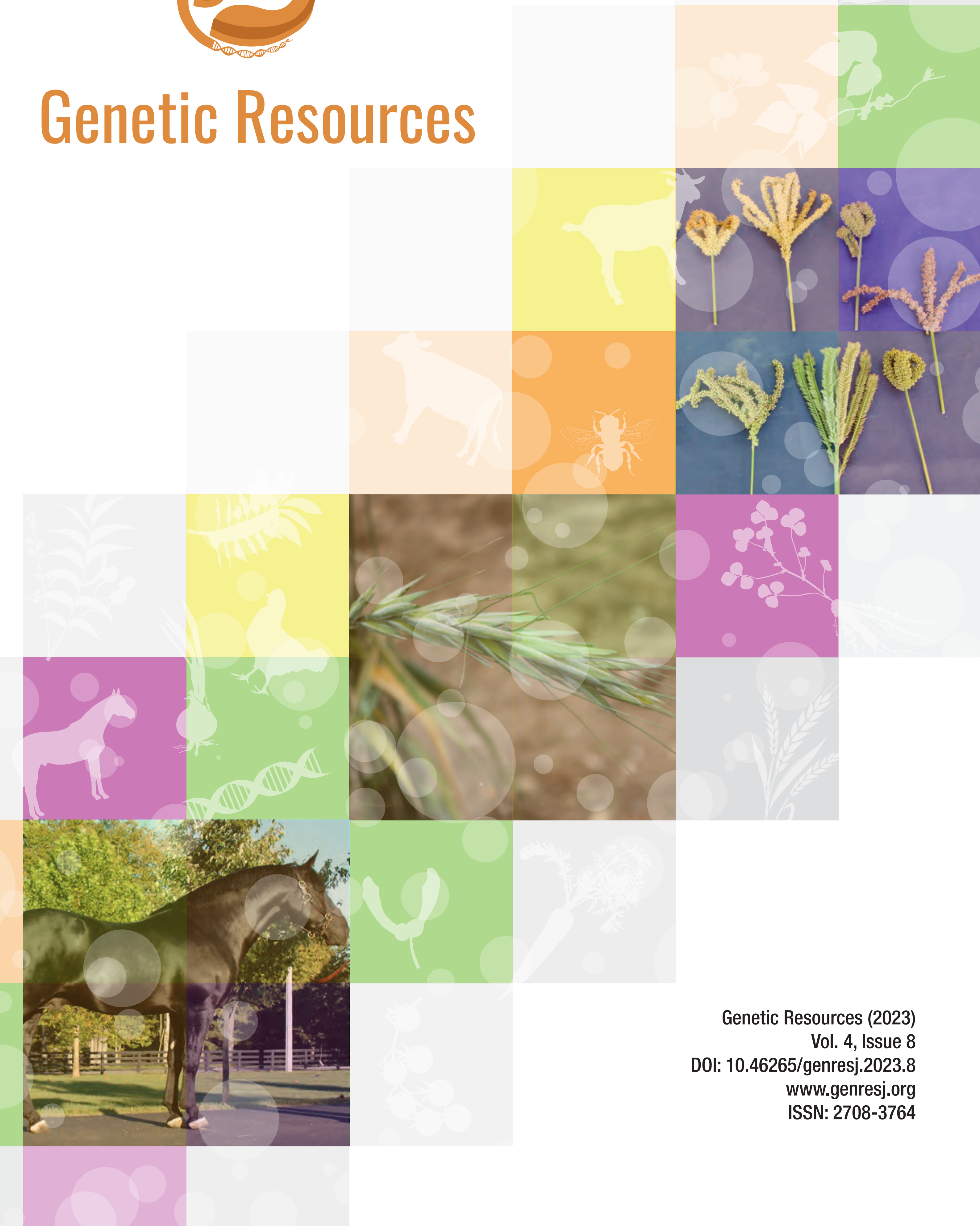
Genetic Resources



Genetic Resources (2023)
Vol. 4, Issue 8
DOI: 10.46265/genresj.2023.8
www.genresj.org
ISSN: 2708-3764

Focus and Scope of *Genetic Resources*

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

Cover photos (from left to right):
Irish Draught stallion, credit: Sabu Varghese; Triticum polonicum, credit: Bioversity International/Nora Capozio; Finger millet diversity, credit: Krishna Ghimire.

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

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

Editorial Office:

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

Submissions to www.genresj.org

Editorial Board:

Genetic Resources

TABLE OF CONTENTS

Genetic Resources (2023), 4 (8)
DOI: 10.46265/genresj.2023.8
www.genresj.org
ISSN: 2708-3764

Original Articles

Short Communications

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2023), 4 (8), 1–14
DOI: 10.46265/genresj.MYZA2446
https://www.genresj.org
ISSN: 2708-3764

Phenotypic diversity among finger millet (*Eleusine coracana* (L.) Gaertn.) landraces of Nepal

Krishna Hari Ghimire [*,a,b], **Madhav Prasad Pandey** [b], **Bal Krishna Joshi** [a], **Surya Kanta Ghimire** [b], **Hira Kaji Manandhar** [b] and **Devendra Gauchan** [c]

[a] *National Agriculture Genetic Resources Centre (Genebank), NARC, Khumaltar, Lalitpur, Nepal*
[b] *Agriculture and Forestry University, Faculty of Agriculture, Chitwan, Rampur, Nepal*
[c] *Alliance of Bioversity International and CIAT, Khumaltar, Lalitpur, Nepal*

Abstract: Finger millet (*Eleusine coracana* (L.) Gaertn.) is the fourth most important crop in Nepal having multiple benefits but is still neglected by mainstream research and development. The main option to boost its productivity is developing superior varieties through enhanced use of germplasm in breeding programmes. With the objective of enhancing utilization of landraces conserved *ex situ*, a total of 300 finger millet accessions collected from 54 districts were characterized in three hill locations of Nepal for two consecutive years (2017–2018). Nine qualitative and 17 quantitative traits were recorded, and combined mean data were subjected to multivariate analysis to assess agromorphological diversity. Shannon–Weaver diversity indices (H') showed high diversity (0.647–0.908) among the accessions for qualitative traits except for finger branching and spikelet shattering whereas high diversity (0.864–0.907) was observed for all quantitative traits. The first five principal components (PC) explained 61.8% of the total phenotypic variation with two PCs explaining 37.5% variation mainly due to flowering and maturity days, plant height, flag leaf length, grain and straw yield, ear weight, ear exsertion and number of fingers per head. Genotypes were grouped into four clusters with 16, 66, 107 and 111 accessions based on quantitative traits. The correlation between the traits indicated that accessions with early flowering, tall plants, long leaves, high tillers, large ears and bold grains could be given priority for further evaluation in multiple locations. Potential landraces identified for each trait could either be deployed to wider areas as varieties or used as trait donors in finger millet breeding.

Keywords: Eleusine coracana, finger millet, multivariate analysis, phenotypic diversity, Nepal

Citation: Ghimire, K. H., Pandey, M. P., Joshi, B. K., Ghimire, S. K., Manandhar, H. K., Gauchan, D. (2023). **Phenotypic diversity among finger millet (*Eleusine coracana* (L.) Gaertn.) landraces of Nepal**. *Genetic Resources* 4 (8), 1–14. doi: 10.46265/genresj.MYZA2446.

Introduction

Finger millet (*Eleusine coracana* (L.) Gaertn.) is an allotetraploid species (2n = 4x = 36; genome constitution AABB) of the grass family Poaceae (Dida *et al*, 2007). It evolved from a wild species (*E. coracana* ssp. *africana* (Kenn.-O'Bryne) (AABB)) which is the natural cross between wild species (*E. indica* (L.) Gaertn. (AA)) and extinct unknown species (BB) (Liu *et al*, 2014). East Africa is considered its primary centre of diversity (Phillips, 1972) where nine out of ten known species of the genus *Eleusine* – *coracana, africana, indica, floccifolia* (Spreng), *intermedia* (Chiov.) (S.M.Phillips), *multiflora* (Hochst. ex A.Rich), *jaegeri* (Pilg.), *kigeziensis* (S.M.Phillips) and *semisterilis* (S.M.Phillips) – are found, except *E. tristachya* (Lam.) (Hilu and De-Wet, 1976). Ploidy and hybridization barriers suggest that tetraploid (2n = 4x = 36) species *E. coracana* (AABB) and *E. africana* (AABB) are in the primary gene pool, diploid (2n = 2x = 18) species *E. indica* (AA), *E. tristachya* (AA) and *E. floccifolia* (BB) formed the secondary gene pool and the rest of the species are in the tertiary gene pool (Sood *et al*, 2019). It was domesticated about 5,000 years ago in eastern Africa (Ethiopian highlands)

---

[*] Corresponding author: Krishna Hari Ghimire (ghimirekh@gmail.com)

Received: 12.04.2023 | Accepted: 20.06.2023 | Published online: 11.07.2023

and introduced into the Indian subcontinent 3,000 years ago (Hilu *et al*, 1979; Upadhyaya *et al*, 2006).

Globally, finger millet ranked fourth in importance among millet crops after sorghum, pearl millet and foxtail millet (Upadhyaya *et al*, 2007). In many countries, precise data on cultivation area and production of finger millet are not available because the production statistics of this crop had often been combined with other millets (Upadhyaya *et al*, 2010). It is cultivated on 3.8 million ha (12% of the total millet area) with coverage in more than 25 countries in Africa (e.g. Uganda, Tanzania, Kenya, Ethiopia, Rwanda, Zaire, Eritrea and Somalia) and Asia (e.g. India, Nepal, Sri Lanka, Myanmar, China and Japan) (Upadhyaya *et al*, 2010; Bora, 2013; Kumar *et al*, 2016; Vetriventhan *et al*, 2016; Hittalmani *et al*, 2017). It is grown in a wide range of environments from the tropical coastal regions of India (Upadhyaya *et al*, 2006) to the high mountains (3,130 masl) of Nepal (Bastola *et al*, 2015; Gaihre *et al*, 2021). Having the C4 photosynthetic pathway (Hittalmani *et al*, 2017; Parvathi *et al*, 2019), it is a hardy crop grown in marginal land and stress environments with very low or minimum input (Goron *et al*, 2015). Finger millet (kodo in Nepali) is the fourth most important cereal crop in Nepal after rice, maize and wheat in terms of area and production, and occupies an average of 7.7% (265,401ha) of the total cultivated area covered by cereal crops and accounts for 2.9% (326,443t) of total cereal production with average yield of 1.23t/ha (MOALD, 2022).

Nutritionally, its importance is well recognized because of its high content of calcium (0.34%), dietary fiber (18%), protein (6–13%), minerals (2.5–3.5%), phytates (0.48%) and phenolic compounds (0.3–3%) (Chandra *et al*, 2016). It is enriched with calcium, iron, zinc, proteins and calories (O'Kennedy *et al*, 2006; Upadhyaya *et al*, 2011). The crop is also valued for its health beneficial effects like anti-diabetic, anti-tumorigenic, antioxidant and antimicrobial properties (Devi *et al*, 2011; Kumar *et al*, 2016; Nakarani *et al*, 2021). Besides food and nutrition, it is an integral component of agrotourism in Nepal due to local specialties made from it, such as dhindo (thick porridge) and high quality raksi (home-made wine) (Ghimire *et al*, 2017; Joshi *et al*, 2020; Gaihre *et al*, 2021).

Vetriventhon *et al* (2020) reported a total of 36,873 finger millet accessions (including landraces, improved cultivars, wild and weedy relatives) conserved *ex situ* at global level and this number is ever increasing. Most of these collections are yet to be characterized and utilized in breeding. The International Crop Research Institute for Semi-Arid Tropics (ICRISAT) has developed a finger millet core collection of 622 accessions, including 70 accessions from Nepal, based on agromorphological diversity from their entire collection of 5,940 accessions (Upadhyaya *et al*, 2006). Two studies reported on the characterization of Nepalese finger millet accessions (Bhattarai *et al*, 2014) but the landraces used in those studies were not properly represented in the entire genebank collection. Characterization of collected landraces is the most important avenue to open the door for their utilization. However, less utilization of local genetic resources for crop improvement programmes is evident in Nepal due to lack of information about the desirable accessions in the genebank resulting from poor characterization and evaluation data. Three out of six finger millet varieties notified in the country were improved from native landraces, which include Okhle-1, Kabre kodo-1 and Rato kodo. The present study describes the characterization of finger millet accessions conserved at the National Agriculture Genetic Resources Centre (NAGRC, Genebank) of Nepal, grouping of accessions with similar characters using a range of multivariate statistical tools and identifying potential landraces to be utilized in finger millet improvement programmes.

Materials and methods

Plant materials and experimental sites

This study used 300 finger millet accessions (Supplemental Table 1) received from NAGRC which include 295 landraces collected from 54 districts of 6 provinces, and 5 released varieties (Okhle-1, Dalle kodo-1, Kabre kodo-1, Kabre kodo-2 and Shailung kodo-1) of Nepal. A total of 295 landraces were selected from nearly 1,000 accessions of 54 districts based on proportions and representation, so that there was minimum repetition in local name and at least one accession from each district. Experiments were conducted at three mountain locations of Nepal, namely Agriculture Research Station (ARS) Vijayanagar, Jumla (2,350 masl); NAGRC Khumaltar, Lalitpur (1,360 masl) and Hill Crops Research Programme (HCRP), Kabre, Dolakha (1,740 masl). Geo-coordinates of experimental locations and collection sites of landraces were mapped (Figure 1). All three sites had coarse textured sandy loam soil.

