001899

Cattoli, G., Susta, L., Terregino, C., and Brown, C. (2011). Newcastle disease: A review of field recognition and current methods of laboratory detection. *Journal of Veterinary Diagnostic Investigation* 23(4), 637–656. doi: https://doi.org/10.1177/1040638711407887

Chang, C.-S., Tixier-Boichard, M., Chazara, O., Lee, Y.-P., Chen, C.-F., Chang, P.-C., Chen, J.-W., and Bed'hom, B. (2011). Different immune responses to three different vaccines following H6N1 low pathogenic avian influenza virus challenge in Taiwanese local chicken breeds. *BMC Proceedings* 5(S4). doi: https://doi.org/10.1186/1753-6561-5-s4-s33

Dana, N., Dessie, T., van der Waaij, L. H., and van Arendonk, J. A. (2010). Morphological features of indigenous chicken populations of Ethiopia. *Animal Genetic Resources/Ressources génétiques animales/Recursos genéticos animales* 46, 11–23. doi: https://doi.org/10.1017/s2078633610000652

Dessie, T. and Ogle, B. (2001). Village Poultry Production Systems in the Central Highlands of Ethiopia. *Tropical Animal Health and Production* 33, 521–537. doi: https://doi.org/10.1023/A:1012740832558

Dunnington, E. A. and Siegel, P. B. (1996). Long-Term Divergent Selection for Eight-Week Body Weight in White Plymouth Rock Chickens. *Poultry Science* 75(10), 1168–1179. doi: https://doi.org/10.3382/ps.0751168

Eleroğlu, H., Yildirim, A., Şekeroğlu, A., Çoksöyler, F. N., and Duman, M. (2014). Comparison of growth curves by growth models in slow-growing chicken genotypes raised the organic system. *International Journal of Agriculture and Biology* 16(3), 529–535. url: http://www.fspublishers.org/published_papers/54943_..pdf.

Fiengül, T. and Kiraz, S. (2005). Non-Linear Models for Growth Curves in Large White Turkeys. *Turk J Vet Anim Sci* 29, 331–337. url: https://journals.tubitak.gov.tr/veterinary/issues/vet-05-29-2/vet-29-2-22-0306-39.pdf.

Gavora, S., Spencer, J., and J (1983). Breeding for immune responsiveness and disease resistance. *Animal Blood Groups and Biochemical Genetics* 14(3), 159–180. doi: https://doi.org/10.1111/j.1365-2052.1983.tb01070.x

Gebreselassie, G., Meseret, R., Mulalem, Z., Hailay, H., Minister, B., and Gebru, B. (2015). Comparative production performance evaluation of exotic and indigenous chickens under farmers management practice in Tigray, Northen Ethiopia. *Scientific Journal of Biological Sciences* 4(12), 181–186. url: http://www.sjournals.com/index.php/sjbs/article/view/221.

Gondwe, T. N. and Wollny, C. B. A. (2005). Evaluation of the Growth Potential of Local Chickens in Malawi. *International Journal of Poultry Science* 4(2), 64–70. doi: https://doi.org/10.3923/ijps.2005.64.70

Gross, W. G., Siegel, P. B., Hall, R. W., Domermuth, C. H., and DuBoise, R. T. (1980). Production and Persistence of Antibodies in Chickens to Sheep Erythrocytes. *Poultry Science* 59(2), 205–210. doi: https://doi.org/10.3382/ps.0590205

Habimana, R., Okeno, T. O., Ngeno, K., Mboumba, S., Assami, P., Gbotto, A. A., Keambou, C. T., Nishimwe, K., Mahoro, J., and Yao, N. (2020). Genetic diversity and population structure of indigenous chicken in Rwanda using microsatellite markers. *PLOS ONE* 15(4), e0225084–e0225084. doi: https://doi.org/10.1371/journal.pone.0225084

Habimana, R., Okeno, T. O., Ngeno, K., Mboumba, S., Assami, P., Keambou, C. T., and Eastern, B. (2018). Four gene pools of Rwandese indigenous chicken based on microsatellite markers. In Proceedings of the World Congress on Genetics Applied to Livestock Production, volume 11.874, 1-4.

Halima, H., Neser, F. W. C., van Marle-Koster, E., and de Kock, A. (2007). Phenotypic variation of native chicken populations in northwest Ethiopia. *Tropical Animal Health and Production* 39(7), 507–513. doi: https://doi.org/10.1007/s11250-007-9032-2

Hassan, M. K., Afify, M. A., and Aly, M. M. (2004). Genetic Resistance of Egyptian Chickens to Infectious Bursal Disease and Newcastle Disease. *Tropical Animal Health and Production* 36(1), 1–9. doi: https://doi.org/10.1023/b:trop.0000009524.47913.d4

Hauser, L., Carvalho, G. R., and Pitcher, T. J. (1995). Morphological and genetic differentiation of the African clupeid *Limnothrissa miodon* 34 years after its introduction to Lake Kivu. *Journal of Fish Biology* 47(sa), 127–144. doi: https://doi.org/10.1111/j.1095-8649.1995.tb06049.x

Hirwa, C. D., Kugonza, D. R., Kayitesi, A., Murekezi, T., Semahoro, F., Uwimana, G., and Habimana, R. (2019). Phenotypes, production systems and reproductive performance of indigenous chickens in contemporary Rwanda. *International Journal of Livestock Production* 10(10), 213–231. doi: https://doi.org/10.5897/ijlp2019.0618

Hyánková, L., Knížetová, H., Dědková, L., and Hort, J. (2001). Divergent selection for shape of growth curve in Japanese quail. 1. Responses in growth parameters and food conversion. *British Poultry Science* 42(5), 583–589. doi: https://doi.org/10.1080/00071660120088371

Ibe, S. (1993). Growth performance of normal frizzle and naked-neck chickens in a tropical environment. *Nigerian Journal of Animal Production* 20(1). url: https://www.ajol.info/index.php/njap/article/view/126757.

James, J. W. (2009). Genotype by Environment Interaction in Farm Animals. In *Adaptation and Fitness in Animal Populations,* ed. van der Werf, J., Graser, H. U., Frankham, R., and Gondro, C., (Dordrecht: Springer), 151-167.

Karkach, A. (2006). Trajectories and models of individual growth. *Demographic Research* 15, 347–400. doi: https://doi.org/10.4054/demres.2006.15.12

Lamont, S. J., Laan, M. H. P.-V. D., Cahaner, A., van der Poel, J. J., and Parmentier, H. K. (2009). Selection for disease resistance: direct selection on the immune response. In *Poultry genetics, breeding and biotechnology,* ed. Muir, W. M. and Aggrey, S. E. 399-418.

Li, Y.-C., Korol, A. B., Fahima, T., Beiles, A., and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11(12), 2453–2465. doi: https://doi.org/10.1046/j.1365-294x.2002.01643.x

Lwelamira, J. (2012). Genotype-Environmental (G x E) Interaction for Body Weights for Kuchi Chicken Ecotype of Tanzania Reared On-Station and On-Farm. *International Journal of Poultry Science* 11(2), 96–102. doi: https://doi.org/10.3923/ijps.2012.96.102

Lwelamira, J. and Katule, A. M. (2005). Genetic determination of immune responses to Newcastle disease virus vaccine in chickens. *Bulletin of Animal Health and Production in Africa* 52(3), 186–197. doi: https://doi.org/10.4314/bahpa.v52i3.32669

Lwelamira, J., Kifaro, G. C., Gwakisa, P. S., and Msoffe, P. L. M. (2008). Association of LEI0258 microsatellite alleles with antibody response against newcastle disease virus vaccine and body weight in two Tanzania chicken ecotypes. *African Journal of Biotechnology* 7(6), 714–720. url: https://www.ajol.info/index.php/ajb/article/view/58502.

Magothe, T. M., Muhuyi, W. B., and Kahi, A. K. (2010). Influence of major genes for crested-head, frizzle-feather and naked-neck on body weights and growth patterns of indigenous chickens reared intensively in Kenya. *Tropical Animal Health and Production* 42(2), 173–183. doi: https://doi.org/10.1007/s11250-009-9403-y

Mahoro, J., Muasya, T. K., Mbuza, F., Habimana, R., and Kahi, A. K. (2017). Characterization of indigenous chicken production systems in Rwanda. *Poultry Science* 96(12), 4245–4252. doi: https://doi.org/10.3382/ps/pex240

Mahoro, J., Muasya, T. K., Mbuza, F., Mbuthia, J., and Kahi, A. K. (2018). Farmers' breeding practices and traits of economic importance for indigenous chicken in RWANDA. *Tropical Animal Health and Production* 50(1), 121–128. doi: https://doi.org/10.1007/s11250-017-1411-8

Martin, A., Dunnington, E. A., Gross, W. B., Briles, W. E., Briles, R. W., and Siegel, P. B. (1990). Production Traits and Alloantigen Systems in Lines of Chickens Selected for High or Low Antibody Responses to Sheep Erythrocytes. *Poultry Science* 69(6), 871–878. doi: https://doi.org/10.3382/ps.0690871

Maruyama, K., Vinyard, B., Akbar, M. K., Shafer, D. J., and Turk, C. M. (2001). Growth curve analyses in selected duck lines. *British Poultry Science* 42(5), 574–582. doi: https://doi. org/10.1080/00071660120088380.

