Genetic Resources

Genetic Resources (2021) Vol. 2, Issue 3
DOI: 10.46265/genresj.2021.3
www.genresj.org
ISSN: 2708-3764

Focus and Scope of *Genetic Resources*

Genetic Resources is an open access journal disseminating global knowledge and tools used by the community of practitioners of plant, 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:
Inflorescences and bulblets of asexual (left) and sexual (right) types of *Allium ampeloprasum*. Kik et al, pp. 1-10.

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 Alliance of Bioversity International and CIAT
Via di San Domenico 1
00153 Rome, Italy

Submissions to www.genresj.org

Editorial Board:

Genetic Resources

TABLE OF CONTENTS

Genetic Resources (2021), 2 (3)
DOI: 10.46265/genresj.2021.3
https://www.genresj.org
ISSN: 2708-3764

Review

Original Articles

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2021), 2 (3), 1–10
DOI: 10.46265/genresj.DMAT2233
https://www.genresj.org
ISSN: 2708-3764

Collecting and regenerating populations of the *Allium ampeloprasum* complex from Greece

Chris Kik [*,a], **Liesbeth De Groot** [a], **Gerlof Bottema** [b], **Marc Op 't Hof** [c], **Niels De Visser** [d], **Peter Willems** [e], **Toon Van Doormalen** [f], **Stelios Samaras** [g], **Elias Polemis** [h] **and Dimitris Tzanoudakis** [i]

[a] *Centre for Genetic Resources, Wageningen University & Research, Wageningen, 6700 AA, the Netherlands*
[b] *Enza Zaden Research and Development B.V, Haling 1/E, 1602 CB, Enkhuizen, the Netherlands*
[c] *Hazera Seeds B.V, Schansereind 27, 4921 PM, Made, the Netherlands*
[d] *Rijk Zwaan Zaadteelt en Zaadhandel B.V, Burgemeester Crezéelaan 40, 2678 KX, De Lier, the Netherlands*
[e] *Bejo Zaden B.V, Trambaan 1A, 1749 CZ, Warmenhuizen, the Netherlands*
[f] *Nunhems Netherlands B.V, Napoleonsweg 152, 6083 AB, Nunhem, the Netherlands*
[g] *Hellenic Agricultural Organization-DEMETER, Directorate General of Agricultural Research, Institute of Plant Breeding and Phytogenetic Resources, Thermi-Thessaloniki, 57001, Greece*
[h] *Laboratory of General and Agricultural Microbiology, Agricultural University of Athens, Athens, 11855, Greece*
[i] *Section of Plant Biology, Department of Biology, University of Patras, Patras, 26500, Greece*

Abstract: Collecting expeditions are of prime importance to acquire genetically unique material, as for many crops and their wild relatives, large gaps are present in collections worldwide. This is also true for the three species of the *Allium ampeloprasum* complex, native to Greece, which are considered as the crop wild relatives of cultivated leek (*Allium porrum*). Therefore, a collecting expedition was carried out in Greece in 2009. A total of 62 populations of *A. ampeloprasum*, 20 populations of *A. bourgeaui*, 19 populations of *A. commutatum* and three mixed species populations were sampled. The sampled populations were mostly small (less than 50 plants), but sometimes large populations (more than 10,000 plants) were encountered, especially for *A. commutatum*. Two different reproduction systems were observed in *A. ampeloprasum*, which is probably due to ploidy level differences. The sexual type was predominantly found along cultivated fields, whereas the asexual type occurred in abandoned fields together with *Sarcopoterium spinosum* (L). Spach and *Cistus* spp. Regeneration protocols were developed for these species as the phenology of cultivated leek is different from its wild relatives. Regenerating *A. ampeloprasum* was more difficult compared to the other two species. Ten years after the collecting mission only one-third of the collected material has been regenerated. This is partly due to the characteristics of the material and partly because the Dutch collectors and the Greek competent national authorities on Access and Benefit Sharing were not able to conclude a specific arrangement which also involved the commercial use of the material.

Keywords: Collecting expedition, regeneration, mode of reproduction, access and benefit sharing, leek, crop wild relatives

Citation: Kik, C., De Groot, L., Bottema, G., Op 't Hof, M., De Visser, N., Willems, P., Van Doormalen, T., Samaras, S., Polemis, E., Tzanoudakis, D. (2021). Collecting and regenerating populations of the *Allium ampeloprasum* complex from Greece. *Genetic Resources* 2 (3), 1–10. doi: 10.46265/genresj.DMAT2233.

Introduction

Crop wild relatives (CWR) are nowadays considered important gene reservoirs for the genetic improvement of their related crops. However, the occurrence of CWR in genebanks worldwide is often poor and the wild relatives of cultivated leek (*Allium porrum* L.), namely *Allium ampeloprasum* L., *A. bourgeaui* Rech.f. and *A. commutatum* Guss., are no exception to this (Keller and Kik, 2018). The breeding of new leek cultivars takes

*Corresponding author: Chris Kik (chris.kik@wur.nl)

Received: 05.03.2020 | Accepted: 18.01.2021 | Published online: 24.03.2021

many years as the crop is a cross-fertilizing (segmental) autotetraploid (2n=4x=32). Major challenges in leek breeding are the identification and subsequent introgression of disease (e.g *Phytophthora porri*) and pest resistances (e.g. *Thrips tabaci*) as these organisms cause major yield reductions when a crop is infected. Until present no adequate resistances have been found to many diseases and pests in cultivated leek germplasm.

The three CWR of leek form the *Allium ampeloprasum* complex (von Bothmer, 1970) which can be crossed successfully with their cultivated crop species leek (Kik *et al*, 1997; Kik, 2002a). Genecological aspects of the three cross-fertilizing wild species have been treated by von Bothmer (1970, 1974), taxonomic issues by de Wilde-Duyfjes (1976), Mathew (1996) and Hirschegger *et al* (2010). Breeding aspects of leek have been treated by de Clercq and Bockstaele (2002).

The geographic distribution of the three CWR species of leek is quite different (Hirschegger *et al*, 2010). *A. bourgeaui* is predominantly found on the Greek Aegean islands, but populations of the species have also been recorded from the Aegean coasts of both East continental Greece and West Asia minor. *A. commutatum* has a broader distribution range compared to *A. bourgeaui*, as it occurs also along the Adriatic coast and on the Mediterranean islands and islets west of the Italian peninsula (Sicily, Sardinia and Corsica). *A. ampeloprasum* has a very broad distribution range and can be found in the entire Mediterranean region and as far as Afghanistan in the East.

The ecological preferences of the three species differ to a large extent. *A. ampeloprasum* is a ruderal species which occurs in disturbed places such as cultivated and abandoned fields; population sizes range from small to large. *A. bourgeaui* is a chasmophyte and can be found on limestone cliffs and inland rocky slopes and occurs mostly in small populations. *A. commutatum* occurs in coastal habitats and can be found on islets and shores and is able to dominate the area (von Bothmer, 1974).

The first collecting expeditions focusing on the three species of the *A. ampeloprasum* complex took place between 1964-1969 (von Bothmer, 1970) as part of a larger floristic and biogeographical project (1957-1969) aimed at studying the effects of genetic and reproductive drift in small populations in the central Aegean archipelago (Runemark, 1969, 1970). In case of the *Allium* study, a total of 66 *A. ampeloprasum*, 56 *A. bourgeaui* and 64 *A. commutatum* populations were sampled (von Bothmer, 1974). In 1982 Q.P. van der Meer (IVT, Wageningen, the Netherlands) received from R. von Bothmer (University of Lund, Sweden) 21, 16 and 39 populations of *A. ampeloprasum, A. bourgeaui* and *A. commutatum* respectively from this expedition. After 20 years only ten, one and 20 populations of *A. ampeloprasum, A. bourgeaui* and *A. commutatum* respectively were still present in this working collection (Kik, 2002b). In the global GENESYS database (https://www.genesys-pgr.org/) no *Allium* material collected between 1964 and 1969 in Greece is included and only a few populations of *A. bourgeaui* and *A. commutatum* are present in this database. Given the low numbers of leek CWR accessions in public genebanks and considering the near absence of *in situ* management of the three *Allium* species, a new collecting expedition was clearly needed.

The aim of the present paper is to report on a collecting expedition in Greece which took place in 2009, describing the ecologies of the three *Allium* species, recording the geographic locations of the populations sampled, reporting on the challenges in collecting, regeneration and utilization, and providing two newly developed regeneration protocols for the three wild relatives of leek.

Material and Methods

Germplasm acquisition

A field collecting form, based on the modified FAO multi-crop passport descriptor list (Alercia *et al*, 2015), was used to document the passport data of the populations sampled (for collecting details: https://www.wur.nl/en/Research-Results/Statutory-research-tasks/Centre-for-Genetic-Resources-the-Netherlands-1/Genebank/Special-collections.htm).

All sampled material received a collecting number, in this case TKxxx (i.e. Tzanoudakis Kik followed by a number). Latitude, longitude and altitude were determined via GPS (Garmin, eTrex series Venture HC) with an inaccuracy of 1-5 meters, using WGS84 as map datum and ddd.dddd° as position format. Digital pictures were taken of all collecting sites (Kik, 2009).

Considering the peculiarities of the geography of Greece, the organisation of the mission (continental and insular areas to be visited, period and duration of the expedition) was mainly based on the information from the rich *Allium* collection (mainly herbarium specimens) deposited at the Botanical Museum of the University of Patras (UPA, Patras, Greece) and on the field experience of the Greek participants, especially of the senior author (D.T.) who is an expert on the taxonomy and geographical distribution of the genus *Allium* in Greece. The collecting mission took place from July 2 to August 21 2009. Collecting took place in the coastal area of the Peloponnese on the mainland and on the islands Karpathos, Crete, Andros, Kythera and Lesbos (Figure 1).

The three *Allium* species could easily be recognized in the surrounding vegetation because of the height of their flower stalks (ca. 1 m). As a rule of thumb, at least three individual plants needed to be seen at a first glance at a specific location to start an area survey. The area explored by foot per population varied from ca. 0.1 to 2 ha.

The inflorescences of all three species were collected in linen bags (20 x 35 cm). For each population collected, a label with the collecting number was put inside the bag and a label attached to the rope which tied up a bag. Whenever possible, around 20 inflorescences

Figure 1. Overview of the collecting sites of the leek CWR collecting mission 2009. The blue, yellow and red dots indicate *Allium ampeloprasum, A. commutatum* and *A. bourgeaui*, respectively. For collecting site details: https://www.wur.nl/en/Research-Results/Statutory-research-tasks/Centre-for-Genetic-Resources-the-Netherlands-1/Genebank/Special-collections.htm

per population were randomly collected and bulked in a linen bag. In almost all cases, the seed was not visible from the outside. Therefore, the quality of the seed was inspected by opening the green fruit capsule (ovary) manually and by checking if the seed was black or white. If fruit capsules in a number of inflorescences contained black (mature) seed, a seed sample was taken. The populations collected at the southeastern coast of Crete always had black seeds as in this part of Greece summer and high temperatures arrive early. Together with the inflorescence, 15-20 cm of the flower stalk was also collected to provide the maturing seeds with nutrients. If the seeds were still translucent/white, 5-10 bulblets (small bulbs inside the foliage leaves of the storage bulbs; not to be confused with bulbils which are present in the inflorescence) were collected per plant. In a few cases bulbs were also collected.

Upon arrival at CGN the material was dried outdoors in open plastic cages (day/night temperatures: +16/10°C; relative humidity: ~75%) in linen bags for two weeks. Then the material was transferred for three months to a conditioned storage room with a temperature of 15°C and 15% relative humidity. The seed cleaning also took place in this room. Subsequently, the dried seeds were placed in vacuum-sealed, three-layered aluminum foil bags and stored at -18°C, awaiting future regeneration. The bulblets and bulbs sampled during the expedition were planted in mid-November in a conditioned greenhouse.

Ex situ regeneration of the collected seed material

Protocols had to be developed for the regeneration of *Allium ampeloprasum*, *A. bourgeaui* and *A. commutatum* as the life cycle of cultivated leek is quite different from its wild relatives. Firstly, a protocol (Regeneration protocol I) for regeneration only under the climatic conditions of the Netherlands was developed. However, as regeneration in the Netherlands took three years and plants were lost during the regeneration process, an alternative protocol (Regeneration protocol II) was developed which involved a two stage regeneration process: the first stage took place under the climatic conditions of the Netherlands (Wageningen) and the second step under the climatic conditions of south-east Spain (Cartagena). This protocol took two years and fewer plants were lost.

Regeneration protocol I

1. Sowing during the last week of August to first week of September of 300 seeds in sowing trays with normal potting soil in a heated glasshouse (day/night temperatures: +15/12°C).
2. After the first seedlings have appeared in the heated glasshouse, the sowing trays should be placed for 3-4 days at 4°C to enhance germination; subsequently, the sowing trays are transferred back to the heated glasshouse for 2-3 weeks and the entire procedure of placing the sowing trays at 4°C should be repeated if not enough seeds have germinated. This procedure may be repeated 3-4 times. In the meantime, emerging seedlings are planted in 24 cm diameter pots (3 plants per pot) and placed in a heated greenhouse; care should be taken that plants do not receive too much water.
3. During next May the watering is gradually decreased to a minimum and in the summer months no watering takes place. Watering will start again in September.
4. Every year in August the bulbs are uprooted and bulblets are removed from the main bulb and bulbs replanted again at the end of August or beginning of September.

5. When the main bulb has a minimum diameter of 4-5 cm flowering is possible. It usually takes about 2.5 years after sowing until the bulb has reached this size.
6. When the plants start flowering in April/May, flowering plants are isolated in cages together with plants from the same population. Regeneration takes place with a minimum of 12 plants per population, using blow flies for pollination. Bulbs that did not produce a flowering stalk can be maintained in the same pot in the greenhouse for the next season.
7. In August, when the seeds have a black colour, the seed stalks are placed upside down in linen bags for drying during 2 weeks, subsequently the seeds are threshed and placed in a paper bag in a drying room (15°C and 15% RH).

Regeneration protocol II

1. Sowing during the last week of August to first week of September of 300 seeds in sowing trays with normal potting soil in a heated glasshouse (day/night temperatures: +15/12°C).
2. After the first seedlings have appeared in the heated glasshouse, the sowing trays should be placed for 3-4 days at 4°C to enhance germination; subsequently, the sowing trays are transferred back to the heated glasshouse for 2-3 weeks and the entire procedure of placing the sowing trays at 4°C should be repeated if not enough seeds have germinated. This procedure may be repeated 3-4 times. In the meantime, emerging seedlings are planted in 24 cm diameter pots (3 plants per pot) and placed in a heated greenhouse; care should be taken that plants do not receive too much water.
3. During May the watering is gradually decreased to a minimum and in the summer months no watering takes place.
4. The bulbs (diameter ca. 1 cm) are harvested in July-August and sent to Cartagena (Spain), where they are planted in the field by the end of August or beginning of September. Individual plants are planted 30 cm apart from each other (within and between rows) to allow sufficient bulb development.
5. When the main bulb has a minimum diameter of 4-5 cm flowering is possible. In this environment/these conditions the bulb usually reaches this size 1.5 years after sowing.
6. When the plants start flowering in April/May, flowering plants are isolated by placing cages around plants from the same population. Regeneration takes place with a minimum of 12 plants per population. Blow flies are used for pollination.
7. In August, when the seeds have a black colour, the seed stalks are placed upside down in linen bags for drying during one week, subsequently the seeds are threshed and sent to the Netherlands for further processing.

Figure 2. *Allium ampeloprasum* L. (TK072; Mylopotamos, Kythera) in its natural habitat.

Specific collecting mission arrangements

An agreement on the basis of which the collecting mission could take place, the so-called Mutually Agreed Terms (MAT), was negotiated between the Greek Ministry of Rural Development and Food, the University of Patras and CGN for the collecting mission. An authorization (prior informed consent; PIC) to collect the three leek CWRs in Greece was issued by the Greek competent national authorities on Access and Benefit Sharing (CNA-ABS) and countersigned by CGN, as a basis for collecting and subsequent distribution of the material. In this arrangement it was agreed that the material collected would be regenerated by CGN and a fair share of each successfully regenerated population will be sent to the Greek genebank. Another agreement was concluded between CGN and a number of breeding companies. In this agreement it was stipulated that breeding companies would co-finance the expedition and help to regenerate the material sampled. Furthermore, an embargo period of five years after the successful regeneration of an accession was negotiated before the material would become available in public databases.

Results

Ecogeographic description of species of the *Allium ampeloprasum* complex

Allium ampeloprasum

Allium ampeloprasum was mainly found adjacent to cultivated fields or abandoned agricultural fields (Figure 2). Contrary to its name ('ampelos' in Greek means 'vineyard' and 'prason' means 'leek') the species was only observed growing between grapevines on a few occasions. The populations sampled often consisted of 1-15 plants and were found in altitudes ranging from 5 to 567 meters above sea level (masl).

Two types of *A. ampeloprasum* plants were observed (von Bothmer, 1974): a sexual type (type

I) with a small inflorescence (ca. 10 cm in diameter), small bulblets (bulbs in between foliage leaves which embrace the large renewal bulb; < 7 mm length) and with seeds; and an asexual type (type II) with a large inflorescence (ca. 20 cm in diameter), larger and more flattened bulblets, and with shrivelled, non- viable seeds (Figure 3).

Allium bourgeaui

The populations sampled consisted of 1-50 plants collected at altitudes ranging from 2 to 522 masl. According to Mathew (1996) *A. bourgeaui* comprises three subspecies, subsp. *bourgeaui* Rech.f., subsp. *cycladicum* Bothmer and subsp. *creticum* Bothmer. The three subspecies are distinguished from each other on the basis of the colour and shape of papillae of the perianth. As collecting took place when the flowering period was over and seeds had set, it was not possible to recognize the three subspecies on the basis of these traits. Identification of the three subspecies is nevertheless possible as the three subspecies have a distinct and scarcely overlapping distribution pattern (Mathew, 1996). More precisely, ssp. *creticum* can only be found in Crete (N=5 populations sampled), ssp. *bourgeaui* only on the East Aegean islands of Rhodos, Karpathos and Kasos (N=6; Figure 4) and ssp. *cycladicum* on Cyclades, Ikaria and the eastern part of the Greek mainland (N=9). The habitats in which these subspecies were found ranged from rocky gorges and steep hillsides to more accessible locations like field borders and hillsides along roads.

Figure 3. The two *Allium ampeloprasum* types differ in their mode of reproduction. Inflorescences and bulblets of both types are shown. On the left the asexual type (type II) and on the right the sexual type (type I).

Figure 4. *Allium bourgeaui* ssp.*bourgeaui* Rech.f. (TK012; Achata beach, Karpathos) in its natural habitat.

Allium commutatum

The population size of this species varied considerably between locations: from 1-15 plants (Diakofti, Kythera) to over 100,000 plants (Methoni, Monemvasia, Kythera Chora castles; Figure 5), but it mostly occured in relatively large (>50 plants) populations. When it occured in large populations, it formed mats, where almost no other plant species could be found. These large populations were also found on islets near the coast. During the expedition the species was collected on islets in front of the coast (Andros), but also on coastal shores (Xerokambos, Crete) and close to old castles (e.g. Monemvasia, Peloponnese).

One population was found on a steep rocky inland hill on the western Peloponnese (Lake/Lagoon of Kaiafa; ca. 1 km from the actual shoreline). However, this can be considered a relict population as in the past, due to the fluctuation of the sea level, the present location of the population was actually a coastal area.

Only a number of islets in front of the Southwestern coast of Andros were visited during this mission due to the relative inaccessibility of these locations. Inhabitants of the islands have knowledge of the presence of *Allium* species on these islets as these islets are often called 'Prasoudha' or 'Prasonisi' (islet of prason).

Mixed *Allium* species locations

On three occasions we observed that next to *A. commutatum* or *A. bourgeaui* also *A. ampeloprasum* was present (Table 1). The *A. commutatum* / *A. ampeloprasum* combination was found on a sandy beach plain (Kythera). The *A. bourgeaui* / *A. ampeloprasum* combination was found on a clearly disturbed rocky slope along a road (Karpathos) and on a (probably disturbed) rocky slope along a road (Peloponnese). The mixed populations were located at altitudes between 2 and 177 masl. Plants with intermediate morphological characters suggesting hybridization between the two species have not been observed in the localities concerned.

Figure 5. *Allium commutatum* Guss (TK076; Methoni Castle, Peloponnese) in its natural habitat.

Collecting: number of populations collected

In total, around 4,000 km were travelled in Greece by car, boat and plane and 104 populations were collected: 62 single species populations of *A. ampeloprasum*, 20 of *A. bourgeaui*, 19 of *A. commutatum* and three mixed populations, which consisted of two *Allium* species (Table 1).

Collecting: number of seeds, bulblets and bulbs sampled

From 65 single species populations enough black mature seed (criterium: > 0.5 g = ca. 200 seeds) were harvested to be sure that ca. 80 plants per population could be obtained for regeneration. For 36 single species populations this was not the case as the number of mature seeds obtained was below the threshold. For the three mixed populations enough seed could be collected for each population (Figure 6).

A. ampeloprasum clearly differed from the other two species as 42% (26 out of 62) of the populations did not yield enough seeds, compared to 25% and 26% for *A. bourgeaui* and *A. commutatum* respectively. Possible reasons for this could be a) a lower number of plants made up the population; b) immature seeds were collected; and c) presence of plants with different reproduction systems occurring in some populations of *A. ampeloprasum*. In nine *A. ampeloprasum* populations plants of the two reproduction types (see above) were observed and care was therefore taken to select plants with smaller inflorescences (i.e. sexual plants) and with black matured seed present within green fruit capsules. In three *A. ampeloprasum* and three *A. bourgeaui* populations, only a few plants could be harvested, thus seed sampling could only be taken from fewer than five plants, thus yielding only low seed amounts. From two A. *bourgeaui* populations vegetative material (bulblets/bulbs) had been collected, which allowed for the multiplication of plants of these populations. When collecting immature seeds, which occurred in four populations of *A. ampeloprasum*, two of *A. bourgeaui* and five of *A. commutatum*, translucent white seeds present within the green fruit capsules were collected, assuming that these seeds could still mature on the stalk. Therefore, inflorescences with ca. 20 cm flower stalk were collected. However, this precaution proved to be not enough to obtain mature black seeds for these 11 populations. However, as a precautionary measure vegetative material was also sampled for these 11 populations and thus these samples resulted in viable accessions.

Regeneration

Initially, a protocol (Material and Methods: Regeneration protocol I) was developed for regeneration in the Netherlands. However, the next generation could only be harvested three years after sowing the originally collected seeds. Therefore, a second protocol (Material and

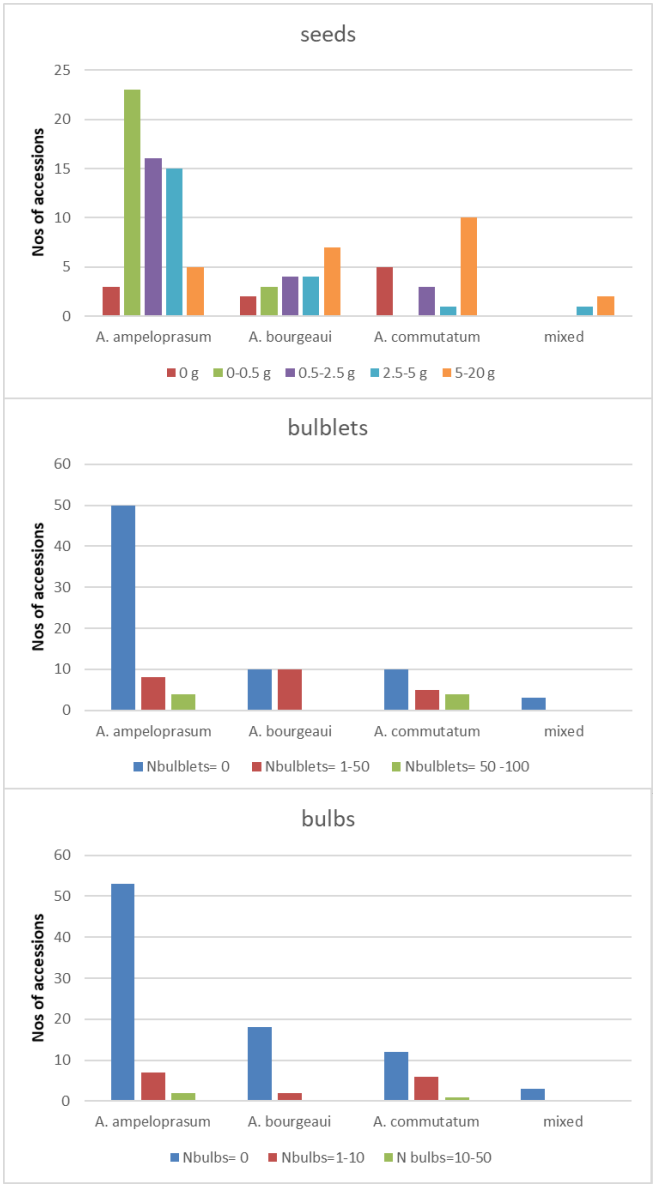

Figure 6. The quantity of seeds and the number of bulblets and bulbs sampled in the various leek CWR populations. When a category on the X-axis is presented as an interval the lower value is excluded from the interval and the higher value is included.

Table 1. The number of collected leek CWR populations per location visited in Greece. Mixed species populations: populations in which two species are present.

Location	*A. ampeloprasum*	*A. bourgeaui*	*A. commutatum*	Mixed Species Populations		Total
				A. ampeloprasum/ A. commutatum	*A. ampeloprasum/ A. bourgeaui*	
Karpathos	9	7			1	17
Crete	25	5	3			33
Andros	1	4	6			11
Peleponnese	4	4	6		1	15
Kythera	10		4	1		15
Lesbos	13					13
Total	62	20	19	1	2	104

Methods: Regeneration protocol II) was developed in which the first stages of the regeneration took place the Netherlands and the later stages in Cartagena (Spain). This procedure shortened the production of the next generation of seeds by one year.

As the collecting took place in 2009, the regeneration of the material started in 2010. In 2013 the first eight populations were successfully regenerated and, subsequently these populations were included as accessions in the Dutch genebank (CGN) and placed in the public domain (https://www.wur.nl/en/Research-Results/Statutory-research-tasks/Centre-for-Genetic-Resources-the-Netherlands-1/Genebank/Special-collections.htm). This was only in 2019, due to an embargo period of five years after the successful regeneration of a population, which had been agreed upon by the breeding companies co-financing the mission and CGN on the material and associated information (Table 2). In the coming years all accessions which were successfully regenerated will be placed in the public domain.