General methodology

The experiments were laid out in alpha lattice design with 300 entries and two replications having 15 blocks within replications and 20 plots in each block. Each plot was constituted by 20 plants in a single row of 2m length with 25cm spacing between rows. During 2017 and 2018 respectively, seeding was done on 24 and 19 April at Jumla, 3 June and 26 May at Dolakha and 17 June and 7 June at Khumaltar. Direct seeding was done with the application of chemical fertilizers at the rate of 20:10:10 kg/ha N:P_2O_5:K_2O as basal doses. Thinning was applied within 25–30 days after seeding to maintain a plant-to-plant spacing of 10cm within rows. Manual weeding was done as per requirement but no irrigation and pesticides were applied.

Data recording

Morphological data of nine qualitative and 17 quantitative traits were recorded as per the standard descriptors of finger millet (IBPGR, 1985). The qualitative traits were recorded from a single replication of 2017 at Khu-

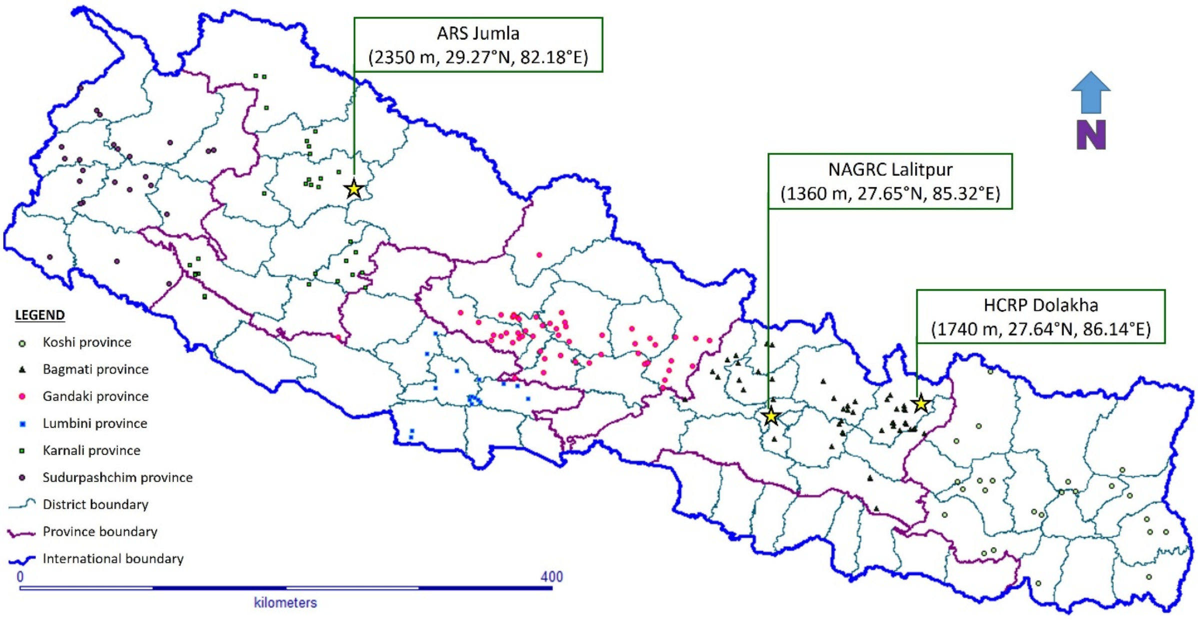


Figure 1. Map of Nepal showing collection sites of characterized finger millet accessions in six provinces, coded as indicated in the legend. The three experimental sites where field trials were conducted are indicated by stars.

maltar only, based on observations as per the descriptor states. Observations on days to 50% flowering, days to 80% maturity, grain yield (kg/ha) and straw yield (t/ha) were based on whole plot data whereas measurements on other quantitative traits such as plant height (cm), tillers per hill (n), flag leaf length (cm), flag leaf width (cm), flag leaf sheath length (cm), ear exsertion (cm), ear head length (cm), ear head width (cm), fingers per head (n), length of the longest finger (cm), width of the longest finger (cm), weight per head (g) and weight of 1,000 grains (g) were made from five randomly selected plants. Observations for days to maturity of those accessions which did not set grains due to extremely low temperature after flowering at Jumla were considered as missing values whereas grain yield of those accessions were estimated as zero.

Data analysis

The frequency of each descriptor state for all qualitative traits were tabulated with their proportions whereas observations for all quantitative traits of each year and locations were subjected to unbalanced analysis of variance (ANOVA) using regression model with the software GenStat version 15 (VSN International, 2015). The combined mean data were subjected to descriptive statistical analysis such as minimum, maximum, mean, standard error and coefficient of variation using Minitab-17 (MINITAB, 2010). Hierarchical clustering of observations and construction of dendrogram were done based on unweighted pair group method with arithmetic mean (UPGMA) by using FactoMineR package (Sebastien *et al*, 2008) and factor map visualization was made by using Factoextra package (Kassambara and Mundt, 2020) of R statistical software (R Core Team, 2020). Correlation analysis (Pearson's coefficient with probability) and principal component analysis (PCA) were done using Minitab-17 (MINITAB, 2010). Standardized Shannon-Weaver diversity indices (H') (Shannon and Weaver, 1949) were calculated for each trait with Microsoft Excel (Ghimire *et al*, 2018a,b) .

Results

Diversity index and frequency distribution of qualitative traits

Based on Shannon-Weaver diversity index (H'), we observed very low diversity for finger branching and grain shattering (0.122) but it was highest for plant pigmentation (0.908) followed by grain covering (0.793), ear shape (0.790) and seed colour (0.722) (Table 1). The data showed that the predominant ear shapes in Nepalese finger millet are open (42.7%) and semi-compact (41.7%) types. More than two-thirds (68%) of accessions had intermediate ear size followed by large ear size (29.7%). Less than 2% of the accessions had branching in fingers and grain shattering. Similarly, one-third of the accessions had pigmented plants while very few accessions were highly susceptible to lodging (3.3%). Seed colour varied from white to purple-brown (Figure 2). The predominant seed colour in the collection was light-brown (53%) followed by purple-brown (37%).

Descriptive statistics and diversity indices of quantitative traits

Range, mean, standard error of mean (SE), coefficient of variation (CV) and Shannon-Weaver diversity index (H') of each quantitative trait are presented in Table 2. A wide range of variation in agronomic performance

Table 1. Shannon-Weaver diversity indices (H'), descriptor states and frequency of nine qualitative traits.

Qualitative traits	H'	Descriptor states	Frequency (n)	Proportion (%)
Ear shape	0.790	Droopy	38	12.7
		Open	128	42.7
		Semi-compact	125	41.7
		Compact	9	3.0
Ear size	0.647	Small	7	2.3
		Intermediate	204	68.0
		Large	89	29.7
Finger branching	0.122	Absent	295	98.3
		Present	5	1.7
Grain covering	0.793	Exposed	67	22.3
		Intermediate	196	65.3
		Enclosed	37	12.3
Lodging susceptibility	0.654	Low	209	69.7
		Intermediate	81	27.0
		High	10	3.3
Plant pigmentation	0.908	Not pigmented	203	67.7
		Pigmented	97	32.3
Seed colour	0.722	White	11	3.7
		Light-brown	159	53.0
		Copper-brown	19	6.3
		Purple-brown	111	37.0
Spikelet density	0.676	Sparse	89	29.7
		Intermediate	201	67.0
		Dense	10	3.3
Spikelet shattering	0.122	Absent	295	98.3
		Present	5	1.7

Figure 2. Seed colour variation (white to purple-brown) on different finger millet accessions (In case of mixed seeds, colour of the majority of the seed is considered, e.g. colour of NGRC04793 is recorded as white).

was observed among the evaluated accessions. The early maturing accessions started flowering at 75 days after seeding and the late maturing accessions flowered at 140 days whereas the average plant height ranged from 61 to 119cm. The average adjusted grain yield and straw yield ranged from 230 to 3,494kg/ha and 2.0 to 20.2t/ha, respectively. The CV varied from 6.2% for flag leaf length to 33.3% for grain yield. H’ ranged from 0.864 to 0.907 suggesting high diversity in finger millet accessions for all quantitative traits.

Clustering observations

A UPGMA hierarchical clustering divided the entire 300 accessions into four clusters (Figure 3). The number of accessions in each cluster and cluster characteristics for each quantitative trait are presented in Table 3. Cluster 4 was the largest cluster with 111 (37%) accessions having the highest cluster means for grain yield (1,858 kg/ha), straw yield (10.6t/ha), plant height (100cm), 1,000-grain weight (2.4g), weight per head (6.5g) and ear length (6.6cm). Cluster 3 was the second largest cluster with 107 (35.6%) accessions having the lowest mean grain yield (1,147kg/ha) but the longest mean flowering days (122) and maturity days (161 days). Cluster 1 was the smallest cluster with 16 (5.3%) accessions characterized by the lowest cluster mean for straw yield (4.9t/ha), weight per head (4.5g), finger length (5.0cm), ear length (5.5cm), fingers per head (5.6), plant height (78cm), flowering (85 days) and maturity (132 days). Non-significant difference was observed between cluster mean and overall mean for flag leaf width.

Principal component analysis

The contribution of various traits in total phenotypic variation among 300 finger millet accessions was evaluated by principal component analysis (PCA). The first five principal components with eigenvalue ~1 or more, explained 61.8% of the total variation (Table 4). The first principal component (PC-1) explained 22.6% of the total variation which was positively attributed to days to maturity (0.421), days to flowering (0.401), straw yield (0.338), fingers per head (0.303), finger length (0.295), ear length (0.290), plant height (0.280), leaf length (0.255), weight per head (0.252) and ear width (0.229). The second component (PC-2) explained an additional 14.9% of the total variation.The maximum variation in this PC was primarily due to the lower grain yield (-0,496), sheath length (-0.389), plant height (-0.313), 1,000-grain weight (-0.278), ear exsertion (-0.256), and tillers per plant (-0.255) but higher value of days to flowering (0.314) and maturity (0.267). The third component (PC-3), which explained 9.1% of the total variation, differentiated the accessions by higher finger length and ear length but lower leaf length.

PCA using cluster means showed that the first three components explained 100% of the total variability with 60.5 and 28.2% contribution by PC-1 and PC-2, respectively (Table 4). Most of the traits occupied the right side of the bi-plot and thus contributed with positive loadings to the variation explained by PC-1 (Figure 4b). A clear-cut elbow on the fourth component in the scree plot (Figure 4a) as well as the two-dimension scatterplot of PC1 and PC2 (Figure 5) revealed strong support for the clustering result since we can see the visible groupings of accessions as per the clusters (Figure 3).

Correlation between traits

The Pearson’s correlation coefficients between traits are presented in Table 5. A total of 136 trait associations were estimated among the 17 quantitative traits. Out of these, associations between days to 50% flowering and days to 80% maturity (0.93) as well as between ear length and length of the longest finger (0.63) had high estimates. This indicates that only one trait from each of these pairs could be recorded and assessed during future characterization work. Grain yield was positively correlated with plant height, productive tillers per plant, flag leaf length, leaf sheath length, ear exsertion, ear length, finger length, ear weight and 1,000-grain weight but negatively correlated with days to flowering and maturity. When selecting for grain yield, we should consider these traits strongly associated with grain yield while when selecting for straw yield, our focus should be on taller plant height and late maturity since straw yield had strong positive association with days to flowering, maturity and plant height.

Promising landraces

Promising trait-specific donors were identified (Table 6) based on combined mean data of six environments (Supplemental Table 2). Some landraces were good for multiple traits (highlighted in Table 6) and some others for particular traits. For instance, NGRC06490 was high yielding with higher number of tillers; NGRC04849 was high yielding with higher 1,000-grain weight; NGRC04871 was high yielding with taller plant, longer ears and higher 1,000-grains weight; Kabre kodo-2 was high yielding with higher 1,000-grain weight and higher number of tillers; and NGRC04818 was high yielding with taller plant, higher 1,000-grain weight and higher weight per head. Landraces NGRC04849 and NGRC06490 produced 121.6% and 120.1% higher grain yield respectively, than the overall mean (1,577kg/ha) and 10.4% and 9.7% higher yield respectively, compared to Kabre kodo-2, a newest and best among the five released varieties. Mean days to flowering was 112 but 11 accessions flowered before 90 days after seeding. Early flowering genotypes may be selected for drought-prone lowlands as well as higher altitudes since they can escape both drought and cold stress during reproductive stage. These landraces showed potential to be efficiently utilized in breeding programmes for the improvement of finger millet.

Table 2. Variability statistics and Shannon-Weaver diversity indices (H’) of 17 quantitative traits. SE, standard error; CV, coefficient of variation.

Trait	Minimum	Maximum	Mean	SE	CV (%)	H’
Days to 50% flowering (n)	75	140	112.2	0.70	10.8	0.870
Days to 80% maturity (n)	122	174	154.4	0.59	6.6	0.871
Plant height (cm)	61	119	95.4	0.44	8.0	0.864
Tillers per hill (n)	2.5	5.9	4.4	0.03	11.1	0.898
Flag leaf length (cm)	22	32	27.2	0.10	6.2	0.880
Flag leaf width (cm)	0.5	1.2	0.91	0.01	9.7	0.906
Flag leaf sheath length (cm)	11	22	15.1	0.08	8.7	0.872
Ear exsertion (cm)	7.2	14	10.8	0.07	10.7	0.899
Ear head length (cm)	4.5	8.8	6.2	0.04	11.0	0.880
Ear head width (cm)	3.0	6.5	4.2	0.03	13.3	0.867
Fingers per head (n)	4.1	8.9	7.0	0.04	9.4	0.878
Length of the longest finger (cm)	4.0	9.8	5.9	0.06	16.9	0.875
Width of the longest finger (cm)	0.48	0.93	0.68	0.005	11.9	0.902
Weight per head (g)	3.1	10	5.9	0.06	17.6	0.897
Weight of 1,000 grains (g)	1.8	3.2	2.3	0.01	9.5	0.882
Grain yield (kg/ha)	230	3,494	1,577	30.3	33.3	0.884
Straw yield (t/ha)	2.0	20.2	9.3	0.17	30.7	0.907

Table 3. Number of accessions and characteristics of each cluster in comparison with population mean. Sd, standard deviation; *, significant (p = 0.01–0.05) and **, highly significant ($p \leq 0.001$) difference with overall mean.