Mathur, P. K. and Horst, P. (1994). Methods for evaluating genotype-environment interactions illustrated by laying hens. *Journal of Animal Breeding and Genetics* 111(1-6), 265–288. doi: https://doi.org/10.1111/j.1439-0388.1994.tb00467.x

Mazimpaka, E., Tukei, M., Shyaka, A., and Gatari, E. N. (2018). Poultry production and constraints in Eastern Province of Rwanda: case study of Rukomo sector, Nyagatare district. *Tropical Animal Health and*

Production 50(4), 753–759. doi: https://doi.org/10.1007/s11250-017-1491-5

Mbuza, F., Denis, M., Janvier, M., and Xavier, R. (2016). Characterization of low cost village Poultry production in Rwanda. *International Journal of Livestock Production* 7(9), 76–82. doi: https://doi.org/10.5897/ijlp2016.0300

McAinsh, C. V., Kusina, J., Madsen, J., and Nyoni, O. (2004). Traditional chicken production in Zimbabwe. *Poultry Science Journal* 60(2), 233–246. doi: https://doi.org/10.1079/WPS2OO416

Mignon-Grasteau, S. and Beaumont, C. (2000). Les courbes de croissance chez les oiseaux. *INRAE Productions Animales* 13(5), 337–348. (in French). doi: https://doi.org/10.20870/productions-animales.2000.13.5.3802

Ministry of Agriculture and Animal Resources (MINAGRI) (2012). Strategic and investment plan to strengthen the animal genetic improvement in Rwanda. url: http://extwprlegs1.fao.org/docs/pdf/rwa170087.pdf.

Missohou, A., Sow, R. S., and Ngwe-Assoumou, C. (1998). Caractéristiques morphobiométriques de la poule du Sénégal. *Animal Genetic Resources Information* 24, 63–69. doi: https://doi.org/10.1017/s1014233900001140

Msoffe, P. L. M., Minga, U. M., Olsen, J. E., Yongolo, M. G. S., Juul-Madsen, H. R., Gwakisa, P. S., and Mtambo, M. M. A. (2001). Phenotypes including immunocompetence in scavenging local chicken ecotypes in Tanzania. *Tropical Animal Health and Production* 33, 341–354. doi: https://doi.org/10.1023/A:1010544221028

Mwalusanya, N. A., Katule, A. M., Mutayoba, S. K., Mtambo, M. M. A., Olsen, J. E., and Minga, U. M. (2002). Productivity of local chickens under village management conditions. *Tropical Animal Health and Production* 34(5), 405–416. doi: https://doi.org/10.1023/A:1020048327158

Ndri, A., Koua, B., Ahouchi, V., and AdepoGourene, A. (2018). Body weight and growths curve parameters evaluation of three chicken genotypes (Gallus gallus domesticus) reared in claustration. *Journal of Advanced Veterinary and Animal Research* 5(3), 188–188. doi: https://doi.org/10.5455/javar.2018.e265

Ngeno, K., Bebe, B. O., and Kahi, A. K. (2010). Estimation of Growth Parameters of Indigenous Chicken Populations Intensively Reared in Kenya. *Egerton Journal Science of Technology* 11, 13–28. url: http://hdl.handle.net/123456789/59.

Ngeno, K., Magothe, T. M., Okeno, T. O., Bebe, B. O., and Kahi, A. K. (2013). Heritabilities and Correlations Between Body Weights and Growth Curve Parameters of Indigenous Chicken Populations Reared Intensively in Kenya. *Research Journal of Poultry Science* 6(3), 43–52. doi: https://doi.org/10.36478/rjpscience.2013.43.52

Ngeno, K., van der Waaij, E., Kahi, A. K., and van Arendonk, J. (2014). Morphological features of indigenous chicken ecotype populations of Kenya. *Animal Genetic Resources / Ressources Génétiques Animales / Recursos Genéticos Animales* 55, 115–124. doi: https://doi.org/10.1017/s2078633614000356

Ojedapo, L. O. (2013). Evaluation of Body Weight and Other Linear Parameters of Marshall Broiler for Repeatability Estimates. *International Journal of Applied Agriculture and Apiculture Research* 9(1-2), 175–181. url: https://www.ajol.info/index.php/ijaaar/article/view/96946.

Padhi, M. K. (2016). Importance of Indigenous Breeds of Chicken for Rural Economy and Their Improvements for Higher Production Performance. *Scientifica* 2016, 2604685. doi: https://doi.org/10.1155/2016/2604685

Palladino, M. A., Gilmour, D. G., Scafuri, A. R., Stone, H. A., and Thorbecke, G. J. (1977). Immune response differences between two inbred chicken lines identical at the major histocompatibility complex. *Immunogenetics* 5(1), 253–259. doi: https://doi.org/10.1007/bf01570480

Parmentier, H. K., Walraven, M., and Nieuwland, M. G. (1998). Antibody responses and body weights of chicken lines selected for high and low humoral responsiveness to sheep red blood cells. 1. Effect of Escherichia coli lipopolysaccharide. *Poultry Science* 77(2), 248–255. doi: https://doi.org/10.1093/ps/77.2.248

Pitcovski, J., Cahaner, A., Heller, E. D., Zouri, T., Gutter, B., Gotfried, Y., and Leitner, G. (2001). Immune Response and Resistance to Infectious Bursal Disease Virus of Chicken Lines Selected for High or Low Antibody Response to Escherichia coli. *Poultry Science* 80(7), 879–884. doi: https://doi.org/10.1093/ps/80.7.879

Rauw, W. M. (2012). Immune response from a resource allocation perspective. *Frontiers in Genetics* 3, 1–14. doi: https://doi.org/10.3389/fgene.2012.00267

Reyer, H., Hawken, R., Murani, E., Ponsuksili, S., and Wimmers, K. (2015). The genetics of feed conversion efficiency traits in a commercial broiler line. *Scientific Reports* 5(1), 16387–16387. doi: https://doi.org/10.1038/srep16387

Ricklefs, R. E. (1985). Modification of Growth and Development of Muscles of Poultry. *Poultry Science* 64(8), 1563–1576. doi: https://doi.org/10.3382/ps.0641563

Rouzic, A. L., Álvarez Castro, J. M., and Carlborg, Ö. (2008). Dissection of the Genetic Architecture of Body Weight in Chicken Reveals the Impact of Epistasis on Domestication Traits. *Genetics* 179(3), 1591–1599. doi: https://doi.org/10.1534/genetics.108.089300

SAS Institute (2011). Statistical Analysis Software (SAS) Release 9.3. url: https://support.sas.com/documentation/93/.

Sayed, E., Ahmed, M. S., Alyousef, A. S., and M, Y. (2016). Characterization of Newcastle disease antibody response and some related performance indicators of two local Saudi chicken lines and two

cross lines during the rearing period. *Journal of Animal and Plant Sciences* 26(5), 1236–1241.

Shapiro, B. I., Gebru, G., Desta, S., and Nigussie, K. (2017). Rwanda Livestock Master Plan. url: http://extwprlegs1.fao.org/docs/pdf/rwa172923.pdf.

Siegel, P. B., Gross, W. B., and Cherry, J. A. (2009). Correlated responses of chickens to selection for production of antibodies to sheep erythrocytes. *Animal Blood Groups and Biochemical Genetics* 13(4), 291–297. doi: https://doi.org/10.1111/j.1365-2052.1982.tb01573.x

Star, L., Frankena, K., Kemp, B., Nieuwland, M. G. B., and Parmentier, H. K. (2007). Natural Humoral Immune Competence and Survival in Layers. *Poultry Science* 86(6), 1090–1099. doi: https://doi.org/10.1093/ps/86.6.1090

Tadelle, D., Kijora, C., and Peters, K. J. (2003). Indigenous Chicken Ecotypes in Ethiopia: Growth and Feed Utilization Potentials. *International Journal of Poultry Science* 2(2), 144–152. doi: https://doi.org/10.3923/ijps.2003.144.152

Taha, A. E., El-Edel, M. A., El-Lakany, H. F., and Shewita, R. S. (2012). Growth Performance and Immune Response Against Newcastle and Avian Influenza Vaccines in Egyptian Chicken Strains. *Global Veterinaria* 9(4), 434–440. url: https://www.idosi.org/gv/GV9(4)12/10.pdf.

Wakchaure, R. and Ganguly, S. (2016). Disease Resistance in Livestock: A Review. *Journal of Immunology and Immunopathology* 18(2), 94–94. doi: https://doi.org/10.5958/0973-9149.2016.00015.0

Weigend, S. and Romanov, M. N. (2001). Current strategies for the assessment and evaluation of genetic diversity in chicken resources. *World's Poultry Science Journal* 57(3), 275–288. doi: https://doi.org/10.1079/wps20010020

Witter, R. L. (1998). Control strategies for Marek's disease: a perspective for the future. *Poultry Science* 77(8), 1197–1203. doi: https://doi.org/10.1093/ps/77.8.1197

Yang, Y., Mekki, D. M., Lv, S. J., Wang, L. Y., Yu, J. H., and Wang, J. Y. (2006). Analysis of fitting growth models in Jinghai mixed-sex yellow chicken. *International Journal of Poultry Science* 5(6), 517–521.

Genetic Resources

OTHER ARTICLE

Genetic Resources (2020), 1 (2), 51–68
DOI: 10.46265/genresj.KZEW5023
https://www.genresj.org
ISSN: 2708-3764

A historical appraisal of the tropical forages collection conserved at CIAT

Rainer Schultze-Kraft [*,a], **Michael Peters** [b] **and Peter Wenzl** [c]

[a] *Emeritus, International Center for Tropical Agriculture (CIAT)*
[b] *Tropical Forages Program, Africa Hub, International Center for Tropical Agriculture (CIAT), Nairobi, Kenya*
[c] *Genetic Resources Program, Headquarters, International Center for Tropical Agriculture (CIAT), Cali, Colombia*

Abstract: A report on the tropical forage germplasm collection conserved in the CIAT genebank is presented. Emphasis is firstly on the assembling of the collection during 1972−1993 through about 70 major and minor collecting missions in tropical America, Africa and Southeast Asia. Along with introductions from existing collections, currently some 1,600 accessions of 134 grass species and 21,000 accessions of 637 legume species are being maintained. Secondly, information on the utilization of the collection, with emphasis on cultivar development based on selection of accessions from the CIAT collection, is presented. Worldwide, a total of 44 grass and 34 legume cultivars derived from germplasm maintained at CIAT are reported. Information on germplasm distribution and knowledge sharing during the last four decades is also presented as well as a brief discussion on future needs.