The percentage of populations that have been successfully regenerated until present is 39% (42/107). The mean germination percentages of the populations regenerated until present are 78%, 65% and 89% for *A. ampeloprasum, A. bourgeaui* and *A. commutatum*, respectively. Eight populations (seven *A. ampeloprasum* and one *A. bourgeaui*) have been donated to the Czech genebank in Olomouc as no seeds could be produced from these populations, but enough bulbs/bulblets were present for vegetatively maintaining these populations. During regeneration, up to now, ten populations of *A. ampeloprasum* (15% losses; 10/65), three of *A. bourgeaui* (14% losses; 3/22) and two of *A. commutatum* (10% losses; 2/20) were lost due to no germination or decay of bulbs.

Discussion

Ecogeography of the species from the *Allium ampeloprasum* complex

During the collecting mission more *A. ampeloprasum* populations were collected than *A. bourgeaui* and *A. commutatum* populations. This is probably due to the habitat preferences of the three species: *A. ampeloprasum* has a broad ecological amplitude, whereas the other two species have a narrow ecological amplitude. In terms of generalist and specialist species (Fried *et al*, 2010), *A. ampeloprasum* can be considered as a generalist and *A. bourgeaui* and *A. commutatum* as specialists. Greece can be considered as an important biodiversity centre for *Allium* subgenus *Allium* which is represented in Greece by more than 100 species (Tzanoudakis, 2001; Dimopoulos *et al*, 2013). *In situ* management of these species is currently not taking place in Greece, but in case of *A. bourgeaui* and *A. commutatum* this should be considered due to their specific environmental requirements.

The amount of seeds that could be collected from natural populations of the three species varied to a large extent among these species. It was observed that in many populations of *A. ampeloprasum* (42%) only a few seeds could be harvested. The reasons for this are unclear, but it is possible that this is due the presence of a polyploid series *A. ampeloprasum*, which is much larger compared to the polyploid series in *A. commutatum* and *A. bourgeaui*, and which might affect seed production (von Bothmer, 1970, 1974). In this context, the presence of plants with different modes of reproduction (sexual and asexual) in *A. ampeloprasum* in cultivated and abandoned fields, especially observed in Kythera, was an interesting phenomenon, probably reflecting different selective forces acting on the species. These forces are probably related to human activities, changes in agricultural methods and land use in general. The sexual type was observed on its own in cultivated fields only, probably due to its generative capacity which allowed the species to co-evolve with the cultivated crops (potatoes, tomatoes, cucumbers, onions). Ploughing in these cultivated fields occurs in early spring and irrigation takes place by running water from springs or by underground well-water. The asexual type was mainly collected from places more dry and remote from villages on larger pieces of land formerly cultivated (ploughed during winter, not irrigated) with crops like cereals (*Triticum, Avena, Hordeum*) or legumes (*Vicia, Lathyrus*). After abandoning these fields, species like *Sarcopoterium, Cistus* and others perennial species dominate.

Preliminary flow cytometry of a few sexual and asexual individuals showed that the asexual type had higher 2C peaks compared to the sexual type, indicating a higher ploidy level than the tetraploid level (Kik, unpubl. results). Most probably, the sexual type is a (segmental) autotetraploid with 2n=4x=32 and not a diploid (2n=2x=16; Guenaoui *et al* (2013), whereas the asexual type has an increased ploidy level (2n>32), which might impair sexual reproduction and leads to strong(er) vegetative growth and thus stronger competitive strength.

No cultivation of leek was observed during the collecting mission. However, instead of cultivated leek, wild leek species, i.e. only *Allium ampeloprasum* and *A. commutatum*, as *A. bourgeaui* is difficult to collect, is used as a condiment in the traditional Greek kitchen (salads, soups, pies etc; Stavridakis (2006). The absence of leek cultivation implies that it is highly likely that no hybridisation and subsequent introgression of genes from the cultivated species into the three wild relatives of leek has taken place. Consequently, no replacement of wild genes has taken place in the three species, which increases their value as a gene reservoir for research and breeding.

Collecting and regenerating the species from the *Allium ampeloprasum* complex

Collecting

The theoretical framework concerning the sampling of populations is well developed for collecting a natural population, using probability-based sampling approaches (Volk *et al*, 2007). However, in collecting expeditions one deals with more than one population, raising the question of how to balance the number of plants to be collected in populations versus the number of populations to be sampled in a given area to adequately acquire a significant percentage of the genetic variation present in these populations. Until present only a few studies have addressed sampling in more than one population and simulation approaches were used to study the effects of various factors on the collecting of genetic variation present in these populations. In this context Hoban and Scharlbaum (2014) argued that it was not realistic to focus on obtaining 95% of the variation present in the populations sampled. A more practical approach would be to collect around 75% of the variation present in these populations. This would require sampling seeds from 20-30 plants per population in case of a diploid species and even less in case of a tetraploid species (Bray, 1983). Therefore, in the case of the three species of the *Allium ampeloprasum* complex, the aim was to collect seeds from 20-25 plants per population.

Another issue which needs to be taken into account is how to collect in populations where a strong vegetative propagation can be assumed, for example in and around Monemvasia, Methoni and Kythera Chora castles, where large populations (> 100,000 plants) of *A. commutatum* are present. To sample these populations appropriately, knowledge about the clonal structure of the populations is essential. However, in case of *A. commutatum* no literature on this issue has been found. The way this problem has been approached in the current collecting expedition was to collect seeds from plants that were meters apart from each other, hoping that they were genetically different from each other.

Regeneration

Bulk sampling took place within populations instead of sampling seeds from individual plants and keeping their progenies apart. This was done for logistic reasons, as it proved unfeasible to isolate between 10 and 15 bi-parental crosses per population rather than using one isolation cage for all the plants of a population for each of the 104 collected populations. Bulk sampling and subsequent bulk regeneration per population involves a risk as genetic variation in the subsequent generations could be lost due to random sampling effects (Gale and

Table 2. Overview of the regeneration process of leek CWR. 1) year when material will be publicly available; 2) number of populations; 3) 107 = 104 (populations) + 3 (mixed species populations that were subdivided during the regeneration process, based on their species identity and thus resulting in three extra populations for regeneration).

	year[1]	*A. ampeloprasum*	*A. bourgeaui*	*A. commutatum*	Total
succesfully regenerated, already available	2019	2[2]	3	3	8
	2020	1	4	6	11
succesfully regenerated, available in	2021	2	3	1	6
	2022	0	0	5	5
	2023	2	2	1	5
	2024	3	1	0	4
	2025	1	1	1	3
to be regenerated		37	4	1	42
lost		10	3	2	15
donated to Czech genebank		7	1	0	8
Total		65	22	20	107[3]

Lawrence, 1984; Cross and Wallace, 1994). However, the decline in genetic variation in tetraploid species is less compared to diploid species (Bray, 1983). Furthermore, the loss of genetic variation after a number of generations also depends upon the frequency of genes in the collected samples and the ratio of effective population size versus total population size (Ne/N) in case of cross-fertilizing species, such as the three species of the *Allium ampeloprasum* complex. However, very important in maintaining the genetic variation in a sample is to keep the number of regenerations as low as possible and to use as many plants as possible per population. Therefore, good storage conditions are essential in this context (van Treuren *et al*, 2013).

Regenerating the CWR of leek proved to be more difficult than expected, especially as it was thought that the regeneration protocol used for cultivated leek could be applied for its CWR species. However, due to differences in flowering phenologies this was not the case, i.e. the CWRs of leek cannot be regenerated in the same way as leek. Leek is sown in March and grows until autumn in the Netherlands, whereas the CWRs grow during the winter. Leek does not grow in winter, but resumes growth in spring and flowers/sets seed in June-July. The CWRs of leek, on the other hand, are sown in August/September and grow from autumn until spring. At the end of spring the foliage dies back when the bulb has been formed. After the summer (August/September) the bulb sprouts and the growth resumes. When the bulb is large enough, flowering and subsequent seed setting takes place next May/June.

In Regeneration protocol I, the regeneration takes place in the Netherlands and takes three years, whereas in Regeneration protocol II, the regeneration takes place in the Netherlands and in Spain and takes two years. The difference between both protocols is attributable to the better growing conditions in Spain compared to the Netherlands, which favour bulb growth from the end of August to May. A relatively large number of seeds had to be sown as germination percentages varied between 55-83%, depending upon the species. *A. ampeloprasum* was the species where most losses occurred, so for this species it may be suggested to sow twice as many seeds (N= 600).

Access and Benefit-Sharing issues

Access and benefit sharing (ABS) is an important aspect in the two international agreements that regulate the exchange of plant genetic resources (PGR), namely the Convention on Biodiversity (CBD) and the International Treaty on Plant Genetic Resources for Food and Agriculture (IT-PGRFA). In this context, national focal points (NFPs) and competent national authorities for ABS (CNA-ABS) have been appointed in the countries that ratified one or both agreements; Greece and the Netherlands ratified both the two agreements. PIC and MAT were concluded before the collecting mission took place, and in the MAT an article was present referring to the Bonn guidelines.

In this context, a problem was encountered which appeared after the collecting mission when the Dutch collectors and the Greek CNA-ABS were not able to conclude a specific arrangement in which it was stipulated that the material collected could also be used for commercial purposes. The consequence was that breeding companies, who were involved in this project, preferred to focus their contributions primarily on material that it could certainly be used in their own breeding programmes. This meant that only a small number of Greek populations could be regenerated annually as the Dutch national genebank CGN did not have sufficient capacity to carry out a large number of regenerations per year, which resulted in a low speed of the regeneration for these three species. Another consequence of not being able to conclude a specific arrangement was that the utilization of the material collected is currently limited to research purposes. Therefore, collecting missions should ensure the appropriateness of the terms and conditions of ABS in all its aspects well before initiating expeditions.

Acknowledgements

The Dutch Ministry of Agriculture, Nature and Food Quality is gratefully acknowledged for funding this research which was carried out in the framework of the Programme Genetic Resources (WOT-03). In addition, we would like to thank Bejo Zaden B.V., Enza Zaden Research and Development B.V., Nunhems Netherlands B.V., Rijk Zwaan Zaadteelt en Zaadhandel B.V., Hazera Seeds B.V. and Syngenta Seeds B.V. for supporting the collecting and regeneration of the material collected. Furthermore, the manuscript was clearly improved by the contributions of two anonymous reviewers.

Author contributions

All authors have contributed substantially to the study design, execution, data analysis and interpretation, drafting and revision of the submitted manuscript.

Conflict of interest statement

The authors declare that they have no conflict of interest, that the work submitted is their own, that copyright has not been breached in seeking its publication, and that the work submitted has not previously been published and is not being considered for publication elsewhere.

References

Alercia, A., Diulgheroff, S., and Mackay, M. (2015). FAO/Bioversity Multi-Crop Passport Descriptors V.2.1 [MCPD V.2.1]. url: https://hdl.handle.net/10568/69166.

Bray, R. A. (1983). Strategies in gene maintenance. In *Genetic Resources of Forage Plants,* ed. McIvor, J. G. and Bray, R. A., (Melbourne, Australia: CSIRO), 157-168.

Cross, R. J. and Wallace, A. R. (1994). Loss of genetic diversity from heterogeneous self-pollinating

genebank accessions. *Theoretical and Applied Genetics* 88(6-7), 885–890. doi: https://doi.org/10.1007/bf01254001

de Clercq, H. and Bockstaele, E. V. (2002). Leek: advance in agronomy and breeding. In *Allium Crop Science: Recent Advances,* ed. Rabinowitch, H. D. and Currah, L., (Wallingford, UK: CABI International), 459-491.

de Wilde-Duyfjes, B. E. E. (1976). A revision of the genus Allium L. (Liliaceae) in Africa. Ph.D. thesis, Landbouwhogeschool Wageningen, Wageningen, the Netherlands.

Dimopoulos, P., Raus, T., Bergmeier, E., Constantinidis, T., Iatrou, G., Kokkini, S., Strid, A., and Tzanoudakis, D. (2013). Vascular plants of Greece; an annotated checklist. Volume 31 of *Englera*. (Berlin, Athens: Botanic Garden and Botanical Museum Berlin-Dahlem and Hellenic Botanical Society), 1-371.

Fried, G., Petit, S., and Reboud, X. (2010). A specialist-generalist classification of the arable flora and its response to changes in agricultural practices. *BMC Ecology* 10(1), 20. doi: https://doi.org/10.1186/1472-6785-10-20

Gale, J. S. and Lawrence, M. J. (1984). The decay of variability. In *Crop Genetic Resources: Conservation and Evaluation,* ed. Holden, J. H. W. and Williams, J. T., (UK: Unwin Hyman), 77-101.

Guenaoui, C., Mang, S., Figliuolo, G., and Neffati, M. (2013). Diversity in *Allium ampeloprasum*: from small and wild to large and cultivated. *Genetic Resources and Crop Evolution* 60(1), 97–114. doi: https://doi.org/10.1007/s10722-012-9819-5

Hirschegger, P., Jakše, J., Trontelj, P., and Bohanec, B. (2010). Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section Allium (Allium: Alliaceae). *Molecular Phylogenetics and Evolution* 54(2), 488–497. doi: https://doi.org/10.1016/j.ympev.2009.08.030

Hoban, S. and Scharlbaum, S. (2014). Optimal sampling of seeds from plant populations for ex situ conservation of genetic biodiversity, considering realistic population structure. *Biol Cons* 177, 90–99. doi: https://doi.org/10.1016/j.biocon.2014.06.014

Keller, E. R. J. and Kik, C. (2018). *Allium* Genetic Resources. In *The Allium Genomes,* ed. Shigyo, M., Khar, A., and Abdelrahman, M. (Springer, Cham), 23-52. Compendium of Plant Genomes.

Kik, C. (2002a). Exploitation of wild relatives for the breeding of cultivated *Allium* species. In *Allium crop science: recent advances,* ed. Rabinowitch, H. D. and Currah, L., (Wallingford, UK: CABI International), 81-100.

Kik, C. (2002b). The *Allium* collections at Plant Research International, with special reference to the vegetatively maintained leek (von Bothmer) collection. In Maggioni, L., Keller, J., and Astley, D., in European collections of vegetatively propagated *Allium.* Report of a workshop, IPGRI, Rome, 50-52.

Kik, C. (2009). Report of a leek (*Allium porrum*) wild relatives collecting expedition in Greece. volume 2009/15 of *CGN report*, CGN, Wageningen.

Kik, C., Samoylov, A. M., Verbeek, W. H. J., and van Raamsdonk, L. W. D. (1997). Mitochondrial DNA variation and crossability of leek (*Allium porrum*) and its wild relatives from the *Allium ampeloprasum* complex. *Theoretical and Applied Genetics* 94(3-4), 465–471. doi: https://doi.org/10.1007/s001220050438

Mathew, B. (1996). A review of *Allium* sect. *Allium* (Richmond, Surrey, UK: Royal Botanic Gardens, Kew), 1-176.

Runemark, H. (1969). Reproductive drift, a neglected principle in regenerative biology. *Bot Notiser* 122, 90–128.

Runemark, H. (1970). The Role of small populations for the differentiation in plants. *Taxon* 19(2), 196–201. doi: https://doi.org/10.2307/1217954

Stavridakis, K. (2006). Wild Edible Plants of Crete (Greece: Privately Published), 1-413.

Tzanoudakis, D. (2001). The genus Allium in Greece: a genus representative of the floristic richness and diversity of the area. In Proceedings 2nd Balkan Bot Congress.

van Treuren, R., de Groot, E. C., and van Hintum, T. J. L. (2013). Preservation of seed viability during 25 years of storage under standard genebank conditions. *Genetic Resources and Crop Evolution* 60(4), 1407–1421. doi: https://doi.org/10.1007/s10722-012-9929-0

Volk, G. M., Lockwood, D. R., and Richards, C. M. (2007). Wild plant sampling strategies: the roles of ecology and evolution. In *Plant Breeding Reviews,* ed. Janick, J. volume 29, 285-313.

von Bothmer, R. (1970). Cytological studies in *Allium* I. Chromosome numbers and morphology in *Allium* sect *Allium* from Greece. *Bot Notiser* 123, 518–550.

von Bothmer, R. (1974). Studies in the Aegian Flora XXI. Biosystematic studies in the *Allium ampeloprasum* complex. *Opera Botanica* 34, 1–104.

Genetic Resources

REVIEW

Genetic Resources (2021), 2 (3), 11–24
DOI: 10.46265/genresj.HXSV9592
https://www.genresj.org
ISSN: 2708-3764

Sheep genetic resources in Bulgaria with focus on breeds with coloured wool

Doytcho Dimov * **and Atanas Vuchkov**

Department of Animal Sciences, Agricultural University of Plovdiv, Plovdiv, Bulgaria

Abstract: This review describes sheep genetic resources in Bulgaria with coloured wool and evaluates the country's potential for coloured wool production. In 2018, Bulgaria counted 125,422 animals belonging to 18 native sheep breeds, six of which are in danger of extinction. Native sheep breeds in Bulgaria can be divided into three groups depending on the fleece colour of the animals: i) sheep breeds with fully pigmented fleece in all animals; ii) sheep breeds with animals with fully pigmented fleece or fully white fleece; iii) sheep breeds with spotted coloured fleece. Colouration in populations of the native Bulgarian sheep breeds can be explained by the phenotypic expression of several alleles of coat colour genes: *Extension*^{D}, *Agouti*^{a} and *Pigmented head*^{T}, while white colour is due to the presence of the dominant *Agouti*^{wt} allele.

Based on the relative share of the animals with coloured wool, the country's potential for annual coloured wool production from native breeds was estimated at a minimum of 133,791 kg of unwashed wool. Recent tendencies in lifestyle changes of the Bulgarian people are gradually leading to increased interest in naturally coloured wool products, which may benefit the conservation of endangered native breeds by promoting relevant breeding programmes.

Keywords: native sheep breeds, coloured wool, genetic loci

Citation: Dimov, D., Vuchkov, A. (2021). Sheep genetic resources in Bulgaria with focus on breeds with coloured wool. *Genetic Resources* 2 (3), 11–24. doi: 10.46265/genresj.HXSV9592.

Introduction

Historically, the development of the wool textile industry in Bulgaria has gone through many stages. In the first half of the 20th century, through a number of laws, all governments in Bulgaria encouraged local textile products (Savov, 1964). For example, The wool-textile industry had received a number of tax reliefs and aids in the import of machinery and raw materials (including imports of high-quality wool and yarn) for the production of textile products, which has led to an increase in the production of fine worsted white wool fabrics.

During the wars, and especially during the First World War, the wool-textile industry continued to work, with much of the raw material used being local coarse wool. When purchasing local coarse wool, the textile industry in Bulgaria preferred to buy coloured wool for the production of military fabric and blankets because there was no need to dye the wool. The main supply of this wool was from the areas where Karnobat, Copper-red Shumen, Dubenska and other native sheep breeds with coloured wool have been raised (Savov, 1964).

Following state policy since 1950, these local breeds have been crossed with rams of merino breeds to create new breeds of sheep for the production of fine merino white wool needed for the production of thin fabrics for finer woolen products. By 1983, Bulgarian production reached 33,000 tonnes of wool, a large part of which was merino white wool. After 1990, as a result of the economic and political crisis of the socialist system in Eastern European countries, the annual production of 28,000 tonnes of wool in 1990 collapsed to 6,500 tonnes by 2005 and recently to about 2,946 tonnes in 2017 (FAO, 2017). In general, this decline in production was observed in many wool-producing countries around the world including Bulgaria.

By the middle of the 20th century, the manufacturing of handmade textile products had declined and nowa-

*Corresponding author: Doytcho Dimov (doytcho.dimov@gmail.com)

Received: 22.03.2020 | Accepted: 26.05.2021 | Published online: 15.06.2021

days, they are produced only by a small number of artisans, as well as by some people continuing folk traditions.

In recent years, efforts to protect the environment and tendencies towards a more biological and ecological lifestyle have increased consumers' interest in woolen handicrafts and textile products from naturally coloured animal fibres. There is a growing concern to add value to sheep production in view of possible decreases in agricultural premiums and growing demand for organically produced agricultural products (Chaupin, 2014). Renewed interest in local, natural resources due to concerns over the carbon footprint of long-distance transportation of goods fuels a growing movement for 'sustainable fashion' that takes into consideration the ecological impact of garment production (Chaupin, 2014). In this context, the naturally coloured wool obtained from some sheep breeds acquires a new meaning.

The colour of the wool is not included as a selection trait in any breeding programme in Bulgaria and therefore there has so far been no scientific research on this feature in the country. Until now, there were no studies about genetic polymorphisms in native sheep breeds in Bulgaria in which the coat colour was subject to genetic assessments.

The aim of this review is to describe and characterize sheep genetic resources with coloured wool in Bulgaria and to evaluate the country's potential for coloured wool production, as well as to give some explanations about genetic background of sheep colour variations.

National gene pool of native sheep breeds

According to official statistics, in Bulgaria there currently are 18 native, eight locally adapted and seven introduced sheep breeds (Table 1) and the number of sheep in the country in the last eight years was about 1.3 million. According to total population size, Bulgarian native sheep breeds can be classified in several categories of risk status using FAO criteria (FAO, 2013). Six of the native sheep breeds are 'endangered' and five breeds are 'vulnerable'. The remaining seven breeds have the risk status 'not at risk' (Table 1).

In the country, there are nine breeding organisations for native sheep breeds which perform breeding programmes approved by the Ministry of Agriculture, Food and Forests (MAFF), all of which put emphasis on preserving native sheep breeds as genetic resources and overcoming the risk of extinction by increasing their population size. The breeding programmes of almost all native sheep breeds are currently not focused on colour of wool or its yield and qualities in the context of the emerging interest in the textile industry and for other coloured wool products.

Depending on the colour of the fleece, the native sheep breeds in Bulgaria can be divided into three groups: 1) breeds with fully pigmented fleece; 2) breeds with both fully pigmented and fully white fleece and 3) breeds with spotted coloured fleece.

Table 1. Native sheep breeds in Bulgaria and their population size (including ewes and rams) and risk status according to FAO criteria in 2018. (Data source: Executive Agency on Selection and Reproduction in Animal Breeding (EASRAB; https://www.iasrj.eu/).

Breed №	Breed name	Population size	Risk status
1	Copper-red Shumen	16,382	Not at risk
2	Karnobatska	1,301	Endangered
3	Karakachanska	11,621	Not at risk
4	Dubenska	12,660	Not at risk
5	White Maritza	886	Endangered
6	Kotlenska	5,998	Vulnerable
7	Black-headed Pleven	18,452	Not at risk
8	Native Starozagorska	834	Endangered
9	Tetevenska	4,642	Vulnerable
10	Koprivshtenska	3,375	Endangered
11	Zapadnostaroplaninska	4,407	Vulnerable
12	Srednorodopska	8,716	Not at risk
13	Srednostaroplaninska	10,091	Not at risk
14	Patch-faced Maritza	8,037	Not at risk
15	Elin-Pelinska	6,599	Vulnerable
16	Replianska	5,982	Vulnerable
17	Sakarska	2,238	Endangered
18	Breznishka	3,201	Endangered
Total number of native breed sheep in 2018		125,422	

The genetic basis of coloured wool

The coat colour of sheep is often not only a visual trait in the morphological characterization of sheep breeds, but sometimes also an important trait for selection. Classic genetics, which relied on 'by eye' classification of coat colour type, was helpful in breeding schemes for many years (Cieslak *et al*, 2011). Genetic control of coat colour is performed by 11 identified loci: *Agouti* (*A*), *Albino* (*C*), *Australian Piebald* (*AsP*), *Brown* (*B*), *Extension* (*E*), *Pigmented head* (*Ph*), *Roan* (*Rn*), *Spotting* (*S*), *Sur Bukhara* (*SuB*), *Sur Surkhandarya* (*SuS*) and *Ticking* (*Ti*) (Sponenberg *et al*, 1996; Sponenberg, 1997).

The colour diversity results from the presence and biochemical activity of melanocytes, cells that are specialized in producing melanins (Koseniuk *et al*, 2018) . In sheep, as in other mammals, there are two types of melanins: eumelanin and pheomelanin. Eumelanin causes either a black or brown shade of wool. Pheomelanin results in a red, tan or fawn colour of wool (Lundie, 2011). It is widely recognized that the production of eumelanin and pheomelanin are genetically controlled by the *Extension* and *Agouti* loci (Searle, 1968). The *Agouti* locus encodes for the agouti signalling protein (ASIP; Bultman *et al*, 1992). This small paracrine signalling molecule interacts with the product of the *Extension* locus that encode for the melanocortin 1 receptor (MC1R). Many authors have identified two alleles at the *Extension* locus: the dominant black allele (E^D) is responsible for black

colour in a few coloured breeds, and the wild type allele *(E^+) is* widely distributed in most breeds in which the *Agouti* locus is most likely responsible for the majority of colour variation in wool (Searle, 1968; Sponenberg, 1997). There are assumptions that there is a third allele in the *Extension* locus but, unlike in other mammals, this is not yet well documented in sheep and its potential role in the phenotypic manifestation of wool colour is not well understood. Fontanesi *et al* (2010) studied the polymorphisms of the *MC1R* gene in nine Italian sheep breeds and identified one non-synonymous mutation (C199T) causing an amino acid substitution (R67C) in a highly conserved position in the first intracellular loop of the MC1R protein. The authors proposed that this mutation, identified only in the Valle del Belice breed, may represent the recessive *eallele* in the ovine *Extension* series, but this is not completely recognized in sheep. A different study to assess the genetic polymorphisms of the *MC1R* gene in eleven local Greek sheep breeds (Stamatis *et al*, 2017) did not find the putative *e* allele proposed by Fontanesi *et al* (2010). Yang *et al* (2013) investigated the variability in the MC1R protein and its possible association with coat colour in ten Chinese sheep breeds and also did not detect the recessive allele *e*. There is an epistatic interaction between the *Extension* and *Agouti* loci (Searle, 1968). The presence of the (E^D) allele in the homozygous (E^DE^D) or heterozygous ($E^DE^{+)}$ state leads to phenotypic manifestation of the black colour, while the *E^+allele* in the homozygous state (E^+E^+) is a prerequisite for the phenotypic manifestation of the *Agouti locus* alleles. The wild type allele E^+ is the most common *Extension* allele in most European breeds of sheep in which segregation at the *Agouti* locus accounts for the majority of colour variation (Sponenberg, 1997).