Trait	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Overall	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Number of accessions	16		66		107		111		300	
Days to 50% flowering (n)	85**	6.56	101**	6.13	122**	7.14	116**	6.39	112	12.1
Days to 80% maturity (n)	132**	5.86	145**	6.01	161**	5.53	158**	5.71	154	10.2
Plant height (cm)	78**	8.92	–	–	93**	5.90	100**	5.69	95.4	7.6
Tillers per hill (n)	–	–	–	–	4.2**	0.44	4.5**	0.45	4.4	0.49
Flag leaf length (cm)	24.4**	1.07	26.8*	1.46	–	–	27.9**	1.44	27.2	1.68
Flag leaf width (cm)	–	–	–	–	–	–	–	–	0.91	0.09
Flag leaf sheath length(cm)	13.8**	0.91	16.0**	1.18	14.4**	1.17	–	–	15.1	1.31
Ear exsertion (cm)	10.1*	1.18	–	–	10.3**	1.08	11.2**	1.10	10.8	1.15
Ear head length (cm)	5.5**	0.45	5.8**	0.53	–	–	6.6**	0.63	6.2	0.68
Ear head width (cm)	3.8**	0.56	4.0**	0.48	–	–	4.4**	0.50	4.2	0.56
Fingers per head (n)	5.6**	0.88	6.8**	0.46	7.3**	0.52	–	–	7.0	0.66
Length of longest finger (cm)	5.0**	0.70	5.4**	0.80	–	–	6.6**	0.96	5.9	1.01
Width of longest finger (cm)	–	–	–	–	0.67*	0.09	–	–	0.68	0.08
Weight per head (g)	4.5**	0.72	5.6**	0.80	5.6**	0.84	6.5**	0.96	5.9	1.03
Weight of 1,000 grains (g)	–	–	–	–	2.2**	0.19	2.4**	0.22	2.3	0.22
Grain yield (kg/ha)	1,181**	379.8	1,756**	359.5	1,147**	385.1	1,858**	473.9	1,577	524.4
Straw yield (t/ha)	4.9**	2.22	7.3**	1.98	10.1**	2.35	10.6**	2.42	9.3	2.85

Discussion

Genetically diverse accessions conserved in *ex situ* genebanks are tremendous genetic resources for breeding high-yielding and stable crop varieties to ensure global food security. Enormous morphological and genetic diversity exists among finger millet accessions but their utilization in breeding programmes is very weak in most countries including Nepal because this crop has received very little attention for characterization, evaluation and pre-breeding activities. Only about 10% of genetic resources including finger millet stored in genebanks have been utilized in crop improvement programmes, which is mainly due to a lack of information about the desirable accessions resulting from the poor characterization and evaluation data (Hodgkin *et al*, 2003; Nguyen and Norton, 2020).

Although NAGRC holds nearly 1,000 finger millet accessions, we characterized only 300 accessions which might not represent the total genetic diversity of the

Table 4. Eigenvalues and eigenvectors under five principal components (PC) for entry means and cluster means.

Eigen analysis	**Entry means**					**Cluster means**				
	PC-1	**PC-2**	**PC-3**	**PC-4**	**PC-5**	**PC-1**	**PC-2**	**PC-3**	**PC-4**	**PC-5**
Eigenvalue	3.763	2.691	1.461	1.237	0.998	10.28	4.79	1.93	0	0
Proportion	0.226	0.149	0.091	0.083	0.069	0.605	0.282	0.113	0	0
Cumulative (%)	22.6	37.5	46.6	54.9	61.8	60.5	88.7	100	100	100
Eigenvectors										
Days to 50% flowering	0.401	0.314	-0.044	-0.057	0.088	0.296	-0.114	0.141	0.083	-0.081
Days to 80% maturity	0.421	0.267	-0.046	-0.086	0.044	0.300	-0.099	0.116	0.057	-0.454
Plant height	0.280	-0.313	-0.214	0.065	0.166	0.279	0.189	0.124	0.042	0.308
Tillers per hill	0.032	-0.255	0.103	-0.383	0.091	-0.123	0.419	-0.032	0.148	-0.007
Flag leaf length	0.255	-0.102	-0.318	0.206	-0.343	0.304	0.079	0.097	-0.005	-0.256
Flag leaf width	-0.014	-0.022	-0.142	-0.442	-0.780	-0.130	0.007	0.654	-0.054	-0.001
Flag leaf sheath length	-0.021	-0.389	-0.283	0.042	0.147	0.147	0.392	0.150	0.149	0.305
Ear exsertion	0.124	-0.256	-0.124	0.396	-0.136	0.260	0.252	0.034	0.177	-0.456
Ear head length	0.290	-0.145	0.456	-0.017	-0.005	0.274	0.022	-0.343	-0.068	-0.168
Ear head width	0.229	-0.006	0.258	0.206	-0.343	0.266	-0.230	-0.101	-0.102	0.298
Fingers per head	0.303	0.118	-0.340	-0.132	0.101	0.294	0.015	0.240	-0.090	-0.071
Length of longest finger	0.295	-0.094	0.468	-0.100	0.070	0.275	-0.065	-0.322	0.025	0.039
Width of longest finger	-0.010	-0.087	0.218	0.519	-0.200	-0.098	0.340	-0.425	0.197	-0.023
Weight per head	0.252	-0.218	-0.088	-0.159	0.095	0.297	0.141	0.013	-0.030	0.327
Weight of 1,000 grains	0.041	-0.278	0.220	-0.255	-0.074	-0.124	0.412	0.122	0.287	-0.088
Grain yield	0.073	-0.496	-0.061	-0.048	0.036	0.131	0.414	-0.009	-0.776	0.016
Straw yield	0.338	0.131	-0.073	0.097	-0.010	0.305	-0.089	0.066	0.403	0.295

Table 5. Correlation coefficients among grain yield and other associated quantitative traits in finger millet based on combined mean data. FD, days to 50% flowering; MD, days to 80% maturity; PH, plant height; T/P, number of tillers per plant; LL, flag leaf length; LW, flag leaf width; SL, flag leaf sheath length; EE, ear exsertion; EL, ear length; EW, ear width; F/H, number of fingers per head; FL, length of longest finger; FW, width of longest finger; W/H, weight per head; TW, 1000-grain weight; GY, grain yield; SY, straw yield; **, significant at 1% level; *, significant at 5% level.

Trait	**FD**	**MD**	**PH**	**T/P**	**LL**	**LW**	**SL**	**EE**	**EL**	**EW**	**F/H**	**FL**	**FW**	**W/H**	**TW**	**GY**
MD	0.93**															
PH	0.21**	0.26**														
T/P	–0.06	–0.06	0.11													
LL	0.27**	0.32**	0.33**	0.01												
LW	–0.03	0.04	0.02	0.02	0.07											
SL	–0.26**	–0.25**	0.33**	0.05	0.16**	–0.01										
EE	–0.04	–0.06	0.33**	0.04	0.20**	–0.05	0.21**									
EL	0.24**	0.29**	0.30**	0.10	0.13*	0.00	–0.03	0.10								
EW	0.25**	0.26**	0.10	0.05	0.22**	0.00	–0.14*	0.09	0.37**							
F/H	0.53**	0.50**	0.31**	0.00	0.28**	0.02	0.03	0.02	0.14*	0.17**						
FL	0.34**	0.38**	0.29**	0.10	0.11	0.01	–0.04	0.06	0.63**	0.27**	0.11					
FW	–0.08	–0.09	0.01	–0.01	0.05	–0.10	0.01	0.06	0.07	0.06	–0.10	0.00				
W/H	0.20**	0.23**	0.39**	0.17**	0.22**	0.05	0.07	0.12*	0.32**	0.12*	0.27**	0.24**	0.02			
TW	–0.08	–0.06	0.10	0.20	0.04	0.08	0.12*	0.02	0.12*	0.04	–0.18**	0.15*	0.04	0.16*		
GY	–0.26**	–0.13*	0.46**	0.33**	0.23**	0.09	0.35**	0.23**	0.19**	0.02	–0.03	0.16*	0.10	0.30**	0.36**	
SY	0.59**	0.58**	0.25**	–0.03	0.30**	–0.03	–0.13*	0.13*	0.21**	0.15*	0.26**	0.28**	0.00	0.21**	0.04	–0.01

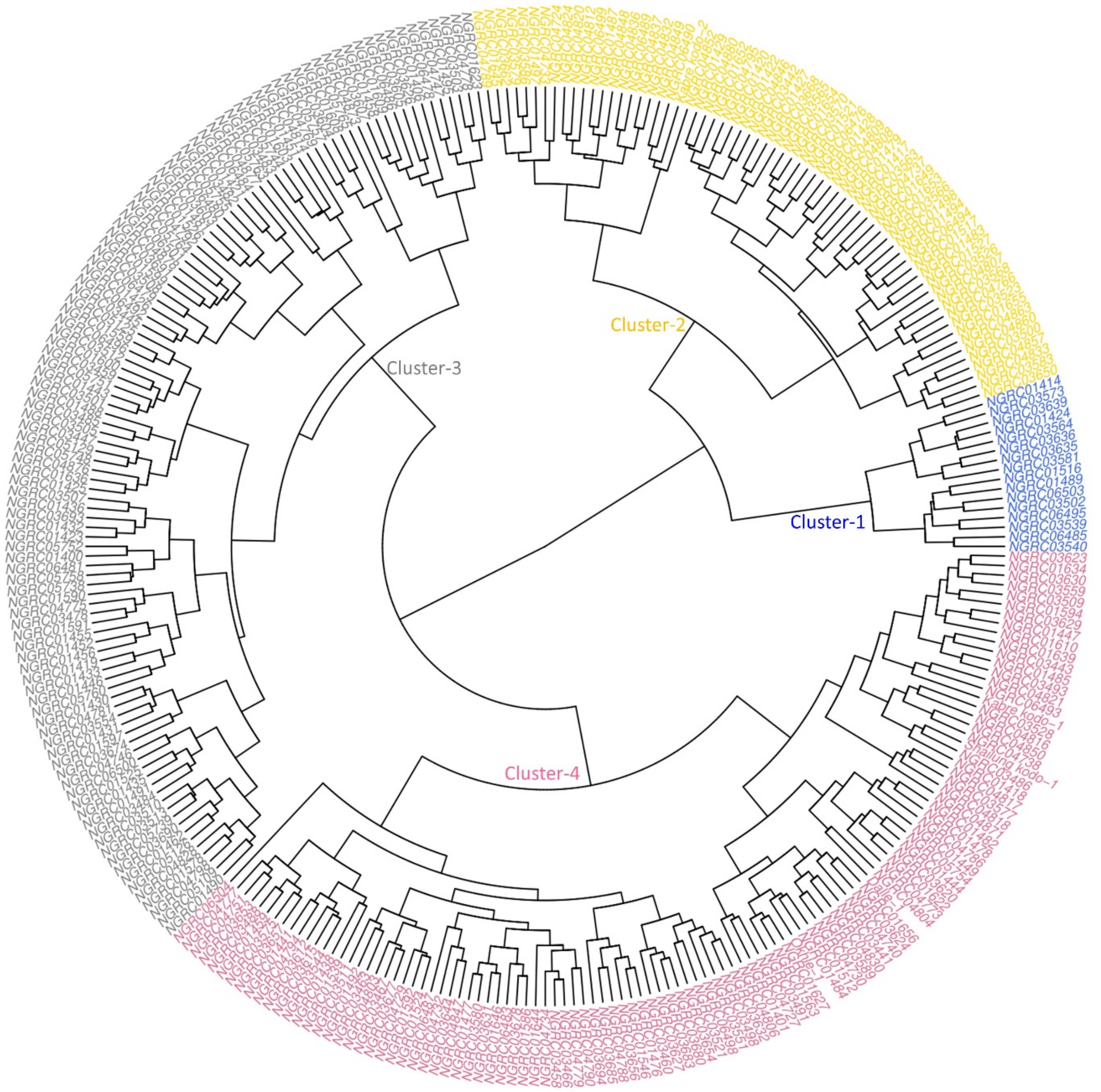


Figure 3. UPGMA hierarchical clustering divided 300 finger millet accessions into four clusters. Landraces are coded with their accession number whereas released varieties are with their name.

country. Shannon-Weaver diversity index (H') considers both richness and evenness of the phenotypic classes of the qualitative traits but emphasizes the normality in observation of quantitative traits (Ghimire *et al*, 2018b; Upadhyaya *et al*, 2010). We estimated as low, medium and high diversity if the values of H' were $\leq$ 0.400, 0.401–0.600 and $\geq$ 0.601, respectively (Eticha *et al*, 2005). Very low H' for grain shattering and finger branching was observed since $>$ 98% of the accessions didn't have shattering type of grains and branching type of fingers. According to the findings of Dasanayaka and Kaluthanthri (2017), a very small proportion of Sri Lankan finger millet accessions exhibited shattering and finger branching traits. The rest of the qualitative traits (H' = 0.645–0.908) as well as all 17 quantitative traits (H' = 0.864–0.907) showed very high polymorphism. Similar diversity indices were reported in the entire global collections of 5,940 accessions (3,567 from 14 African countries, 2,163 from five South-Asian countries, seven from the USA, 22 from three European countries and 181 from unknown origin), a core collection of 622 accessions and a mini-core collection of 80 accessions (Upadhyaya *et al*, 2006, 2010) as well as in an East African collection of 1,993 accessions (Reddy *et al*, 2009). Phenotypic proportions of qualitative traits were calculated as the frequencies of each descriptor state for the traits. Observations on ear shape with droopy, open, compact and semi-compact suggested the presence of all cultivated races (elongata, plana, compacta and vulgaris) of *E. coracana* ssp. *coracana* (Upadhyaya *et al*, 2006; Bharathi, 2011; Sood *et al*, 2019; Backiyalakshmi *et al*, 2021). Our observations for other qualitative traits were similar as in core and mini-