Keywords: tropics, wild species, grasses, legumes, germplasm, collection, utilization, genebank

Citation: Schultze-Kraft, R., Peters, M., Wenzl, P. (2020). **A historical appraisal of the tropical forages collection conserved at CIAT**. *Genetic Resources* 1 (2), 51–68. doi: 10.46265/genresj.KZEW5023.

© Copyright 2020 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Research related to forage genetic resources at the International Center for Tropical Agriculture (CIAT), Cali, Colombia, has been a continuing activity since the inception of the center in 1969. CIAT was the third of the international agricultural research centers established within the Consultative Group on International Agricultural Research (CGIAR), one of its missions being the development of beef cattle production in the lowlands of tropical America (Lynam and Byerlee, 2017). When looking at the development of the CIAT forages collection and its achievements, two phases can be distinguished: the first phase focused on assembling the collection and some initial, however intensive, characterization and utilization of the collected materials (1972−1993); and the second phase consisted of continuing utilization of germplasm, diversity studies and routine germplasm management and its optimization (1993−2020).

This paper summarizes the assembling of the collection during the first phase as well as its utilization and germplasm distribution. Furthermore, research on forage diversity and knowledge sharing, including training, are addressed. Data presented were compiled mainly from information accessible via the CIAT website (https://ciat.cgiar.org/), CIAT annual reports (accessible at https://cgspace.cgiar.org/handle/10568/35699), unpublished reports on germplasm collecting missions, research bulletins of CIAT´s national partner institutions and the like.

It should be noted that in the past years there have been changes in plant nomenclature for numerous taxa of particular interest, even at the genus level (Cook and Schultze-Kraft, 2015). In this paper we are still referring to the earlier used names. Supplemental Table 1 lists new names of species mentioned in this report, following the taxonomy of GRIN, the database of the USDA Genetic Resources Information Network (https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysearch).

*Corresponding author: Rainer Schultze-Kraft (Rainer@Schultze-Kraft.de)

Received: 28.09.2020 | Accepted: 29.11.2020 | Published online: 30.12.2020

Assembling the forage germplasm collection

At the beginning of research at CIAT in the early 1970s, a focus of the then Beef Production Systems Program was forage-based livestock production on acid, low-fertility soils in humid and sub-humid lowlands of tropical America, particularly savanna areas. The lack of edaphic adaptation of available, mainly Australian commercial pasture grass and legume cultivars, was soon identified as the main constraint to their use in Neotropical savannas. In the case of legumes, diseases, such as anthracnose in the promising genus *Stylosanthes*, were an additional restriction. It was consequently recognized that available species and genotype pools of grasses and legumes needed to be broadened for screening for adaptation to abiotic (mainly soil) and biotic (pests and diseases) constraints.

Missions were thus initiated in 1972/73 to collect germplasm of wild species with forage potential throughout tropical America. The objective was to create a diverse germplasm pool that can be tapped for cultivar development, either by identifying suitable accessions for direct use or through genetic improvement. These collecting missions ranged from short excursions, particularly within the Center´s host country, Colombia, to field expeditions over several weeks. Another source of germplasm material was through opportunistic collecting undertaken by CIAT scientists during field visits. The missions were largely funded by CIAT, at the beginning of the collecting phase, also with support from the former International Board for Plant Genetic Resources (IBPGR; later International Plant Genetic Resources Institute (IPGRI), now part of the Alliance of Bioversity International and CIAT). There was a focus on acid-soil regions and plant genera of known value. Emphasis was on legumes, in many cases including associated rhizobia, taking into account that the Neotropics are the main center of diversification of the Fabaceae (Leguminosae) family. The particular value of legumes lies in their ability of symbiotic fixation of atmospheric nitrogen and the subsequent provision of protein-rich forage to livestock.

From 1979 onwards, collecting missions within the new Tropical Pastures Program expanded to Southeast Asia, a minor yet important center of legume diversification (e.g. the genera *Pueraria* and *Desmodium*), and in 1984/85 and 1989 also to Africa, with focus on grasses (particularly the genus *Brachiaria*). The latter took into account that Sub-Saharan Africa is the main center of diversification of those genera with forage potential in the Poaceae family.

All missions were organized as joint ventures in association with national research institutions. Table 1 provides an overview of the countries where the collecting efforts were undertaken; the main genera collected are summarized in Table 2. The germplasm collectors who participated in the missions are acknowledged in Supplemental Table 2.

Alongside the collecting activities, the introduction of germplasm in the form of donations from existing collections held by national and international research and development (R&D) institutions around the globe has been an ongoing activity since the inception of the CIAT forages collection. It became the sole source of additions to the collection, when germplasm collecting with the involvement of CIAT scientists came to an end in the 1990s. A total of more than 9,000 accessions were received as donations; among them, in 2006, a significant part of the former Australian Tropical Forages Collection of CSIRO (Commonwealth Scientific and Industrial Research Organisation) (Table 3).

Two factors contributed to the cessation of collecting missions: (1) a sustained decline in funding for tropical-forages research and (2) the fact that hardly any tropical-forage species are included in Annex 1 of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA; FAO, 2001), which streamlines germplasm distribution through the use of a Standard Material Transfer Agreement (SMTA). Instead, most tropical forage species are regulated by legal frameworks that require case-by-case negotiations of individual material transfer agreements, which would create unmanageable legal overheads given the frequent germplasm exchange in agricultural research.

Some concluding remarks on the assembling of the forage germplasm collection

With more than 22,000 accessions from a total of 75 countries of origin, the CIAT collection is the largest tropical forages germplasm collection worldwide. Its particular value lies in its focus on: (1) plants adapted to acid, low-fertility soils; (2) legumes; and (3) the large and diverse collection of *Brachiaria*. This grass genus comprises currently the economically most important tropical forage species worldwide. The CIAT *Brachiaria* collection stems mainly from the Center´s collecting activities in the 1980s in East Africa and has become an important source of germplasm for selection and breeding programs throughout the tropics.

It is recognized that there are still important gaps in terms of countries and regions where germplasm has been collected. The collection is likely far from being representative of the geographic diversity of tropical Poaceae and Fabaceae (Leguminosae).

Table 4 provides a summary of the current tropical forages collection conserved at the CIAT genebank. Differences of numbers in Table 4 in comparison with Tables 2 and 3 are due to initial misidentifications and losses of accessions due to a range of reasons, including a limited number (or low viability) of collected seeds, lack of seed setting in the environments available for regeneration, and insufficient funds for regenerating large numbers of accessions between the 1990s and the launch of the CGIAR Research Program on Genebanks in the 2010s.

Table 1. Forage germplasm collecting missions conducted by CIAT with national research institution partners during 1972−1993. See Supplemental Table 2 for collector names

Year	Region and countries (No. of missions)	Genera collected
1972-73	**South America**: Colombia (2), Venezuela (1)	Legumes: Mainly *Stylosanthes*, also *Centrosema, Desmodium* & others
1974-75	**South America**: Bolivia (1), Brazil (4), Colombia (1), Venezuela (1)	Legumes: Mainly *Stylosanthes*, also *Centrosema, Desmodium* & others
1976-77	**South America**: Brazil (1), Colombia (3), Venezuela (1)	Legumes: Mainly *Stylosanthes*, also *Centrosema, Desmodium* & others
1978-79	**Mesoamerica**: Panama (1) **South America**: Brazil (1), Colombia (1), Venezuela (1) **Southeast Asia**: Thailand (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Desmodium, Pueraria* & others
1980-81	**South America**: Brazil (3), Colombia (2), Venezuela (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others
1982-83	**South America**: Colombia (1), Peru (1) **Southeast Asia**: Malaysia (1), Papua New Guinea (1), Thailand (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Desmodium, Pueraria* & others
1984	**South America**: Brazil (1), Colombia (3), Venezuela (1) **Africa**: Ethiopia (1), Kenya (1) **Southeast Asia**: China (1), Indonesia (1), Thailand (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Grasses: Mainly *Brachiaria* Legumes: *Desmodium, Pueraria* & others
1985	**Mesoamerica**: Panama (1) **South America**: Colombia (1), Venezuela (1) **Africa**: Burundi (1), Rwanda (1), Tanzania (1), Zimbabwe (1) **Southeast Asia**: Indonesia (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Centrosema, Desmodium, Stylosanthes* & others Grasses: Mainly *Brachiaria* Legumes: *Desmodium, Pueraria* & others
1986	**Mesoamerica**: Costa Rica (1), Mexico (1) **South America**: Colombia (3), Venezuela (2) **Southeast Asia**: Indonesia (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Desmodium, Pueraria* & others
1987-88	**South America**: Brazil (1), Colombia (5) **Southeast Asia**: China (1), Thailand (1)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Desmodium, Pueraria* & others
1989-90	**Mesoamerica**: Honduras (1) **South America**: Colombia (1) **Africa**: Cameroon (1)	Legumes: *Centrosema, Desmodium* & others Legumes: *Centrosema, Desmodium, Stylosanthes* & others Grasses: *Hyparrhenia, Andropogon* & others
1991-93	**South America**: Colombia (2) **Southeast Asia**: Thailand (1), Vietnam (2)	Legumes: *Centrosema, Desmodium, Stylosanthes* & others Legumes: *Desmodium, Pueraria* & others

Table 2. Most frequently collected genera with numbers of species and samples (accessions) and target countries, obtained through collecting during 1972−1993 (CIAT missions with national research institution partners and opportunistic collecting) and conserved at the CIAT genebank.