The predominant allele at the *Agouti* locus is A^{Wt}, which is dominant and produces white fleece in most wool breeds (Notter and Sponenberg, 2002). The most common alternative allele at this locus is the so-called *non-agouti* allele (A^a). This allele is recessive and in sheep of most breeds this results in a completely black colour at birth. Animals of genotype A^aA^a are generally black although the final expression of colour can be influenced by other genes and depends on modifiers at independent loci (Notter and Sponenberg, 2002). Between the A^{Wt} and A^a alleles, there are other intermediate *Agouti* alleles that cause different patterns of pigmentation. Some of them can cause reversals of pigmentation type, so that areas that are eumelanic in one pattern are pheomelanic in another and vice versa (Sponenberg, 1997).

The *Spotted* locus (*S*) is responsible for white spotting. According to the Committee on Genetic Nomenclature of Sheep and Goats (COGNOSAG) nomenclature there are two alleles in this locus (Broad *et al*, 1999). The wild type allele (S^+) results in a solid coloured animal. The spotted allele (S^s) in eumelanic animals usually produces some white markings on the crown of the head, tip of the tail and bottom of the legs. When the *S^s allele* occurs in a pheomelanic background colour it produces a piebald animal when heterozygous and mainly white animals when homozygous (Lauvergne and Raffier, 1975).

Four alleles are known at the *Pigmented Head* (*Ph*) locus: wild type (Ph^+), Persian (Ph^P), Turkish (Ph^T) and Afgan lethal (Ph^{afl}) (Lundie, 2011). Since the Ph^P and Ph^{afl}alleles cannot be found in Bulgarian sheep breeds and the *Ph^+allele* is practically impossible to distinguish from other alleles that cause the same solid-coloured animals, the most interesting allele for Bulgarian sheep breeds is Ph^T, which is used in the selection of the Patch-faced Maritza sheep breed (see below).

In some wool sheep breeds animals with brown coat colour occasionally appear. Brown colour is controlled by the *Brown* gene, for which there are two alleles (Notter and Sponenberg, 2002; Lundie, 2011). The wild type allele (B^+) is dominant and the most common allele at this *locus*, resulting in black eumelanin pigment. The *Brown* allele (B^b) is a rare recessive allele that when homozygous results in brown (chocolate shade) eumelanin pigment (Lundie, 2011). The *Brown* locus controls tyrosinase-related protein 1 (TRP-1) which has an important role within melanocytes (Jackson, 1994). Brown colour is quite common in the Northern Short-tailed group of breeds of Europe and found more rarely in other European breeds such as the Merino and Romney (Lundie, 2011).

The discussion below of native sheep breeds in Bulgaria includes some hypotheses about genetic loci and alleles affecting coat colour in sheep populations, following the recommendations of COGNOSAG.

Breeds with fully pigmented fleece

Copper-red Shumen

The Copper-red Shumen sheep is an ancient native sheep breed in Bulgaria with a population size in 2018 of 16,382 (Table 1). According to the style of the fleece there are two types: 'kabarliavi', with open fleece of semi-rough, mixed wool, and 'rudavi', with a homogeneous, thin wool (Staikova *et al*, 2015). The wool yield varies from 2.711 to 3.454 kg and the average staple length is 14.13 cm (Staikova, 2005). Adult animals of the breed are almost 100% coloured with red-brown or chocolate-coloured wool (Figure 1). The face, ears and legs are covered with short black hairs. The lambs at birth are solid black and develop the typical copper-red colour only after weaning at three months of age and later after shearing. Some individuals display lighter colour while, in others, it can be more intense and in different shades (Staikova *et al*, 2015). Animals with white colour are very rare in this breed. Based on our observations and interviews with specialists, when mating this breed with white animals, all offspring in F1 were black, which can be attributed to the phenotypic expression of the dominant E^D allele, suggesting that the colour genotype of the typical animals is E^DE^D or E^DE^+. Only 4-7% of animals

a

b

Figure 1. Copper-red Shumen sheep breed. a) ewes b) rams. (Photographs A. Vuchkov (a) and D. Dimov (b))

in the population are born black with white spots on the top of the head and tip of the tail, attributed to the phenotypic expression of the *Spotted* allele S^s. Thus, the genotype which causes the colour can be described as $E^D E^D S^+ S^s$ or $E^D E^+ S^+ S^s$. The change in colour at an early age from black to reddish-brown may be due to the action of other gene modifiers, which are not known at this stage.

Karnobatska

The Karnobat native sheep is an autochthonous sheep breed, known in the past for its delicious meat (Stefanova and Iliev, 2005) and soft pigmented wool (Hlebarov, 1933). The tail is thin, but at the base there is an enlargement of subcutaneous fat which reaches the hock joints. The length of the tail is between 24 and 26 cm. The live weight of the ewes varies from 42 to 47 kg (Iliev, 2012). Wool yield in recent years varied by about 3 to 3.6 kg in ewes, and 5 to 5.6 kg in rams Iliev (2007).

The lambs are born completely black, and subsequently the fleece lightens to different shades of brown (Staikova and Iliev, 2017). The front of the head and legs are covered with short black hairs, the body is covered with coloured wool fibres (Figure 2). Animals with almost white coat colour are rarely encountered, but these are usually removed from the population. When crossing with white sheep of other breeds in F1 mostly black animals are born, which is reason to assume that the colour of the coat is determined by the dominant E^D allele as in the Copper-red Shumen breed. The change in colour of the wool from black to reddish-brown is probably due to the action of gene modifiers, and probably a certain influence is exerted by environmental factors, such as age or sunlight.

Breeds with both fully pigmented and fully white fleece

Karakachanska

The Karakachan sheep is a typical mountain breed, small in height with a compact body, extremely mobile and very lively temperament. The fleece is open and has a hairy conformation (Figure 3).

a

b

Figure 2. Karnobat native sheep breed. a) ewe b) group of ewes. (Photograph: A. Vuchkov)

The wool is coarse, mixed with long fibres and relative low density. The wool yield is 2.7 kg, wool length 16 cm and fibre diameter is around 80 μm. Pigmented sheep are dominant in the population, and white, non-pigmented sheep are relatively rare (Barić, 1952; Hlebarov, 1942; Milić, 1954; Sedefchev and Sedefchev, 2011). Animals are typically one colour, with black head and legs. The lamb's fleece is black, changing with age to shades of grey-black or grey.

According to our observations, the pigmentation in this breed may be due to the dominant allele E^D in the *Extension* locus. The extremely rare white colour in this breed is due to the dominant allele A^{wt} in the *Agouti* locus, which is expressed in the absence of E^D in the *Extension* locus. Expression of the recessive S^s allele in the *Spotted* locus was also observed in about 10% of the animals with white spots on the top of the head and the tip of the tail.

White Maritza

The White Maritza sheep breed is a multi-purpose native breed for milk, meat and wool. This native lowland sheep breed, contrary to notions of native sheep breeds, has a good milk yield (110.57 l) and high body weight of ewes (71.71 kg) and rams (100.12 kg) (Dimov, 2011b). The size of this population is comparatively small and in 2018 was estimated at about 886 sheep including 40 rams and therefore considered endangered. The fleece looks almost closed and has staple conformation with staple length of 11.19 cm (Dimov, 2015). The average fiber diameter is 32.1 μm (Dimov and Djorbineva, 1999) and according to the Bradford system for wool classification, this means 48^{th} quality (USDA, 2018).

The typical fleece colour of Maritza sheep is white (Figure 4a), but sheep with coloured fleece (Figure 4c) can also be found. The wool around the head is not shorn and a so-called 'cowl' is formed over the

Figure 3. Karakchan sheep breed. a) flock of ewes with lambs b) black lamb (genotype E^D) c) black and white lamb d) black lamb with white spots caused by S^s allele at *Spotted* locus. (Photographs: A. Vuchkov)

Figure 4. White Maritza sheep breed with different colour genotypes. a) genotype $E^+E^+A^{Wt}A^{Wt}$, b)genotype $E^+E^+A^aA^aS^+S^s$, c) black mother genotype $E^+E^+A^aA^aS^+S^s$ and white lamb genotype $E^+E^+A^{Wt}A^a$, d) piebald ram, genotype $E^+E^+A^aA^aS^sS^s$. (Photographs: D. Dimov)

years. This cowl has decorative character and satisfies the aesthetic criteria of the shepherds. There are also piebald sheep (1% of the population, Figure 4d). There is no doubt that white colour in over 90% of the animals in this population is caused by the dominant allele (A^{wt}) at the *Agouti* locus, and the genotype of animals with white fleece can be described as $E^+E^+A^{Wt}A^{Wt}$or $E^+E^+A^{Wt}A^a$ (Figure 4). The 'black' colour of animals in this population is recessive and it is due to a *non-agouti* allele (A^a) in a homozygous state A^aA^a (Figure 4b). The lambs are born completely black or dark brown and subsequently the wool acquires a reddish-brown colour, but the front and legs remain black.

The genotype of pigmented animals in the population of this breed can be described as $E^+E^+A^aA^a$. In most cases, pigmented animals in this population have a white spot on the head and tip of the tail, which could be perceived as phenotypic expression of the S^s allele at the *Spotted* locus. Then, genotype of such pigmented animals can be described as $E^+E^+A^aA^aS^+S^s$ (Figure 4b) or $E^+E^+A^aA^aS^sS^s$ (Figure 4d).

Native Starozagorska

The native Starozagorska is a native sheep breed that has taken its name from the area where it was created (Stara Zagora). The geographic area where the breed is kept mainly covers the plains of the districts in Southern Bulgaria: Stara Zagora, Sliven, Haskovo and Yambol. This is a typical large, lowland sheep with a long, narrow, unfleeced neck and head and a convex profile. The ears are long, wide and lop-eared and the belly and legs are naked and unfleeced. The fleece is semi-open and the wool is soft and mostly white with a fibre thickness of 39.6 μm Hlebarov (1937). Typical

Native Starozagorska sheep breed are white, but they also appear pigmented (reddish-brown) with a white spot on an oblong white strip on the face called 'brezi' by native people (Figure 5). There are sheep breeders who prefer to breed coloured sheep (Djorbineva, 2008). Pigmented animals turn black and, with age, the colour of the wool acquires a reddish-brown colour, but the covering hairs on the front and legs remain intensely black. The average live weight of ewes is 64 kg and 80 to 90 kg in rams (Djorbineva, 1984). Recently, there is information about animals with higher body weight. The wool yield for the ewes is 2.8 kg, and 3.5 kg for the rams. The staple length is 9 to 10 cm.

The white colour in this breed is dominant and the genotype of animals with white coat can be either $E^{+}E^{+}A^{Wt}A^{Wt}$ or $E^{+}E^{+}A^{Wt}A^{a}$ (Figure 5a). The genotype of the coloured variety in this breed can be $E^{+}E^{+}A^{a}A^{a}S^{+}S^{s}$ (Figure 5b). At its most extreme piebald animals may be found ($E^{+}E^{+}A^{a}A^{a}S^{s}S^{s}$).

Dubenska

The Dubenska sheep breed is a native sheep breed found in the central part of the Sredna Gora region. Dubenska sheep are medium-sized and their tail is long, reaching below the hocks, sometimes to the ground. The sheep are well-fleeced; the head is fleeced up to the line of the eyes, the belly is totally fleeced, the forelimbs are fleeced up to the carpal joints, and the hind limbs are fleeced until under the hocks and, in some cases, they are completely fleeced. Until the 1930s, the predominant share of the Dubian sheep population was pigmented (brown-black, downy, reddish-brown) and white animals were rarely found (Figure 6a, b). This was related to the production of a coarse, woolen natural brown cloth called 'Shiak', which was known from the Karlovo district in the past, but is no longer in high demand. In the modern population, the white sheep are beginning to dominate and the proportion of coloured sheep is about 60%. The wool is mostly 46-48th quality in the Bradford system for wool classification (USDA, 2018).

The white colour of about 40% of the animals in the population is due to the dominant allele A^{wt} of the *Agouti* locus, but the remaining 60% of the animals are pigmented and this is due to the recessive *non-agouti* allele A^{a}. The frequency of these two alleles in the population depends on the farmer's choice for white or pigmented lambs. Keeping in mind that it is phenotypically difficult to distinguish the manifestation of the E^{D} allele in the *Extension* locus and the recessive allele A^{a} in the *Agouti* locus, it would be very interesting to perform DNA analysis to prove the presence or absence of the dominant allele E^{D} in the *Extension* locus for this breed.

Black-headed Pleven

The Black-headed Pleven is one of the most popular native sheep breeds in Bulgaria. This breed has a large population (18,452 sheep) and can mainly be found in the northern part of the country, but there are also flocks in Southern Bulgaria.

The head is long and narrow with long ears. The lambs are born black, sometimes with a white spot on the crown of the head and tip of the tail, but then the wool on the body fades to white (Figure 7a, b). The colour of the fleece is usually white (Figure 7d) but black hairs can often be found among the white fibres (Figure 7c), making the fleece look faded and grey (Ivanov, 1942). About 5% of the sheep have a fully coloured fleece. The tail is long and covered with wool that corresponds to the quality of the wool on the body. The live weight of ewes is about 55 to 70 kg, and of rams 80 to 100 kg. The wool yield in the ewes is about 2.8 to 3.5 kg and for the rams 4 to 4.5 kg. The staple length is about 12 to 16 cm and fibre thickness is 30.09 to 32.97 μm (Georgiev, 1990).

a b

Figure 5. Native Starozagorska sheep breed. a) genotype $E^{+}E^{+}A^{Wt-}$, b) genotype $E^{+}E^{+}A^{a}A^{a}S^{+}S^{s}$. (Photographs: D. Dimov)

The coat colour of the Black-headed Pleven sheep breed is very similar to that of many other breeds that have black faces and legs and white fleece (e.g. Tsigai, Suffolk breeds). It has been suggested that this is due to another gene called *Dark Brown* (Sponenberg, 1997), but this colouring is still poorly understood. All crosses of Black-headed Pleven with white wool sheep breeds look piebald or similar due to the Ph^T allele in the *Pigmented head* locus, which excludes the presence of dominant E^D allele in the *Extension* locus. Probably this lightening of the wool is due to the interaction of the *Agouti* and the *Brown* loci or to the action of other gene modifiers.

Kotlenska

The animals of the Kotlenska sheep breed are relatively small, the fleeced tail reaching the hock joints. The body, including the abdomen, is well fleeced. The fleece is mostly semi-open white or coloured with shades of grayish to black. The wool is roughly mixed with a fibre length of 14 to 18 cm. Average wool yield is 1.9 kg for ewes, and 3.25 kg for rams (Tzochev *et al*, 2017).

Tetevenska

The Tetevenska sheep breed can be found in the region of the Central Balkan Mountains. At the age of 3.5 years, the ewes of this breed reach 43.47 kg. The wool yield of the Tetevenska sheep breed is about 3 kg with a variation of 3.03 to 3.28 kg. They have a non-uniform mixed coarse or semi-rough wool and have a fibre length of 12.93 cm. White coloured fleece dominates in the population, but also coloured sheep with a different nuance of the fleece colour (black, red, grey) can be found (Genkovski, 2002). An approved breeding programme for preservation of the Tetevenska sheep breed was carried out by the approved breeding association with headquarters in the town of Troyan. According to the association's data, by 2018 the population size was 4.642 animals (ewes and rams), with a risk status of vulnerable. From the point of view of coat colour, the Tetevenska sheep breed is poorly studied.

Koprivshtenska

This native breed is at risk of extinction. Its population size at 2018 was 3,375 (Table 1). The wool of the sheep is mostly uniform, but there are also animals with non-uniform wool. The fibre thickness of uniform wool varies in a very wide range from 29.1 to 40.0 μm. The fleece is white, but in around 40% of the animals, the colour is dark brown. The weight of the fleece averages 3.45 kg and the fibre length 10 to12 cm (EASRAB, 2013).

Figure 6. Dubenska sheep breed. a) flock, b) old ewe, c) lambs, d) lamb with black fleece variety (genotype $E^+E^+A^aA^aS^+S^s$). (Photographs D. Dimov)

Figure 7. Black-headed Pleven sheep breed. a) male hogged, b) ram, c) black fibres visible in white fleece, d) fleece with only white fibres. (Photographs: A. Vuchkov)

Zapadnostaroplaninska

This is a popular native sheep breed from the western part of the country. According to its population size of 4,407 in 2018 (Table 1), the breed's risk status is endangered. Average fleece weight is about 3.0 kg with 10 to 12 cm fibre length and 31.1 to 40.0 μm fibre thickness. The fleece is usually closed and predominantly white, but about 10% of the sheep in the population have coloured fleece (EASRAB, 2013). In terms of wool and hair colour, the population is quite diverse and fully pigmented or completely white sheep can be found, but there are also piebald animals. From the point of view of coat colour, the Zapadnostaroplaninska sheep breed is poorly studied.

Srednorodopska

This breed is typical in the Rodopa Mountain in the southern part of the country, where sheep are capable of long transitions in the mountains on steep and rocky terrains. The population size is 8,716, which defines the risk status as not endangered (Table 1). The wool is rough and around 55% of the animals have coloured fleece. The wool yield is 1.75 kg (EASRAB, 2013).

Srednostaroplaninska

This native sheep breed is typical of the Central Balkan mountains. Some flocks can be found in settlements in lowlands close to the mountains. The wool is coarse mixed and the fleece is made up of three types of fibres, fine, coarse and medium. The fleece is coloured with shades of grey, fawn and red. The wool yield is 2.75 kg per ewe and 3.5 kg per ram (EASRAB, 2013).

Sakarska

This breed is a typical, rare native breed with endangered risk status named after the Sakar Mountain in Southeastern Bulgaria. Today it is widespread in other parts of Southern Bulgaria. The colour of the fleece is mostly white, but 5% of the animals in the population have pigmented fleece. The weight of the fleece is 2.9 kg and fibres are coarse (EASRAB, 2013).

Breeds with spotted coloured fleece

Patch-faced Maritza

The area in which the Patch-faced Maritza sheep breed was created and became popular is located to the west and north of the city of Plovdiv (central part of Southern Bulgaria). In recent years, the Patch-faced Maritza sheep have gained popularity and have spread to other areas including Sofia, Stara Zagora, Haskovo, Yambol, Burgas, Kardzhali, Sliven and Blagoevgrad. The elongated form of the body, the limbs and the tail are characteristic of the Patch-faced Maritza sheep breed (Figure 8a). The head and legs are long and unfleeced, also the tail is thin and long. The belly is also usually unfleeced but, in

some animals, it is poorly fleeced. The wool is uniform with a fibre thickness of 35.01 μm which, according to the Bradford classification, means 46^{th} quality (Dimov and Djorbineva, 1999; USDA, 2018). Black spots with varying sizes are found on the face, the legs and on certain areas of the body, e.g. the root of the tail, the chest and abdomen (Figure 8b, d). Coloured wool usually grows on the pigmented areas of the body, which is why animals with larger spots on the skin of the body look piebald. However, the main fleece mass has a white colour. The pigmentation on the face, shape and size of the ears, and certain features in the body, are crucial in the selection of male and female lambs for breeding (Figure 8c). The weight of the fleece after shearing is 2.8 kg, with a fibre length of 11.2 cm. Our studies show that the live weight of ewes can reach 74.47 kg (Figure 8b) and of rams up to 121.14 kg (Dimov *et al*, 2016). The level of nutrition and flock husbandry has a strong influence on the live weight of the ewes and the rams, thus other reports of different flocks recorded sheep live weight in the range of 68.21 kg to 90.82 kg (Dimov *et al*, 2016).

The breed standard of Patch-faced Maritza sheep is associated with specific pigmented noses, eye patches, ears, root of the tail and distal parts of the legs (Dimov, 2011a). This phenotype is due to the *turkish* allele located in the *Pigmented head* locus (Sponenberg, 1997; Lundie, 2011). The genetics of the *turkish* allele (Ph^T) was first reported in the Akkaraman sheep breed in Turkey (Mason, 1967). In the homozygous state, the *turkish* allele causes an intense black colour of the specific spots on the head, ears, body and legs (Figure 8a, b, c).

Occasionally, in the population of Patch-faced Maritza sheep breed there are piebald animals with smaller or larger coloured spots on the body (Figure 8d). Fully pigmented animals with a white spot on the crown of head and tip of the tail, due to the S^s allele of the *Spotted* locus, are also relatively rare.

Figure 8. Patch-faced Maritza sheep breed. a) male hogged (colour genotype $E^+E^+A^aA^a$ Ph^tPh^t), b) flock (colour genotype homozygous for turkish allele Ph^t), c) lambs (homozygous for turkish allele Ph^t), d) piebald ram. (Photographs: D. Dimov)

The *turkish* allele (Ph^T) is the main reason for the coat colour also of the Breznishka sheep breed and probably of the Elin-Pelinska sheep breed (Figure 9a). In the population of Elin-Pelinska sheep breed, the colour pattern of the head (Figure 9b) is different from that of Patch-faced Maritza, but it is probably also regulated by the *turkish* allele Ph^T.

Other native sheep breeds listed in Table 1 have a white fleece colour and do not contribute to the production of coloured wool nor are they well studied.

Production potential for natural coloured wool in Bulgaria

Considering the great genetic diversity of native sheep breeds in Bulgaria and especially those with coloured wool, we assessed the country's potential for producing coloured wool based on the average wool yield and the percentage of animals with pigmented wool in their populations EASRAB (2013). This is essential for the prospect of business strategies that could be based on the use of naturally coloured wool.

Based on the population sizes of the native sheep breeds and the relative share of the animals with coloured wool, the country's potential for coloured wool production was estimated at 133,791 kg of wool (Table 2), which can be considered the minimum amount of naturally coloured wool that can be produced in the country. In addition, some breeds also produce grey wool, caused by black fibres present in the otherwise white fleece, which adds a grey hue to the wool.

a

b

Figure 9. Elin-Pelinska sheep breed. a) Ram, b) Typical colour pattern of the head. (Photographs: D. Dimov)

Primary wool processing

According to the European Directive 1069/2009 Art. 10, greasy wool produced in sheep farms is classified as 'animal by-product' of the 3^{rd} category European Commission (2009). This circumstance has added a series of regulations and administrative requirements to handling of grease wool (Chaupin, 2014). Pursuant to the Waste Management Act in Bulgaria (WMA, 2012), the primary processing of grease wool and in particular its washing must be carried out in specialised licensed enterprises. An essential requirement for obtaining a permit in the country for wool washing production is the availability of a wastewater purification plant. At this moment, only one company in the country meets the requirements of the Waste Management Act and can lawfully carry out all wool treatments of sorting, scouring, carding and combing, focusing, however, on large batches of wool. At present, there are no suitable enterprises for the legal processing of small batches of wool produced by small producers.

The market for naturally coloured wool products

Current trends in lifestyle, focusing on protection of the environment, reducing carbon emissions, protecting traditional production systems and sustainable rural development, are gradually increasing interest in naturally coloured wool products, providing a unique opportunity for breeders of native sheep breeds to add value to the annual wool production in their farms.

Current purchase prices of greasy wool announced by the main Bulgarian wool processing company KOLHIDA-SLIVEN JSC are presented in Table 3.

Unfortunately, for many years, the low purchase prices of greasy wool led to a sceptical attitude of farmers to the wool as a product that could bring additional income. At present, there are no farmer initiatives to produce wool products from coloured wool. The market for naturally coloured wool products is heterogeneous. In the country's markets, naturally coloured wool from native breeds can be seen in handmade socks, hats, sweaters, vests (Figure 10), usually produced by artisans and not the farmers themselves. There are informal communities of artists in the country who produce various souvenirs and articles, but this market is weak and, at this stage, artworks made of naturally coloured wool are mainly sold around the major tourist centres in the country.

Table 2. Potential of native sheep breeds in Bulgaria for coloured wool production. Data source: Executive Agency on Selection and Reproduction in Animal Breeding (EASRAB; https://www.iasrj.eu/).

Breed name	Population size	Wool Yield (kg)	Coloured animals (%)	Coloured animals (total)	Potential for coloured wool production (kg)
Copper-red Shumen	16,382	2.50	100	16,382	40,955
Karnobatska	1,301	3.25	100	11,301	4,228
Karakachanska	11,621	2.75	100	11,621	31,958
Dubenska	12,660	2.75	60	7,596	20,889
White Maritza	886	3.50	8.5	75	264
Kotlenska	5,998	1.90	70	4,199	7,977
Black-headed Pleven	18,452	3.15	8	1,476	4,650
Native Starozagorska	834	3.75	2	17	63
Tetevenska	4,642	2.50	10	464	1,161
Koprivshtenska	3,375	3.45	40	1,350	4,658
Zapadnostaroplaninska	4,407	3.00	10	441	1,322
Srednorodopska	8,716	1.75	55	4,358	8,389
Srednostaroplaninska	10,091	2.75	10	1,009	2,775
Patch-faced Maritza	8,037	2.60	20	1,607	4,179
Sakarska	2,238	2.90	5	112	325
Elin-Pelinska	6,599	2.50	0	0	
Replianska	5,982	2.50	0	0	
Breznishka	3,201	2.70	0	0	
Country potential	**125,422**			**62,008**	**133,791**

Figure 10. Artisan products made from naturally coloured wool available in Bulgaria. a) handmade wool socks, b) woolen works of art and naturally coloured yarn. (Photographs: D. Dimov)

Table 3. Purchase prices of greasy wool from native sheep breeds in Bulgaria in 2019.

Colour of the wool	price (€/kg)
White	0.66
Grey	0.41
Coloured	0.31
Trim	0.12

Conclusions

The wide variety of native sheep breeds in Bulgaria, in whose populations there are animals with coloured wool, can be used to satisfy the new trends in the textile industry. In this review, we have highlighted the phenotypic expression of several genetic alleles responsible for wool colour (E^D, A^a, Ph^T) in the populations of native Bulgarian sheep breeds. The white colour in populations of several breeds is due to the presence of the dominant A^{wt} allele in the *Agouti* locus. Bulgaria's potential for the production of coloured wool was estimated at 133,791 kg, which can be considered as a minimum at this stage, and with an appropriate market demand, this quantity can be increased. Bulgarian native sheep breeds have the potential to produce different types of wool, which is a prerequisite for various applications in the textile, fashion and art industries. In the immediate future, the market for naturally coloured wool products could be enlarged by increasing public awareness for their consumption and promoting fashion trends which respect environmental protection.

Supplemental data

Supplemental file: Abstract in French.

Author contributions

The authors contributed equally to the writing of this review.

Conflict of interest statement

The authors declare that no conflict of interest exists.