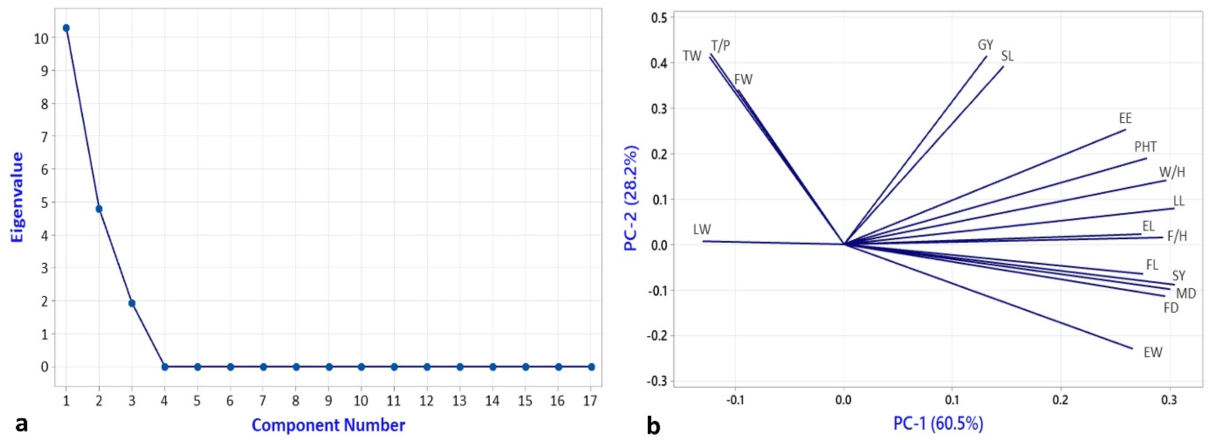


Figure 4. Scree plot (**a**) and loading plot (**b**) of PC-1 and PC-2 for 300 finger millet accessions based on cluster means, showing association between the traits. EW, ear head width; FD, days to flowering; MD, days to maturity; SY, straw yield; FL, length of the longest finger; F/H, fingers per head; EL, ear head length; LL, flag leaf length; W/H, weight per head; PHT, plant height; EE, ear exsertion; SL, flag leaf sheath length; GY, grain yield; FW, width of the longest finger; T/P, tillers per plant; TW, weight of 1,000 grains; LW, flag leaf width.

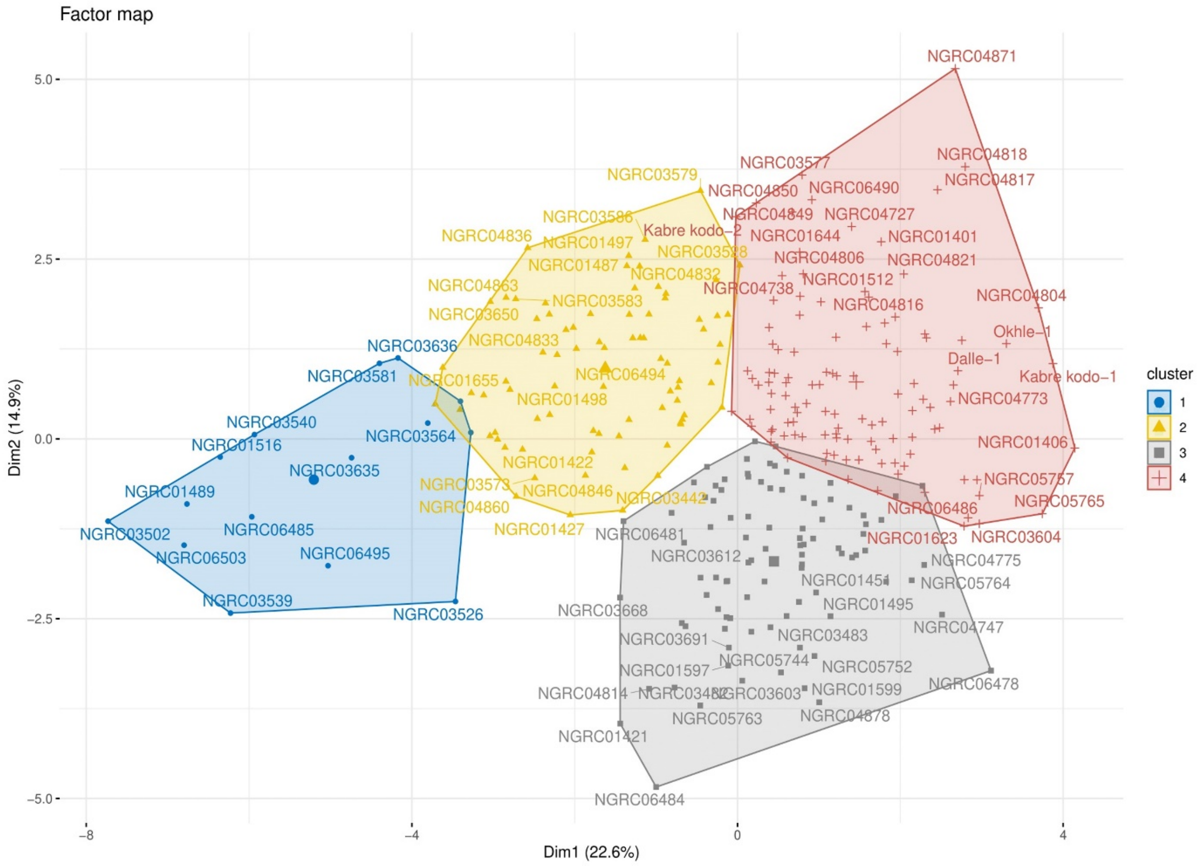


Figure 5. Scatter plot of the first two dimensions of principal component analysis using entry means for 300 finger millet accessions. Landraces are coded with their accession number whereas released varieties are with their name. Cluster centroid in respective cluster is indicated with larger symbol. The representative accessions of each cluster are indicated with accession numbers and are colour-coded as in Figure 3.

Table 6. Promising finger millet trait donors selected based on combined mean data of six environments. Accessions with multiple promising traits are highlighted in bold.

Trait	Promising accessions
Days to flowering (< 90 d)	NGRC03581, NGRC03540, NGRC03502, NGRC01516, NGRC01489, NGRC03635, NGRC03636, NGRC03539, NGRC06503, NGRC06485, NGRC03650
Days to maturity (> 130 d)	NGRC03502, NGRC03540, NGRC01489, NGRC03581, NGRC03636, NGRC03539, NGRC06503
Plant height (< 70cm)	NGRC06503, NGRC03502, NGRC03639, NGRC04814
Plant height (> 110cm)	Dalle-1, **NGRC04871**, NGRC03511, **NGRC04818**, NGRC05764
Number of tillers/plant (> 5.4)	**NGRC06490**, **Kabre kodo-2**, NGRC04852, NGRC01609, NGRC03579, NGRC05739
Flag leaf length (> 30cm)	NGRC05109, NGRC01490, NGRC03605, NGRC04852, NGRC06493, Kabre kodo-1, NGRC03678, NGRC03690, NGRC04746, NGRC01401
Flag leaf sheath length (> 17.5cm)	NGRC01655, **NGRC04871**, NGRC04724, NGRC06498, **NGRC04818**, NGRC03528, NGRC06504, NGRC04863
Ear exsertion (> 14cm)	NGRC06493, NGRC01527, NGRC01610, NGRC03693
Ear length (> 7.5cm)	**NGRC04871**, NGRC01406, NGRC04821, NGRC01447, NGRC04817, NGRC01458, NGRC01609, NGRC01451
Weight per head (> 8g)	**NGRC04818**, Dalle-1, NGRC04804, NGRC01446, NGRC04806, NGRC01401, NGRC01639, NGRC01487
1000-grain weight (> 2.8g)	**Kabre kodo-2**, NGRC04816, **NGRC04849**, NGRC01418, NGRC04873, NGRC04850, NGRC04824, **NGRC04871**, **NGRC04818**
Grain yield (> 2,769kg/ha)	**NGRC04849**, **NGRC06490**, **NGRC04871**, **Kabre kodo-2**, NGRC06487, NGRC04727, NGRC04836, NGRC04806, **NGRC04818**
Straw yield (> 13t/ha)	Kabre kodo-1, NGRC01456, NGRC01452, Okhle-1, NGRC01455, NGRC01539, NGRC05758

core collections of global accessions (Upadhyaya *et al*, 2010), in Sri Lankan accessions (Dasanayaka and Kaluthanthri, 2017; Kumari *et al*, 2018), in north-west Indian accessions (Kumar *et al*, 2019) and in global collections from ICRISAT (Malambane and Jaisil, 2015). Descriptive statistics such as range, mean and CV of various quantitative variables suggested wide variability within 300 accessions of the Nepalese collection. Mean flowering data ranged from 75 to 140 days which was narrower than in the global collections which included wild species (41–164 days) (Bharathi, 2011) but wider than in the core collection (51–96 days), the mini-core subsets (51–93 days) (Upadhyaya *et al*, 2010) and in a global collection of 314 accessions (51–97 days) (Backiyalakshmi *et al*, 2021). Since our research was conducted at higher altitudes, we observed mean flowering time of 112 days (pooled mean of three sites) which is higher than in studies conducted at lower elevations, which reported 79 days (Bhattarai *et al*, 2014), 89 days (Bastola *et al*, 2015) and 74 days (Backiyalakshmi *et al*, 2021). Our range for average plant height (61–119cm) was narrower compared to the range of 23–155cm observed in a much larger sample (537) of Nepalese accessions (Bhattarai *et al*, 2014), 55–240cm in east African accessions (Reddy *et al*, 2009), 84–143cm in a global collection (Backiyalakshmi *et al*, 2021) and 45–180cm in a global composite collection of 1,000 accessions (Bharathi, 2011), but wider than the range of 73–113cm in the mini-core collection (Upadhyaya *et al*, 2010).

Principal component analysis partitions the total variation into components and measures how each component contributes to the total phenotypic variation. The important traits in the evaluation of our landraces are days to maturity and flowering, straw yield, fingers per head, finger length, plant height, grain yield and 1,000-grain weight. The scatter plot (Figure 5) suggested that PC-2 has efficiently divided the accessions based on grain yield. High-yielding landraces and all released varieties (Shailung kodo-1, Dalle-1, Kabre kodo-1, Kabre kodo-1 and Okhle-1) were in the upper half of the scatter plot. Clustering of observations gives the average of a variable in the cluster, the average of the variable for the whole data set, the associated standard deviations and the p-value to test the hypothesis: the cluster mean is equal to the overall mean. Rejection of this hypothesis ($p < 0.05$) means there is significant difference between the clusters for multiple quantitative traits. The representative genotypes of each cluster illustrate the overall characters of that cluster. Pearson's correlation coefficient suggested that grain yield is positively correlated with plant height, tillers, leaf length, ear exsertion, ear and finger length, ear weight and 1,000-grain weight. This suggests that our focus when selecting landraces for grain yield should be on accessions with early flowering, taller height, high tillering, big ears and bold grains. Previous findings suggested a positive correlation between grain yield vs. days to flowering and maturity (Bharathi, 2011; Lule *et al*, 2012; Bastola *et al*, 2015; Patel *et al*, 2017), but we observed contrasting results i.e. negative correlation of grain yield with flowering and maturity days. This can be explained because the early flowering genotypes escaped cold stress in higher altitudes like Jumla where many late maturing accessions yielded zero due to susceptibility to low temperatures. Early maturity is, therefore, a very important trait for mountain farmers cultivating millets in dry and rainfed conditions to escape drought as well as cold stress during the grain-filling stage. Straw yield had a strong positive association with days to flowering, maturity, plant height, leaf length, ear length, finger length, ear weight and number of fingers which is in support of the findings of Bharathi (2011); Bastola *et al* (2015); Patel *et al* (2017); Backiyalakshmi *et al* (2021). This means the accessions with taller plant height and late maturity produced higher straw yield which is a very important trait for farmers who are growing livestock in their farm and use millet straw for feed.

Although finger millet is the fourth most important cereal in Nepal, its cultivation area has been shrinking over the last decade (MOALD, 2022). One of the reasons behind this is the limited options of high-yielding varieties accessible to farmers. Average grain yield in our research ranged from 230–3,494kg/ha but 15 accessions produced an average grain yield of more than 2,500kg/ha, which is comparable to commercial varieties. Research conducted by Anuradha *et al* (2022) on Indian genotypes and by Wolie and Belete (2012) on Ethiopian genotypes also reported some promising genotypes with more than 3,000kg/ha grain yield. Since Nepal is a member of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), exotic high-yielding genetic resources through the Multilateral System could be introduced and used in crossings between Asian and African genotypes to develop high-yielding lines exploiting the wide genetic variability to improve finger millet productivity of the country.

Conclusion

We observed wide phenotypic variation within Nepalese finger millet landraces conserved at NAGRC for various qualitative and quantitative traits. We identified landraces with high potential for functional agronomic traits such as higher grain and straw yield, early maturity, bold grains, which could be utilized as trait-specific donors in paving the breeding pathway of finger millet. These selected landraces are currently being tested under national coordinated varietal trials by HCRP throughout the country. Furthermore, these landraces could be deployed to similar hilly environments without further delay to enrich the varietal options for finger millet-growing farmers of Nepal.