Genus	No. of species	No. of accessions	Countries of origin[1]
Legumes			
Aeschynomene	36	821	BDI, BRA, CHN, CMR, COL, CRI, ETH, HND, IDN, KEN, MEX, MYS, PAN, PER, PNG, THA, VEN, VNM
Alysicarpus	6	193	BDI, BRA, CHN, CMR, COL, ETH, HND, IDN, KEN, MYS, PAN, PNG, TGO, THA, VEN, VNM, ZWE
Cajanus	5	75	BRA, CHN, COL, ECU, IDN, PAN, PNG, THA, VNM
Calopogonium	5	446	BOL, BRA, CHN, COL, CRI, ECU, HND, IDN, KEN, MYS, PAN, PER, THA, VEN, VNM
Canavalia	14	185	BRA, CHN, COL, CRI, ECU, HND, MEX, PAN, PER, THA, VEN
Centrosema	33	1,677	BRA, CHN, COL, CRI, DOM, ECU, GTM, HND, IDN, MEX, MYS, PAN, PER, THA, VEN, VNM
Chamaecrista	16	257	BDI, BRA, CHN, COL, CUB, ECU, ETH, HND, KEN, MEX, MYS, NGA, PAN, PER, TGO, VEN, VNM, ZWE
Crotalaria	16	192	BRA, CHN, CMR, COL, CRI, ECU, ETH, HND, IDN, MEX, PAN, PER, PNG, THA, VEN, VNM
Desmanthus	9	110	BRA, COL, ECU, HND, MEX, PAN, PER, VEN
Desmodium	56	2,085	AUS, BDI, BLZ, BOL, BRA, CHN, CMR, COL, CRI, ECU, ETH, GUF, HND, IDN, IND, KEN, MEX, MYS, PAN, PER, PNG, PYF, SLV, TGO, THA, VEN, VNM, ZWE
Dioclea	15	191	BRA, COL, HND, MEX, PAN, PER, VEN
Flemingia	8	130	CHN, CMR, COL, HND, IDN, MYS, PAN, PNG, THA, VNM
Galactia	14	462	BOL, BRA, CHN, COL, CRI, CUB, ECU, HND, MEX, PAN, PER, PNG, VEN, VNM
Indigofera	12	147	BDI, BOL, BRA, CHN, COL, ECU, ETH, KEN, PAN, PER, PNG, TGO, THA, VEN, VNM
Macroptilium	10	445	ATG, BLZ, BOL, BRA, CHN, COL, CUB, DOM, ECU, GTM, HND, MEX, PAN, PER, SLV, THA, VEN,
Phyllodium	5	130	CHN, IDN, PNG, THA, VNM
Pueraria	6	187	BRA, CHN, COL, CRI, ECU, HND, IDN, MYS, PAN, PNG, THA, VEN, VNM
Rhynchosia	17	303	BOL, BRA, CMR, COL, CRI, ECU, ETH, HND, IDN, KEN, MEX, PAN, PER, PNG, RWA, THA, VEN

Continued on next page

Table 2 continued

Genus	No. of species	No. of accessions	Countries of origin[1]
Stylosanthes	24	2,263	ATG, AUS, BOL, BRA, CHN, CMR, COL, CRI, DOM, ECU, ETH, GUY, HND, KEN, MEX, MYS, NGA, PAN, PER, RWA, TGO, VEN, VNM, ZWE
Tadehagi	3	99	CHN, IDN, PNG, THA, VNM
Tephrosia	14	112	BDI, BRA, CHN, COL, ECU, ETH, KEN, MYS, PAN, PNG, THA, VEN, ZWE
Teramnus	9	304	BDI, BOL, BRA, CHN, COL, CRI, ECU, ETH, HND, IDN, KEN, MEX, PAN, PER, THA, VEN, ZWE
Uraria	6	109	CHN, IDN, MYS, PNG, THA, VNM
Vigna	39	481	BDI, BRA, CMR, COL, CRI, ECU, ETH, HND, IDN, KEN, MEX, MYS, PAN, PER, PNG, TGO, THA, VEN, ZWE
Zornia	12	812	AUS, BRA, CHN, CMR, COL, ETH, HND, KEN, MEX, MYS, NGA, PAN, PER, TGO, THA, VEN, ZWE
59 further legume genera with a total of:		693	
Total of legume samples		**12,909**	
Grasses			
Andropogon	2	20	BDI, TGO, ZWE
Brachiaria	17	472	BDI, CMR, COL, ECU, ETH, KEN, RWA, TGO, TZA, ZWE
Hyparrhenia	12	35	BDI, ETH, KEN, NGA, TGO, TZA, ZWE
Panicum	5	20	BDI, ETH, COL, KEN, RWA, ZWE
Paspalum	11	55	BDI, BRA, CHN, COL, DOM, IDN, KEN, MYS, PAN, PER, PHL, VEN, ZWE
Pennisetum	5	23	BDI, CHN, CMR, ETH, KEN
24 further grass genera with a total of:		127	
Total of grass samples		**752**	

[1]Country abbreviations: ATG = Antigua and Barbuda; BDI = Burundi; BLZ = Belize; BOL = Bolivia; BRA = Brazil; CHN = China, People's Republic; CMR = Cameroon; COL = Colombia; CRI = Costa Rica; CUB = Cuba; DOM = Dominican Republic; ECU = Ecuador; ETH = Ethiopia; GTM = Guatemala; GUF = French Guyana; GUY = Guyana; HND = Honduras; IDN = Indonesia; IND = India; KEN = Kenya; MEX = Mexico; MYS = Malaysia; NGA = Nigeria; PAN = Panama; PER = Peru; PHL = Philippines; PNG = Papua New Guinea; PYF = French Polynesia; RWA = Rwanda; SLV = El Salvador; TGO = Togo; TZA = Tanzania; THA = Thailand; VEN = Venezuela; VNM = Vietnam; ZWE = Zimbabwe.

Utilization

Germplasm collecting and introduction activities were the responsibility of the former CIAT Beef Production Systems Program (renamed Tropical Pastures Program (TPP) in 1979) until the late 1980s. Maintenance of the collection passed on to the CIAT Genetic Resources Unit (GRU) after its foundation in 1977 (in 2009 renamed Genetic Resources Program, GRP). Maintenance encompasses activities such as seed testing, seed increase, germplasm preservation at different cold storage levels, safety back-ups of the collection, maintenance of living collections, seed distribution, etc. Standard Operating Procedures (SOP) for each of these areas are available upon request. The collection is conserved in the GRP´s genebank at CIAT Headquarters, Cali, Colombia. At the time of writing (2020), 90% of CIAT's tropical forages collection is backed-up in the Svalbard seed vault and 77% at CIMMYT (https://ciat.cgiar.org/what-we-do/crop-conservation-and-use/tropical-forage-diversity/) .

For effective utilization of the germplasm collection, the close interaction between the genebank and CIAT´s TPP (today: Tropical Forages Program, TFP) played a key role, particularly in the 1980s, in view of the TPP's germplasm-focused research agenda and its alliances with national R&D institutions. The main research approaches were:

- Multidisciplinary germplasm characterization and evaluation by the TPP with involvement of specialists in the fields of agronomy, plant pathology, entomology, plant and animal nutrition, legume rhizobiology, soils, pasture establishment, pasture utilization, farming systems, economics, plant breeding, and seed production;
- Stepwise categorization of promising accessions as germplasm evaluations proceeded from small-plot observations to ultimately animal production experiments under grazing;
- Multi-site testing of germplasm in different eco-climatic zones at: CIAT-Quilichao (Colombia), Colombia-Llanos (with ICA, at Carimagua), Brazil-Cerrados (with EMBRAPA, at Planaltina), Peru-Humid Tropics (with IVITA, at Pucallpa), and later Central America (with Ministerio de Agricultura y Ganadería, in Costa Rica);
- Networking: multi-location testing of elite germplasm within the networks:
  - RIEPT (Red Internacional de Evaluación de Pastos Tropicales), with national research institution partners in Latin America and the Caribbean);
  - RABAOC/WECAFNET (Reseau de Recherche en Alimentation du Bétail en Afrique Occidentale et Centrale/West and Central African Forage Evaluation Network), with national research institution partners in West and Central Africa and in cooperation with ILCA/ILRI (International Livestock Centre for Africa/International Livestock Research Institute);
  - SEAFRAD (South East Asian Forage and Feed Resources Network), with national research institution partners in Southeast Asia; this network developed in the 1990s into the ACIAR (Australian Centre for International Agricultural Research) funded Forages for Smallholders project;
- Development, publication and use of network-wide common research methodologies (see section Publications below);
- Publication of research results in RIEPT reports (accessible at https://cgspace.cgiar.org/discover?scope=%2F&query=riept) and the TPP's journal Pasturas Tropicales (accessible at https://tropicalgrasslands.info/index.php/tgft/pages/view/Pasturas).

As a result of the germplasm evaluation and selection research conducted by CIAT and its partners, a number of grass and legume accessions that had been provided by the CIAT forages collection were developed into cultivars by national R&D institutions worldwide, in a few cases after incorporating them in national breeding programs (Table 5). Several issues should be pointed out:

- In relation to cultivars selected from genebank accessions, formal release and/or registration has been and is the exclusive responsibility of national institutions. In the case of bred lines developed at CIAT (not included in Table 5), cultivar release/registration is done by the private sector, with authorization by CIAT.
- Some cases are mentioned where CIAT accessions were adopted by end users without a formal and documented cultivar release; the real number is probably much higher. On the other hand, it is most likely that not all released cultivars were or are actually used by farmers to a major extent. The use of materials from tropical forages genebanks is further discussed by Hanson *et al* (2020) and Duncan *et al* (2020).
- Several important grass varieties, which had been developed by institutions in Brazil [(e.g. *Brachiaria brizantha* cv. Marandu (CIAT 6294) and *Panicum maximum* cvv. Tanzânia (CIAT 16031), Mombaça (CIAT 9692) and Tobiatã (CIAT 6299)] and subsequently introduced to CIAT, are not included in Table 5. Nonetheless, CIAT, within its network evaluation activities in Latin America and SE Asia, was instrumental for eventual release and adoption of these cultivars by non-Brazilian end-users.
- Mentioning of cultivar releases of the accessions *Brachiaria decumbens* CIAT 606 and *B. humidicola* CIAT 679 is restricted to countries other than Brazil: Both accessions represent early Australian cultivars ('Basilisk' and 'Tully', respectively) and,

Table 3. Forage germplasm donations received during 1972−2006 from national and international institutions and conserved at the CIAT genebank, highlighting main genera and number of samples (accessions).