References

Barić, C. (1952). Grada I eksterijer kucovlaške ovce. *Poljoprivrednaznanstvena smotra* 13. (in Croatian).

Broad, T. E., Dolling, C. H., Lauvergne, J.-J., and Millar, P. (1999). Revised COGNOSAG guidelines for gene nomenclature in ruminants 1998. *Genetics Selection Evolution* 31(3), 263–263. doi: https://doi.org/10.1186/1297-9686-31-3-263

Bultman, S. J., Michaud, E. J., and Woychik, R. P. (1992). Molecular characterization of the mouse agouti locus. *Cell* 71(7), 1195–1204. doi: https://doi.org/10.1016/s0092-8674(05)80067-4

Chaupin, M. T. (2014). Developing a market for European wool. In *Timeless coloured sheep,* ed. du Toit, D. 87-89.

Cieslak, M., Reissmann, M., Hofreiter, M., and Ludwig, A. (2011). Colours of domestication. *Biological Reviews* 86(4), 885–899. doi: https://doi.org/10.1111/j.1469-185x.2011.00177.x

Dimov, D. (2011a). Breeding program of Patch-faced Maritza sheep breed. (Intelexpert-94 Publishing house), (in Bulgarian).

Dimov, D. (2011b). Milk yield and body weight of White Maritza sheep. In Proceedings of 8th Global conference on the conservation of Animal Genetic Resources, , 113-115.

Dimov, D. (2015). White Maritza sheep - origin, state and perspectives for development (Plovdiv: Intelexpert-94 Publishing house).

Dimov, D. and Djorbineva, M. (1999). Study on thickness of wool fibres of White and Patch Faced Maritza sheep. *Scientific Works* XLIV(3), 191–194.

Dimov, D., Vuchkov, A., and Ivanov, I. (2016). Study on ewe live weight of Patch-faced Maritza sheep breed. In 2nd International Symposium for Agriculture and Food, 67-72.

Djorbineva, M. (1984). Variability of selection traits of local Stara Zagora sheep breed and possibilities for their improvement. Ph.D. thesis, Research institute of cattle breeding and sheep breeding, Agricultural Academy, Stara Zagora. (in Bulgarian).

Djorbineva, M. (2008). The Stara Zagora sheep breed and breeding association's breeding activities. *Sheep Breeding* (2), 7–13.

EASRAB (2013). Livestock breeds in the republic of Bulgaria. Catalogue 4th edition.

European Commission (2009). Regulation laying down health rules as regards animal by-products and derived products not intended for human consumption and repealing Regulation (EC) No 1774/2002 (Animal by-products Regulation). url: https://eur-lex.europa.eu/eli/reg/2009/1069/.

FAO (2013). In vivo conservation of animal genetic resources volume 14 of *FAO Animal Production and Health Guidelines*. (Rome: FAO). url: http://www.fao.org/3/i3327e/i3327e00.htm.

FAO (2017). FAOSTAT. url: http://www.fao.org/faostat/en/. accessed date: 2021-05-26

Fontanesi, F., Beretti, V., Riggio, S., Olio, D., Calascibetta, V., Russo, B., and Portolano (2010). Sequence characterization of the melanocortin 1 receptor (MC1R) gene in sheep with different coat colours and identification of the putative e allele at the ovine Extension locus. *Small Ruminant Research* 91(2-3), 200–207. doi: https://doi.org/10.1016/j.smallrumres.2010.03.015

Genkovski, D. (2002). Comparative characteristics of different breeds and varieties of sheep, reared under the conditions of the Central Balkan Mountains. Ph.D. thesis, Research Institute of Mountain Stockbreeding and Agriculture, Trojan. National Centers of Agrarian Sciences, Sofia, Bulgaria.

Georgiev, D. (1990). Genealogical structure, phenotypic and genotypic characterization of selection trait in sheep of Black-head Pleven sheep breed by the example of the flock of village Riben to PJK "G.Dimitrov". Ph.D. thesis, Agricultural academy, Sofia, Bulgaria.

Hlebarov, G. (1933). The Karnobat sheep volume 31 of *Scientific works*. (Sofia: Bulgarian Agricultural Society), 104p.

Hlebarov, G. (1937). Sheep breeds and varieties in Bulgaria. In Yearbook of Sofia University, Faculty of Aronomy and Forestry, Sofia.

Hlebarov, G. (1942). Karakachan sheep. In Yearbook of Sofia University, Faculty of Agriculture and Forestry, Sofia.

Iliev, M. (2007). Characteristics of Karnobat local sheep. In Scientific and practical conference with international participation "AGROECO 2007", 265-269.

Iliev, M. (2012). Characteristics of productivity in sheep from Karnobat local breed. *Animal Sciences* 3, 11–16.

Ivanov, P. (1942). Black-head Pleven sheep breed. Study on the exterior and productive performances (Pleven) 116p.

Jackson, I. J. (1994). Molecular and Developmental Genetics of Mouse Coat Color. *Annual Review of Genetics* 28(1), 189–217. doi: https://doi.org/10.1146/annurev.ge.28.120194.001201

Koseniuk, A., Ropka-Molik, K., Rubi, D., and Smołucha, G. (2018). Genetic background of coat colour in sheep. *Archives Animal Breeding* 61(2), 173–178. doi: https://doi.org/10.5194/aab-61-173-2018

Lauvergne, J. J. and Raffier, P. (1975). Génétique de la couleur de la toison des races ovines françaises solognote, bizet et berrichonne. *Annales de Génétique et de Sélection Animale* 7(3), 263–276. (in French). doi: https://doi.org/10.1186/1297-9686-7-3-263

Lundie, R. S. (2011). The genetics of colour in fat-tailed sheep: a review. *Tropical Animal Health and Production* 43(7), 1245–1265. doi: https://doi.org/10.1007/s11250-011-9850-0

Mason, I. L. (1967). Sheep breeds of the Mediterranean (Rome: FAO).

Milić, L. (1954). Prilog poznavanju osobina crne vlaške (karakačanske) ovce I način stočarenja s njom (Skopje: Zavod za stočarstvo NR Makedonije), (in Serbian).

Notter, D. R. and Sponenberg, D. P. (2002). Genetic control of colour in Dorper sheep and their Crosses. *Sheep & goat research journal* 17(2), 48–51.

Savov, T. (1964). Development of sheep farming in Bulgaria till 9 September 1944, ed. and others. (Sofia: Publishing House of the Bulgarian Academy of Sciences).

Searle, A. G. (1968). Comparative Genetics of Coat Colour in mammals (London, UK: Logos Press).

Sedefchev, S. and Sedefchev, A. (2011). Comparison of exterior parameters of in situ conserved Karakachan sheep with the breed type from the first half of the twentieth century. *Agrarni Nauki* 3(6), 147–153. (in Bulgarian).

Sponenberg, D. P. (1997). Genetics of colour and hair texture. In *The genetics of sheep,* ed. Piper, L. and Ruvinsky, A., (New York: CAB International), 51-86.

Sponenberg, D. P., Dolling, C. H. S., Lundie, R. S., Rae, A. L., Renieri, C., and Lauvergne, J. J. (1996). Coat colour loci (Category 1). In *Mendelian Inheritance in Sheep 1996 (Mis 1996),* ed. Lauvergne, J. J., Dolling, C. H. S., and Renieri, C., (France, Italy: COGNOSAG and University of Camerino), 13-57.

Staikova, G. (2005). Study on the value of the productive traits in sheep from the Karakachan breed and Copper-Red Shumen strain. Ph.D. thesis, Agricultural institute, Shumen, Bulgaria.

Staikova, G. and Iliev, M. (2017). Karnobat sheep breed. *Bulgarian Journal of Animal Husbandry* 3, 3–15.

Staikova, G., Stancheva, N., and Dimitrova, I. (2015). Copper-red Shumen sheep breed. *Bulgarian Journal of Animal Husbandry* 52(5), 72–80. (in Bulgarian).

Stamatis, C., Giannoulis, T., Galliopoulou, E., Billinis, C., and Mamuris, Z. (2017). Genetic analysis of melanocortin 1 receptor gene in endangered Greek sheep breeds. *Small Ruminant Research* 157, 54–56. doi: https://doi.org/10.1016/j.smallrumres.2017.09.018

Stefanova, S. and Iliev, M. (2005). Comparative evaluation of the slaughter qualities and the main tissue structures in the carcass of lambs from the Karnobat thin-wool and Karnobat local breeds. In and others, Sat. Selection and agrotechnics of field crops. Livestock. Balkan Scientific Conference, 586-592.

Tzochev, I., Slavov, R., Karadjov, H., Bonchevska, S., and Nikolova, T. (2017). Breeding program of Kotlenska sheep breed in Bulgaria.

USDA (2018). Wool Grades and Standards. url: https://www.ams.usda.gov/grades-standards/wool-grades-and-standards. accessed date: 2021-05-26

WMA (2012). Waste Management Act. url: https://www.me.government.bg/en/library/wastem-anagement-act-342-c25-m258-1.html.

Yang, G.-L., Fu, D.-L., Lang, X., Wang, Y.-T., Cheng, S.-R., Fang, S.-L., and Luo, Y.-Z. (2013). Mutations in MC1R Gene Determine Black Coat Color Phenotype in Chinese Sheep. *The Scientific World Journal* 2013, 1–8. doi: https://doi.org/10.1155/2013/675382

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2021), 2 (3), 25–35
DOI: 10.46265/genresj.EQFQ1540
https://www.genresj.org
ISSN: 2708-3764

Genetic diversity and population structure among indigenous and imported goat breeds in Kenya

Ruth W Waineina [*,a,b], **Kiplangat Ngeno** [a], **Tobias O Okeno** [a] **and Evans D Ilatsia** [b]

[a] *Department of Animal Sciences, Animal Breeding and Genomics Group, Egerton University, P.O. Box 536-20115, Egerton, Kenya*
[b] *Dairy Research Institute, Agricultural and Livestock Organization, P.O. Box 25-20117, Naivasha, Kenya*

Abstract: Population structure and relationship information among goats is critical for genetic improvement, utilization, and conservation. This study explored population structure and level of introgression among four goat breeds in Kenya: the indigenous Galla (n = 12) and three imported breeds, the Alpine (n = 29), Toggenburg (n = 31), and Saanen (n = 24). Genetic diversity was analyzed using four indices (polymorphic SNPs, mean allele frequency, observed and expected heterozygosity and inbreeding coefficient) within each breed. Population structure assessed using model-based clustering (ADMIXTURE) revealed four breeds according to their geographic regions in Kenya. Kenyan Alpine goats were the most admixed breed with about 10 % of its genome derived from Galla, 10 % and 6 % from Saanen and Toggenburg respectively. The association of Galla with other breeds was anticipated since the Galla breed was used as the founder population for crossbreeding with Saanen, Alpine and Toggenburg breeds. The relationship information evaluated by computing Reynolds genetic distance revealed five distinctive clusters: Alpine, Galla, Saanen, Toggenburg and some mixture of Alpine and Toggenburg. Saanen and Galla breeds seem to be the most genetically distinct among the sampled populations. The genetic variation among the goat populations observed will provide a good opportunity for sustainable utilization, conservation, and future genetic resource improvement programmes in goat breeds in Kenya.

Keywords: Admixed breeds, Breed relationship, Crossbred goats, Gene introgression

Citation: Waineina, R. W., Ngeno, K., Okeno, T. O., Ilatsia, E. D. (2021). Genetic diversity and population structure among indigenous and imported goat breeds in Kenya. *Genetic Resources* 2 (3), 25–35. doi: 10.46265/genresj.EQFQ1540.

Introduction

Goats are known to be the most adaptable and widespread species of domestic animals, thriving across various geographical conditions, ranging from mountains to deserts and the tropics, Africa included. The importance of goats in supporting rural household economies in developing countries is well documented (Deshingkar *et al*, 2008; Herrero *et al*, 2013). They are an important source of food and nutritional security through the supply of milk and meat, income generation through sale of surplus stock, and insurance against unforeseen risks in addition to having important, non-tangible cultural values (Herrero *et al*, 2013; Mbuku *et al*, 2015; Ogola *et al*, 2010). Recent studies have shown that goat farming is one of alternative climate-smart agricultural practices that could build farmers resilience to climate change-related challenges (Ojango *et al*, 2016). The diminishing land sizes in the medium to high potential areas for agriculture due to human population pressure, expansion of urban areas and climate change-related challenges, call for alternative farming practices such as intensive dairy goat production, which offers more multi-functionality, flexibility, and adaptability to varied production conditions (Scarpa *et al*, 2003).

In Kenya, dairy goat production has mainly been supported by imported breeds such as Toggenburg, Anglo-Nubian, German Alpines, Saanen and Boer, and crossbreeds between imported and selected local breeds such as Galla and small East African goat (Ahuya *et al*, 2006; Bett *et al*, 2007; Krause, 2006). Galla

*Corresponding author: Ruth W Waineina (waineinaruth@gmail.com)

Received: 29.12.2020 | Accepted: 27.05.2021 | Published online: 17.06.2021

in Kenya is also referred to as Boran/Somali goat. They are indigenous in arid and semi-arid northern Kenya, and pure Galla are maintained in various Government breeding and conservation stations in the Country. Their characteristics, such as resistance to dehydration, preference for browsing and a wide range of feeding habits (Chenyambuga *et al*, 2004) have allowed them to adapt to the massive arid and semi-arid regions in the country, and could potentially be advantageous traits for goat breeding programmes. The imported breeds were introduced to various parts of the country by the Government of Kenya and Non-Governmental Organizations, with the aim of increasing goat productivity through appropriate husbandry and disease interventions (Peacock, 2005) and targeted breeding strategies such as crossbreeding (Bett, 2009; Peacock *et al*, 2011). The crossbreeds were kept in different geographical locations as isolated populations and subjected to separate breeding objectives for several decades.

Crossbreeding has been the strategy of choice to improve the productivity of goats under various production systems (Ahuya *et al*, 2009; Peacock *et al*, 2011). This has resulted in an increase in population sizes of crossbred goats especially in the areas the breeds were introduced (Mburu *et al*, 2014; Peacock *et al*, 2011). However, increase in population sizes did not necessarily correspond to enhanced productivity of the goats but rather reflected large numbers of households striving to support their livelihoods through goat farming (Aziz, 2010; Bett *et al*, 2011; Mburu *et al*, 2014).

In Kenya, there has been limited technical capacity on the farmers' side on how to manage the rather complex crossbreeding programmes, a fact that may have had a bearing on the sustainability of such initiatives in the long term (Aziz, 2010; Bett *et al*, 2011; Mburu *et al*, 2014). The net result of this has been the unsystematic crossing of the existing population, poor flock management, lack of records to support decision making and general lack of simplified breeding programmes to guide in genetic improvement of goats in the country (Kosgey and Okeyo, 2007). Currently, crosses of imported and local goats are reared as dairy goats in different parts of the country under different production systems. There is a huge source of genetic diversity in the current goat populations in Kenya. This is a result of unsystematic crossbreeding and lack of record keeping by most of the smallholder farmers. This calls for the need to characterize, conserve and sustainably utilize goats under various production systems in Kenya. It is important to determine genetic diversity in populations because it provides the basis for natural and artificial selection (Qanbari and Simianer, 2014).

To measure and describe genetic diversity of animal genetic resources, phenotypic and molecular characterization tools are used as a starting point to understand the animal resources and make use of them sustainably (FAO, 2011). Characterization starts with the gathering of all information on breed origin, development, structure, population, quantitative and qualitative characteristics in defined management and climatic conditions (Gizaw *et al*, 2011; Rege and Okeyo, 2011).

Molecular characterization, using genetic markers to detect polymorphisms in nuclear DNA, is a powerful tool which can be applied in breeding programmes. For instance, it can be used to characterize the genetic variability within, and genetic distance between, populations, as well as for genomic selection, parentage verification and genetic diversity preservation (Groeneveld *et al*, 2010).

Microsatellite markers and single-nucleotide polymorphisms (SNP) are the most commonly used markers in animal breeding related fields (FAO, 2011). Microsatellite markers have several limitations, for example in the detection of null alleles (Hoshino *et al*, 2012) and homoplasy (Jarne and Lagoda, 1996; Anmarkrud *et al*, 2008), while SNPs have several advantages over microsatellites, including being highly reproducible and informative (Vignal *et al*, 2002) and the fact that SNPs can represent either neutral or functional genetic diversity (Kohn *et al*, 2006). A SNP microarray with more than 50,000 SNPs (GoatSNP50 Bead Chip, Illumina, Inc. San Diego, CA 92122 USA), which was developed from SNP loci detected by whole-genome sequencing of six different goat breeds according to Tosser-Klopp *et al* (2014) is available. This has made SNP markers the most popular and advanced technology in molecular breed characterization in goats. Additionally, its robustness, low genotyping costs, automatic allele calling and capability to interrogate the goat genome at high resolution (Ajmone-Marsan *et al*, 2014) demonstrate practicality in implementing genomic characterization in goats.

There has been no deliberate effort to understand the genetic diversity and population admixture among the goat breed populations in Kenya by use of SNP makers. This study, therefore, investigated genetic diversity, population structure and admixture among goat breeds in Kenya. The results from this study will facilitate management efforts in conserving and utilizing the various goat genetic resources sustainably.

Materials and Methods

Study area

Blood samples were collected from goats in three counties in Kenya: Nyeri (Mukurweini Sub-County), Meru (Central Imenti Sub-County) and Homa Bay (Homa Bay Town Sub-County) located in the Central, Eastern and Western regions of Kenya, respectively (Figure 1). These areas were selected because they represent the entry points of different imported dairy goat breeds in Kenya. Mukurweini Sub-County lies in the Upper midlands, also known as the main coffee zone, at an altitude of 1460-1710 metres above sea level (masl) and receives 950-1500 mm of mean annual rainfall.

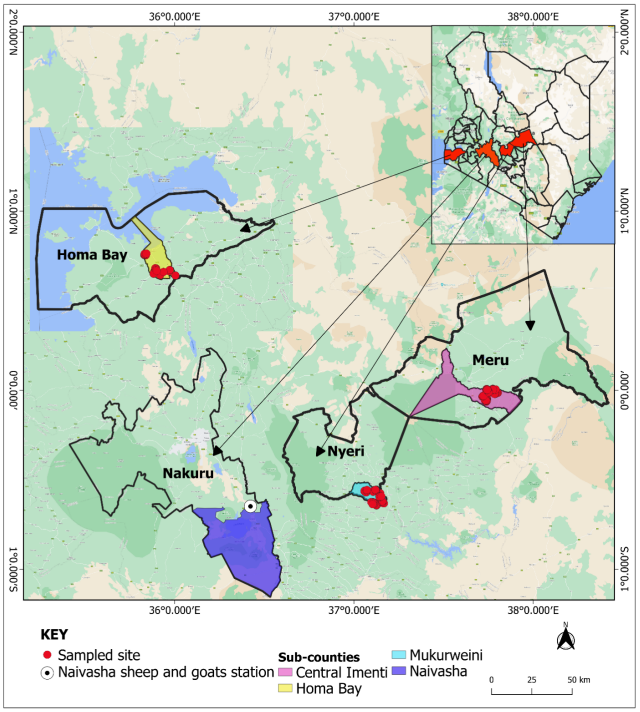

Figure 1. Map of Kenya showing area sampled within the selected sub-counties

Central Imenti is in the upper highlands, at an altitude ranging between 1830-2210 masl and has an average annual precipitation of 800-2600 mm. The Homa Bay Town Sub-County lies in the lower midlands, at 1166 masl and receives an annual rainfall of 1226 mm.

Animal resources and sampling

Goat keeping households were purposively selected based on the following criteria: 1) being a member of dairy goat farmer group and 2) having more than two mature does which matched the breed kept by the farmer group in the said county of study. The herd structures between the breeds and within the county/breed varied among the selected households. Therefore, when a farm had only two mature does, only one doe was sampled. Where more than two does were available, the relationship of the does was confirmed by the farmer to avoid sampling closely related does. Sampling of full and half siblings was avoided. To ensure the representativeness of sampling for each breed, unrelated animals were selected from various farms across the designated counties. The Galla goat breed, however, did not follow the criteria because they were from the government breeding station where breeding records were used to identify the animal to be sampled. Therefore, to minimize sampling from closely related animals within the Galla population, pedigree data were used to select against full and half sibling animals.

A total of 96 animals including three imported breeds (31 Toggenburg, 29 Alpine, 24 Saanen) and one indigenous breed (12 Galla) were incorporated in this study. The Toggenburg and Alpine were found in Eastern and Central Kenya under Meru Goat Breeders Association (MGBA) and Dairy Goat Association of Kenya (DGAK), respectively. Saanen goats were found in Homa Bay under Nyanza Goat Breeders Association (NGBA). All samples were collected from a total of 53 farms in the three counties, Nyeri (18), Meru (19) and Homa Bay (16). Galla goats were sampled from the sheep and goats government station in Naivasha.

Whole blood (10 ml) was collected from the jugular vein into Vacutainer tubes with Ethylenediaminetetraacetic acid (EDTA) as an anticoagulant. The blood was stored at -20°C for two months before genomic DNA extraction. The procedure of blood sampling followed FAO guidelines (FAO, 2012). From each animal, duplicate samples were collected and kept separately during transportation and storage. For each sample, the following information was collected: sex of the animal, basic pedigree information, size of the flock, breed, any relevant phenotypic feature, and a photograph of the goat.

The study was conducted in strict accordance with the recommendations of the Institute of Primate Research (IPR) Ethical Guidelines on Animal Care and Use of Laboratory Animals (https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf).

The protocol was approved by the committee on the ethics of Animal Experiment of Egerton University of Egerton in Kenya (ISERC/03/2020). A qualified veterinary officer collected the whole blood following FAO guidelines (FAO, 2012) to reduce pain and discomfort to a minimum.

DNA extraction and genotyping

Genomic DNA was extracted from the whole blood using the DNeasy Blood and Tissue Kit (Qiagen®, Hilden, Germany) according to the manufacturer's instructions. Ten μl of each ten randomly selected samples were subjected to a preliminary estimate of the DNA quality and quantity on a 1 % agarose gel electrophoresis. Secondary quantification and purity analysis of the DNA were confirmed using one μl for each sample on both Nanodrop Spectrophotometer (Nanodrop® ND-2000) and Qubit ® dsDNA BR (Broad-Range) Assay Kit on the Qubit 3.0 fluorimeter (Invitrogen). The extraction and quality control check of genomic DNA was done at Kenya Agricultural and Livestock Research Organization biotechnology laboratory in Kabete, Kenya.

The DNA samples were genotyped using the GoatSNP50 Bead Chip (Illumina, Inc. San Diego, CA 92122 USA), developed by the International Goat Genome Consortium (IGGC), which features 53347 SNPs across the whole genome with inter-SNP spacing of approximately 40 kb (Tosser-Klopp *et al*, 2014). The genotyping was outsourced to Neogen Europe Limited in Scotland (https://genomics.neogen.com/en/).

SNP quality control and data analysis

The SNP genotype quality control process was applied to raw reads for both merged (all breeds) and then per breed using PLINK v1.07 (Purcell *et al*, 2007).

Table 1. Goat breed, number of goats and SNPs excluded and remaining after quality control processes on genotyping data. N, number of animals; MIND, genotype missing (< 0.1), GENO, SNP missing (< 0.15), MAF, minor allele frequency (< 0.05); HWE, Hardy-Weinberg equilibrium (P-value < 0.001).

Breed	N	Excluded SNPs					Remaining SNPs	Remaining samples
		MIND	GEN	HWE	MAF	TOTAL		
Saanen	24	0	2453	38	3658	6149	47198	24
Alpine	29	1	2496	47	2586	5129	48218	28
Galla	12	0	2413	26	8249	10688	42659	12
Toggenburg	31	1	2345	50	4690	7085	46262	30
Merged	96	2	2235	663	644	3542	49805	94

First, individuals with a missing genotype call rate of more than or equal to 10 % were removed from further analysis using the mind function in PLINK with default settings. The remaining individuals were then exposed to further filtering. SNPs with less than 95 % call rate, Minor Allele Frequency (MAF) of less than 0.05 and P<0.001 Hardy Weinberg Equilibrium (HWE) were excluded from downstream analyses. The SNP data set used for downstream analysis is accessible from the Mendeley Digital Repository (https://doi.org/10.17632/hhb9rhdzzt.1).

Basic genetic diversity indices, which include the proportion of polymorphic markers, inbreeding coefficient, observed (H_o) and expected (H_e) heterozygosity were calculated within breeds using PLINK (Purcell *et al*, 2007). The proportion of polymorphic SNPs (P_N) offers the fraction of the total SNPs that showed both alleles within each population. The P_N was calculated as the proportion of SNPs with more than 1 % MAF within each breed. The MAF is the approximate frequency of the second most common allele per breed. The output from PLINK for observed and expected heterozygosity per animal within breed was subjected to further calculation to get an average estimate of H_o and H_e per breed. The heterozygosity values were calculated by getting the average of all SNPs (that is the sum of all heterozygosity values averaged over the total number of SNPs passed the quality control).

The population structure and relatedness were estimated by principal components analysis (PCA) using the R package SNPRELATE (Zheng *et al*, 2012) and admixture proportion inference using model-based clustering ADMIXTURE 1.3.0 software (Alexander *et al*, 2009). The PCA analysis allowed for visual investigation and solid quantitative summaries. The admixture analysis inferred the proportions of ancestry within the populations by use of prior defined K-values matching the assumed number of ancestral populations. The admixture procedure employs a maximum-likelihood based method by converging the ancestry proportions and allele frequencies that maximize the likelihood function. The most optimal population structure was determined by cross-validation error procedure (McVean, 2009) with assumed ADMIXTURE runs from K = 2 to K = 4. The K-value with the lowest CV error was selected as the optimal value. A phylogenetic tree based on Reynolds genetic distances representing relationships among goat breeds was visualized using iTOL software (Letunic and Bork, 2019).

Results

Quality control procedure on the 53,347 SNPs included on the SNP chip excluded a total of 3,542 SNPs retaining 49,805 SNPs for downstream analyses as shown in Table 1. Of the excluded SNPs, 2,235 had less than 0.1 missing per SNP, 663 SNPs significantly deviated from HWE ($P < 0.001$) and 644 SNPs had MAFs lower than 0.05. The Galla breed had the highest number of SNPs excluded in total (10,688), whereas Alpine revealed the lowest number of SNPs excluded (5,129). It is worth noting that some SNPs were left out due to more than a single criterion.