Acknowledgements

This research was funded by UNEP/GEF and technically coordinated by the Alliance of Bioversity International and CIAT. We acknowledge the logistic support received from NAGRC Lalitpur, HCRP Dolakha and ARS Jumla for the field experiments.

Authors' contribution

KHG conceived the idea, conducted the field experiments, analyzed the data and drafted the manuscript. All co-authors guided in the field experiment and contributed to writing and revising the manuscript.

Conflict of interest

We declare that there is no conflict of interest among authors regarding this work.

Supplemental data

Supplemental Table 1. List of the 300 finger millet accessions used in the study with their local names and collection sites. The accessions with asterisk (*) are released varieties whereas others are landraces.

Supplemental Table 2. Entry means of 300 finger millet accessions for 17 quantitative traits (combined over 6 environments). The accessions with asterisk (*) are released varieties whereas others are landraces. FD, days to flowering; MD, days to maturity; PHT, plant height; T/P, tillers per plant; LL, flag leaf length; LW, flag leaf width; SL, flag leaf sheath length; EE, ear exsertion; EL, ear head length; EW, ear head width; F/H, fingers per head; FL, length of the longest finger; FW, width of the longest finger; W/H, weight per head; TW, weight of 1000 grains; GY, grain yield; SY, straw yield.

References

Anuradha, N., Patro, T. S. S. K., Singamsetti, A., Sandhyarani, Y., Triveni, U., Nirmalakumari, A., Govanakoppa, N., Lakshmipathy, T., and Tonapi, V. A. (2022). Comparative study of AMMI- and BLUP-based simultaneous selection for grain yield and stability of finger millet. *genotypes. Front. Plant Sci* 12(786839). doi: https://doi.org/10.3389/fpls.2021.786839

Backiyalakshmi, C., Babu, C., Reddy, D. N., Padmakumar, V., Prasad, K. V. S. V., Azevedo, V. C. R., and Vetriventhan, M. (2021). Assessing Forage Potential of the Global Collection of Finger Millet (*Eleusine coracana* (L.) Gaertn.) Germplasm Conserved at the ICRISAT Genebank. *Agronomy* 11, 1706–1706. doi: https://doi.org/10.3390/agronomy11091706

Bastola, B. R., Pandey, M. P., Ojha, B. R., Ghimire, S. K., and Baral, K. (2015). Phenotypic diversity of Nepalese finger millet (*Eleusine coracana* (L.) Gaertn) accessions at IAAS, Rampur, Nepal. *Int. J. Appl. Sci. Biotechnol* 3(2), 285–290. doi: https://doi.org/10.3126/ijasbt.v3i2.12413

Bharathi, A. (2011). Phenotypic and genotypic diversity of global finger millet (*Eleusine coracana* (L.) Gaertn) composite collection. Ph.D. thesis, Tamil Nadu Agricultural University, Coimbatore, India.

Bhattarai, M., Ghimire, K. H., Joshi, B. K., and Bhatta, M. R. (2014). Characterization of finger millet (*Eleusine coracana* Gaertn.) germplasm with agro-morphological markers. In Khadka, Y. G., Giri, Y. P., Sah, B. P., Mahto, B. N., Khatiwada, S. P., Bhatta, M., Khatri-Chhetri, B. B., Gautam, A. K., Ansari, A. R., Ranjit, J. D., Gauchan, D., Shrestha, R., and Sapkota, B., Proceedings of the 27th National Summer Crops Workshop, 18-20 April 2013, Chitwan, Nepal, volume 2.

Bora, P. (2013). Nutritional properties of different millet types and their selected products. University of Guelph, Ontario, Canada. url: https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/7738/Bora_Pragyani_201312_Msc.pdf?sequence=1.

Chandra, D., Chandra, S., Pallavi, and Sharma, A. K. (2016). Review of Finger millet (*Eleusine coracana* (L.) Gaertn): A power house of health benefiting nutrients. *Food Science and Human Wellness* 5, 149–155. doi: http://dx.doi.org/10.1016/j.fshw.2016.05.004

Dasanayaka, P. S. and Kaluthanthri, D. V. S. (2017). Genetic variability characterization of some ex situ conserved finger millet (*Eleusine coracana* L. Gaertn. ssp. *coracana*) germplasm accessions in Sri Lanka using qualitative morphological markers. *Int. J. Adv. Res* 5(5), 1381–1389. doi: https://dx.doi.org/10.21474/IJAR01/4259

Devi, P. B., Bharathi, V. R., Sathyabama, S., Malleshi, N. G., and Priyadarisini, V. B. (2011). Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. *J. Food Sci. Technol* 51, 1021–1040. doi: https://doi.org/10.1007/s13197-011-0584-9

Dida, M. M., Srinivasachary, Ramakrishnan, S., Bennetzen, J. L., Gale, M. D., and Devos, K. M. (2007). The genetic map of finger millet *Eleusine coracana*. *Theor Appl Genet* 114, 321–332. doi: https://doi.org/10.1007/s00122-006-0435-7

Eticha, F., Bekele, E., Belay, G., and Borner, A. (2005). Phenotypic diversity in tetraploid wheats collected from Bale and Wello regions of Ethiopia. *Plant Genet Res* 3(1), 35–43. doi: https://doi.org/10.1079/PGR200457

Gaihre, S., Gauchan, D., and Timsina, K. P. (2021). Prospect and potentiality of finger millet in Nepal: Nutritional security and trade perspective. *J Agric. Nat. Res* 4(2), 63–74. doi: https://doi.org/10.3126/janr.v4i2.33657

Ghimire, K. H., Bhandari, B., Gurung, S. B., Dhami, N. B., and Baniya, B. K. (2017). Diversity and utilization status of millets genetic resources in Nepal. In Joshi, B. K., Khatri-Chhetri, H. B., and Acharya, A. K., Proceedings of 2nd National Workshop on Conservation and Utilization of Agricultural Plant Genetic Resources in Nepal, 22-23 May 2017 Dhulikhel. NAGRC, FDD, DoA and MoAD, Nepal.

Ghimire, K. H., Joshi, B. K., Dhakal, R., and Sthapit, B. R. (2018a). Diversity in proso millet (*Panicum miliaceum* L.) landraces collected from Himalayan Mountains of Nepal. *Genet Resour Crop Evol* 65, 503–512. doi: https://doi.org/10.1007/s10722-017-0548-7

Ghimire, K. H., Joshi, B. K., Gurung, R., and Sthapit, B. R. (2018b). Nepalese foxtail millet [*Setaria italica* (L.) P. Beauv.] genetic diversity revealed by morphological markers. *Genet Resour Crop Evol* 65, 1147–1157. doi: https://doi.org/10.1007/s10722-017-0602-5

Goron, T. L., Bhosekar, V. K., Shearer, C. R., Watts, S., and Raizada, M. N. (2015). Whole plant acclimation responses by finger millet to low nitrogen stress. *Front. Plant Sci* 6, 652–652. doi: https://doi.org/10.3389/fpls.2015.00652

Hilu, K. W. and De-Wet, J. M. J. (1976). Domestication of *Eleusine coracana*. *Econ. Bot* 30(3), 199–208. doi: https://doi.org/10.1007/BF02909728

Hilu, K. W., De-Wet, J. M. J., and Harlan, J. R. (1979). Archaeobotanical studies of *Eleusine coracana* ssp. *coracana* (Finger millet). *Amer. J. Bot* 66(3), 330–333. doi: https://doi.org/10.2307/2442610

Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., Lohithaswa, H. C., and Mohanrao, A. (2017). Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) Gaertn) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18, 1–16. doi: https://doi.org/10.1186/s12864-017-3850-z

Hodgkin, T., Rao, V. R., Cibrian-Jaramillo, A., and Gaiji, S. (2003). The use of ex situ conserved plant genetic resources. *Plant Genet Res* 1(1), 19–29. doi: https://doi.org/10.1079/PGR200313

IBPGR (1985). Descriptors for finger millet (Rome, Italy: International Board for Plant Genetic Resources), 23p. url: https://hdl.handle.net/10568/72880.

Joshi, B. K., Joshi, D., and Ghimire, S. K. (2020). Genetic Diversity in Finger Millet Landraces Revealed by RAPD and SSR Markers. *Nepal J Biotechnol* 8(1), 1–11. doi: https://doi.org/10.3126/njb.v8i1.30204

Kassambara, A. and Mundt, F. (2020). Factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.7. url: https://CRAN.R-project.org/package=factoextra.

Kumar, A., Chawla, H. S., Jeena, A. S., and Rohit (2019). Morphological characterization of finger millet germplasm collected from Uttarakhanda hills for qualitative traits. *Int. J. Curr. Microbiol. App. Sci* 8(9), 2105–2109. doi: https://doi.org/10.20546/ijcmas.2019.809.243

Kumar, A., Metwal, M., Kaur, S., Gupta, A. K., Puranik, S., Singh, S., Singh, M., Gupta, S., Babu, B. K., Sood, S., Yadav, and R (2016). Nutraceutical value of finger millet [*Eleusine coracana* (L.) Gaertn] and their improvement using omics approaches. *Front. Plant Sci* 7. doi: https://doi.org/10.3389/fpls.2016.00934

Kumari, W. M. R., Pushpakumara, D. K. N. G., Weerakoon, W. M. W., Senanayake, D. M. J. B., and Upadhyaya, H. D. (2018). Morphological Characterization of Local and Introduced Finger millet (*Eleusine coracana* (L.) Gaertn) Germplasm in Sri Lanka. *Tropical Agricultural Research* 29(2), 167–183. doi: https://doi.org/10.4038/tar.v29i2.8287

Liu, Q., Jiang, B., Wen, J., and Peterson, P. M. (2014). Low-copy nuclear gene and McGISH resolves polyploid history of *Eleusine coracana* and morphological character evolution in *Eleusine*. *Turk. J. Bot* 38, 1–12. doi: https://doi.org/10.3906/bot-1305-12

Lule, D., Tesfaye, K., Fetene, M., and Villiers, S. (2012). Multivariate analysis for quantitative traits in finger millet (*Eleusine coracana* ssp. *Coracana*) population collected from eastern and southern Africa: detection for patterns of genetic diversity. *Int. J. Agric. Res* 7(6), 303–314. doi: https://doi.org/10.3923/ijar.2012.303.314

Malambane, G. and Jaisil, P. (2015). Morphological variability for qualitative and quantitative traits in finger millet (*Eleusine coracana* L. Gaertn). *Journal of Advances in Agriculture* 5(1), 527–537. doi: https://doi.org/10.24297/jaa.v5i1.4263

MINITAB (2010). Minitab 17 Statistical Software (State College, Pennsylvania: Minitab, Inc). url: www.minitab.com.

MOALD (2022). Statistical Information on Nepalese Agriculture 2077/78 (2020/2021). Government of Nepal, Ministry of Agriculture and Livestock Development, Kathmandu, Nepal. url: https://moald.gov.np/publication-types/agriculture-statistics/.

Nakarani, U. M., Singh, D., Suthar, K. P., Karmakar, N., Faldu, P., and Patil, H. E. (2021). Nutritional and phytochemical profiling of nutracereal finger millet (*Eleusine coracana* L.) genotypes. *Food Chemistry* 314 (Pt 2):128271. doi: https://doi.org/10.1016/j.foodchem.2020.128271

Nguyen, G. N. and Norton, S. L. (2020). Genebank Phenomics: A Strategic Approach to Enhance Value and Utilization of Crop Germplasm. *Plants* 9, 817–817. doi: https://doi.org/10.3390/plants9070817

O'Kennedy, M. M., Grootboom, A., and Shewry, P. R. (2006). Harnessing sorghum and millet biotechnology for food and health. *J. Cereal Sci* 44(3), 224–235. doi: https://doi.org/10.1016/j.jcs.2006.08.001

Parvathi, M. S., Nataraja, K. N., Reddy, Y. A. N., Naika, M. B. N., and Gowda, M. V. C. (2019). Transcriptome analysis of finger millet (*Eleusine coracana* (L.) Gaertn) reveals unique drought responsive genes. *Journal of Genetics* 98, 46–46. doi: https://doi.org/10.1007/s12041-019-1087-0

Patel, S. N., Patil, H. E., and Popat, R. C. (2017). Genetic diversity study in finger millet (*Eleusine coracana* L.) genotypes: A multivariate analysis approach. *Int. J. Pure App. Biosci* 5(5), 183–189. doi: http://dx.doi.org/10.18782/2320-7051.5430

Phillips, S. M. (1972). A Survey of the Genus *Eleusine* Gaertn. (Gramineae) in Africa. *Kew Bulletin* 27(2), 251–270. doi: https://doi.org/10.2307/4109450

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing (Vienna, Austria). url: https://www.R-project.org/.

Reddy, V. G., Upadhyaya, H. D., Gowda, C. L. L., and Singh, S. (2009). Characterization of eastern African finger millet germplasm for qualitative and quantitative characters in ICRISAT. *Journal of SAT Agricultural Research* 7(1). url: http://oar.icrisat.org/id/eprint/2557.

Sebastien, L., Josse, J., and Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25(1), 1–18. doi: https://doi/org/10.18637/jss.v025.i01

Shannon, C. E. and Weaver, W. (1949). The mathematical theory of communication (Chicago: The University of Illinois), 3-24.