Period	Total no. of accessions	Main genera	Main donor institutions[1]
1972−75	176	*Centrosema, Desmodium, Stylosanthes, Vigna*	INIAP (Ecuador), CSIRO (Australia), QDPI (Australia), IDRC projects (West Indies, Belize), Univ. Florida (USA), FAO-David (Panama)
1976−80	1,424	*Andropogon, Centrosema, Desmodium, Panicum, Stylosanthes, Zornia*	Instituto de Pesquisas IRI (Brazil), IDRC projects (West Indies, Belize), EPAMIG (Brazil), INIAP (Ecuador), EMGOPA (Brazil), Univ. Florida (USA), EMBRAPA (Brazil), CSIRO (Australia)
1981−85	1,971	*Brachiaria, Calopogonium, Centrosema, Crotalaria, Desmanthus, Desmodium, Galactia, Leucaena, Macroptilium, Macrotyloma, Panicum, Pueraria, Stylosanthes, Vigna, Zornia*	ORSTOM (France), CSIRO (Australia), EMBRAPA (Brazil), EMGOPA (Brazil), INTA (Argentina), QDPI (Australia), EPAMIG (Brazil), FAO-Kitale (Kenya), IBPGR-SE Asia
1986−90	1,238	*Alysicarpus, Brachiaria, Centrosema, Desmodium, Pseudarthria, Pueraria, Stylosanthes, Uraria, Vigna*	EMBRAPA (Brazil), CSIRO (Australia), Instituto de Zootecnia (Brazil), FONAIAP (Venezuela), Univ. Florida (USA), USDA (USA), IBPGR-SE Asia
1991−95	212	*Arachis, Cajanus, Panicum*	EMBRAPA (Brazil), ICRISAT (India), ICA (Colombia)
1996−00	218	*Arachis, Calliandra, Cratylia, Paspalum, Stylosanthes*	EMBRAPA (Brazil), OFI (UK)
2001−06	3,858	*Aeschynomene, Centrosema, Desmanthus, Desmodium, Lablab, Macroptilium, Stylosanthes, Vigna*	CSIRO (Australia)
2007−20	0		
Total	**9,157**		

[1]Abbreviations of main donor institutions: CSIRO = Commonwealth Scientific and Industrial Research Organisation; EMBRAPA = Empresa Brasileira de Pesquisa Agropecuária; EMGOPA = Empresa Goiana de Pesquisa Agropecuária; EPAMIG = Empresa de Pesquisa Agropecuária de Minas Gerais; FAO = Food and Agriculture Organization of the United Nations; FONAIAP = Fondo Nacional de Investigaciones Agropecuarias; IBPGR = International Board for Plant Genetic Resources; ICA = Instituto Colombiano Agropecuario; ICRISAT = International Crops Research Institute for the Semi-Arid Tropics; IDRC = International Development Research Centre; INIAP = Instituto Nacional de Investigaciones Agropecuarias; INTA = Instituto Nacional de Tecnología Agropecuaria; OFI = Oxford Forestry Institute; ORSTOM = Office de la Recherche Scientifique et Technique Outre-Mer; QDPI = Queensland Department of Primary Industries; USDA = United States Department of Agriculture.

unlike the other countries where CIAT-coordinated network evaluations were instrumental for selection and release, adoption in Brazil was an entirely CIAT-independent process.

- *Brachiaria* spp. accessions that were used by CIAT internally to produce *Brachiaria* breeding lines and from which hybrid cultivars were developed by globally operating seed companies (Grupo Papalotla, Dow AgroSciences) within public-private-partnership (PPP) agreements, are not included either. Information on those hybrid cultivars (e.g. 'Mulato', Mulato II', 'Cayman', 'Cobra', 'Camello') is available in Cook *et al* (2020).
- Table 5 also provides information on the timespan between germplasm acquisition (collection or introduction) and cultivar release.

In the early 1990s, CIAT gradually changed its germplasm utilization research priority from development of grass/legume pastures for acid soils to selection of multipurpose plants for smallholders. This adjustment took into account the need for forage plants to maintain and restore soil fertility, including in mixed (crop-livestock) production systems, and to contribute to increased small-farmer livelihoods. Since 2006/07 the research focus of the CIAT Tropical Forages Program was further refined and includes the adaptation of forages to climate change and their potential contribution to ecosystem services, including the mitigation of greenhouse gas emissions. Within these new developments, species selection is benefiting from the broad diversity represented in the CIAT forages collection.

Table 4. The tropical forages collection conserved at the CIATgenebank (as of July 2020).

Genus	**No. of accessions**	**No. of** species
Grasses:		
Brachiaria	592	22
Panicum	544	11
Paspalum	140	19
Andropogon	89	2
Hyparrhenia	44	12
Pennisetum, Cenchrus, Echinochloa, Axonopus, Setaria, Eragrostis and *Chloris* (more than 10 accessions each) and 19 other genera	202	68
Total grasses	**1,611**	**134**
Legumes:		
Stylosanthes	4,198	31
Desmodium	3,484	71
Centrosema	2,849	33
Aeschynomene	1,182	33
Macroptilium	1,052	11
Vigna	1,026	36
Zornia	947	14
Galactia	561	13
Calopogonium	550	4
Rhynchosia	384	13
Teramnus, Chamaecrista, Desmanthus, Crotalaria, Alysicarpus, Pueraria, Canavalia, Dioclea, Leucaena, Indigofera, Flemingia, Uraria, Arachis, Clitoria, Lablab, Tephrosia, Phyllodium, Cajanus and *Tadehagi* (between 100 and 384 accessions each) and 59 other genera	4,848	378
Total legumes	**21,081**	**637**
Grand total	**22,692**	**771**

Table 5. Formally and informally released cultivars developed from accessions provided by the CIAT tropical forages collection.

Species	CIAT Accession no.	Year of acquisition[1]	Cultivar name	Country	Year of registration/ release	Comments
Legumes:						
Aeschynomene americana	CIAT 7026	C: 1978	Lee	Australia	1984	Collected in Panama in cooperation with IDIAP (Instituto de Investigación Agropecuaria de Panamá)
Arachis pintoi	CIAT 17434	I: 1983	Maní Forrajero Perenne	Colombia	1992	Introduced as CPI 58113 from CSIRO; = cv. Amarillo, released 1987 in Australia
			Pico Bonito	Honduras	1993	
			Maní Mejorador	Costa Rica	1994	
	CIAT 18744	I: 1984	Porvenir	Costa Rica	1998	Introduced as BRA-012122 from EMBRAPA
	Mix of CIAT 17434 and CIAT 18744	I: 1983 resp. 1984	Maní Forrajero	Panama	1997	See information on CIAT 17434 and CIAT 18744 above
	CIAT 22160	I: 1993	Reyan No. 12	PR China	2004	Introduced as BRA-031143 from EMBRAPA
Centrosema acutifolium	CIAT 5277	C: 1979	Vichada	Colombia	1987	Collected in Colombia
Centrosema macrocarpum	CIAT 25522; = mix of 12 accessions	C: 1980-84	Ucayali[2]	Peru SE Asia[3]	1992 1990s	Individual accessions collected in cooperation with national institution partners in Colombia (8 accessions), Venezuela (3) and Brazil (1)
Centrosema pubescens	CIAT 15160	C: 1984	Barinas[2]	SE Asia[3]	1990s	Collected in Venezuela in cooperation with FONAIAP
Chamaecrista rotundifolia	CIAT 21565	C: 1989	Minyin 2	PR China	2011	Collected in Colombia in cooperation with ICA; released after introduction as ATF 3248 from Australia to China
Codariocalyx gyroides	CIAT 3001	I: 1975	Belize[2]	SE Asia[3]	1990s	Introduced as CF-29 from IDRC-Belize
			Cora Cora[2]	Colombia	1990s	
Cratylia argentea	Mix of CIAT 18516 and CIAT 18668	I. resp. C.: 1985, 1984	Veranera Veraniega	Colombia Costa Rica	2002 2001	CIAT 18516 introduced from EMGOPA, CIAT 18668 collected in Brazil in cooperation with EMBRAPA

Continued on next page

Table 5 continued

Species	CIAT Accession no.	Year of acquisition[1]	Cultivar name	Country	Year of registration/ release	Comments
C. argentea (cont.)	Mixture of unknown CIAT accessions		Cratilia[2]	Venezuela	2000s	
				Bolivia	2000s	
Desmodium heterocarpon subsp. *ovalifolium*[4]	CIAT 13651	C: 1984	Maquenque	Colombia	2002	Collected in Thailand in cooperation with TISTR (Thailand Institute of Scientific and Technological Research)
	CIAT 350	I: 1973	Itabela	Brazil	1989	Commercial cover crop variety, introduced from FAO Seed Exchange Unit, Rome
			Reyan No. 16	PR China	2005	
Desmodium strigillosum	CIAT 13158	C: 1982	Reyan No. 27	PR China	2010s	Collected in Thailand in cooperation with TISTR
Flemingia macrophylla	CIAT 17403	C: 1982	Chumphon[2]	SE Asia[3]	1990s	Collected in Thailand in cooperation with TISTR
Leucaena leucocephala	CIAT 21888; mix of CIAT 17481, 17482, 17491 and 17492	C. resp. I.: 1982	Romelia	Colombia	1992	CIAT 17481 and 17482 collected as naturalized populations in Brazil; CIAT 17491 and 17492 introduced as K8 and K72, respectively, from University of Hawaii at Manoa
Stylosanthes capitata	CIAT 10280; mix of 5 accessions	C: 1975-77	Capica	Colombia	1983	All accessions (CIAT 1315, 1318, 1342, 1693 and 1728) collected in Brazil in cooperation with EMBRAPA and CSIRO
			Alfalfa Criolla[2]	Venezuela	1990s	
Stylosanthes guianensis var. *guianensis*	CIAT 184	C: 1973	Pucallpa	Peru	1985	Collected in Colombia
			Reyan No. 2	PR China	1991	Also known as 'Pi Hua Dou 184', 'Zhuhuacao'
			Reyan No. 5	PR China	2000	Selected from CIAT 184
			Stylo 184[2]	SE Asia[3]	1990s	
	CIAT 136	C: 1973	Reyan No. 7	PR China	2000	Collected in Colombia
	Unidentified CIAT accession number		Reyan No. 13	PR China	2003	Origin: mislabeled seed bags or a physical contaminant in a sample ("CIAT 1044") of a different species