Genetic diversity

The four indices of genetic diversity (polymorphic SNPs, mean allele frequency, observed and expected heterozygosity and inbreeding coefficient) were calculated within each breed (Table 2). The assessment of the proportion of SNPs that exhibited both alleles within each breed indicated high levels of diversity. The percentage of within-breed polymorphic SNPs ranged from 94.6% to 80.7%. The highest values of polymorphic loci were found in Alpine (94.6%) and Saanen (92.2%) while the lowest proportion was found in the Galla breed (80.7%). Across all loci, the lowest MAF was found in Galla (0.291) and the highest in Alpine (0.323).

Results revealed differences in genetic diversity between breeds. The expected heterozygosity was, in all cases, higher than the observed heterozygosity ($H_e > H_o$). The Alpine had the lowest observed heterozygosity ($H_o = 0.558 \pm 0.026$) while Toggenburg had the highest ($H_o = 0.580 \pm 0.032$). Inbreeding coefficients for all the breeds were negative and ranged between -0.013 (Toggenburg) and -0.042 (Galla).

Population structure analysis

Principal components analysis was used to cluster goats and explore the association among individuals and breed groups. In Figure 2, the principal component 1, which accounts for 15.2% of the total variance, separated Galla breed from the other three breeds. The

Table 2. Proportion of polymorphic SNPs(P_N), mean allele frequency (MAF), observed (H_o) and expected (H_e) heterozygosity and inbreeding coefficient (F) for the goat breeds. N, number of animals.

Breed	N	P_N [%]	MAF	$H_o \pm$ SD	$H_e \pm$ SD	F
Toggenburg	30	89.9	0.297	0.580±0.032	0.580±0.001	-0.013
Alpine	28	94.6	0.323	0.558±0.026	0.564±0.001	-0.015
Galla	12	80.7	0.291	0.563±0.025	0.580±0.000	-0.042
Saanen	24	92.2	0.311	0.559±0.019	0.573±0.001	-0.034

principal component 2 accounts for 14.1% of the total variance, split the goat breeds into four clusters (Alpine, Saanen, Galla and Toggenburg clusters). One outlier, was, however, observed for the Saanen population mixing with Alpine population.

To examine admixture between the breeds, model-based clustering was performed and the most likely number of genetic population (cluster or K) between the four goat breeds were deduced using ADMIXTURE cross-validation procedure (McVean, 2009). The K-value with the lowest CV error was K = 4 and was selected as an optimal number of ancestral populations (Figure 3).

A population structure plot (Figure 4) showed proportions of ancestral populations for all breeds (Alpine, Galla, Saanen and Toggenburg) for K = 2 to K = 4. At K = 2, Galla goats were separated from the other three goat breeds (Toggenburg, Saanen and Alpine). Moreover, Galla goats largely do not carry ancestral components present in Saanen, Alpine and Toggenburg goats (shown in light blue, Figure 4). At K = 3, Alpine and Saanen goats carry ancestral components largely absent from either the Galla or Toggenburg goats. At K = 4, Galla had the lowest level of admixture, whereas Toggenburg, Alpine and Saanen demonstrated some signs of admixture with Galla.

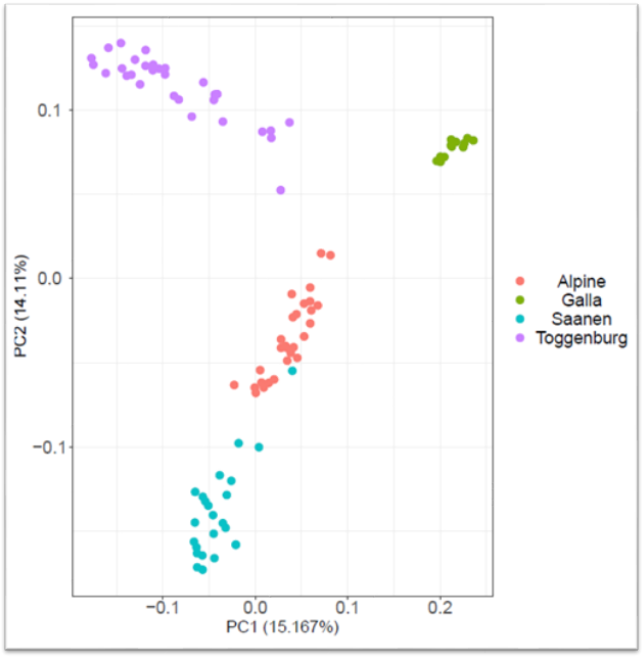

Figure 2. Principal components analysis plot based on SNP array data of goat breeds

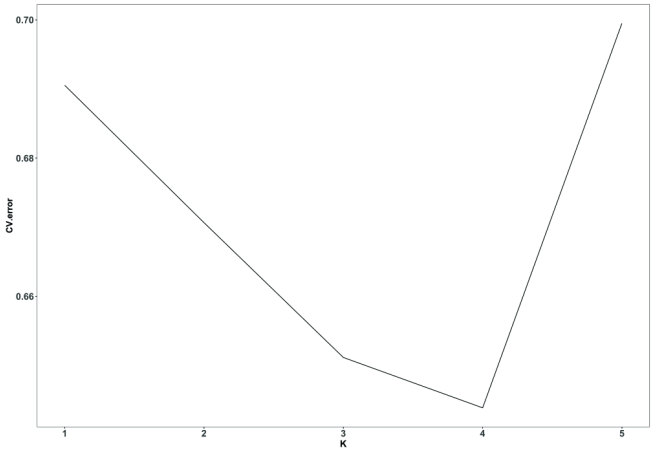

Figure 3. A cross-validation plot indicating the choice of optimal K-value

The proportions of individuals in each breed in the four most likely clusters estimated by ADMIXTURE is shown in Table 3. 71 % of Alpine breed were allocated to cluster one, 97 % of Galla were assigned to cluster two with one percent (1 %) of its genome assigned to cluster one, three and four, 84 % percent of Saanen were in cluster three with seven percent (7 %) of its genome assigned to cluster one. On the other hand, 78 % of Toggenburg were assigned to cluster four with 17 %, and three percent (3 %) of its genome allocated to cluster two and one respectively.

Breed relationships were evaluated by computing the genetic distance between all pairwise combinations of individuals (D) from the average proportion of shared alleles. Based on the calculated Reynolds genetic distances, a phylogenetic tree was constructed to represent breed clustering (Figure 5). The results revealed five clusters for the four populations (Alpine, Saanen, Toggenburg and Galla). Some Toggenburg goats were found to be grouped together with Alpine, forming the fifth cluster.

Discussion

Genetic diversity

Livestock has been exposed to various forces that contributed to the genetic diversity underlying phenotypic dissimilarities ever since domestication. These forces include natural selection, artificial selection for specific traits, migration, genetic drift and inbreeding (Andersson and Georges, 2004; Groeneveld *et al*, 2010). Genetic drift plays an important role during short-term evolu-

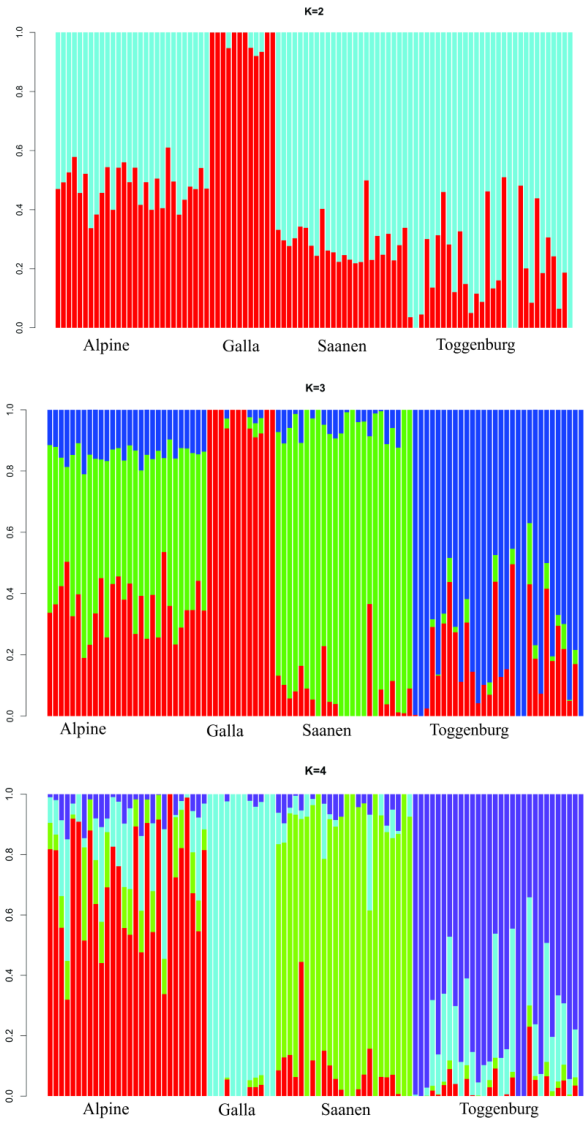

Figure 4. Population structure plots showing proportions of ancestral populations for individuals of sampled goat breeds (Alpine,Galla, Saanen and Toggenburg) for K= 2 to K = 4

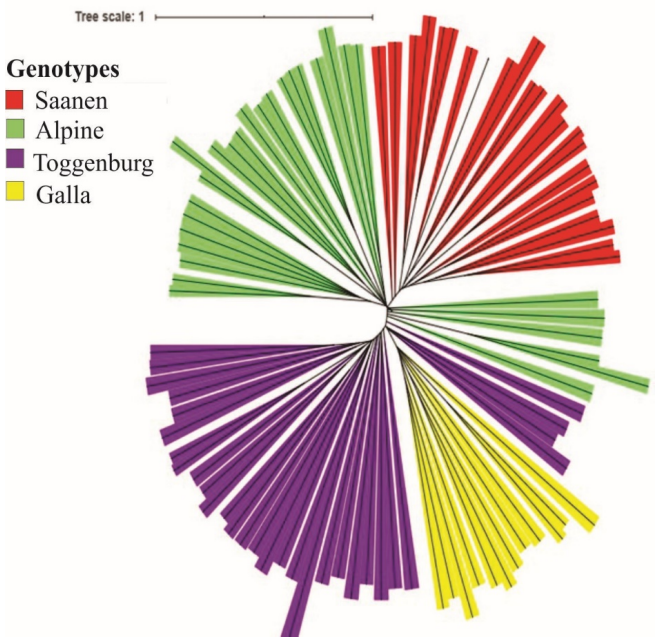

Figure 5. Phylogenetic tree based on Reynolds genetic distances representing breed relationships among goat breeds.

tion in situations where populations are reproductively isolated (Laval *et al*, 2002).

Genotyping with the GoatSNP50 Bead Chip revealed some levels of diversity within the goat breeds in this study. In each breed, fewer than 80% of SNPs exhibited polymorphisms, and heterozygosity ranged from 0.558 to 0.580 (Table 2). A large number of polymorphic SNPs were detected for Alpine (94.6) and Saanen (92.2) breeds; this was expected because sequenced data from Alpine and Saanen were included in the 50K SNPs panel discovery (Tosser-Klopp *et al*, 2014). Other results from various authors using different numbers of samples and goat breeds showed >93 % of polymorphism (Visser *et al*, 2016; Onzima *et al*, 2018). However, it is difficult to compare and conclude the estimates of SNPs stated as polymorphic by other authors, because the number of samples genotyped per breed and proportion of genotyped samples used for SNP discovery varied.

The diversity amongst the four breeds showed Galla had the lowest diversity among sampled individuals in comparison with the other breeds. The polymorphic variance is dependent on the history of each breed. As opposed to Galla goats, the three introduced breeds showed greater diversity, likely as a result of crossbreeding with local goats (Galla and small East African). Therefore, each breed may contain genetic contributions from various breeds, thus revealing higher polymorphisms than Galla goats. In contrast, Galla goats sampled from the government breeding station with detailed pedigree records still maintained levels of diversity (P_N = 81%). The Galla population had been introduced in the Naivasha Sheep and Goat Station in the early 1970s during a sheep and goat

Table 3. Proportion of individuals of goat breeds in each of the four clusters estimated by ADMIXTURE. The diagonal indicates the inferred cluster. N, number of animals sampled

Predefined populations	Inferred clusters N	1	2	3	4
Alpine	28	**0.708±0.197**	0.136±0.128	0.101±0.071	0.055±0.048
Galla	12	0.013±0.020	**0.970±0.044**	0.008±0.013	0.009±0.015
Saanen	24	0.073±0.095	0.049±0.073	**0.838±0.149**	0.040±0.039
Toggenburg	30	0.030±0.047	0.173±0.142	0.015±0.019	**0.782±0.193**

project funded by FAO (Palian and Racokzi, 1976). The population is registered with the Galla Goat Breeders Society of Kenya (GGBSK) and the Kenya Stud Book. The population is inspected every year by inspectors from GGBSK using the Galla goat breed standards. Therefore, the station has maintained pure breed Galla goats which produce meat, milk and reproduce under harsh conditions while maintaining or conforming to the set standards of excellence defined by the GGBSK, where the objective is genetic improvement of target traits while controlling the level of inbreeding. The negative inbreeding coefficient in this study can effectively be taken as zero values, which means that there is no inbreeding observed in the reference populations. It could also mean that many heterozygotes were observed although the sample size for the four breeds was small. The increased heterozygosity could be due to random mating within the herd rather than random differences between herds.

The observed heterozygosity was lower than the expected heterozygosity ($H_o < H_e$) in the three breeds apart from Toggenburg, which recorded the same value for both observed and expected heterozygosity. The difference between the observed and expected heterozygosity was small, which may not be due to inbreeding but a Wahlund effect (Garnier-Géré and Chikhi, 2013). The observed heterozygosity in the current study was from a sample of individuals from a structured population even though all sub-divisions were in Hardy-Weinberg equilibrium. Over and above the semi and intensive production systems practised by smallholders, there is the presence of artificial selection, gene flow and non-random mating, hence not holding the law of HWE in these populations. In this study the observed and expected heterozygosity for Alpine (H_o = 0.558; H_e = 0.564), Saanen (H_o = 0.559; H_e = 0.573) and Toggenburg (H_o = 0.580; H_e = 0.580) were higher than those stated in Canada for Alpine (H_o = 0.385; H_e = 0.388), Saanen (H_o = 0.379; He = 0.382) and Toggenburg (H_o = 353; H_e = 336) (Brito *et al*, 2017). Moreover, Saanen in Italy recorded the same trend as in Canada (H_o = 0.41; H_e = 0.41) (Nicoloso *et al*, 2015). Differences in effective population sizes, length of isolation, selection and breeding management practices in the various production system may be the cause of these variances.

Toggenburg and Galla breeds had the highest expected heterozygosity. This could be explained by the types of crossbreeding programmes practised by farmers keeping Toggenburg breeds resulting in an admixed population. Organized breeding strategies using artificial selection are practised for Galla goats under the government breeding station resulting in genetic variability and lack of inbreeding for the populations.

Population structure and relationship

Principal component and population structure analyses confirmed distinctiveness among the goat breeds (Saanen, Galla, Toggenburg and Alpine) according to their geographic regions in Kenya. This can be explained by the demographic history of these breeds that have been reared for a long time in separate geographic locations (Ahuya *et al*, 2009; Peacock *et al*, 2011). Although goats from each breed clustered separately, model-based clustering revealed some signs of admixture and genetic links between Alpine, Toggenburg, Saanen and Galla.

The results (Figure 4 and Table 3) of this study indicate that Kenyan Alpine goats were the most admixed breed with about 14 % of its genome derived from Galla, while ten and six percent of its genome is resulting from Saanen and Toggenburg respectively. It is worth noting that Saanen were introduced in the sampling region (Nyeri County) already before the Alpine were imported in the late 1970s. Therefore, the 10 % of Saanen genes in the Alpine genome may be a result of Saanen being one of the Kenyan Alpine ancestors. According to Waineina *et al* (2019), lack of breeding stock was one of the challenges Alpine farmers were encountering, thus driving them to source breeding animals from local markets, friends, neighbours and commercial farms notwithstanding their undefined genetic composition. Furthermore, the increase in demand for dairy goats in the country has resulted in several farms setting up nucleus flocks with a significant proportion of the crossbred flocks as a source of breeding material for distribution to lower cadre farmers (Ahuya *et al*, 2006; Bett, 2009; Ogola *et al*, 2010). Through such arrangements, most of the breed-types have migrated to other areas apart from their original entry in the country (Mburu *et al*, 2014; Peacock, 2007). Toggenburg and Alpine goats shared some linkage with Galla goats, 17 % and 14 %, respectively. This was expected because Galla goat was used as the founder population for crossbreeding with Alpine and Toggenburg breeds (Ahuya *et al*, 2009; Bett *et al*, 2011; Mburu *et al*, 2014; Peacock *et al*, 2011; Shivairo *et al*, 2013).

As expected, Galla was the least admixed breed, in agreement with the history of this breed as the first indigenous goat for which a breed society was formed in Kenya. Moreover, the particular population in this study has been managed in seclusion within the government farm, and only animals registered within the society are allowed into the population. The Galla breed displayed isolation by distance and seemed to be at equilibrium under dispersal and genetic drift. In comparison with the other breeds in this study, Galla arrived in their current locations long before these breeds were introduced in Kenya and that is why there has been sufficient time for isolation by distance to take effect and, that long distance gene dispersal is sufficiently common to prevent genetic divergence.

The phylogenetic analysis categorized the breeds into five clusters (Figure 5). The outcomes show a clear differentiation of Galla, Saanen, some Toggenburg and Alpine. A group of some Alpine and Toggenburg, however, remained clustered together, which may

be attributed to the adjacent regions of the breeds (Figure 1). Lack of differentiation in some of Alpine and Toggenburg breeds signified a high level of genetic resemblance and low divergence, which may be a result of gene flow among Alpine and Toggenburg breeds. Common ancestry, short domestication history, lack of selection pressure and movement of the goats may play a role in lack of differentiation in varied geographically separated populations. Furthermore, in Kenya, as well as other parts of Africa, goats are also used for religious and other cultural ceremonies such as payment of dowry and gifts (Herrero *et al*, 2013; Mbuku *et al*, 2015; Ogola *et al*, 2010). Therefore, some of the Alpine and Toggenburg breeds clustering together may be a result of movement of breed animals between the communities in those two regions due to the forementioned cultural ceremonies. As mentioned earlier, one of the criteria for selecting the goat keeping households in this study was them being members of a dairy goat farmer group association (DGAK, MGBA, NGBA). The associations are responsible for buck rotation among the group members, maintaining the purtity of the breed and providing technical backstoping. However, the results indicate a need to technically strengthen the Dairy Goat Association of Kenya for Alpine and the Meru Goat Breeders Association for Toggenburg, because urgent management efforts are essential to improve on breeding aspects, utilization and conservation of the various goat genetic resources.

All Saanen goats formed one cluster in the phylogenetic analysis. Indeed, the long distance (over 450 km) between the regions where Saanen and the rest of the breeds are kept may be the barrier to gene flow from other breeds. Through adaptive hitchhiking, natural selection can play an essential role in shaping this variability (Andolfatto, 2001). Therefore, the observed genetic divergence of Saanen from Alpine, Toggenburg and Galla breeds could have been contributed by random genetic drift and natural selection for adaptation to their environment/region.

Genetic uniqueness can be determined from the magnitude of genetic distances and phylogenetic relationships between populations if supporting indications such as genetic history, records of production, reproduction and on adaptation are lacking (Eding and Laval, 1999; Tosser-Klopp *et al*, 2014; Zheng *et al*, 2012). Embracing this principle with respect to the results of this study, Saanen and Galla breeds seem to be the most genetically distinct among the populations sampled, and can be categorised as important genetic resources. It will be interesting to enlarge this breed level investigation in later studies through addition of all Kenyan goat breeds to better appreciate the genetic relationship among them.

Conclusion

The study revealed clear divergence between some goat breeds, which provides a wide prospect on the current genetic diversity of goats in Kenya. This will be vital in planning breeding strategies for genetic resources that should be sustainably utilized and conserved. Of the breeds studied, Galla breed displayed isolation by distance and seemed to be at equilibrium under dispersal and genetic drift. This shows that stronger efforts of genetic conservation and sustainable management of its gene pool have been undertaken. However, further studies are required for the on-farm Galla population. The most admixed breeds were Alpine and Toggenburg. Therefore, there is need to technically strengthen the Dairy Goat Association of Kenya for urgent management efforts that are essential for genetic improvement, utilization and conservation of the various goat genetic resources. Additional studies on phenotypic similarities and performance evaluation of the breeds in this study could add value to the information generated from this study to form the basis for future genetic resource conservation and improvement of goat breeds in Kenya.

Acknowledgements

The authors wish to acknowledge the African Development Fund (AfDB), USAID and the Government of Kenya under the Kenya Climate Smart Agriculture Project (KCSAP CGS/CRGs-AD-2019) for their financial support to the first author. We sincerely thank Egerton University (Njoro, Kenya) for providing technical support to undertake the study, Kenya Agricultural and Livestock Research Organizations for granting the study leave for the first author, dairy goat farmers and Naivasha Sheep and Goat Breeding Station for making their animals available for sample collection.

Author contributions

RW, KN, TO and EI conceived the study, RW analyzed the data and drafted the manuscript. All authors read and approved the manuscript.

Conflict of interest statement

The authors declare that there is no conflict of interest with any organization concerning the material discussed in the manuscript.

References

Ahuya, C. O., Ojango, J. M. K., Mosi, R. O., Peacock, C. P., and Okeyo, A. M. (2009). Performance of Toggenburg dairy goats in smallholder production systems of the eastern highlands of Kenya. *Small Ruminant Research* 83(1-3), 7–13. doi: https://doi.org/10.1016/j.smallrumres.2008.11.012

Ahuya, C. O., Okeyo, A. M., and Murithi, F. M. (2006). Productivity of cross-bred goats under smallholder production systems in the Eastern highlands of Kenya. In Small stock in development: Proceedings of a workshop on enhancing the contribution of small livestock to the livelihoods of resource-poor communities, 7p.

Ajmone-Marsan, P., Colli, L., Han, J. L., Achilli, A., Lancioni, H., Joost, S., Crepaldi, P., Pilla, F., Stella, A., Taberlet, P., Boettcher, P., Negrini, R., and Lenstra, J. A. (2014). The characterization of goat genetic diversity: Towards a genomic approach. *Small Ruminant Research* 121(1), 58–72. doi: https://doi.org/10.1016/j.smallrumres.2014.06.010

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19(9), 1655–1664. doi: https://doi.org/10.1101/gr.094052.109

Andersson, L. and Georges, M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. *Nature Reviews Genetics* 5(3), 202–212. doi: https://doi.org/10.1038/nrg1294

Andolfatto, P. (2001). Adaptive hitchhiking effects on genome variability. *Current Opinion in Genetics & Development* 11(6), 635–641. doi: https://doi.org/10.1016/S0959-437X(00)00246-X

Anmarkrud, J. A., Kleven, O., Bachmann, L., and Lifjeld, J. T. (2008). Microsatellite evolution: Mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. *BMC Evolutionary Biology* 8(1), 138–138. doi: https://doi.org/10.1186/1471-2148-8-138

Aziz, M. A. (2010). Present status of the world goat populations and their productivity. *Lohmann Information* 45(2), 42–52. url: https://www.lohmann-information.com/content/l_i_45_artikel17.pdf.

Bett, R. C. (2009). Design and evaluation of breeding strategies for low input dairy goat production systems in Kenya.

Bett, R. C., Kosgey, I. S., Bebe, B. O., and Kahi, A. K. (2007). Genetic improvement of the Kenya Dual Purpose Goat: Influence of economic values and prospects for a practical breeding programme. *Tropical Science* 47(3), 105–119. doi: https://doi.org/10.1002/ts.204

Bett, R. C., Kosgey, I. S., Kahi, A. K., and Peters, K. J. (2011). Definition of breeding objectives and optimum crossbreeding levels for goats in the smallholder production systems. *Small Ruminant Research* 96(1), 16–24. doi: https://doi.org/10.1016/j.smallrumres.2010.11.008

Brito, L. F., Kijas, J. W., Ventura, R. V., Sargolzaei, M., Porto-Neto, L. R., Cánovas, A., Feng, Z., Jafarikia, M., and Schenkel, F. S. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics* 18(1), 229–229. doi: https://doi.org/10.1186/s12864-017-3610-0

Chenyambuga, S. W., Hanotte, O., Hirbo, J., Watts, P. C., Kemp, S. J., Kifaro, G. C., Gwakisa, P. S., Petersen, P. H., and Rege, J. E. O. (2004). Genetic Characterization of Indigenous Goats of Sub-saharan Africa Using Microsatellite DNA Markers. *Asian-Australasian Journal of Animal Sciences* 17(4), 445–452. doi: https://doi.org/10.5713/ajas.2004.445

Deshingkar, P., Farrington, J., Rao, L., Akter, S., Sharma, P., Freeman, H. A., and Reddy, J. (2008). Livestock and poverty reduction in India: findings from the ODI Livelihood Options Project. url: https://hdl.handle.net/10568/281.

Eding, J. H. and Laval, G. (1999). Measuring genetic uniqueness in livestock. In *Genebanks and the conservation of farm animal genetic resources,* ed. Oldenbroek, J. K., (Lelystad, the Netherlands: DLO-Institute for Animal Science and Health (ID-DLO), Instituut voor Veehouderij en Diergezondheid, Research Branch Zeist), 33-58.

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

FAO (2012). Phenotypic characterization of animal genetic resources volume 11 of *FAO Animal Production and Health Guidelines*. (Rome) . url: http://www.fao.org/3/i2686e/i2686e00.htm.

Garnier-Géré, P. and Chikhi, L. (2013). Population Subdivision, Hardy–Weinberg Equilibrium and the Wahlund Effect. In *eLS,* ed. John Wiley & Sons, L.

Gizaw, S., Komen, H., Hanote, O., van Arendonk, J., Kemp, S., Haile, A., Okeyo, A. M., and Dessie, T. (2011). Characterization and conservation of indigenous sheep genetic resources: A practical framework for developing countries. In ILRI Research Report, volume 27, ILRI, Nairobi, Kenya.

Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D., Negrini, R., Finlay, E. K., Jianlin, H., Groeneveld, E., and and, S. W. (2010). Genetic diversity in farm animals - a review. *Animal Genetics* 41, 6–31. doi: https://doi.org/10.1111/j.1365-2052.2010.02038.x

Herrero, M., Grace, D., Njuki, J., Johnson, N., Enahoro, D., Silvestri, S., and Rufino, M. C. (2013). The roles of livestock in developing countries. *Animal* 7, 3–18. doi: https://doi.org/10.1017/s1751731112001954

Hoshino, A. A., Bravo, J. P., Nobile, P. M., and Morelli, K. A. (2012). Microsatellites as tools for genetic diversity analysis. In *Genetic Diversity in Microorganisms,* ed. Caliskan, M.