Sood, S., Joshi, D. C., Chandra, A. K., and Kumar, A. (2019). Phenomics and genomics of finger millet: current status and future prospects. *Planta* 250, 731–751. doi: https://doi.org/10.1007/s00425-019-03159-6

Upadhyaya, H. D., Gowda, C. L. L., Pundir, R. P. S., Reddy, G. V., and Singh, S. (2006). Development of core sub set of finger millet germplasm using geographical origin and data on 14 quantitative traits. *Genet. Resour. Crop Evol* 53, 679–685. doi: https://doi.org/10.1007/s10722-004-3228-3

Upadhyaya, H. D., Gowda, C. L. L., and Reddy, G. V. (2007). Morphological diversity in finger millet germplasm introduced from southern and eastern Africa. *J. SAT Agric. Res* 3(1), 1–3. url: http://oar.icrisat.org/2656/1/Morphological_diversity_in_finger.pdf.

Upadhyaya, H. D., Ramesh, S., Sharma, S., Singh, S. K., Varshney, S. K., Sarma, N. D. R. K., Ravishankar, C. R., Narasimhudu, Y., Reddy, V. G., Sahrawat, K. L., Dhanalakshmi, T. N., Mgonja, M. A., Parzies, H. K., Gowda, C. L. L., and Singh, S. (2011). Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. *Field Crops Research* 121, 42–52. doi: https://doi.org/10.1016/j.fcr.2010.11.017

Upadhyaya, H. D., Sarma, N. D. R. K., Ravishankar, C. R., Albrecht, T., Narasimhudu, Y., Singh, S. K., Varshney, S. K., Reddy, V. G., Singh, S., Dwivedi, S. L., Wanyera, N., Oduori, C. O. A., Mgonja, M. A., Kisandu, D. B., Parzies, H. K., and Gowda, C. L. L. (2010). Developing a mini-core collection in finger millet using multilocation data. *Crop Sci* 50, 1924–1931. doi: https://doi.org/10.2135/cropsci2009.11.0689

Vetriventhan, M., Upadhyaya, H. D., Dwivedi, S. L., Pattanashetti, S. K., and Singh, S. K. (2016). Finger and foxtail millets. In *Genetic and genomic resources for grain cereals improvement,* ed. Singh, M. and Upadhyaya, H. D., (San Diego: Academic Press), 291-319.

Vetriventhon, M., Azevedo, V. C. R., Upadhyaya, H. D., Nirmalakumari, A., Kane-Potaka, J., Anitha, S., Ceasar, S. A., Muthamilarasan, M., Bhat, B. V., Hariprasanna, K., Bellundagi, A., Cheruku, D., Backiyalakshmi, C., Santra, D., Vanniarajan, C., and Tonapi, V. A. (2020). Genetic and genomic resources, and breeding for accelerating improvement of small millets: current status and future interventions. *Nucleus* 63, 217–239. doi: https://doi.org/10.1007/s13237-020-00322-3

VSN International (2015). Genstat for Windows 18th Edition (Hemel Hempstead, UK: VSN International). url: https://www.genstat.co.uk/.

Wolie, A. and Belete, K. (2012). Genetic divergence and variability studies in some Ethiopian finger millet germplasm collections. *Scholarly J. Agric. Sci* 3(4), 110–116. url: https://www.researchgate.net/publication/315892606.

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2023), 4 (8), 15–28
DOI: 10.46265/genresj.MRBT4299
https://www.genresj.org
ISSN: 2708-3764

Is the ecosystem services concept relevant to capture the multiple benefits from farming systems using livestock biodiversity? A framework proposal

Anne Lauvie [*,a], Gisèle Alexandre [b], Valérie Angeon [c], Nathalie Couix [d], Olivia Fontaine [a], Claire Gaillard [e], Michel Meuret [a], Catherine Mougenot [f], Charles-Henri Moulin [a], Michel Naves [b], Marie-Odile Nozières-Petit [a], Jean-Christophe Paoli [g], Lola Perucho [g], Jean Michel Sorba [g], Emmanuel Tillard [a] and Etienne Verrier [h]

[a] *UMR Systèmes d'élevage méditerranéens et tropicaux (SELMET), INRAE – CIRAD – Institut Agro Montpellier – Univ. Montpellier, 34060 Montpellier cedex 01, Montpellier, France*

[b] *ASSET (Agroécologie, génétique et SystèmeS d'Elevage Tropicaux), INRAE, 97170 Petit-bourg, France*

[c] *UR Ecodéveloppement, INRAE, 84914 Avignon cedex 9, France*

[d] *UMR AGIR, INRAE-Université de Toulouse, 31326 Castanet Tolosan cedex, France*

[e] *AgroSup Dijon, INRAE UMR Territoires, 21079 Dijon Cedex, France*

[f] *Arlon Campus Environnement, Université de Liège, 6700 Arlon, Belgium*

[g] *INRAE Selmet-LRDE, 20250 Corte, France*

[h] *Université Paris-Saclay, AgroParisTech, INRAE UMR GABI, 91120 Palaiseau, France*

Abstract: Local breeds are key components of livestock farming systems. They are part of livestock biodiversity and this diversity has been threatened since the second half of the 20th century by their replacement with animals from specialized breeds. The multiple benefits of farming systems using local breeds – provision of goods, landscape and environmental management, and uses related to cultural and heritage dimensions – have long been recognized and used to argue for their conservation. However, the notion of ecosystem services is rarely used to analyze those benefits. This article presents a qualitative approach to the provision of ecosystem services by farming systems that use livestock biodiversity. Based on diverse case studies of breeds from several species, we propose an analytical framework that accounts for how a service is qualified, who is concerned by the services identified, the role of the breed in the process of service provision, and interactions between services. Finally, the framework considers the links between the provision of services and the management of the breeds. We discuss to what extent the notion of ecosystem services is useful in dealing with the multiple benefits from farming systems using local breeds.

Keywords: Livestock biodiversity, ecosystem services, inductive approach, interdisciplinary, local breeds

Citation: Lauvie, A., Alexandre, G., Angeon, V., Couix, N., Fontaine, O., Gaillard, C., Meuret, M., Mougenot, C., Moulin, C., Naves, M., Nozières-Petit, M., Paoli, J., Perucho, L., Sorba, J. M., Tillard, E., Verrier, E. (2023). Is the ecosystem services concept relevant to capture the multiple benefits from farming systems using livestock biodiversity? A framework proposal. *Genetic Resources* 4 (8), 15–28. doi: 10.46265/genresj.MRBT4299.

Introduction

Livestock biodiversity has been threatened since the second half of the 20^{th} century, as many local breeds have been replaced by a few specialized breeds and hybrids, which, among other traits, have been intensively selected to increase production, and have become mainstream breeds (Audiot, 1995; FAO, 2015). However, local breeds are both a resource for and a product of livestock farming. Globally, FAO (2015) reported

---

[*] Corresponding author: Anne Lauvie (anne.lauvie@inrae.fr)

Received: 22.03.2023 | Accepted: 15.06.2023

a total of 8,774 breeds from 38 domesticated animal species, resulting from long-term livestock farming in diverse territories with different goals. The first challenge for livestock biodiversity faced by stakeholders and the scientific community, is its conservation, mainly focused on breed conservation and within-breed genetic variability management (Hall and Bradley, 1995; Ajmone-Marsan and Consortium GLOBALDIV, 2010). Beyond this conservation challenge, adding value to local breeds has been underlined as a key process for farm animal biodiversity, in particular through the production of quality food products (Verrier *et al*, 2005; Mathias *et al*, 2010; Ligda and Casabianca, 2013). Various authors in the field of local breeds management have stressed that the supply of animal products is associated with other benefits: provision of a high diversity of food and other goods, services related to landscape and environmental management, uses related to cultural and heritage dimensions (Audiot, 1995; Alexandre *et al*, 2002; Gandini and Villa, 2003; Rege and Gibson, 2003; Verrier *et al*, 2005; Berland *et al*, 2006; Fontaine *et al*, 2008; Naves *et al*, 2011; Leroy *et al*, 2018; Hall, 2019). These material and immaterial benefits have been used as one of the arguments to underline the importance of local breeds' conservation (FAO, 2015).

Few authors use ecosystem service approaches to analyze the material and immaterial benefits provided by the raising of local breeds (Hoffmann *et al*, 2014). However, from being a simple metaphor to raise public awareness (Norgaard, 2010; Barnaud *et al*, 2011), since the Millennium Ecosystem Assessment (2005), the ecosystem services concept is being increasingly used (Kull *et al*, 2015; Droste *et al*, 2018). A prolific literature highlights the diversity of ecosystem services produced or used by agricultural ecosystems (Zhang *et al*, 2007) and how it is linked to human practices (Lescourret *et al*, 2015). Up to now, when the ecosystem services approach was applied to livestock biodiversity, it was mainly at a global scale, through the generic inventory of ecosystem services (Leroy *et al*, 2018), or to focus on specific dimensions of them, like the cultural and heritage values of breed diversity in the Alpine area (Marsoner *et al*, 2018).

Hall (2019) hypothesized that the underuse of this framework by stakeholders and the scientific community is partly due to a lack of recognition of livestock biodiversity by the ecosystem services community, and suggested it would be useful and beneficial that linkages be strengthened among several scientific communities. However, Velado-Alonso *et al* (2021) highlighted the interest in considering both cultural and ecological dimensions of the relationships between livestock breeds and ecosystem services.

Moreover, as underlined by Beudou *et al* (2017), the most commonly used ecosystem services approaches are quantitative and, as a consequence, neglect the social dimension of livestock farming systems, while qualitative approaches are relevant to comprehend the complex processes and interrelations underlying the production of ecosystem services (Barnaud *et al*, 2018). Indeed, we can wonder if a qualitative ecosystem service approach could help better understand the dynamics underlying the multiple benefits obtained from farming systems using livestock biodiversity.

The purpose of the present paper is to propose a qualitative approach to ecosystem services provision by different farming systems that exploit livestock diversity. Considering that the ecosystem services notion applies at the ecosystem scale, we do not look here at ecosystem services provided by livestock biodiversity but ecosystem services provided by farming systems using livestock biodiversity (Martin-Collado *et al*, 2019). In this paper, we present a framework for the analysis of ecosystem services provided by farming systems using livestock biodiversity focusing on the processes at play and the interactions supporting them. Then, we discuss to what extent the notion of ecosystem services is appropriate to deal with the multiple benefits obtained from farming systems using livestock biodiversity.

Materials and methods

Our work is based on empirical research on nine ruminant breeds in six different regions. All the case studies are located in France, including in French overseas territories, Reunion Island (Indian Ocean) and Guadeloupe Island (French West Indies). The locations were chosen to represent diverse (i) biogeographical and climatic conditions, (ii) species, (iii) main uses and (iv) population sizes (Figure 1). We analyzed data and empirical knowledge on those breeds and the associated livestock farming systems, taken from previous or ongoing studies (Table 1).

All the breeds concerned are local except one, the Montbéliarde breed, which is the second most important dairy cattle breed in France. For this breed, we repurposed work carried out in its cradle of origin (located in the east of France, in the Bourgogne Franche Comté Region), which allowed us to have a wide variety of situations on the fourth criterion mentioned above (Gaillard *et al*, 2018).

We developed an inductive approach (Woo *et al*, 2017) based on several working meetings that brought together researchers from different animal sciences, and social and economic sciences (the authors of the present paper). We share a research position that relies on the importance of a qualitative approach, tackling empirical issues to build realistic conceptualizations. We also share a vision of ecosystem services as social constructs: Barnaud *et al* (2018) illustrated that in such a vision, “an open landscape does not ‘naturally’ or ‘intrinsically’ provide a cultural ecosystem service but someone, in a given geographical, cultural, and historical context, attributes a specific patrimonial or aesthetic value to such landscape.” Indeed, even for the services that result from ecological dynamics, the way humans qualify them as services, recognizing them as such, has a social dimension. Consequently, our aim is not to describe ‘objective’ ecosystem services and

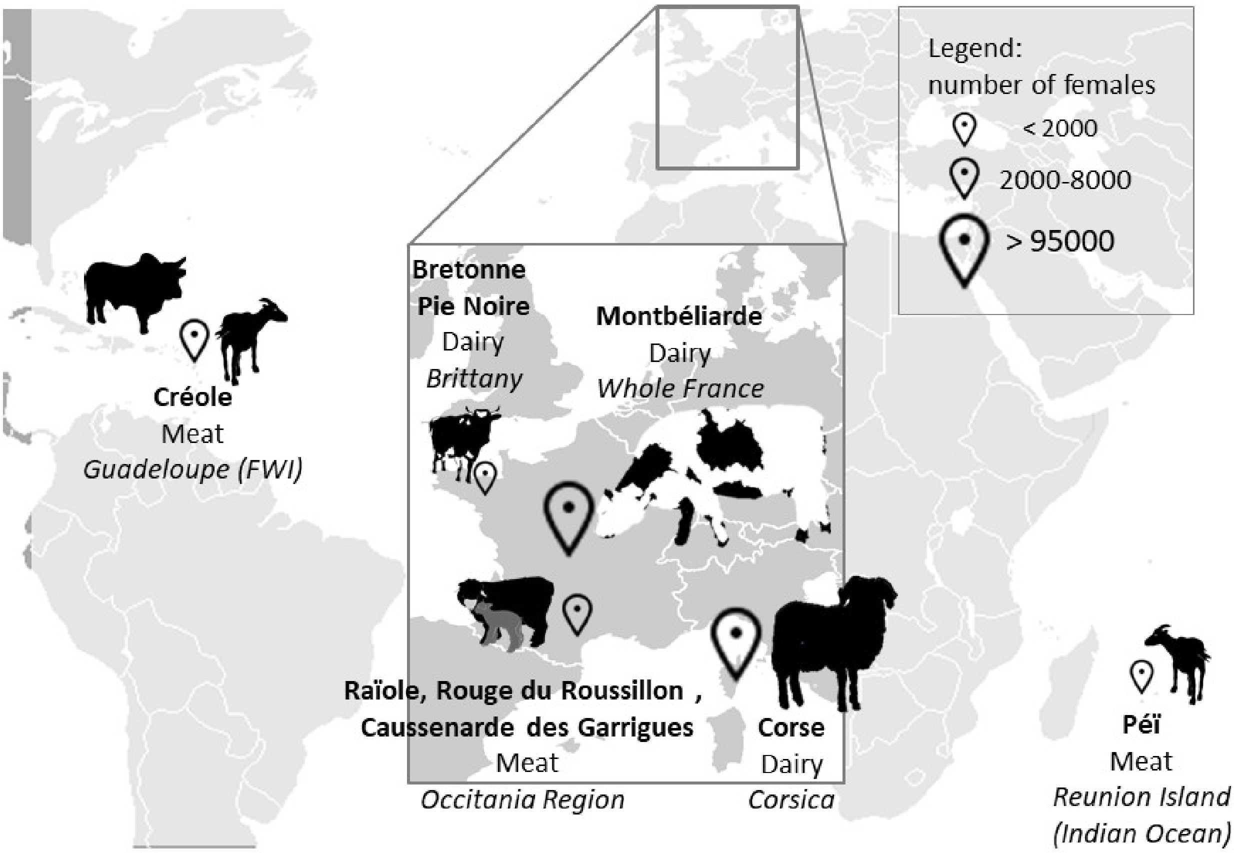


Figure 1. The six regions and the nine breeds analyzed in the present study (source of the population size estimates: INRA (2014)).