Continued on next page

Table 5 continued

Species	CIAT Accession no.	Year of acquisition[1]	Cultivar name	Country	Year of registration/ release	Comments
S. guianensis var. *guianensis* (cont.)	CIAT 2340	C: 1980	Ubon stylo	Thailand	2002	Selection (GC 1480) used for the 4-line mix cv. Ubon stylo; see below; original accession CIAT 2340 collected in Colombia
Stylosanthes guianensis var. *pauciflora*[5]	CIAT 11833 (= cross CIAT 10136 × CIAT 2031)	C: 1974 C: 1978 Cross: 1991?	Ubon stylo	Thailand	2002	3 selections from CIAT 11833 (GC 1463, GC 1517, GC 1579) used for the 4-line mix cv. Ubon stylo; see above; original accessions CIAT 10136 and CIAT 2031 collected in Brazil in cooperation with EMBRAPA
	CIAT 1283	C: 1975	Reyan No. 10	PR China	2001	Collected in Brazil in cooperation with EMBRAPA
Stylosanthes seabrana	CIAT 10033	C: 1981	Unica	Australia	2000	Collected in Brazil in cooperation with EMBRAPA
Grasses:						
Andropogon gayanus	CIAT 621	I: 1973	Carimagua 1	Colombia	1980	Original seed sample introduced from Shika Research Station, Nigeria
			Planaltina	Brazil	1980	
			Sabanero	Venezuela	1983	
			Veranero	Panama	1983	
			San Martín	Peru	1984	
			Llanero	Mexico	1986	
			Andropogon	Cuba	1988	
			Veranero	Costa Rica	1989	
			Otoreño	Honduras	1989	
			Gamba	Nicaragua	1989	
			ICTA-Real	Guatemala	1992	
Brachiaria humidicola	CIAT 679	I: 1976	Humidícola	Colombia	1992	CIAT 679 is cv. Tully (Koronivia grass), released 1981 in Australia
			INIAP-NAPO 701	Ecuador	1985	

Continued on next page

Table 5 continued

Species	CIAT Accession no.	Year of acquisition[1]	Cultivar name	Country	Year of registration/ release	Comments
B. humidicola (cont.)	CIAT 679 (cont.)		Chetumal	Mexico	1991	
			Humidícola	Panama	1990	
			Aguja	Venezuela	1989	
	CIAT 26149	C: 1985	BRS Tupi	Brazil	2012	Collected in Burundi in cooperation with ISABU (Institut des Sciences Agronomiques du Burundi)
Brachiaria humidicola (former species name for the accession concerned: *B. dictyoneura*)	CIAT 6133	I: 1978	Llanero	Colombia	1987	Introduced as *B. dictyoneura* CPI 59610 from CSIRO
			Yanero[2]	SE Asia[3]	1990s	
			Ganadero	Venezuela	1992	
			Gualaca	Panama	1992	
			Brunca	Costa Rica	1994	
			Reyan No. 14	PR China	2004	
Brachiaria decumbens	CIAT 606	I: 1973	Brachiaria	Cuba	1986/87	CIAT 606 is cv. Basilisk, released 1966 in Australia
			Señal	Panama	1986	
			Chontalpo	Mexico	1989	
			Barrera	Venezuela	1989	
			Peludo	Costa Rica	1991	
			Reyan No. 3	PR China	1991	
Brachiaria hybrid	CIAT 16309 (*B. brizantha*)	C: 1984	BRS Ipyporã	Brazil	2017	*B. brizantha* CIAT 16309/ILCA 13619 (B4 at Embrapa), collected in Ethiopia in cooperation with ILCA, is one of the parental lines in the hybridization *B. brizantha* × *B. ruziziensis* that led to this Brazilian cultivar
Brachiaria ruziziensis	Unidentified CIAT accession number		Reyan No. 15	PR China	2005	Origin: mislabeled seed bags or a physical contaminant in a sample (“CIAT 6095”) of a different species
Brachiaria brizantha	CIAT 16125/ILCA 13372	C: 1984	BRS Piatã	Brazil	2007	Collected in Ethiopia in cooperation with ILCA

Continued on next page

Table 5 continued

Species	CIAT Accession no.	Year of acquisition[1]	Cultivar name	Country	Year of registration/ release	Comments
B. brizantha (cont.)	CIAT 16467/ILCA 12751	C: 1984	BRS Paiaguás	Brazil	2013	Collected in Kenya in cooperation with ILCA
	CIAT 26110	C: 1985	Toledo	Colombia	2002	Collected in Burundi in cooperation with ISABU
				Costa Rica	2001	
			Xaraés, MG5 Vitória	Brazil	2000/01	
	CIAT 6387	I: 1981	Serengeti[2]	SE Asia[3]	1990s	Introduced as K-75232A-E from FAO-Kitale, Kenya
	CIAT 16835	C: 1985	Karanga[2]	SE Asia[3]	1990s	Collected in Zimbabwe in cooperation with Grasslands Research Station Marondera
	CIAT 16315/ILCA 13635	C: 1984	Capiporã	Brazil	2003	Collected in Ethiopia in cooperation with ILCA
	CIAT 16488/ILCA 13079	C: 1984	Arapoty	Brazil	2003	Collected in Kenya in cooperation with ILCA
Panicum maximum	CIAT 6901	I: 1988	Reyan No. 8	PR China	2000	Introduced as K 71 from ORSTOM; origin: Kenya
	CIAT 6172	I: 1988	Reyan No. 9	PR China	2000	Introduced from Ecuador
	CIAT 6799	I: 1983	Agrosavia Sabanera	Colombia	2018	Introduced as G 27 from ORSTOM; origin: Angola
	CIAT 16051	I: 1983	Agrosavia Michaya	Colombia	2020[6]	Introduced as T 90 from ORSTOM; origin: Tanzania
Pennisetum purpureum × *P. glaucum*	CIAT 6263	I: 1979	Reyan No. 4	PR China	1998	King Grass; introduced from IDIAP, Panama

[1]C = year of collection; I = year of introduction of a donated accession; [2]no formal release; [3]several countries; [4]formerly known as *D. ovalifolium*; [5]taxon var. *pauciflora* not recognized by GRIN; [6]release programmed.

Distribution

The basis for effective and efficient utilization of CIAT's forage germplasm has been, and still is, the production and distribution of germplasm samples for plant introduction work by the Center´s GRU/GRP. Since 1980 a total of 93,222 germplasm samples, representing 14,695 accessions, were distributed to 110 countries. Approximately 41% of these samples went to CIAT's TPP/TFP. The remainder was distributed to users in Colombia (33%), Brazil (9%), Peru, China, Venezuela (4% each), Australia, Mexico, USA (3% each), and 102 other countries. Approximately 47% of all externally distributed samples went to national agricultural research institutes, 23% to universities, 16% to farmers, with the remainder going to CGIAR centers, NGOs, companies, regional organizations, and other genebanks worldwide (Figure 1). Seed samples were predominantly used for agronomic evaluations (64%), but also for basic research (13%), capacity building (12%) and other purposes such as applied research, genetic improvement and conservation, in alignment with the ITPGRFA and the SMTA. The list of accessions included in each seed shipment is linked to available passport data and more recently to accession-specific Digital Object Identifiers (DOI).

For evaluation trials and on-farm experimentation with accessions that were considered as highly promising by CIAT, its network partners and/or development agencies, seed in larger amounts has been/is being provided by the Tropical Forages Program (Tropical Pastures Program until 1992). Between 1982 and 2014, the TPP/TFP produced and distributed almost 64 tons of grass and legume seeds, either free of charge for research purposes or at cost price for larger quantities, representing a total of 27,325 samples, to 88 countries worldwide.

Some concluding remarks on the utilization and distribution of the CIAT forage germplasm collection

A strong multidisciplinary research program (TPP/TFP) at CIAT and its alliance with national partners and international networks was instrumental for effective and efficient utilization of the collection. This includes the capacity to supply sufficiently large seed samples for research beyond small-plot level and eventually for on-farm testing.

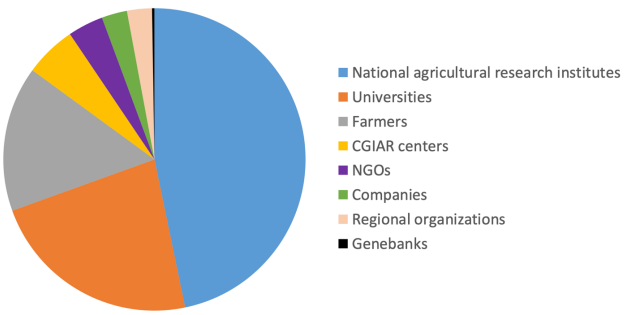

Figure 1. Historical distribution of 51,850 tropical-forage samples to different categories of germplasm users outside CIAT since the inception of CIAT's tropical forages collection.

The PPP arrangements between CIAT and the seed industry are proving to be conducive to increased adoption of improved-forage technology based on breeding lines.

The number of legume species (15), of which cultivars were released (33), contrasts with the number of grass species (9) and released cultivars (44). We suggest that this reflects, on the one hand, the particular focus of CIAT and its research partners on legumes in the past; on the other hand it reflects, to some extent, that the user community of tropical forages seems to be more receptive to grasses as they are easier to manage and have a better developed commercialization pathway.