Jarne, P. and Lagoda, P. J. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10), 424–429. doi: https://doi.org/10.1016/0169-5347(96)10049-5

Kohn, M. H., Murphy, W. J., Ostrander, E. A., and Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in Ecology & Evolution* 21(11), 629–637. doi: https://doi.org/10.1016/j.tree.2006.08.001

Kosgey, I. S. and Okeyo, A. M. (2007). Genetic improvement of small ruminants in low-input, smallholder production systems: Technical and infrastructural issues. *Small Ruminant Research* 70(1), 76–88. doi: https://doi.org/10.1016/j.smallrumres.2007.01.007

Krause, A. K. (2006). Breeding programmes for small ruminants in the tropics with special reference to the crossbreeding programme of the Dairy Goat Association of Kenya (DGAK). Ph.D. thesis, Humboldt University, Berlin, Germany.

Laval, G., SanCristobal, M., and Chevalet, C. (2002). Measuring genetic distances between breeds: use of some distances in various short term evolution models. *Genetics Selection Evolution* 34(4), 1–27. doi: https://doi.org/10.1186/1297-9686-34-4-481

Letunic, I. and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* 47(W1), W256–W259. doi: https://doi.org/10.1093/nar/gkz239

Mbuku, S. M., Okeyo, A. M., Kosgey, I. S., and Kahi, A. K. (2015). Optimum crossbreeding systems for goats in low-input livestock production system in Kenya. *Small Ruminant Research* 123(1), 55–61. doi: https://doi.org/10.1016/j.smallrumres.2014.10.001

Mburu, M., Mugendi, B., Makhoka, A., and Muhoho, S. (2014). Factors Affecting Kenya Alpine Dairy Goat Milk Production in Nyeri Region. *Journal of Food Research* 3(6), 160–160. doi: https://doi.org/10.5539/jfr.v3n6p160

McVean, G. (2009). A Genealogical Interpretation of Principal Components Analysis. *PLoS Genetics* 5(10), e1000686–e1000686. doi: https://doi.org/10.1371/journal.pgen.1000686

Nicoloso, L., Bomba, L., Colli, L., Negrini, R., Milanesi, M., Mazza, R., et al. (2015). Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* 47(1). doi: https://doi.org/10.1186/s12711-015-0140-6

Ogola, T. D. O., Nguyo, W. K., and Kosgey, I. S. (2010). Economic contribution and viability of dairy goats: implications for a breeding programme. *Tropical Animal Health and Production* 42(5), 875–885. doi: https://doi.org/10.1007/s11250-009-9501-x

Ojango, J. M. K., Audho, J., Oyieng, E., Recha, J., Okeyo, A. M., Kinyangi, J., and Muigai, A. W. T. (2016). System characteristics and management practices for small ruminant production in "Climate Smart Villages" of Kenya. *Animal Genetic Resources/Ressources génétiques animales/Recursos genéticos animales* 58, 101–110. doi: https://doi.org/10.1017/s2078633615000417

Onzima, R. B., Upadhyay, M. R., Mukiibi, R., Kanis, E., Groenen, M. A. M., and Crooijmans, R. P. M. A. (2018). Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal Genetics* 49(1), 59–70. doi: https://doi.org/10.1111/age.12631

Palian, B. and Racokzi, G. (1976). Sheep and goat development project, Kenya. Breeding and research - Technical report 3.

Peacock, C. (2005). Goats: Unlocking their potential for Africa's farmers. In Proceedings of the Seventh Conference of Ministers Responsible for Animal Resources, 23p.

Peacock, C. (2007). The goats model. A proven approach to reducing poverty among smallholder farmers in Africa by developing profitable goat enterprises and sustainable support services. url: https://www.farmafrica.org/downloads/resources/WP9%20The%20Goat%20Model.pdf.

Peacock, C., Ahuya, C. O., Ojango, J. M. K., and Okeyo, A. M. (2011). Practical crossbreeding for improved livelihoods in developing countries: The FARM Africa goat project. *Livestock Science* 136(1), 38–44. doi: https://doi.org/10.1016/j.livsci.2010.09.005

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., and Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* 81(3), 559–575. doi: https://doi.org/10.1086/519795

Qanbari, S. and Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livestock Science* 166, 133–143. doi: https://doi.org/10.1016/j.livsci.2014.05.003

Rege, J. and Okeyo, A. (2011). Improving our knowledge of tropical indigenous animal genetic resources. In Animal Genetics Training Resource Version 3 Training Module 2, ILRI, Nairobi, Kenya; Uppsala, Sweden.

Scarpa, R., Ruto, E. S., Kristjanson, P., Radeny, M., Drucker, A. G., and Rege, J. E. (2003). Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates. *Ecological Economics* 45(3), 409–426. doi: https://doi.org/10.1016/s0921-8009(03)00094-6

Shivairo, R. S., Matofari, J., Muleke, C. I., Migwi, P. K., and Lugairi, E. (2013). Production Challenges and Socio-Economic Impact of Dairy Goat Farming amongst Smallholder Farmers in Kenya. *Food Science and Quality Management* 17, 54–61.

Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., Donnadieu-Tonon, C., Eggen, A., Heuven, H. C. M., Jamli, S., Jiken, A. J., Klopp, C., Lawley, C. T., McEwan, J., Martin, P., Moreno, C. R., Mulsant, P., Nabihoudine, I., Pailhoux, E., Palhière, I., Rupp, R., Sarry, J., Sayre, B. L., Tircazes, A., Wang, J., Wang, W., and Zhang, W. (2014). Design and Characterization of a 52K SNP Chip for Goats. *PLoS ONE* 9(1), e86227–e86227. doi: https://doi.org/10.1371/journal.pone.0086227

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* 34(3), 275–305. doi: https://doi.org/10.1186/1297-9686-34-3-275

Visser, C., Lashmar, S. F., Marle-Köster, E. V., Poli, M. A., and Allain, D. (2016). Genetic Diversity and Population Structure in South African, French and Argentinian Angora Goats from Genome-Wide SNP Data. *PLOS ONE* 11(5), e0154353–e0154353. doi: https://doi.org/10.1371/journal.pone.0154353

Waineina, R. W., Okeno, T. O., Ngeno, K., and Ilatsia, E. D. (2019). Performance of Dairy Goat Breeds in different Production Systems in Kenya. In Proceedings of the Tanzania Society of Animal Production (TSAP), 42nd Scientific Conference, 23-25.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28(24), 3326–3328. doi: https://doi.org/10.1093/bioinformatics/bts606

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2021), 2 (3), 36–50
DOI: 10.46265/genresj.OBVW6791
https://www.genresj.org
ISSN: 2708-3764

Biodiversity assessment of African locust bean (*Parkia biglobosa)* accessions from Savanna and Forest zones of Nigeria as revealed by seed storage proteins and RAPD markers

Pamela E Akin-Idowu [*,a], **Ayodeji O Aduloju** [a], **Omolara I Akinyoola** [a], **Dorcas O Ibitoye** [b], **Uterdzua Orkpeh** [b], **Usifo G Adebo** [a,c] **and Yemisi O Olagunju** [a]

[a] *Biotechnology Unit, National Horticultural Research Institute, P.M.B. 5432, Jericho Reservation Area, Idi-Ishin, Ibadan, Nigeria*
[b] *Genetic Resources Unit, National Horticultural Research Institute, P.M.B. 5432, Jericho Reservation Area, Idi-Ishin, Ibadan, Nigeria*
[c] *School of Agriculture and Food, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Melbourne, Australia*

Abstract: Understanding the level and distribution of genetic diversity in African locust bean (*Parkia biglobosa*) would strengthen breeding and conservation programmes towards domestication and sustainable use of this species. Sixteen accessions of *P. biglobosa* were assessed for variability based on seed morphology, seed protein and DNA profiling. Significant variation in seed characteristics were observed across locations. Seed protein profiling by SDS-PAGE revealed homogeneity as most bands were found common in all accessions, indicating that the protein profiles are highly conserved. Protein profiling separated the 16 accessions into four major clusters at 0.93 similarity coefficient. Most accessions grouping into Cluster 1 had a similarity coefficient of close to 100% and were from the Derived Savanna suggesting the presence of duplicates. Accessions NH/2016/P14, NH/2016/P03 and NH/2016/P04 grouped into clusters II, III and IV; respectively. Sixteen RAPD markers generated a total of 256 bands of which 63.67% were polymorphic. Gene diversity ranged from 0.41 to 0.93 and Polymorphic Information Content (PIC) from 0.39 to 0.93. The RAPD-based dendrogram separated accessions into six groups at 0.68 similarity coefficient. Based on a polymorphic seed storage protein marker a genetically distinct accession NH/2016/P04 could be exploited for breeding purposes. The homogeneity of alleles and narrow genetic base as revealed by RAPD and SDS-PAGE analyses suggests possible loss of intraspecific genetic diversity. Thus, intensification of germplasm collections across the different agroecological zones and characterization using specific markers will give a better understanding of diversity of *P. biglobosa* in order to enhance selection towards conservation, breeding and sustainable utilization.

Keywords: African locust bean, genetic diversity, agroecological zone, RAPD, SDS-PAGE, *Parkia biglobosa*

Citation: Akin-Idowu, P. E., Aduloju, A. O., Akinyoola, O. I., Ibitoye, D. O., Orkpeh, U., Adebo, U. G., Olagunju, Y. O. (2021). **Biodiversity assessment of African locust bean (*Parkia biglobosa)* accessions from Savanna and Forest zones of Nigeria as revealed by seed storage proteins and RAPD markers**. *Genetic Resources* 2 (3), 36–50. doi: 10.46265/genresj.OBVW6791.

Introduction

In Sub-Saharan Africa several indigenous agroforestry systems exist containing woody species known as multipurpose trees. Africa locust bean [*Parkia biglobosa* (Jacq.) R. Br. ex G. Don] is a well-known indigenous

*Corresponding author: Pamela E Akin-Idowu (elohoidowu@hotmail.com)

Received: 04.05.2020 | Accepted: 11.06.2021 | Published online: 29.06.2021

agroforestry fruit tree species belonging to the subfamily *Mimosoideae* and family *Fabaceae* (Amusa *et al*, 2014; Houndonougbo *et al*, 2020). The species grows in multiple climatic zones and is widely distributed from Senegal and Guinea in West Africa to Uganda in Central Africa (Lompo *et al*, 2018). *Parkia biglobosa* is maintained in the parklands of Africa for non-timber forest foods (Houndonougbo *et al*, 2020; Leakey, 2012; Okoye *et al*, 2014). *P. biglobosa* is a valuable resource for the livelihoods of local people in Sub-Saharan Africa because of its multipurpose function as a source of food, nutrition, medicine and income (Matig *et al*, 2002; Nikiema, 2005; Awodoyin *et al*, 2015; Sankhon *et al*, 2014; Shao, 2002; Sina and Traore, 2002; Teklehaimanot, 2004). The fruit pulp and seeds are both suitable for human consumption. The mealy pulp from the species' fruits is a major source of energy and nutrients including carbohydrates, proteins, lipids, carotenoids, vitamins A, B, C, and oligoelements (Nyadanu *et al*, 2017). The seeds are an important source of plant protein and essential amino acids among rural communities who often have limited access to animal proteins due to high cost (Akin-Idowu *et al*, 2018; Alabi *et al*, 2005) and are also rich in energy value, saccharose, vitamin C, lipids, carbohydrates (Orwa *et al*, 2010) and bioactive components such as phenolic compounds which may contribute in health promoting properties (Dedehou *et al*, 2016; Okoye *et al*, 2014). *P. biglobosa* seeds are fermented to make a food condiment called Dawadawa (Iru) which is rich in dietary protein and serves as an important substitute for animal protein (Lamien *et al*, 2011; Nyadanu *et al*, 2017). The leaves, bark, roots and flowers are also used in the treatment of many diseases such as hypertension, wound healing and malaria (Dedehou *et al*, 2016; Ouedraogo, 1995).

In spite of the importance of this species in traditional agriculture, regular cultivation of the fruit tree has rarely gone beyond conservation of natural stands *in situ* (Hopkins, 1983; Oni *et al*, 1998). Several studies on *P. biglobosa* population structure have shown low regeneration rates and ageing of the stands which may lead to extinction (Padakale *et al*, 2015; Ræbild *et al*, 2012a). Overexploitation, late fruitification for most progenies and climate change are some of the factors suggested to explain the phenomenon (Boffa, 1999; Kwon-Ndung *et al*, 2009; Teklehaimanot, 2004). Improved tree management practices for *P. biglobosa*, such as establishment of new trees or protection of natural regeneration are currently not sufficiently promoted (Ræbild *et al*, 2012b). Thus, African locust bean remains undomesticated despite increasing demand for its use.

The conservation and sustainable use of genetic resources are critical to maintain tree resource availability, especially in the face of significant environmental changes, erosion of cultural heritages and declining cultivation and production activities of neglected indigenous species leading to genetic resources being threatened. The process of plant domestication and conservation requires characterization of genetic resources for identification of cultivars and effective utilization of germplasm.

Genome size variation has been identified as an associate of evolutionary divergence (Dobes *et al*, 2019). Variation in genome size and chromosome number becomes taxonomically significant if associated with some degrees of morphological and ecological differentiation (Murray, 2005). Sina (2006) studied genetic diversity among populations of *P. biglobosa* species in Burkina Faso using enzyme electrophoresis. Dobes *et al* (2019) tested for linkage of relative genome size variation with geography, leaf morphology and population genetic variation in 58 individuals from 15 populations covering most of the distribution of *P. biglobosa* species in Burkina Faso. Most of the variation was found within populations and there was no evidence from the karyological data for structured intraspecific taxonomic heterogeneity (Dobes *et al*, 2019).

In Nigeria, genetic diversity of *P. biglobosa* accessions have been evaluated based on morphological characterization of trees (Gbadamosi *et al*, 2005; Okunlola *et al*, 2011) and seedlings (Adesoye *et al*, 2013). However, these characters are known to be more influenced by the environment. Uyoh *et al* (2011) reported variation in banding pattern of proteins obtained from leaf samples of three accessions of *P. biglobosa* collected from locations within Cross River State, Nigeria. Adesoye and Apo (2015) reported low levels of genetic diversity when 34 accessions of *P. biglobosa* were screened using seed protein electrophoresis. Owing to the limited information on genetic variation and relationship of *P. biglobosa* accessions in Nigeria, characterization using a combination of morphological, biochemical and molecular markers is essential for species identification and genetic variability establishment. This will facilitate the domestication, conservation and management of its genetic resources towards sustainable use.

Morphological characterization has been used to reveal phylogenetic relationships among crop populations, but this has its limitations as it is influenced by environmental factors (Ferreira, 2006). Selection, according to genetic variability using biochemical and molecular markers, has proved advantageous compared with the use of phenotypic markers (Alghamdi *et al*, 2019).

Biochemical analysis, particularly electrophoresis of seed proteins has been effective in studying genetic diversity at intraspecific levels in several legumes including soybean and chickpea (Durán *et al*, 2005; Signor *et al*, 2005; Malik *et al*, 2009; Sammour *et al*, 2007a,b) and for cultivar identification (Krochko and Bewley, 2000; Mustafa and El-Kholy, 2008).

The use of polyacrylamide gel electrophoresis (PAGE) in fractionating seed proteins and evaluating genetic diversity among accessions of several legume species including groundnut, African yam bean, grass pea, buckwheat, bambara groundnut and underground vetches

has also proven effective (la Rosa and González, 2010; Javaid *et al*, 2004; Machuka, 2001; Przybylska *et al*, 2000).

Molecular markers are known as one of the best approaches to study genetic material as well as to assess genetic variation in crop gene pools (Badr, 2008; Mondini *et al*, 2009; Wang *et al*, 2016). A number of PCR based molecular markers such as Random Amplified Polymorphic DNA (RAPD), Inter Simple Sequence Repeats (ISSR) and Amplified Fragment Length Polymporhism (AFLP) have been used to assess genetic variability and cultivar identification in soybean (Barakat, 2004; Baránek *et al*, 2002; Chowdhury *et al*, 2001; El-Kholy, 2013) and other plant populations (Hoque and Hasan, 2012; Lin *et al*, 2009). RAPD is often used successfully to assess genetic diversity among species as it is fast, less technical and expensive and may reveal dominant molecular marker of good potential (Badr *et al*, 2012; Prasanthi *et al*, 2012).

This study aimed to estimate the genetic diversity of *P. biglobosa* accessions across different agro-ecological zones in Nigeria using SDS-PAGE and RAPD markers.

Materials and Methods

Plant material and morphological description

Seeds of *P. biglobosa* were obtained from 16 locations across seven states in Nigeria (Figure 1). Thirteen of the sixteen accessions studied were collected from locations in the Derived Savanna agroecological zone (AEZ); two were collected from the Guinea Savanna and one from Humid Forest (Figure 1,Table 1). The sampling of *P. biglobosa* accessions was not conducted systematically across all AEZs because the species is reported to be widely adapted and naturally distributed throughout the varied savannas of Sub-Saharan Africa (Lompo *et al*, 2016). The seed morphological characteristics of each collected accession, including seed shape, seed coat color and seed coat texture were observed. One hundred (100) seed weight in grams of each accession was also recorded.

Seed preparation and total protein extraction

Dried seeds were milled into flour and defatted with n-hexane at a flour:hexane ratio of 1:10 (w/v) prior to protein extraction (de la Rosa *et al*, 1992).

Seed total protein extraction was carried out by suspending defatted flour samples (100 mg) in 2.0 mL of pre-chilled Tris buffer 50mM Tris-HCl, pH 7.6, 1mM DDT, 150 mM NaCl and 1mM EDTA for 2h at room temperature according to the method of Zarkadas *et al* (2007). The resulting homogenates were centrifuged at 10,000 g for 20 min and supernatants were stored at -20°C until used.

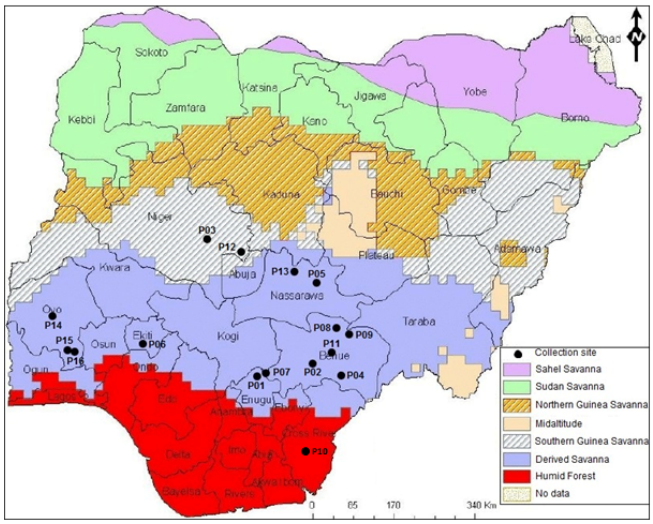

Figure 1. Map of Nigeria showing agroecological zones of the collection sites of sixteen *Parkia biglobosa* accessions

One dimensional SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the procedure of Laemmli (1970). Protein extracts were diluted in a sample buffer that contained 60 mM Tris-HCl (pH 6.8), 2% w/v SDS, 3.33% v/v β-mercaptoethanol, 10% glycerol and 0.05% of bromophenol blue. The samples were heated at 98°C for 5 min before loading onto a vertical slab gel (Mini-PROTEAN II Electrophoresis Cell) using a 4% (w/v) stacking and 12% gradient separating (acrylamide/bis acrylamide) gel containing 0.1% SDS. The separating polypeptide bands were calibrated using a protein molecular weight marker which ranged from 11 to 130 kDa. Electrophoresis was done at 125 V for about 2 h. At the end of the run, the gels were stained with Coomassie brilliant blue R-250 (SigmaAldrich) in methanol/water/acetic acid (40:50:10) and destained in a solution containing methanol/water/acetic acid (50:40:10).

Genomic DNA extraction, RAPD amplification and electrophoresis

For the genetic analysis, ten viable seeds of each of the sixteen accessions of *P. biglobosa* collected from the different locations were grown in a greenhouse. Prior to sowing, the seeds were immersed in concentrated sulphuric acid (H_2SO_4) for 15 minutes to break seed dormancy, rinsed thoroughly in running water and air dried at room temperature. Total genomic DNA was extracted from young leaves of *P. biglobosa* accessions. About 0.2g of a composite of 5 samples per accession was ground in extraction buffer following the cetyltrimethyl ammonium bromide (CTAB) method of Dellaporta *et al* (1983).

The DNA quality was evaluated using 1% agarose gel electrophoresis with known concentrations of undigested lambda DNA (Sigma, St Louis, MO, USA). Quantification of DNA was done using a spectrophotometer

Table 1. *Parkia biglobosa* accessions used in this study and collection sites.

Accession code	Collection site	Agroecological Zone	Geographical Coordinates (DD)	
			Longitude	Latitude
NH/2016/P01	Obollo-Afor, Enugu State	Derived Savanna	6.9153°N	7.5139°E
NH/2016/P02	Oju, Benue State	Derived Savanna	7.0383°N	8.3451°E
NH/2016/P03	Minna, Niger State	Guinea Savanna	9.5836°N	6.5463°E
NH/2016/P04	Vandeikya, Benue State	Derived Savanna	6.7835°N	9.0695°E
NH/2016/P05	Agyaragu, Nasarawa State	Derived Savanna	8.3471°N	8.5679°E
NH/2016/P06	Ikere, Ekiti State	Derived Savanna	7.5000°N	5.2333°E
NH/2016/P07	Orokam, Benue State	Derived Savanna	6.9786°N	7.5572°E
NH/2016/P08	Gbajimba, Benue State	Derived Savanna	7.8197°N	8.8587°E
NH/2016/P09	Kwatan-Sule, Benue State	Derived Savanna	7.8112°N	9.2429°E
NH/2016/P10	Ogoja, Cross River State	Humid Forest	6.6548°N	8.7977°E
NH/2016/P11	Mbatiav, Gboko, Benue State	Derived Savanna	7.2788°N	8.7953°E
NH/2016/P12	Suleja, Niger State	Guinea Savanna	9.2003°N	7.1723°E
NH/2016/P13	Garaku, Nassarawa State	Derived Savanna	8.8444°N	8.1311°E
NH/2016/P14	Iseyin, Oyo State	Derived Savanna	7.9765°N	3.5914°E
NH/2016/P15	NIHORT, Ibadan, Oyo State	Derived Savanna	7.4052°N	3.8499°E
NH/2016/P16	Oje Market, Ibadan, Oyo State	Derived Savanna	7.3891°N	3.9092°E

(Beckman Coulter DU530) at 260 nm. Extracts were diluted to obtain DNA concentrations of 25ng/μl. To generate DNA profiles, 40 decamer oligonucleotide DNA primers were initially screened for polymorphisms and only 16 which were polymorphic were used in the PCR reactions (Table 2).

To ensure reproducibility of RAPD markers, RAPD-PCR amplification was performed in triplicate for each primer and accession. The PCR reaction contained the following reagents with the final concentrations: 1x PCR buffer, dNTPs (0.2 mM), $MgCl_2$ (2.0 mM), primer (25 pmol), template DNA (25ng/μl), Taq polymerase (1.25 U) and nuclease-free water to 50μl. Amplification was carried out in a 2400 Perkin Elmer Gene Amp PCR thermo-cycler as follows: pre-denaturation for 3 min at 94°C; then 40 cycles each consisting of a denaturation step for 1 min at 94°C; an annealing step for 1 min at 36°C; an extension step for 90 sec at 72°C. Amplification was terminated by a final extension of 7 min at 72°C. PCR products were electrophoresed on a 1.5% (w/v) agarose gel in 1X Tris/Borate/EDTA (TBE) buffer at 100 volts for 2h and visualized under UV light after staining with ethidium bromide.

Data analysis

Data on 100-seed weight of all accessions were subjected to ANOVA and means were separated using Duncan's multiple range tests ($p \leq 0.05$). Distribution frequencies of the qualitative traits were calculated using Microsoft Excel 2010. The protein bands and RAPD amplified DNA fragments were scored qualitatively, whereby (0) stands for the absence and (1) stands for the presence of bands in the profile of each accession. All values were pooled together to generate a binary data matrix, which was analyzed using the Numerical Taxonomy and Multivariate Analysis System (Rohlf, 2002). Genetic relationship among accessions was evaluated based on Jaccard's similarity index (Jaccard, 1908). Dendrograms were generated by using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA). Analysis of genetic diversity (GD) was calculated according to the method of Nei (1987). Polymorphic information content (PIC) for each primer was determined from allele frequencies (Nei and Li, 1979).

Results and Discussion

Variation in seed morphology

Improvement of the quality of indigenous fruit trees across Sub-Saharan Africa involves assessment of fruit/seed traits; this has shown significant variation both among and within provenances and has enhanced selection in breeding programmes (Ræbild *et al*, 2012a). In this study, significant differences ($p \leq 0.05$) were observed in the 100-seed weight of *P. biglobosa* collected from 16 different locations (Table 3), which ranged from 17.93g (NH/2016/P09) to 26.11g (NH/2016/P05). The highest 100-seed weight was recorded for accession NH/2016/P05 collected from Agyaragu, Nassarawa State; this was significantly ($p \leq 0.05$) higher than that of all other accessions. The lowest 100-seed weight was recorded for NH/2016/P09 collected from Kwatan-Sule, Benue State. The mean 100-seed weight (21.46g) of all accessions of *P. biglobosa* collected from the Derived Savanna zone was higher than the mean 100-seed weight collected from the Guinea Savanna and Humid Forest (Figure 2a). This variation in seed weight can be attributed to a wide range of factors such as environment (soil, climate), anthropogenic activities and phenotypic plasticity, thus promoting the presence of ecotypes.

Table 2. List of primers used and RAPD polymorphisms among *Parkia biglobosa* accessions.