Table 1. References associated with the described case studies.

Case study	Species	References
Creole	Cattle and goat	Alexandre *et al* (2002); Gautier and Naves (2011); Naves *et al* (2011); Boval *et al* (2012); Gourdine *et al* (2021)
Bretonne Pie Noir	Cattle	Couix *et al* (2016); Lauvie *et al* (2017)
Peï	Goat	Fontaine *et al* (2008)
Montbéliarde	Cattle	Gaillard *et al* (2018)
Raïole, Caussenarde des Garrigues, Rouge du Roussillon	Sheep	Lauvie *et al* (2017); Nozieres-Petit and Lauvie (2018)
Corse	Sheep	Lauvie *et al* (2017); Perucho *et al* (2020)

quantify or assess them (and such a vision is not adapted to do this kind of assessment). Our aim is rather to take into account services as social constructs: such a vision is adapted to a qualitative and comprehensive approach of complex processes in agroecosystems.

First, we organized three online workshops in 2016, during which we shared our knowledge of the case studies (Table 1 and Figure 1) by presenting (i) their main characteristics, and (ii) the ecosystem services produced by the systems that used the breeds concerned. All the researchers were invited to the online workshops and the attendance ranged from 11 to 12, according to the individual availability for each workshop. The ecosystem services were identified from a researcher's point of view, resulting in an interpretation of what could constitute a service, and to whom.

We organized two 2-day workshops in 2017 during which a transverse analysis resulted in a first analytical grid. Eleven researchers participated in the first and nine in the second workshop. In these workshops, we raised the following questions: i) who were the beneficiaries or potential beneficiaries of the ecosystem services identified, ii) who provided the services and iii) the role of the breed in the process of providing the service, thanks to available data and expertise. This analysis raised several theoretical and methodological questions reported in the result section.

We organized four online workshops in 2017 and 2018 to complete the iterative development of the framework, identifying loops between knowledge of the case studies, the questions we wanted to ask and the notions and concepts used to tackle these questions in the literature. This iterative collective back and forth between the case studies and the literature provided the background necessary to discuss to what extent the concept of ecosystem services (and associated concepts) was useful to deal with the multiple benefits of systems using livestock biodiversity. The process is summarized in Figure 2.

Results

Services provided by farming systems using livestock biodiversity: our proposed framework

Our transversal approach enabled us to develop a framework to analyze the provision of ecosystem services in farming systems using livestock biodiversity, focusing on the main elements to be taken into account and their interactions. We developed this framework in the form of a list of items to be considered in studies of ecosystem services produced by farming systems using livestock biodiversity. The main themes of the framework are summed up in Figure 3.

We formulated all the items as questions possibly applicable to other situations. The data from our empirical studies did not necessarily enable us to answer all the questions with the same level of detail, but the questions were considered key to understanding the processes for at least one case, even if empirical data to answer them were not available for all the cases.

Considering ecosystem services as a social construct implies the analysis of each ecosystem service identified, including its temporal dynamics, with these questions:

- How is the service formulated or qualified? The formulation or qualification may be plural, depending on from whose point of view the service is considered. It needs to consider dynamics over time and space.
- What processes underlie service provision?
  - Who is concerned?
    - Who are the beneficiaries and/or the people who help provide the service?
    - Who identifies the service as such?
    - Is the service produced intentionally or not (and by whom)?
    - What are the interactions between the people concerned?
  - What is the role of the breed in the process?
    - Is the effect of the breed linked with biological characteristics and the abilities of the animals of the breed (direct or indirect effects)? If so, which abilities are considered to be involved in the provision of services by the different people concerned (several points of view are possible)?
    - Is the effect of the breed linked with other attributes of the breed that are not directly linked to biological characteristics (e.g. local heritage or image)? If so, which attributes are considered to be at play by the different people involved in service provision (several points of view are possible)?
    - Can the breed be considered a marker of the search for innovation/alternatives in farming systems?
    - Does the breed play a catalytic role in relation to collective action to provide a service?
  - What are the interactions with other services?
    - Do other services result from this service? I.e. does the fact that this service is produced enable the provision of other services? (e.g. cascades or bundles of services)
    - Are there any variations in the generic formulation of the service?
    - Do other services – synergies or trade-offs – interact with it?
- To what extent is the service taken into account in the management of the breed?

In the following sections, for the different items in the framework, we present each item and provide details on:

(i) How the transverse analysis of our case studies enabled the identification of the item

(ii) How creating linkages between our cases and the literature helped build the framework and provided insights into the usefulness of the notion of ecosystem services.

We explore which elements from the literature helped us in the analysis of the case studies. The elements presented in those paragraphs result from linkages between empirical data (from case studies) and theoretical contributions (from the literature and the questions raised by our cross-cutting analysis).

Dynamics of production of a diverse range of services

The first step was to identify each service. We first established that systems that use livestock biodiversity are involved in the production of a wide range of ecosystem services. Table 2 summarizes this diversity by giving examples from our case studies. Our aim here was to illustrate the diversity and put it in perspective with a classification frame proposed in ecosystem services literature, but not to make an exhaustive, generic or ‘objective’ inventory.

Indeed, the temporal (and spatial) dynamics of ecosystem services provision are important in the cases we studied.

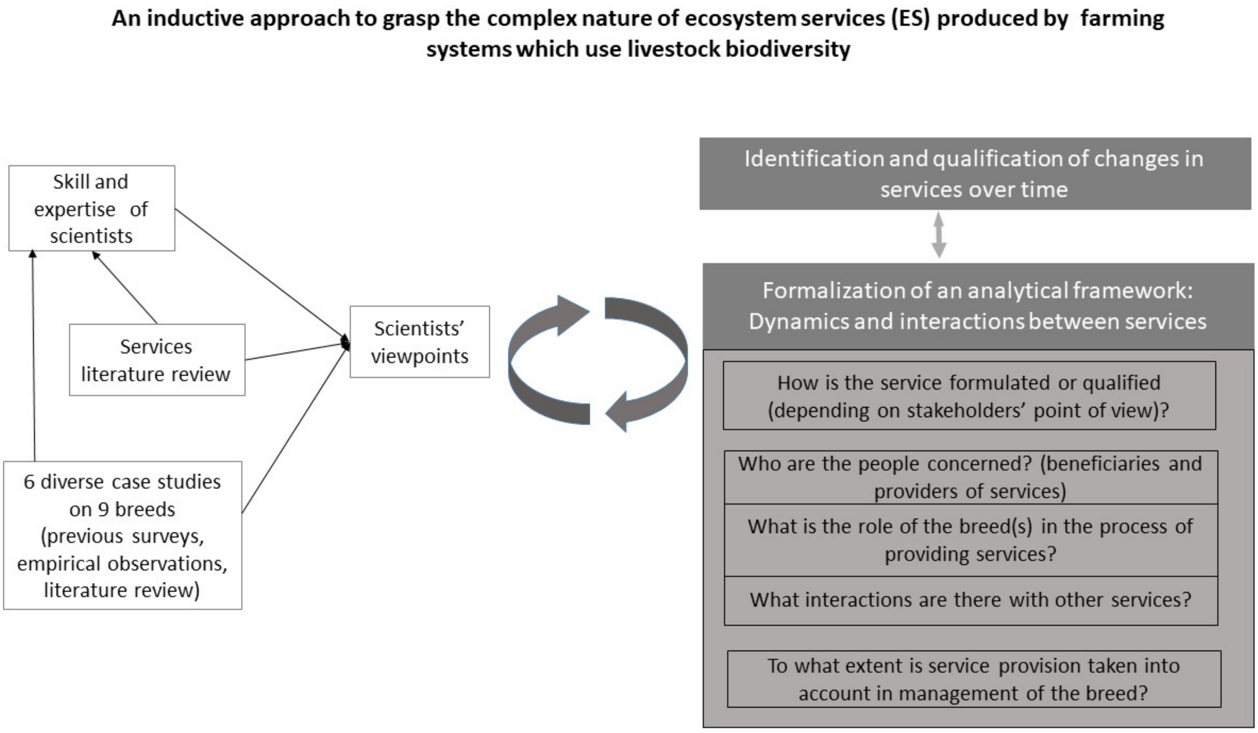


Figure 2. Process used from the six case studies analysis to the building of the framework

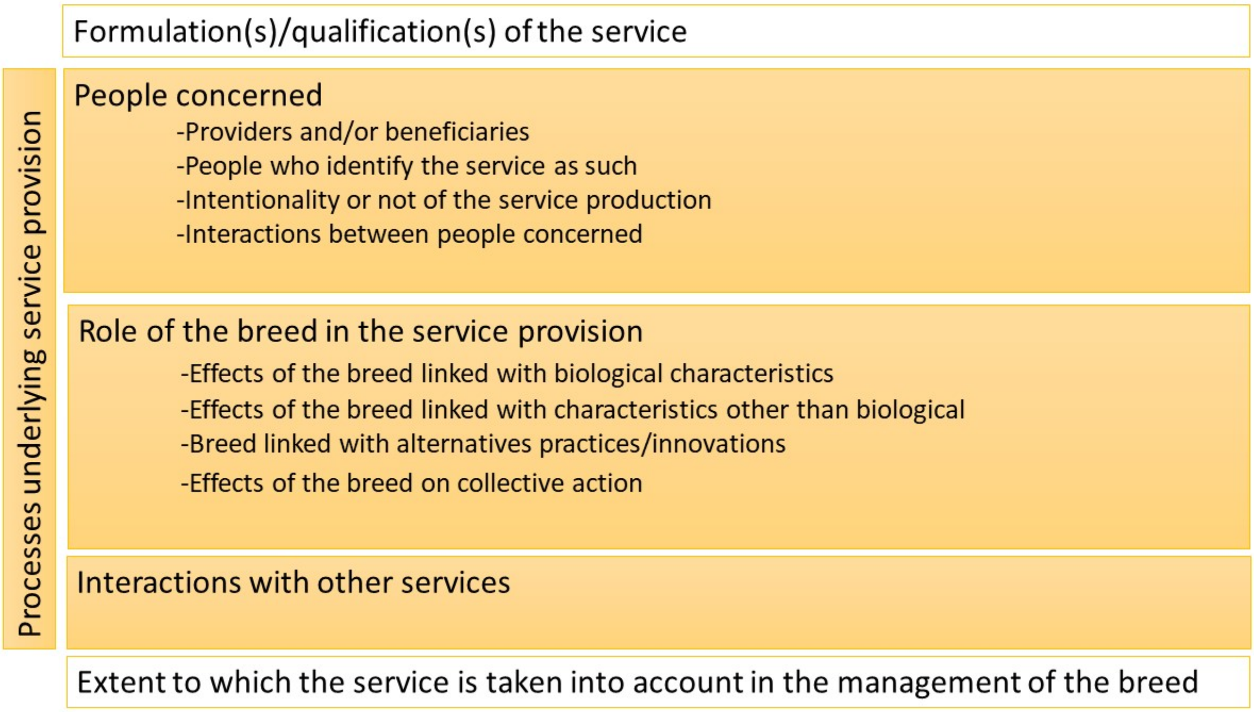


Figure 3. Main themes of the framework

Table 2. The diversity of services identified in the case studies.