Diversity in the forage germplasm collection

Basic plant descriptors have been used during both the germplasm multiplication phase by the GRU/GRP and the primary evaluation phase by the TPP/TFP to provide information on phenotypic diversity within a species and subsequent reduction of accession numbers for evaluation (e.g. the formation of core collections). In addition, molecular marker studies have been used to assess the genetic (= intraspecific) diversity and to elucidate species relationships for species of particular interest. On species that are little-known but of interest, basic floral-biology studies have been performed with the objective of optimizing germplasm management and enabling breeding. The main publications resulting from such research are listed below (section Publications).

The need for basic botanical and genetic studies to expand our knowledge about genera and species of forage interest must be stressed. Any future diversity research is recommended to focus on relevance for: (1) enhanced germplasm management and utilization (e.g. identification of duplicates, establishment of core collections); (2) relationships between traits and geographic origin of populations; and (3) identification of genes responsible for particularly important plant traits.

Sharing of knowledge

Training

During the period 1978−1990, the CIAT Tropical Pastures Program, within its coordinating role in the RIEPT network, held a yearly course 'Programa de Capacitación Científica en Investigación para la Producción de Pastos Tropicales'. The course, aimed at researchers from Latin America and the Caribbean, consisted of an intensive multi-disciplinary phase in which all participants were provided with lectures and practical training in all disciplines represented in the TPP (thus including the field of genetic resources of forage plants and germplasm handling) and a

specialization phase. With an average of 20 participants per course, a total of around 250 researchers were trained during the 13-year period, with about 10 specializing in genetic resources. In addition, several dozen students from both Colombian and foreign universities conducted research for their theses (BSc, MSc and PhD) with focus on genetic diversity of forages under the supervision of TPP/TFP scientists.

From 1990 onwards, training activities in forage germplasm management were essentially taken over by CIAT's GRU/GRP, mainly in the area of germplasm management and in the form of field days for Colombian university students and technicians, with demonstrations on field, greenhouse and laboratory activities. Several hundred students, technicians and researchers participated in this scheme.

Selected Publications

The following list comprises a selection of publications that, with regard to research topics and regions, we consider representative of the CIAT forage germplasm work during the past four decades:

Botanical studies

Schultze-Kraft, R., Williams, R.J. (1990). Una nueva especie de *Centrosema* (DC.) Benth. (Leguminosae: Papilionoideae) del Orinoco. *Caldasia* 16(77), 133–137. https://revistas.unal.edu.co/index.php/cal/article/view/35508

Torres, A.M. (1996). Un herbario de referencia para la colección de germoplasma de forrajes tropicales conservada por el Centro Internacional de Agricultura Tropical. *Pasturas Tropicales Boletín* 18(3), 71–74. https://cgspace.cgiar/handle/10568/87998

Maass, B.L., Torres González, A.M. (1998). Off-types indicate natural outcrossing in five tropical forage legumes in Colombia. *Tropical Grasslands* 32, 124–130. https://www.tropicalgrasslands.info/public/journals/4/Historic/Tropical%20Grasslands%20Journal%20archive/Abstracts/Vol_32_1998/Abs_32_02_98_pp124_130.ht ml

Bystricky, M., Schultze-Kraft, R., Peters, M. (2010). Studies on the pollination biology of the tropical forage legume shrub *Craylia argentea*. *Tropical Grasslands* 44, 246–252. https://www.tropicalgrasslands.info/public/journals/4/Historic/Tropical%20Grasslands%20Journal%20archive/PDFs/Vol_44%20(1_2_3_4)/Vol%2044%20(4)%20Bystricky%20et%20al%20246.pdf

Calles, T., Schultze-Kraft, R. (2010). Re-establishment of *Stylosanthes gracilis* (Leguminosae) at species level. *Kew Bulletin* 65(2), 233–240. doi: 10.1007/s12225-010-9198-z

Calles, T., Schultze-Kraft, R. (2010). *Stylosanthes* (*Leguminosae, Dalbergieae*) of Venezuela. *Willdenowia* 40, 305–329. doi: 10.3372/wi.40.40211

Germplasm collection and biogeography

Schultze-Kraft, R., Reid, R., Williams, R.J., Coradin, L. (1984). The existing *Stylosanthes* collections. In *The biology and agronomy of* Stylosanthes, eds. H.M. Stace & L.A. Edye. (North Ryde, N.S.W.: Academic Press Australia), p. 125–146. doi: 10.1016/B978-0-12-661680-4.50011-1

Schultze-Kraft, R., Pattanavibul, S., Gani, A., He, C., Wong, C.C. (1989). Collection of native germplasm resources of tropical forage legumes in Southeast Asia. In *Proceedings of the XVI International Grassland Congress, Nice, France*, p. 271–272.

Schultze-Kraft, R., Williams, R.J., Coradin, L., Lazier, J.R., Kretschmer Jr, A.E., Franco, M.A., Hernández, C.A. (1989). *1989 world catalog of* Centrosema *germplasm / Catálogo 1989 mundial de germoplasma de* Centrosema. (Cali, Colombia: CIAT and IBPGR), 319 p. https://hdl.handle.net/10568/53945

Costa, N.M.S., Schultze-Kraft, R. (1990). Biogeografia de *Stylosanthes capitata* Vog. e S. *guianensis* Sw. var. *pauciflora*. *Pesquisa Agropecuária Brasileira* 25(11), 1547–1554. https://seer.sct.embrapa.br/index.php/pab/article/view/13685

Schultze-Kraft, R., Williams R.J., Coradin, L. (1990). Biogeography of *Centrosema*. In Centrosema: *Biology, agronomy, and utilization*, eds. R. Schultze-Kraft & R.J. Clements. CIAT Publication No. 92. (Cali, Colombia: CIAT), p. 29–76. https://books.google.com.co/books?id=MMsSFJwy63gC&lpg=PP1&pg=PA29#v=onepage&q&f=false

Coradin, L., Schultze-Kraft, R. (1990). Germplasm collection of tropical pasture legumes in Brazil. *Tropical Agriculture (Trinidad)* 67(2), 98–100. https://journals.sta.uwi.edu/ta/index.asp?action=viewPastAbstract&articleId=1751&issueId=217

Franco, M.A., Ocampo, G.I., Melo, E., Thomas, R. (Comp.). (1993). *Catálogo de cepas de rizobios para leguminosas forrajeras tropicales / Catalogue of rhizobium strains for tropical forage legumes*. Working Document No. 14, 5th Edn. (Cali, Colombia: CIAT), 123 p. http://ciat-library.ciat.cgiar.org/forrajes_tropicales/pdf/4th/wd14.pdf

Flores, A.J., Schultze-Kraft, R. (1994). Recolección de recursos genéticos de leguminosas forrajeras tropicales en Venezuela. *Agronomía Tropical* 44(3), 357–371.

Keller-Grein, G., Maass, B.L., Hanson, J. (1996). Natural variation in *Brachiaria* and existing germplasm collections. In Brachiaria: *Biology, agronomy, and improvement*, eds. J.W. Miles, B.L. Maass, C.B. do Valle, V. Kumble. CIAT Publication No. 259. (Cali, Colombia: CIAT and EMBRAPA-CNPGC), p. 16–42. https://books.google.com.co/books?id=dMF6QpfVdjMC&lpg=PP1&pg=PA16#v=onepage&q&f=false

Diversity studies, including characterization and preliminary evaluation

Schultze-Kraft, R., Costa, N.M.S., Flores, A. (1984). *Stylosanthes macrocephala* M.B. Ferr. et S. Costa – Collection and preliminary agronomic evaluation of a new tropical pasture legume. *Trocpical Agriculture (Trinidad)* 61(3), 230–240. https://jounals.sta.uwi.edu/ta/index.asp?ation=viewPastAbstract&articleId=2322&issueId=243

Pizarro, E.A. (comp.). (1985). *Red de Evaluación de Pastos Tropicales – Resultados 1982−1985*. III Reunión de la RIEPT, Octubre 21−24, 1985. 2 Vols. (Cali, Colombia: CIAT), 1228 p. https://hdl.handle.net/10568/56339

Schultze-Kraft, R., Benavides, G. (1988). Germplasm collection and preliminary evaluation of *Desmodium ovalifolium* Wall. *Genetic Resources Communication No. 12*. (St. Lucia, Qld., Australia: CSIRO, Division of Tropical Crops and Pastures), 20 p. https://www.tropicalgrasslands.info/public/journals/4/GRC/GRC12%5b1792%5d.pdf

Torres González, A.M., Morton, C.M. (2005). Molecular and morphological phylogenetic analysis of *Brachiaria* and *Urochloa* (Poaceae). *Molecular Phylogenetics and Evolution* 37, 36–44. doi: 10.1016/j.ympev.2005.06.003

Andersson, M.S., Peters, M., Schultze-Kraft, R., Gallego, G., Duque, M.C. (2006). Molecular characterization of a collection of the tropical multipurpose shrub legume *Flemingia macrophylla*. *Agroforestry Systems* 68(3), 231–245. doi: 10.1007/s10457-006-9014-9

Andersson M.S., Schultze-Kraft, R., Peters, M., Duque, M.C., Gallego, G. (2007). Extent and structure of genetic diversity in a collection of the tropical multipurpose shrub legume *Cratylia argentea* (Desv.) O. Kuntze as revealed by RAPD markers. *Electronic Journal of Biotechnology* 10(3), 386–399. doi: 10.2225/vol10-issue3-fulltext-2

Research methodologies

Mott, G.O., Jiménez, C.A. (eds.). (1979). *Handbook for the collection, preservation and characterization of tropical forage germplasm resources.* (Cali, Colombia: CIAT), 95 p. https://hdl.handle.net/10568/54059

Paladines, O., Lascano, C.E. (eds.). (1993). *Forage germplasm under small-plot grazing: Evaluation methodologies.* CIAT Publication No. 210. (Cali, Colombia: RIEPT – CIAT), 249 p. https://hdl.handle.net/10568/56518