Primer	Sequence 5' to 3'	Major allele frequency	Allele number	Gene Diversity	Polymorphic Information content (PIC)
OPB10	CTGCTGGGAC	0.13	15	0.93	0.93
OPT07	GGCAGGCTGT	0.13	14	0.92	0.92
OPB04	GGACTGGAGT	0.13	14	0.92	0.92
OPT12	GGGTGTGTAG	0.13	14	0.92	0.92
OPT08	AACGGCGACA	0.13	13	0.91	0.91
OPT06	CAAGGGCAGA	0.19	13	0.91	0.90
OPH07	CTGCATCGTG	0.19	12	0.90	0.89
OPB08	GTCCACACGG	0.19	9	0.87	0.85
OPT20	GACCAATGCC	0.25	9	0.84	0.82
OPH06	ACGCATCGCA	0.38	9	0.80	0.78
OPT09	CACCCCTGAG	0.38	9	0.80	0.78
OPT10	CCTTCGGAAG	0.44	10	0.77	0.76
OPT01	GGGCCACTCA	0.44	7	0.74	0.71
OPT16	GGTGAACGCT	0.56	6	0.64	0.61
OPB05	TGCGCCCTTC	0.56	5	0.63	0.59
OPT05	GGGTTTGGCA	0.75	4	0.41	0.39
Mean		0.31	10	0.81	0.79
Total			163		

Leakey (2017) reported that fruit and/or kernel size was greater in more humid sites for *Adansonia digitata* and *Vitellaria paradoxa*; but greater in drier sites for *Balanite aegyptiaca*. Also, results from tests of *B. aegyptiaca* and *Prosopis Africana* indicated that provenances from drier sites had significantly better aboveground growth than provenances from more humid sites (Leakey, 2017). It is recommended that fruit trees germplasm should be collected in drier sites for future plantings in parklands, especially as this germplasm appears to be better adapted to dry conditions (Leakey, 2017).

The seeds of *P. biglobosa* accessions collected from the different locations in this study had distinct morphology and exhibited variation for seed shape, seed coat colour and seed coat texture (Table 3). Seeds of the sixteen *P. biglobosa* accessions were predominantly round oval in shape, black in colour with smooth seed coat and were mostly obtained from the Derived Savanna zone (Figure 2B, C, D). Seeds collected from the two sites in Guinea Savanna differed in shape and texture as seeds from one site were flat oval, black and smooth (NH/2016/P03) (Figure 3B); while seeds from the other site were round oval, black and rough (NH/2016/P12) (Figure 3C). Seeds collected from the same agroecological zone (Derived Savanna) exhibited distinct variation in colour and texture, NH/2016/P15 seeds were flat oval, brown and smooth (Figure 3E); while seeds of NH/2016/P04 were flat oval, black and wrinkled.

The morphological variation may be due to gene flow across regions because of seed dispersal since *P. biglobosa* is an open pollinated tree crop having main pollinators such as bat, honeybees, rodents and humans (Lassen *et al*, 2012; Lassen, 2016; Lompo *et al*, 2017). This variation offers great potential for future selection and domestication.

Lack of a clear association between genome size and morphological, geographical, or ecological patterns of differentiation has been reported for two species of *Artemisia* (Asteraceae) (Dobes *et al*, 2019; de Xaxars *et al*, 2016).

Seed storage protein analysis

Seed storage protein profiling has been used as an effective tool for varietal identification and determination of phylogenetic relationship in several plant species populations (Emre, 2009; Hameed *et al*, 2012; Machuka, 2001).

In this study, the SDS-PAGE profile of sixteen accessions of *P. biglobosa* showed similar electrophoretic patterns with respect to the number of protein bands and their band intensities (Figure 4). The uniformity in the protein band profiles of all sixteen accessions suggests a low level of genetic diversity, which also corroborates earlier studies on *P. biglobosa* using seed protein electrophoresis (Adesoye and Apo, 2015).

Occurrence of a narrow genetic base from the accessions of *P. biglobosa* would suggest that they resulted not only from a common gene pool, but also from the outcrossing nature of the species over a long time, leading to a low level of inter-population diversity. Low variation was observed in thirteen mungbean varieties using SDS-PAGE (Hameed *et al*, 2012) and also in genetic diversity assessment of groundnut using SDS-PAGE (Javaid *et al*, 2004).

Table 3. Seed characteristics of *Parkia biglobosa* accessions used in this study. Values are mean ± Standard Error. Means within the same column with different letters are significantly ($p \leq 0.05$) different.

Accession code	Weight of 100 seeds (g) Mean ± SE	Shape	Colour	Seed coat texture
NH/2016/P01	$23.11^{d} \pm 0.18$	Round Oval	Black	Smooth
NH/2016/P02	$18.59^{k} \pm 0.16$	Round Oval	Brownish black	Smooth
NH/2016/P03	$21.75^{f} \pm 0.15$	Flat Oval	Black	Smooth
NH/2016/P04	$18.30^{k} \pm 0.32$	Flat Oval	Black	Wrinkle
NH/2016/P05	$26.11^{a} \pm 0.21$	Round Oval	Black	Smooth
NH/2016/P06	$23.38^{d} \pm 0.10$	Flat Oval	Brownish black	Smooth
NH/2016/P07	$19.44^{i} \pm 0.14$	Round Oval	Black	Smooth
NH/2016/P08	$20.87^{g} \pm 0.19$	Round Oval	Black	Rough
NH/2016/P09	$17.93^{l} \pm 0.23$	Round Oval	Brownish black	Smooth
NH/2016/P10	$18.54^{k} \pm 0.13$	Round Oval	Black	Smooth
NH/2016/P11	$21.46^{f} \pm 0.06$	Round Oval	Black	Smooth
NH/2016/P12	$20.32^{h} \pm 0.26$	Round Oval	Black	Rough
NH/2016/P13	$23.74^{c} \pm 0.23$	Round Oval	Black	Rough
NH/2016/P14	$19.00^{j} \pm 0.12$	Flat Oval	Brownish black	Rough
NH/2016/P15	$24.59^{b} \pm 0.11$	Flat Oval	Brown	Smooth
NH/2016/P16	$22.48^{e} \pm 0.13$	Flat Oval	Brownish black	Smooth

The number of clearly visible protein bands ranged from eight to nine among the sixteen accessions of *P. biglobosa* with molecular weights ranging from 11 to 130 kDa. Two major clusters of bands were observed between 11−17 kDa and 34−53 kDa and one minor band was observed above 130 kDa (Figure 4). This is similar to results of Adesoye and Apo (2015) on albumin and globulin fractions of *P. biglobosa* in which two major bands were observed at the 16 and 50 kDa regions.

It has been suggested that high molecular weight proteins facilitate the development of seed hardness (Coelho *et al*, 2007); therefore, the presence of a high molecular weight band observed above the 130 kDa in this study may account for the seed hardness of *P. biglobosa*. The seeds of *P. biglobosa* are very hard and require treatment with sulphuric acid to induce germination.

A polymorphic polypeptide with molecular weight of approximately 24 kDa was absent in NH/2016/P04 but present in all other accessions (Figure 4A) and can be used for its identification. Differences in the presence or absence of bands among accessions are indicative of differences in the genes controlling the different polypeptide subunits (Osanyinpeju and Odeigah, 1998). Also, the presence or absence of polypeptide bands has been found to be linked to the expression of certain characteristics such as wrinkled seeds and insect resistance (Rao and Pernolett, 1981; Odeigah and Osanyinpeju, 1996). The seed shape of NH/2016/P04 was observed to be wrinkled and the absence of the 24 kDa band in this accession may be responsible for this seed texture as it was the only accession expressing this characteristic. Seed storage proteins are non-enzymatic and have the sole purpose of providing proteins (nitrogen and sulphur source) required during germination and establishment of a new plant (Hameed *et al*, 2012). Legume seed storage proteins are mostly 7S and 11S globulins, which tend to be deficient in sulphur-containing amino acids (Tang and Sun, 2010). In mung bean, SDS-PAGE revealed that 11S globulin was composed of two bands: 40 kDa and 24 kDa, 8S vicilin was composed of 60 kDa, 48 kDa, 32 kDa and 26 kDa bands; and basic 7S globulin was composed of 28 kDa, 17 kDa and 16 kDa bands (Hameed *et al*, 2012; Mendoza *et al*, 2001).

In this study, peptides with molecular weights of 40 kDa and 24 kDa were detected that may correspond to 11S globulin, peptides with molecular weights of 48 kDa and 32 kDa were detected that may correspond to the 8S vicilin subunit, while 28 kDa and 16 kDa peptides were also detected that may correspond to the 7S globulin subunit. Barakat (2004) reported low levels of protein polymorphism among six cultivars of soybean in which five major bands were identified at the 72, 36, 32, 20 and 16 kDa regions.

In African yam bean, SDS-PAGE separated identical globulin, albumin and vicilin patterns for all 26 accessions studied (Machuka, 2001). Strelec *et al* (2012) reported that barley varieties could be partially discriminated by albumin/globulin banding patterns using SDS-PAGE, whereas this was not possible with native PAGE which showed more or less identical protein patterns for all varieties.

Through the use of SDS-PAGE on total proteins it is possible to detect a useful band polymorphism to explore the diversity of *P. biglobosa* as was detected in this study, albeit at a low level.

Cluster analysis

A dendrogram was constructed based on protein banding patterns using Jaccard's similarity coefficients

(Table 4). The dendrogram grouped the sixteen accessions into four clusters at 0.93 similarity coefficient (Figure 5). The first cluster grouped together the largest number (thirteen) of accessions which consisted of NH/2016/P01, NH/2016/P02, NH/2016/P05, NH/2016/P16, NH/2016/P15, NH/2016/P06, NH/2016/P13, NH/2016/P12, NH/2016/P11, NH/2016/P10, NH/2016/P09, NH/2016/P08 and NH/2016/P07. The majority of the members of this cluster had similarity coefficient of 1.00 suggesting them to be genetically identical. Cluster II consisted of NH/2016/P14. Accession NH/2016/P03, which was collected in Minna, Niger state and belonging to the Guinea Savanna agroecological zone, was separated alone into Cluster III, displaying pattern of geographic origin. Cluster IV consisted of NH/2016/P04 and was the most divergent of the sixteen tested accessions. The genetic similarity coefficient was 0.85 among NH/2016/P03, NH/2016/P04 and NH/2016/P14. Clustering showed pattern of geographic origin as eleven of the thirteen accessions grouped in Cluster I were collected from the Derived Savanna agro-ecological zone.

Genetic distance values give some idea of the level of genetic variability among selected species or accessions. The variation in genetic diversity among the sixteen *P. biglobosa* accessions ranged from 0.85

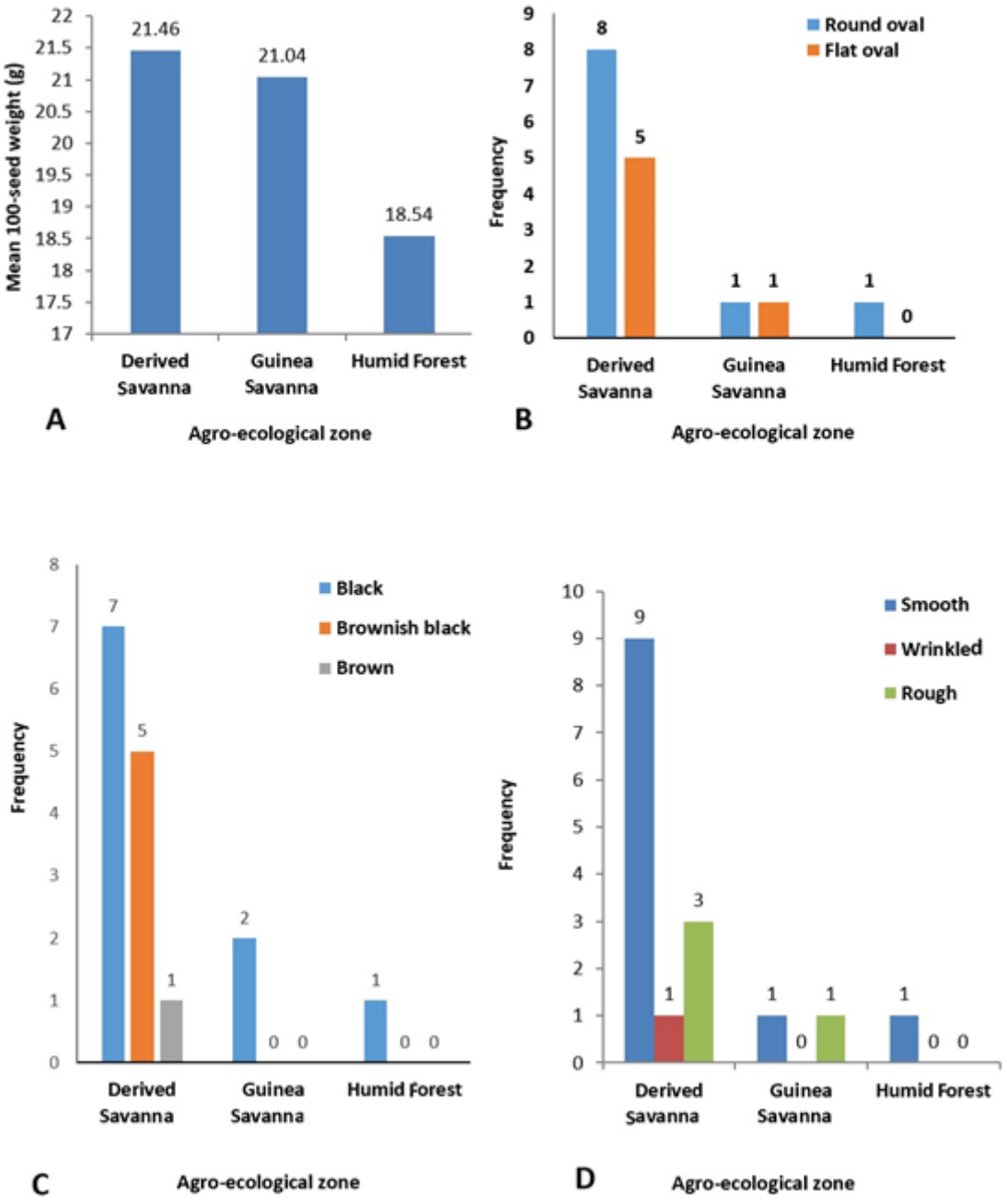

Figure 2. Distribution frequencies of seed characteristics of *Parkia biglobosa* across agro-ecological zones. A) 100-seed weight (g); B) Seed shape; C) Seedcolour; D) Seed coat texture.

Figure 3. Shapes and colour of seeds with seed coat of *Parkia biglobosa*: A) Round oval shape (NH/2016/P02), B) Flat oval shape (NH/2016/P03), C) Black colour (NH/2016/P12), D) Brownish black colour (NH/2016/P14), E) Brown colour (NH/2016/P15).

to 0.92 depicting a low genetic diversity (Table 4). All accessions in Cluster I had a similarity coefficient of 0.92 in any pairwise comparisons with NH/2016/P03, NH/2016/P04 and NH/2016/P14. The low levels of variability observed in *P. biglobosa* is congruent with Sina (2006) who reported that the genetic distances among *P. biglobosa* trees in Burkina Faso were low (between 0 and 0.240), indicating that the populations were similar enough to belong to the same genetic group. Occurrence of a narrow genetic base from the accessions of *P. biglobosa* suggests that they may have resulted from a common gene pool and that fixation of variants through genetic drift has probably occurred.

Amusa *et al* (2014) had earlier reported that close natural populations like *P. biglobosa* exhibited high genetic similarity and low genetic distance attributed to a high rate of exchange of gene flow between populations. Low genetic diversity was also observed in mung bean germplasm based on electrophoresis of seed storage proteins where a dendrogram analysis grouped the tested genotypes into three clusters at 93% homology (Hameed *et al*, 2012). Results of this study is similar to that obtained in cowpea in which seven landraces were separated into different clusters based on their geographical origin (Alghamdi *et al*, 2019). The high level of similarity in the accessions of *P. biglobosa* suggests that the seed storage proteins are highly conserved.

RAPD polymorphism

Genetic polymorphism is an indication of evolutionary adaptation, which has a main role in the survival of species in a changing environment (Stevens *et al*, 2007). In this study, Random Amplified Polymorphic DNA (RAPD) amplification profiles revealed a total of

Table 4. Jaccard's similarity coefficients of seed total protein profiles of *Parkia biglobosa* accessions.

	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15	P16
P01	1.000															
P02	1.000	1.000														
P03	0.923	0.923	1.000													
P04	0.923	0.923	0.846	1.000												
P05	1.000	1.000	0.923	0.923	1.000											
P06	1.000	1.000	0.923	0.923	1.000	1.000										
P07	1.000	1.000	0.923	0.923	1.000	1.000	1.000									
P08	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000								
P09	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000							
P10	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000						
P11	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000	1.000					
P12	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000				
P13	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000			
P14	0.923	0.923	0.846	0.846	0.923	0.923	0.923	0.923	0.923	0.923	0.923	0.923	0.923	1.000		
P15	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.923	1.000	
P16	1.000	1.000	0.923	0.923	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.923	1.000	1.000

256 alleles out of which 163 were polymorphic showing 63.67% polymorphism among the studied accessions. This relatively high polymorphism suggests a significant amount of genetic diversity among the accessions. A similar result was obtained among 11 soybean cultivars in which RAPD amplification profiles revealed 61.7% polymorphism using five primers (El-Kholy, 2013).

Alleles, ranging in size from 250 to 1500 basepairs (bp) and above, were scored for estimation of genetic relationship among the sixteen *P. biglobosa* accessions (Figure 6). An average of 10 alleles per primer was obtained in this study ranging from a minimum of four alleles using primer OPT05 to a maximum of 15 alleles using primer OPB10 (Table 2).

Genetic diversity is one of the important indices for the evaluation of genetic diversity among crop species (Narzary *et al*, 2009; Tamboli *et al*, 2016). The genetic diversity in this study ranged from 0.41 to 0.93, suggesting a significant amount of variation among the 16 accessions evaluated (Table 2). An earlier report on genetic diversity assessment of 23 accessions of *P. biglobosa* from different agro-ecological zones using RAPD indicated weak genetic diversity (expected heterozygosity, H_E = 0.05–0.18 and observed number of alleles, ONA = 1.11–1.65) (Amusa *et al*, 2014). The significance of the polymorphic information content

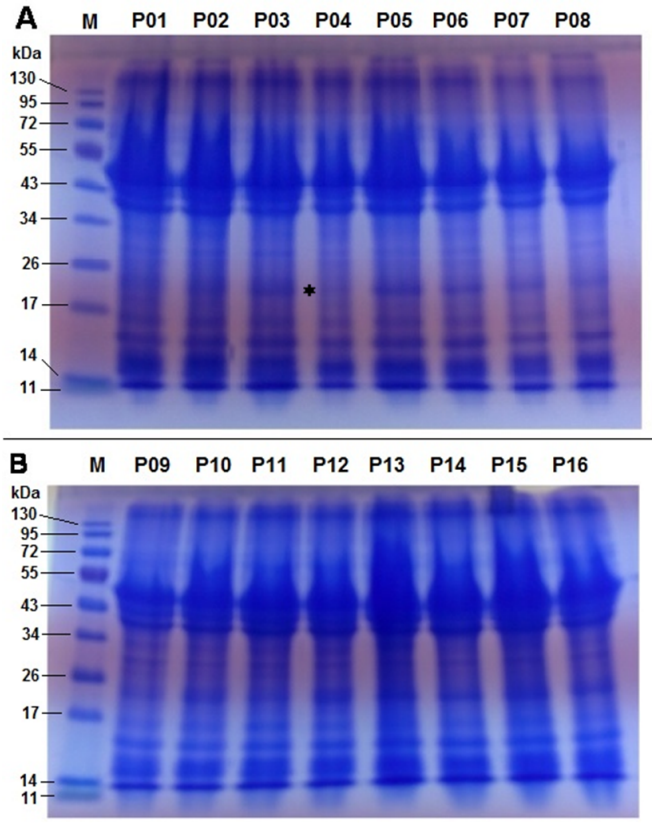

Figure 4. SDS-PAGE of *P. biglobosa* seed total proteins. A) Accessions NH/2016/P01 to NH/2016/08; B) Accessions NH/2016/P09 to NH/2016/P16. M stands for protein size marker. * indicates the missing 24 kDa band in NH/2016/P04.

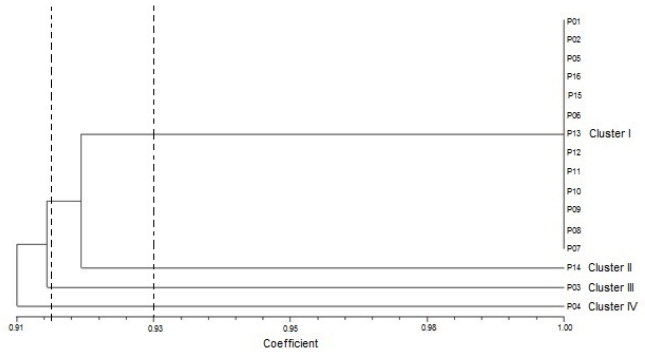

Figure 5. UPGMA dendrogram based on Jaccard's similarity of seed total protein profiles showing close relationship among *Parkia biglobosa* accessions.

(PIC) value is such that they are used to evaluate the amount of genetic diversity and are categorized as high (PIC > 0.05), medium (PIC < 0.05) and low (PIC < 0.25) (Abedian *et al*, 2012). Polymorphic information content ranged from 0.39 for primer OPT05 to 0.93 for primer OPB10 (Table 2).

Fifteen of the sixteen polymorphic primers had PIC values greater than 0.5 indicating that RAPD can develop high-locus polymorphism, which is useful to assess genetic variability of the accessions. All primers showed an average PIC value of 0.79 and nine out of the sixteen primers exceeded the average. This relatively high polymorphism has been observed in studies of genetic diversity on soybean cultivars by Chowdhury *et al* (2001) and Barakat (2004).

Cluster analysis based on UPGMA revealed six distinct clusters at a genetic similarity coefficient of 0.68 on the UPGMA dendrogram (Figure 7). Cluster I consisted of accessions NH/2016/P01, NH/2016/P05, NH/2016/P12, NH/2016/P13, NH/2016/P04, NH/2016/P02, NH/2016/P09, NH/2016/P14 and NH/2016/P15. All the accessions grouped together in Cluster 1 were collected from Derived Savanna except accession NH/2016/P12 which was collected from Guinea Savanna. This suggests that gene flow between the different locations may have occurred as a result of movement of germplasm and exchange of genetic material among farmers and markets. Cluster II consisted of NH/2016/P06; Cluster III consisted of NH/2016/P07 while Cluster IV consisted of NH/2016/P03. Accessions NH/2016/P08, NH/2016/P16 and NH/2016/P11 grouping together in Cluster V were collected from Derived Savanna. Accession NH/2016/P10 collected from Ogoja, Cross River State in the Humid Forest zone grouped alone in Cluster VI. Thus, in the UPGMA analysis of the RAPD profiles, accessions from same agroecological zone grouped together, while in other cases they were placed in different clusters. For most accessions in this study there was similar clustering pattern of geographically closer accessions, indicating significant association between genetic similarity and geographical distance. This does not agree with a report of Adesoye *et al* (2013) who reported that distribution of *P. biglobosa* genotypes among clusters did not correspond with their geographical patterns. The lowest similarity was recorded between NH/2016/P10 and NH/2016/P06, NH/2016/P07, NH/2016/P03. The highest similarity was recorded between accessions NH/2016/P12 and NH/2016/P13, followed by accessions NH/2016/P05 and both NH/2016/P12, NH/2016/P13. Accessions NH/2016/P05 and NH/2016/P13 with narrow genetic distance were collected from the same agro-ecological zone (Derived Savanna). This result corroborates findings of Hamrick and Godt (1989) in which species with small geographic ranges maintained less genetic diversity than geographically widespread species. RAPD clearly distinguished some accessions by accession specific fragments (Figure 6). Therefore, RAPD markers can successfully be used to produce variety specific fingerprints in *P. biglobosa* accessions and are a valuable tool for assessing genetic diversity.

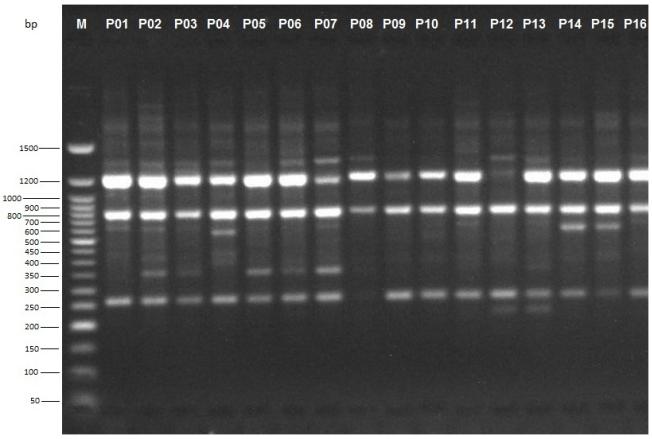

Figure 6. RAPD-PCR analysis of *Parkia biglobosa* accessions as revealed by primer OPB10. Accessions are numbered as in Table 1. M is the molecular size marker.

Conclusion

The low level of variation observed among the accessions as revealed by SDS-PAGE limits the method's applicability to distinguish accessions. Nevertheless, SDS-PAGE was able to detect a polymorphic polypeptide with molecular weight of approximately 24 kDa which was present in all accessions but absent in NH/2016/P04. Thus, SDS-PAGE may not be used to identify accessions based on intraspecific variation;

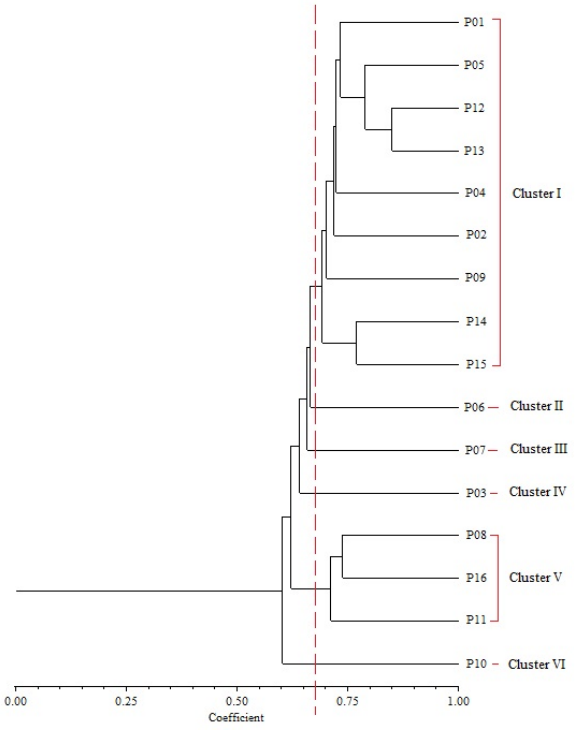

Figure 7. UPGMA dendrogram based on RAPD analysis of *Parkia biglobosa* accessions.

rather it might be more suitable to identify interspecific variation.