Toledo, J.M. (ed.). (1982). *Manual para la evaluación agronómica: Red Internacional de Evaluación de Pastos Tropicales (RIEPT)*. (Cali, Colombia: CIAT), 150 p. https://hdl.handle.net/10568/54148

Schultze-Kraft, R., Mannetje, L.'t. (2000). Evaluation of species and cultivars. In *Field and laboratory methods for grassland and animal production research*, eds. L.'t Mannetje & R.M. Jones. (Wallingford, UK: CAB International), 179–204. doi: 10.1079/9780851993515.0179

Monographs

Schultze-Kraft, R., Clements, R.J. (eds.). (1990). Cen-trosema: *Biology, agronomy, and utilization*. CIAT Publication No. 92. (Cali, Colombia: CIAT), 667 p. https://hd l.handle.net/10568/54383

Toledo, J.M., Vera, R., Lascano, C., Lenné, J.M. (eds.). (1990). Andropogon gayanus *Kunth – a grass for tropicalacid soils*. CIAT Publication No. 90. (Cali, Colombia: CIAT), 381 p. https://hdl.handle.net/10568/54190

Kerridge, P.C., Hardy, B. (eds.). (1994). *Biology and agronomy of forage* Arachis. CIAT Publication No. 240. (Cali, Colombia: CIAT), 209 p. https://hdl.handle.net/10568/54359

Miles, J.W., Maass, B.L., Valle, C.B. do, Kumble, V. (eds.). (1996). Brachiaria: *Biology, agronomy, and improvement.* CIAT Publication No. 259. (Cali, Colombia: CIAT and EMBRAPA-CNPGC), 288 p. https://hdl.handle.net/10568/54362

Cook, B.G., Pengelly, B.C., Schultze-Kraft, R., Taylor, M., Burkart, S., Cardoso Arango, J.A., González Guzmán, J.J., Cox, K., Jones, C., Peters, M. (2020). Tropical Forages: An interactive selection tool. 2nd and revised Edn. (Cali, Colombia and Nairobi, Kenya: CIAT and ILRI). www.tropicalforages.info

Some thoughts about the value of the CIAT tropical forages germplasm collection

Having safeguarded the germplasm is a major achievement in view of increasing world-wide genetic erosion due to habitat destruction and land use changes. The very availability of conserved germplasm represents an option value whose potential impact cannot be calculated and should not be underestimated.

Given that the CIAT tropical forages collection is an in-trust collection, held by CIAT on behalf of the countries of origin of the accessions, another incalculable option value is that all accessions stored in the genebank are available to users under an SMTA at any time. In this context, an important potential benefit to the 75 countries of origin of the germplasm conserved at CIAT's genebank is that it is available for repatriation.

The present report has to content itself with showing the number of released cultivars developed from CIAT's forages collection. An estimation of the total economic benefits derived from the collection is beyond the possibilities of genebank managers or scientists engaged in developing varieties. For this, field studies are suggested to be necessary, which have to involve resource economists for the assessment of acreages planted to the new cultivars, resulting livestock production increases, benefits to the environment (e.g. soil conservation and improvement) and estimates of increased livelihoods of end users.

Some considerations for the future

Although this report is concerned with both past and present, a couple of thoughts related to the future seem to be appropriate. Funding for tropical-forage research and deployment has declined considerably

over the last decades. *Ex situ* conservation of forages is substantially more expensive on a per-accession basis than conservation of other crops such as grains and pulses because cross-pollinating, wild and/or weedy accessions with long, asynchronous life cycles are substantially more expensive to regenerate thaninbred crops with short, synchronous life cycles (Koo *et al*, 2003), yet funding for tropical-forage research and deployment has declined considerably over the last decades. Accordingly, to ensure the long-term conservation and availability of critically important tropical-forage genetic resources, genebanks need to focus on the species with the greatest forage potential. A panel of well-known tropical-forage experts has grouped the species conserved at the CIAT and ILRI collections into priority categories based on the species' proven or suspected forage potential (Pengelly, 2015). Lowpriority accessions of both collections will be archived to focus conservation efforts on more promising species, with a view towards multiplying selected groups of 'bestbet' accessions so that larger quantities of starter seeds can be provided for research projects to shorten the path between the genebank and germplasm use in the field.

Compared with tropical-forage collections at ILRI in Ethiopia (17K accessions), USDA in the US (15K accessions), KARI (Kenya Agricultural Research Institute) in Kenya (15K accessions), the Australian Pastures Genebank (APG) in Australia (13K accessions), EMBRAPA in Brazil (9K accessions), INTA in Argentina, and IGFRI (Indian Grassland and Fodder Research Institute) in India, CIAT's collection is particularly rich in legumes adapted to infertile, acid soils. A full list of accessions, including passport data, photos of plants, flowers and seeds and images of herbarium specimens, can be accessed on the Genesys web portal by filtering for 'COL003' for CIAT and 'forages' as a crop (https://www.genesys-pgr.org).

In 2021, the CIAT genebank will achieve all the CGIAR Genebank Platform's 'performance targets' for its forage collection (Crop Trust, 2016). One of these targets is making more than 90% of all accessions available for immediate distribution, which means the genebank holds sufficient numbers of seeds that have been tested and found to be free of more than 40 different pathogens of quarantine importance. The collection, therefore, will qualify for long-term funding from the Endowment Fund of the Global Crop Diversity Trust to ensure its in-perpetuity conservation.

Although increasing germplasm collections by further collecting is currently difficult to justify, the policy framework may change over time, for example through an expansion of the ITGPRFA's scope to include all plant genetic resources for food and agriculture. Efforts to identify the most important collection gaps, therefore, are indicated to prepare for future collections. As suggested above, the forage germplasm collections currently conserved at the aforementioned genebanks, in addition to those at the CGIAR centers ICARDA (International Center for Agricultural Research in the Dry Areas) and World Agroforestry Centre (ICRAF) cannot be considered adequately representative of the geographic and taxonomic/genetic diversity of the tropical and subtropical Poaceae and Fabaceae (Leguminosae), neither at the level of the individualcollections nor at that of their consolidated total. It is suggested that such gap analyses:

- Consider both neglected areas/regions and genera/species of particular interest with emphasis on potential as forage and for environmental services.
- Be based on ecogeographical surveys paying particular attention to regions where biodiversity loss (including danger of genetic erosion) has been identified and/or is likely to occur.
- Consider the need for germplasm with focus on adaptation to the particular environmental challenges of the future in terms of the effects of climate change, such as increasing soil salinity, rising water table (including flooding), drought, temperature extremes etc. Collecting is suggested in areas of currently similar conditions under the assumption that germplasm occurring there has genetic adaptation to the respective limiting factors.

Such gap analyses can be the basis for future internationally coordinated collecting actions — when they become feasible — but independently may also stimulate immediate plans of action at national levels.

Acknowledgements

The cooperation and engagement of the germplasm collectors (Supplemental Table 2), who contributed to the tropical forages collection conserved at CIAT, are gratefully acknowledged.

Supplemental data

Supplemental Table 1: Recent nomenclatural changes of tropical forage species mentioned in this document under their previous names, based on the taxonomy accepted by the USDA Genetic Resources Information Network (GRIN).

Supplemental Table 2: Plant collectors who participated in the collecting missions mentioned in Table 1.

Author contributions

R.S.K. contributed to the conception and design of the submitted manuscript. All authors contributed to data gathering and analysis and to the drafting, revision and final approval of the submitted manuscript. The views expressed in this publication are those of the authors and do not necessarily reflect the views of CIAT, now part of the Alliance of Bioversity International and CIAT.

Conflict of interest statement

The Authors declare no conflict of interest.

References

Cook, B. G., Pengelly, B. C., Schultze-Kraft, R., Taylor, M., Burkart, S., Arango, J. A. C., Guzmán, J. J. G., Cox, K., Jones, C., and Peters, M. (2020). Tropical Forages: An interactive selection tool. 2nd and revised Edn. url: www.tropicalforages.info.

Cook, B. G. and Schultze-Kraft, R. (2015). Botanical name changes - nuisance or a quest for precision? *Tropical Grasslands - Forrajes Tropicales* 3(1), 34–40. doi: https://doi.org/10.17138/tgft(3)34-40

Crop Trust (2016). Genebank Platform proposal. url: https://www.genebanks.org/wp-content/uploads/2019/09/Genebanks-Platform-Full-Proposal.pdf.

Duncan, A. J., Peters, M., Schultze-Kraft, R., Thornton, P. K., Teufel, N., Hanson, J., and McIntire, J. (2020). The impact of CGIAR Centre research on use of planted forages by tropical smallholders. In *The Impact of the International Livestock Research Institute,* ed. McIntire, J. and Grace, D., (Nairobi, Kenya and Wallingford, UK: ILRI and CABI), 451-479.

FAO (2001). International Treaty on Plant Genetic Resources for Food and Agriculture. url: http://www.fao.org/3/a-i0510e.pdf.

Hanson, J., Schultze-Kraft, R., Peters, M., Wenzl, P., Amri, A., Shehadeh, A., and Yazbek, M. (2020). Forage diversity, conservation and use. In *The Impact of the International Livestock Research Institute,* ed. McIntire, J. and Grace, D., (Nairobi, Kenya and Wallingford, UK: ILRI and CABI), 423-449.

Koo, B., Pardey, P., and Wright, B. (2003). The price of conserving agricultural biodiversity. *Nature Biotechnology* 21(2), 126–128. doi: https://doi.org/10.1038/nbt0203-126

Lynam, J. and Byerlee, D. (2017). *Forever pioneers - CIAT: 50 years contributing to a sustainable food future... and counting.* Publication No. 444, (Cali, Colombia: CIAT), 140 p.

Pengelly, B. C. (2015). Global strategy for the conservation and utilisation of tropical and sub-tropical forage genetic resources. url: https://www.croptrust.org/wp/wp-content/uploads/ 2014/12/Forages-Strategy.pdf.