RAPD-PCR markers were more informative and could be used to discriminate between *P. biglobosa* accessions. Four RAPD primers (OPB10, OPT07, OPB04, OPT12) revealed high degrees of polymorphism among the accessions. The identification of accessions with maximum genetic divergence should be optimized in breeding programmes. The RAPD marker is therefore a more efficient tool to discriminate between accessions of *P. biglobosa*.

Implications of findings to the conservation and sustainable use of *Parkia biglobosa* genetic resources

- The variation observed in qualitative traits and 100-seed weight among the sixteen accessions of *P. biglobosa* across the three agro-ecological zones of Nigeria indicates their potential for domestication through selection towards conservation and breeding.
- Cluster analysis of the seed protein profile identified NH/2016/P04 as genetically distinct as it was clustered alone and was the most divergent of the sixteen tested accessions. The absence of the 24kDa band in accession NH/2016/P04 distinguishes it from all the other accessions. It can be adopted as parental line for heterosis crossing and should be exploited for conservation, breeding and sustainable utilization.
- Four selected RAPD primers (OPB10, OPT07, OPB04 and OPT12) can be used to discriminate the accessions of *P. biglobosa*.
- The highly diversified accessions of *P. biglobosa* indicate potential for domestication through selection in breeding programmes.
- Most accessions from the same agroecological zone clustered together suggesting that there is a strong correlation between genetic similarity and geographic proximity.
- The homogeneity of alleles among the studied *P. biglobosa* accessions suggests possible loss of intraspecific genetic diversity. The weakening gene pool and diversity observed can be enhanced through more germplasm collections, particularly from the diverse agroecological zones of Nigeria for better genetic characterization using more specific markers towards conservation, breeding and sustainable utilization.

Acknowledgements

Authors are grateful to the Virology Unit of the International Institute of Tropical Agriculture (IITA) Ibadan, Nigeria, for providing the facilities to carry out the SDS-PAGE analysis.

Author contributions

P. E. Akin-Idowu designed the experiment. P. E. Akin-Idowu, A. O. Aduloju and O. I. Akinyoola executed the experiment. O. I. Akinyoola , P. E. Akin-Idowu and D. O. Ibitoye performed the analysis. P. E. Akin-Idowu, O. I. Akinyoola, U. G. Adebo, U. Orkpeh, Y. O. Olagunju and D. O. Ibitoye wrote the manuscript. All the authors read and approved its submission.

Conflict of interest statement

The authors declare no conflict of interest.

References

Abedian, M., Talebi, M., Golmohammdi, H. R., and Sayed-Tabatabaei, B. E. (2012). Genetic diversity and population structure of mahaleb cherry (*Prunus mahaleb* L.) and sweet cherry (*Prunus avium* L.) using SRAP markers. *Biochemical Systematics and Ecology* 40, 112–117. doi: https://doi.org/10.1016/j.bse.2011.10.005

Adesoye, A. I. and Apo, K. A. (2015). Evaluating genetic biodiversity of *Parkia biglobosa* (Jacq.) R. Br. Ex Don (African locust bean) accessions from Nigeria using seed protein electrophoresis. *Journal of Crop Science and Biotechnology* 18(3), 171–180. doi: https://doi.org/10.1007/s12892-014-0119-9

Adesoye, A. I., Ogunremi, C. O., and Aina, O. O. (2013). Genetic variation and heritability of seedling traits in Africa locust bean *Parkia biglobosa* (Jacq.) R. Br. ex G. Don. *Legume Research* 36, 976–0571.

Akin-Idowu, P. E., Egbekunle, K. O., Akinyemi, S. O. S., and Orkpeh, U. (2018). Characterization of the seed proteins of African locust bean (*Parkia biglobosa*), African breadfruit (*Treculia africana*) and soybean (*Glycine max*). *Acta Horticulturae* 1225(1225), 419–426. doi: https://doi.org/10.17660/actahortic.2018.1225.59

Alabi, D. A., Akinsulire, O. R., and Sanyaolu, M. A. (2005). Qualitative determination of chemical and nutritional composition of *Parkia biglobosa* (Jacq.) Benth. *Afri J Biotechnology* 4(8), 812–815.

Alghamdi, S. S., Khan, M. A., Migdadi, H. M., El-Harty, E. H., Afzal, M., and Farooq, M. (2019). Biochemical and molecular characterization of cowpea landraces using seed storage proteins and SRAP marker patterns. *Saudi Journal of Biological Sciences* 26(1), 74–82. doi: https://doi.org/10.1016/j.sjbs.2018.09.004

Amusa, O., Adesoye, A., Ogunkanmi, A., Omoche, O., Olowe, O., Akinyosoye, S., and Omodele, T. (2014). Genetic Diversity of *Parkia biglobosa* from Different Agroecological Zones of Nigeria Using RAPD Markers. *International Journal of Biodiversity* 2014, 1–6. doi: https://doi.org/10.1155/2014/457309

Awodoyin, R. O., Olubode, O. S., Ogbu, J. U., Balogun, R. B., Nwawuisi, J. U., and Orji, K. O. (2015). Indigenous Fruit Trees of Tropical Africa: Status, Opportunity for Development and Biodiversity

Management. *Agricultural Sciences* 06(01), 31–41. doi: https://doi.org/10.4236/as.2015.61004

Badr, A. (2008). Molecular approaches in plant systematics and evolution. *Taeckholmia* 28, 127–167.

Badr, A., El-Shazly, H. H., Helail, N. S., and Ghanim, W. E. (2012). Genetic diversity of Artemisia populations in central and north Saudi Arabia based on morphological variation and RAPD polymorphism. *Plant Systematics and Evolution* 298(5), 871–886. doi: https://doi.org/10.1007/s00606-012-0597-5

Barakat, H. (2004). Genetic Fingerprinting and Relationships of Six Soybeans (*Glycine max* L.) Cultivars Based on Protein and DNA Polymorphism. *Int. J. Agri. Biol* 6(5), 877–883.

Baránek, M., Kadlec, M., Raddová, J., Vachůn, M., and Pidra, M. (2002). Evaluation of Genetic Diversity in 19 *Glycine max* (L.) Merr. accessions included in the Czech National Collection of Soybean Genotypes. *Czech J Genet Plant Breed* 38(2), 69–69. doi: http://doi.org/10.17221/6114-CJGPB

Boffa, J. M. (1999). Agroforestry parklands in sub-Saharan Africa volume 34. (Rome: FAO), 1-230. url: http://www.fao.org/3/x3940e/X3940E00.htm.

Chowdhury, A. K., Srinives, P., Tongpmnak, P., and Saksoong, P. (2001). Genetic diversity based on morphology and RAPD analysis in vegetable soybean. *Korea J. Crop Sci* 46(2), 112–120.

Coelho, C. M. M., Bellato, C. M., Santos, J. C. P., Ortega, E. M. M., and Tsai, S. M. (2007). Effect of phytate and storage conditions on the development of the 'hard-to-cook' phenomenon in common beans. *J Sci Food Agric* 87(7), 1237–1243. doi: https://doi.org/10.1002/jsfa.2822

de la Rosa, A. P. B., Gueguen, J., Paredes-Lopez, O., and Viroben, G. (1992). Fractionation procedures, electrophoretic characterization, and amino acid composition of amaranth seed proteins. *Journal of Agricultural and Food Chemistry* 40(6), 931–936. doi: https://doi.org/10.1021/jf00018a002

de Xaxars, G. M., Garnatje, T., Pellicer, J., Siljak-Yakovlev, S., Vallès, J., and Garcia, S. (2016). Impact of dysploidy and polyploidy on the diversification of high mountain *Artemisia* (Asteraceae) and allies. *Alpine Botany* 126(1), 35–48. doi: https://doi.org/10.1007/s00035-015-0159-x

Dedehou, V., Olounlade, P., Alowanou, G., Azando, E., and Adote, S. H. (2016). A review on medicinal plants of *Parkia biglobosa* (Mimosaceae-Fabaceae) and *Pterocarpus erinaceus* (Leguminosae-Papilionoidea). *J. Med. Plants Stud.* 4, 132–137.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter* 1(4), 19–21. doi: https://doi.org/10.1007/bf02712670

Dobes, C., Steccari, I., Kostenberger, S., and Lompo, D. (2019). Relative genome size variation in the African agroforestry tree *Parkia biglobosa* (Fabaceae: Caesalpinioideae) and its relation to geography, population genetics, and morphology. *Genome* 62(10), 665–676. doi: https://doi.org/10.1139/gen-2019-0069

Durán, L. A., Blair, M. W., Giraldo, M. C., Macchiavelli, R., Prophete, E., Nin, J. C., and Beaver, J. S. (2005). Morphological and Molecular Characterization of Common Bean Landraces and Cultivars from the Caribbean. *Crop Science* 45(4), 1320–1328. doi: https://doi.org/10.2135/cropsci2004.0501

El-Kholy, A. S. (2013). Evaluation of genetic variation among Soybean (*Glycine max* L.) Cultivars using SDS-PAGE and RAPD Markers. In and others, Egyptian Journal of Botany 3rd International Conference, 33-48.

Emre, I. (2009). Electrophoretic analysis of some *Lathyrus* L. species based on seed storage proteins. *Genetic Resources and Crop Evolution* 56(1), 31–38. doi: https://doi.org/10.1007/s10722-008-9339-5

Ferreira, E. M. (2006). Molecular analysis of gene banks for sustainable conservation and increased use of crop genetic resources. In *The role of biotechnology in exploring and protecting agricultural genetic resources,* ed. J., R. and A., S., (Rome, Italy: FAO), 121-128.

Gbadamosi, A. E., Faboye, O. O., and Oni, O. O. (2005). Seed morphological traits, germination and early seedling growth of *Parkia biglobosa* (Jacq.) R. Br. Ex G. Don. from four provenances in Nigeria. *Nigerian Journal of Forestry* 35, 129–137.

Hameed, A., Qureshi, M., Nawaz, M., and Iqbal, N. (2012). Comparative seed storage protein profiling of Mungbean genotypes. *Pakistan Journal of Botany* 44(6), 1993–1999.

Hamrick, J. L. and Godt, M. J. W. (1989). Allozyme diversity in plants. In *Plant Population Genetics, Breeding and Genetic Resources,* ed. Brown, D. A. H., Clegg, M. T., Kahler, A. L., and Weir, B. S., (Sunderland, Mass, USA: Sinauer), 43-63.

Hopkins, B. (1983). The taxonomy, reproductive biology and economic potential of *Parkia* (Leguminosae: Mimosoideae) in Africa and Madagascar. *Botanical Journal of the Linnean Society* 87(2), 135–167. doi: https://doi.org/10.1111/j.1095-8339.1983.tb00987.x

Hoque, M. E. and Hasan, M. M. (2012). Molecular Diversity Analysis of Lentil (*Lens culinaris* Medik.) through RAPD Markers. *Plant Tissue Culture and Biotechnology* 22(1), 51–58. doi: https://doi.org/10.3329/ptcb.v22i1.11260

Houndonougbo, J. S. H., Kassa, B., Mensah, S., Salako, V. K., Kakaï, R. G., and Assogbadjo, A. E. (2020). A global systematic review on conservation nd domestication of *Parkia biglobosa* (Jacq.) R. Br. ex G. Don, an indigenous fruit tree species in Sub-Sahara African traditional parklands: current knowledge and future directions. *Genetic Resources and Crop Evolution* 67(4), 1051–1066. doi: https://doi.org/10.1007/s10722-020-00892-w

Jaccard, P. (1908). Nouvelles researches sur la distribution florale. *Bulletin de la Societe vaudoise des sciences naturelles* 44, 223–270.

Javaid, A., Ghafoor, A., and Anwar, R. (2004). Seed storage protein electrophoresis in groundnut for evaluating genetic diversity. *Pakistan Journal of Botany* 36(1), 25–29.

Krochko, J. E. and Bewley, J. D. (2000). Seed storage proteins in cultivars and subspecies of alfalfa (*Medicago sativa* L.). *Seed Science Research* 10(4), 423–434. doi: https://doi.org/10.1017/s0960258500000477

Kwon-Ndung, E. H., Chup, C. D., Binbol, N. A., and Ismaila, A. (2009). Ecological implications of climate change on the genetic diversity and distribution of African locust bean *Parkia biglobosa* in Central Nigeria. *IOP Conference Series: Earth and Environmental Science* 6(37), 372026–372026. doi: https://doi.org/10.1088/1755-1307/6/37/372026

la Rosa, L. D. and González, J. M. (2010). The genetic diversity associated with seed proteins in a collection of Spanish underground vetches (*Vicia sativa* L. subsp. *amphicarpa* (Dorthes) Asch. et Graebn.). *Genetic Resources and Crop Evolution* 57(4), 565–573. doi: https://doi.org/10.1007/s10722-009-9494-3

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227(5259), 680–685. doi: https://doi.org/10.1038/227680a0

Lamien, N., Ekué, M., Ouedraogo, M., and Loo, J. (2011). African locust bean (*Parkia biglobosa*). Conservation and sustainable use of genetic resources of priority food tree species in sub-Saharan Africa. url: http://hdl.handle.net/10532/1681.

Lassen, K. M. (2016). Pollination strategies to increase productivity of the African fruit trees *Vitellaria paradoxa* subsp. *paradoxa* & *Parkia biglobosa*. Ph.D. thesis, University of Copenhagen, Denmark.

Lassen, K. M., Ræbild, A., Hansen, H., Brødsgaard, C. J., and Eriksen, E. N. (2012). Bats and bees are pollinating *Parkia biglobosa* in The Gambia. *Agroforestry Systems* 85(3), 465–475. doi: https://doi.org/10.1007/s10457-011-9409-0

Leakey, R. R. (2012). Living with the trees of life: towards the transformation of tropical agriculture (Wallingford, UK: CABI).

Leakey, R. R. B. (2017). Tree Domestication in Agroforestry: Progress in the 2nd Decade (2003-2012). In Multifunctional Agriculture. Achieving Sustainable Development in Africa, 139-153.

Lin, K. H., Lai, Y. C., Li, H. C., Lo, S. F., Chen, L. F., and Lo, H. F. (2009). Genetic variation and its relationship to root weight in the sweet potato as revealed by RAPD analysis. *Scientia Horticulturae* 120(1), 2–7. doi: https://doi.org/10.1016/j.scienta.2008.09.008

Lompo, D., Kanrad, H., Duminil, J., Gaisberger, H., Vincent, B., Quedraogo, M., and Geburek, T. (2016). Genetic diversity of *Parkia biglobosa* (African locust bean) and its implications for conservation strategies. In Tropentag 2016: Solidarity in a competing world — fair use of resources, 18-21 Sep 2016, Vienna, Austria.

Lompo, D., Vinceti, B., Gaisberger, H., Konrad, H., Duminil, J., Ouedraogo, M., Sina, S., and Geburek, T. (2017). Genetic conservation in *Parkia biglobosa* (Fabaceae: Mimosoideae) - what do we know? *Silvae Genetica* 66(1), 1–8. doi: https://doi.org/10.1515/sg-2017-0001

Lompo, D., Vinceti, B., Konrad, H., Gaisberger, H., and Geburek, T. (2018). Phylogeography of African Locust Bean (*Parkia biglobosa*) Reveals Genetic Divergence and Spatially Structured Populations in West and Central Africa. *Journal of Heredity* 109(7), 811–824. doi: https://doi.org/10.1093/jhered/esy047

Machuka, J. (2001). Heterogeneity in the seed globulin and albumin fractions from African yam bean *Sphenostylis stenocarpa* (Hoechst. Ex A. Rich) harms. *African Crop Science Journal* 9(4), 607–613. doi: https://doi.org/10.4314/acsj.v9i4.27583

Malik, M. F. A., Qureshi, A. S., Ashraf, M., Khan, M. R., and Javed, A. (2009). Evaluation of genetic diversity in soybean (*Glycine max*) lines using seed protein electrophoresis. *Australian Journal of Crop Science* 3(2), 107–112.

Matig, O. E., Gaoué, O. G., and Dossou, B. (2002). Réseau «Espèces Ligneuses Alimentaires». In Compterendu de la première réunion du Réseau tenue 11-13 décembre 2000 au CNSF , 235-235.

Mendoza, E. M. T., Adachi, M., Bernardo, A. E. N., and Utsumi, S. (2001). Mungbean [Vigna radiata(L.) Wilczek] Globulins: Purification and Characterization. *Journal of Agricultural and Food Chemistry* 49(3), 1552–1558. doi: https://doi.org/10.1021/jf001041h

Mondini, L., Noorani, A., and Pagnotta, M. (2009). Assessing Plant Genetic Diversity by Molecular Tools. *Diversity* 1(1), 19–35. doi: https://doi.org/10.3390/d1010019

Murray, B. G. (2005). When does Intraspecific C-value Variation become Taxonomically Significant? *Annals of Botany* 95(1), 119–125. doi: https://doi.org/10.1093/aob/mci007

Mustafa, A. M. A. and El-Kholy, A. S. (2008). Genetic diversity among some accessions of *Vicia faba* L. around the world. In Proceedings of 5th International Conference of Biological Science (Botany) 5, 10-10.

Narzary, D., Mahar, K. S., Rana, T. S., and Ranade, S. A. (2009). Analysis of genetic diversity among wild pomegranates in Western Himalayas, using PCR methods. *Scientia Horticulturae* 121(2), 237–242. doi: https://doi.org/10.1016/j.scienta.2009.01.035

Nei, M. (1987). Molecular evolutionary genetics (New York: Columbia University Press).

Nei, M. and Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76(10), 5269–5273. doi: https://doi.org/10.1073/pnas.76.10.5269

Nikiema, A. (2005). Agroforestry parkland species diversity: uses and management in semi-arid West Africa (Burkina Faso). Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.

Nyadanu, D., Amoah, R. A., Obeng, B., Kwarteng, A. O., Akromah, R., Aboagye, L. M., and Adu-

Dapaah, H. (2017). Ethnobotany and analysis of food components of African locust bean (*Parkia biglobosa* (Jacq.) Benth.) in the transitional zone of Ghana: implications for domestication, conservation and breeding of improved varieties. *Genetic Resources and Crop Evolution* 64(6), 1231–1240. doi: https:// doi.org/10.1007/s10722-016-0432-x

Odeigah, P. G. C. and Osanyinpeju, A. O. (1996). Seed protein electrophoretic characterization of cowpea (*Vigna unguiculata*) germplasm from IITA gene bank. *Genetic Resources and Crop Evolution* 43(6), 485–491. doi: https://doi.org/10.1007/bf00138825

Okoye, T. G., Uzor, P. F., Onyeto, C. A., and Okereke, E. K. (2014). Safe African Medicinal Plants for Clinical Studies. In Toxicological Survey of African Medicinal Plants, 535-555.

Okunlola, A. I., Adebayo, R. A., and Orimogunje, A. D. (2011). Methods of breaking seed dormancy in germination and early seedling growth of African locust bean *Parkia biglobosa* (Jacq.) Benth. . *Journal of Horticulture and Forestry* 3(1), 1–6.

Oni, P. I., Hall, J. B., and Ladipo, D. O. (1998). The ecology of a key African multipurpose tree species, *Parkia biglobosa* (Jacq.) Benth: the current state of knowledge. *Nigerian Journal of Ecology* 1, 59–77.

Orwa, C., Mutua, A., Kindt, R., Jamnadass, R., and Anthony, S. (2010). Agroforestree Database: a tree reference and selection guide version 4.0. CD-ROM.

Osanyinpeju, A. O. and Odeigah, P. G. C. (1998). Variation in seed proteins from mutagen-treated cultivars and selected lines of *Vigna unguiculata* L. Walp. *Plant Breeding* 117(4), 361–365. doi: https://doi.org/10.1111/j.1439-0523.1998.tb01954.x

Ouedraogo, A. S. (1995). *Parkia biglobosa* (Leguminosae) en Afrique de l'Ouest : biosystematique et amelioration. Ph.D. thesis, Wageningen University, Wageningen, Netherlands. (in French).

Padakale, E., Atakpama, W., Dourma, M., Dimobe, K., Wala, K., Guelly, K., and Akpagana, K. (2015). Woody Species Diversity and Structure of *Parkia biglobosa* Jacq. Dong Parklands in the Sudanian Zone of Togo (West Africa). *Annual Research & Review in Biology* 6(2), 103–114. doi: https://doi.org/10.9734/arrb/2015/14105

Prasanthi, L., Geetha, B., Ramyajyothi, B. N., and Reddy, K. (2012). Evaluation of genetic diversity in cowpea, *Vigna unguiculata* (L.) Walp genotypes using random amplified polymorphic DNA (RAPD). *Current Biotica* 6(1), 22–31.

Przybylska, J., Zimniak-Przybylska, Z., and Krajewski, P. (2000). Diversity of seed globulins in *Lathyrus sativus* L and some related species. *Genetic Resources and Crop Evolution* 47(3), 239–246. doi: https://doi.org/10.1023/A:1008750423357

Ræbild, A., Hansen, U. B., and Kambou, S. (2012a). Regeneration of *Vitellaria paradoxa* and *Parkia biglobosa* in a Parkland in Southern Burkina Faso. *Agroforestry Systems* 85, 443–453. doi: https://doi. org/10.1007/s10457-011-9397-0

Ræbild, A., Hansen, U. B., and Kambou, S. (2012b). Regeneration of *Vitellaria paradoxa* and *Parkia biglobosa* in a parkland in Southern Burkina Faso. *Agroforestry Systems* 85(3), 443–453. doi: https://doi. org/10.1007/s10457-011-9397-0

Rao, R. and Pernolett, J. C. (1981). *Pisum sativum* seed globulin as biochemical markers for wrinkled smooth seed character: comparison of genotypes through different protein extraction and electrophoretic procedures. *Agronomie* 1(10), 909–916.

Rohlf, F. J. (2002). NTSYS-PC Numerical taxonomy and multivariate analysis system.

Sammour, R. H., Mustafa, A., Badr, S., and Tahr, W. (2007a). Genetic variations in accessions of *Lathyrus sativus*. *L. Acta Botanical Croatica* 66(1), 1–13.

Sammour, R. H., Radwan, S. A., and El-Kholy, A. (2007b). Genetic variability in *Phaseolus* spp. as revealed by SDS-PAGE Markers. *Seed technology* 29(1), 50–59.

Sankhon, A., Amadou, I., Yao, W. R., Wang, H., Qian, H., and Sangare, M. (2014). Comparison Of Physicochemical And Functional Properties Of Flour And Starch Extract In Different Methods From Africa Locust Bean (*Parkia biglobosa*) Seeds. *African Journal of Traditional, Complementary and Alternative Medicines* 11(2), 264–264. doi: https://doi.org/10.4314/ajtcam.v11i2.6

Shao, M. (2002). Parkia biglobosa: Changes in resource allocation in Kandiga, Ghana. Ph.D. thesis, Michigan Technological University, Houghton, MI, USA.

Signor, C. L., Gallardo, K., Prosperi, J. M., Salon, C., Quillien, L., Thompson, R., and Duc, G. (2005). Genetic diversity for seed protein composition in *Medicago truncatula*. *Plant Genetic Resources* 3(1), 59–71. doi: https://doi.org/10.1079/pgr200562

Sina, S. (2006). Reproduction et diversite genetique chez *Parkia biglobosa* (Jacq) G. Don. Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.

Sina, S. and Traore, S. A. (2002). *Parkia biglobosa* (Jacq.) R. Br. ex G. Don. In *PROTA (Plant Resources of Tropical Africa/Ressources végétales de l'Afrique tropicale),* ed. Oyen, L. P. A. and Lemmens, R. H. M. J., (Wageningen, Netherlands) .

Stevens, L., Salomon, B., and Sun, G. (2007). Microsatellite variability and heterozygote excess in *Elymus trachycaulus* populations from British Columbia in Canada. *Biochemical Systematics and Ecology* 35(11), 725–736. doi: https://doi.org/10.1016/j.bse.2007.05.017

Strelec, I., Has-Schon, E., and Vitale, L. (2012). Comparative electrophoretic patterns of albumins/globulins from dry grains and green malts of barley varieties. *Poljoprivreda* 18(2), 36–43. url: https://hrcak.srce.hr/94628.

Tamboli, A. S., Patil, S. M., Gholave, A. R., Kadam, S. K., Kotibhaskar, S. V., Yadav, S. R., and Govindwar, S. P. (2016). Phylogenetic analysis, genetic diversity and relationships between the recently segregated species of Corynandra and Cleoserrata from the genus Cleome

using DNA barcoding and molecular markers. *Comptes Rendus Biologies* 339(3-4), 123–132. doi: https://doi.org/10.1016/j.crvi.2016.02.005

Tang, C. H. and Sun, X. (2010). Physicochemical and Structural Properties of 8S and/or 11S Globulins from Mungbean [*Vigna radiate* (L.) Wilczek] with various polypeptide constituents. *J. Agric. Food Chem* 58(10), 6395–6402. doi: https://doi.org/10.1021/jf904254f

Teklehaimanot, Z. (2004). Exploiting the potential of indigenous agro forestry trees: *Parkia biglobosa* and *Vitellaria paradoxa* in Sub-Sahara Africa. In New Vistas in Agroforestry, Springer, Dordrecht, Netherlands, 207-220.

Uyoh, E. A., Urua, I. S., Ntui, V. O., and Okpako, E. C. (2011). Flow cytometry analysis of nuclear DNA content, mitotic chromosome number and protein separation by SDS-PAGE in three accessions of African locust bean (*Parkia biglobosa* Benth.). *Journal of Crop Science and Biotechnology* 14(4), 227–232. doi: https://doi.org/10.1007/s12892-011-0043-1

Wang, A., Zhang, P., Liu, X., Liang, J., and Li, W. (2016). Genetic structure and diversity of *Glehnia littoralis*, an endangered medicinal plant in China. *Biochem Syst Ecol* 66, 265–271. doi: https://doi.org/10.1016/j.bse.2016.04.019

Zarkadas, C. G., Gagnon, C., Gleddie, S., Khanizadeh, S., Cober, E. R., and Guillemette, R. J. (2007). Assessment of the protein quality of fourteen soybean [*Glycine max* (L.) Merr.] cultivars using amino acid analysis and two-dimensional electrophoresis. *Food Research International* 40(1), 129–146. doi: https://doi.org/10.1016/j.foodres.2006.08.006