Services identified	Examples given in the case studies
Food products (main products of the systems studied, diversity of meat and dairy products mainly subject to market valuation but can also be subject to non-market valuation)	Without geographical indication, sold through direct sales or through intermediaries, like the meat and dairy products from the Bretonne Pie Noir Protected Designation of Origin (PDO) products like Comté cheese made from the milk of cows belonging to the Montbéliarde or Simmental breeds and the Brocciu whey cheese in Corsica, which can only be made from Corsican sheep and goat milk
Inedible products	Wool from the Raïole breed Skins of goats in Guadeloupe Manure used to fertilize the soil (e.g. by Caussenarde des Garrigues, Raiole and Rouge du Roussillon or Creole cattle in Guadeloupe)
Maintenance of some specificities of the farming systems	The sheep breeds Raiole, Rouge du Roussillon and Caussenarde des Garrigues are associated with pastoral systems in the region concerned, and are reputed for their ability to adapt to those systems
Services linked to the management of habitats or the preservation of associated landscapes	Wildfire prevention (particularly in the Mediterranean region) Birds nesting open areas (recorded in typical Guadeloupean savannahs grazed by tethered Creole cattle (Zoom-Guadeloupe, 2012))
Use of animals from the systems for religious rites	In Réunion and Guadeloupe islands, goats are commonly used for Hindu sacrifices. A diversity of phenotypes are sought. Peï and Créole goats contribute to this diversity, more often in Guadeloupe than in La Réunion, along with other goats from various breeds and crossbred animals.
Educational use	In several cases, farmers have allowed farm visits by schoolchildren
Contribution to heritage and cultures	This dimension can be recognized through the associated landscape, e.g. the pastoral landscapes in the Causses and Cevennes area, which are designated UNESCO world heritage sites Local breed animals, with their specific phenotypic attributes (e.g. colour, horns) also contribute to the aesthetics of the landscape and its original identity, and thus serve as an image vector for agro-tourism Through an informal pathway, they may simply be part of the cultural patrimony, like the Creole society in Reunion Island or Guadeloupe.
Contribution of farming systems to the global dynamics of the territories (by helping maintain an agricultural activity and/or to the image of the territory concerned).	A cow from the Bretonne Pie Noir breed was named 'star cow' at the 2017 Paris agricultural fair, and the Pays de Redon where the cow originated benefited from this event being reported in the press.

For example, the use of Créole cattle in Guadeloupe in oxen drawing (bœufs tirants) contests, has played a role in conserving the breed, thanks to the success of this cultural event. However, the increasing success has led to a modification in the rules to include categories of larger animals, shorten the distance to be covered, and no longer place the yoke directly on the horns. These changes have increased the inclusion of exotic breeds with greater muscle development (Limousin, Charolais, Blonde d'Aquitaine) than that of the local breed. The cultural service is thus developing, but the link between the service and the local breed is dwindling.

A second example of temporal and spatial dynamics of ecosystem services is the Montbéliarde cattle breed in the Comté cheese territory (a protected designation of origin (PDO) cheese). The milk is collected and processed by small-scale processors (called fruitières in French) distributed throughout the territory. This organization and the related livestock farming systems are associated with the production of a specific landscape, which can be considered a service. However, we identified changes in farm structure with an increase in the size of the herds. This trend is due to the increasing demand for Comté cheese (and a readiness to make the most of the attractive price of the milk for Comté), the pleasure involved in raising high-yielding dairy cows, as well as the desire to reorganize labour. The tendency to graze the herd on land close to the farm homestead rather than on more distant pastures (more typical of the Jura landscape), as well as to increase the size and homogeneity of the plant covers used in pastures, has also led to changes in the landscape (Gaillard *et al*, 2018).

We compared data gathered from the case studies with data in the literature, which provided benchmarks for the identification of ecosystem services. The well-known categories suggested by the Millennium Ecosystem Assessment (2005) or the Common International Classification of Ecosystem Services (CICES) (Haines-Young and Potschin, 2018) make it possible to distinguish provisioning, regulation and maintenance, cultural, and support services. Identification of services can be facilitated by using such a classification. We used the provisioning, support and regulation and cultural categories to describe the services in our cases, which helped ensure we covered a wide range of services. However, as we will see hereafter, the services we identified were not always easy to classify into a single category. Moreover, the view of ecosystem services as social constructs (Barnaud *et al*, 2018) underlines the fact that they are not intrinsic properties of ecosystems, and that the classification of specific services provided at a given time by an ecosystem may not still be relevant when a dynamic approach is used.

How can services be qualified and classified?

After the first step of identification of services, the first item of the framework is: How can the services be formulated or qualified?

Our analysis of case studies raised the question of how to qualify (and classify) the services. By using the word qualify we mean describe and attribute one or several qualities to a service by naming it/them. The main products of systems using the breeds concerned are food products. However, systems that produce traditional food products, can be considered to provide both a provisioning service and a cultural service. Indeed, the food products concerned often have a cultural dimension, which can be recognized, for instance, through a PDO. Similarly, a pasture system that has shaped a typical landscape can be considered both as a regulation and a cultural service, like the Causses and the Cévennes, where the pastoral landscapes are UNESCO designated world heritage sites. Many services produced by farming systems using local breeds have a cultural dimension.

Ovaska and Soini (2016) noted the overlapping of ecosystem services categories and for instance, Díaz *et al* (2018), underlined the importance of recognizing "the central and pervasive role that culture plays" in the production of nature's contributions to humans, as well as the importance of applying a context-specific perspective. Increasing research in this field has enabled the development of multiple approaches that might be complementary (Peterson *et al*, 2018) to qualify and classify ecosystem services. Following Barnaud *et al* (2018), we consider ecosystem services as "subjective perceptions, socially situated and constructed". The points of view of the people involved in the situation studied should be taken into account when qualifying and classifying ecosystem services.

People concerned by the production of services

The first question regarding the second item of the framework, is: Who is concerned?

Our analysis of the case studies confirmed the diversity of actors involved in the systems using the breeds concerned and in the processes linked to the provision of ecosystem services. For instance, the range of livestock keepers involved in the use of the Bretonne Pie Noir cattle breed includes both hobby breeders and professional breeders, and both dairy and meat herds (Couix *et al*, 2016). Other stakeholders include researchers, consumers, environmental management associations and restaurant owners. Interactions among actors influence the provision of ecosystem services by systems using the breed.

In several ecosystem services that we identified, the same actors, particularly farmers, can be both beneficiaries and providers. For instance, some farmers who raise Bretonne Pie Noir cattle use areas of ecological interest, like wetlands, as feed resources for their cattle. The farmers thus participate in the production of the management service of those areas through grazing, and at the same time, profit from the vegetation in the areas to feed their herds.

The ecosystem functions are translated into services when they are used, consumed or enjoyed by humans (Fisher *et al*, 2009). This makes ecosystem services beneficiary-dependent (different individuals or collectives benefit from different services (Díaz et al., 2018)). In the framework they proposed for the analysis of social interdependencies underlying ecosystem services dynamics, Barnaud *et al* (2018) identified two other categories of actors together with beneficiaries: providers, and possibly intermediaries between beneficiaries and providers. To grasp the complexity of the actors involved, we suggest adding a step before qualifying actors as a beneficiary, provider or intermediary: identifying what we call the 'people concerned'. The aim of this step is to ensure different levels of concern are included. The people concerned can then be qualified as beneficiaries, producers, intermediaries or, in some cases, may belong to more than one category. Identifying the people concerned is a step in the process of tracing actors who play a direct role in the processes underlying the provision of services and helps grasp the multiple perceptions, values and practices associated with such services (Dendoncker *et al*, 2018; Teixeira *et al*, 2018).

Some of the ecosystem services produced at the scale of the human-driven farming system are produced intentionally, e.g.edible and inedible goods. However, this is not necessarily the case for all ecosystem services. The provision of cultural services, for instance, may sometimes be considered involuntary, as they are inherited from a long history of co-evolution of the breeds, their environment and the human practices involved in the farming system. It is not always easy to determine whether a service is provided intentionally or not, especially when several dimensions of ecosystem services production are interconnected. However, the distinction proposed by Aznar *et al* (2007) may be useful to deal with the general question of whether or not ecosystem services are produced intentionally. Based on the economy of services, this author defines provisioned services as services provided by agriculture which lead a farmer to maintain or change support and/or contribute technical or human capacities. These provisioned services are supplied by the farmer intentionally and are co-built by the farmer and the beneficiaries/applicant. Aznar *et al* (2007) distinguished them from service externalities that are supplied without the intention of the supplier and from service functions which refer to services supplied to humans by nature.

The role of the breed in service production: not only biological characteristics are at play

The second question regarding the second item in the framework, is: What is the role of the breed?

The provision of ecosystem services may be directly linked to specific abilities of the animals of the breeds involved. In Creole cattle, for instance, a signature of selection has been identified in the genomic region that can be linked to the strength and the shape of the horns, directly inherited from their use as draught animals (Naves *et al*, 2011). Resistance to specific sanitary problems, such as internal parasites in Creole goats, or to ticks and associated infectious diseases in Creole cattle, also help provide some ecosystem services (Naves *et al*, 2011). The integrated management of these diseases enabled by the use of resistant animals reduces the need for treatment, in turn reducing the quantity of chemical residues in edible products, but also in animal dung, which may be useful in agroecological or organic production systems.

However, the animals' specific abilities are not sufficient to describe all the characteristics of the breeds that can play a role in the provision of ecosystem services. The Raiole, Caussenarde des Garrigues and Rouge du Roussillon breeds, for instance, show that breeds also play a role in gathering together a group of farmers who exchange breeding animals, technical knowledge or projects to add value to products (Nozieres-Petit and Lauvie, 2018).

The appropriate scale to identify the provision of ecosystem services is the ecosystem (or the farming system as far as farming activity is concerned); as a consequence, the breed is one of the elements in the system that can contribute to their provision (Martin-Collado *et al*, 2019).

Relevant concepts in the literature to better identify the role of the breed are not specifically related to the notion of ecosystem services. The most relevant concepts to identify the role of the animals' specific abilities are two concepts from animal sciences: the concept of animal abilities and the concept of animal performances. Those concepts are used to describe specific biological characteristics of animals that are useful in livestock farming (directly linked with food production, like milk yield, or indirectly linked with food production, like the walking ability of animals, which is useful in pastoral systems).

A global term, often used for local breeds, covering their functional abilities, and not only their productive feature, is the hardiness of animals. Hardiness covers a wide range of abilities which depend on the situation and the point of view, as described by Hubert (2011). Being hardy means being not very demanding and therefore enabling the herd to survive even in harsh conditions (Poussard *et al*, 2016). The hardiness of a breed is difficult to measure and even to define precisely, as it covers a set of different animal functions interacting with the environment. Moreover, these functions are more difficult to evaluate than productive traits (e.g. milk yield, protein content, daily gain, etc.). But such abilities are highly valued in non-intensive farming systems and can be directly linked with the provision of ecosystem services (Naves *et al*, 2011; Marshall *et al*, 2016). They generally result from long-term breeding and from natural selection of the animals in specific environments, and are now being increasingly elucidated thanks to recent genomics technologies (Amills *et al*, 2017). For example, a complex trait such as the adaptation to climatic stress or the ability to walk is very useful in pastoral management, and some specific physical attributes useful in such conditions are present in local breeds (for example, short hair, dark hooves or a hump in cattle) (Naves *et al*, 2011; Flori *et al*, 2012).

However, these concepts from animal sciences are not sufficient to describe other dimensions that are not biological but that may also be involved in the provision of ecosystem services, for instance, the ability to trigger collective action. As a consequence, we propose to use the term ‘breed attributes’ to cover characteristics of breeds that are both biotechnical (abilities and performances) and sociotechnical. In this definition, both the ‘ability to exploit native vegetation in pastoral systems’ and the ‘ability to trigger a collective action’ could be considered as complementary breed attributes and combined in underlying processes to produce a set of ecosystem services.

Relations between services

The third question regarding the second item of the framework is: ‘What are the interactions with other services?’

Our case studies also underline the fact that diverse ecosystem services are interrelated. For instance, in the case of Raïole, Caussenarde des Garrigues and Rouge du Roussillon breeds, we identified interactions among ecosystem services: some services result from others, like the contribution to wildfire prevention that results from the animals’ capacity to graze local resources. Some services are divisions of others: ‘participate in education’ is a subset of the service ‘the social role of livestock farming’. Several ecosystem services are produced jointly in similar farming systems (Nozieres-Petit and Lauvie, 2018). For instance, farms using the Corsican sheep breed produce milk, processed into cheeses and whey cheese (on farm or in industrial dairy plants); and they also help maintain the pastured vegetal resource, thus indirectly helping shape the landscape and preventing wildfires. Creole cattle raised for meat production and manure provision are tethered in natural savannahs, sustaining small-scale family farms, while shaping the typical hilly landscapes of Guadeloupe (mornes). Direct provisional services can be reinvested through the loop of cultural service since they enhance gastronomy, with dishes like goat curry in Guadeloupe, or can provide raw material for the production of musical instruments. The use of goats in Hindu sacrificial rites in Guadeloupe or in Reunion Island illustrates a case of service (with a cultural and religious dimension) that over time has enhanced the development of the entire goat meat sector (service provision).

‘Bundles of services’ is a useful notion in the literature to jointly consider a diverse range of services produced in similar farming systems, and it can help tackle synergies and trade-offs between services (Cord *et al*, 2017; Dumont *et al*, 2019).

Links between services and management of the breeds

The last item in the framework is: What are the links with the management of the breed?

We found examples in the case studies in which the intentional production of a new service engendered changes in the collective management of the breed. For instance, for the Raïole breed, the farmers developed collective marketing of the wool, and during the sale of rams organized by the breeders’ association, they provided information to farmers about the wool quality of rams sold, so that they could consider it in their choice of a breeding animal (Lauvie *et al*, 2017).

Biodiversity and ecosystem services are often closely linked. For instance, Mace *et al* (2012) underlined how biodiversity generally plays a key role in ecosystem services provision, as a regulator of underpinning ecosystem processes, as a final ecosystem service, and as a good subject for valuation. This link between services and management of the breed feeds the question of the links between services and biodiversity dynamics. Indeed, collective management of the breed is an important lever of livestock biodiversity dynamics: through the collective choice of selection criteria that influence the direction of breed management, but also through collective promotion actions that can have an impact on the number of animals, for instance.

Discussion

In this paper, we used several case studies to explore the question of the multiple benefits of farming systems that use livestock biodiversity and developed a framework to address the processes underling the ecosystems services provided by them.

One of our aims was to discuss the extent to which the notion of ecosystem services is useful to deal with these multiple benefits. Services are indeed increasingly taken into account in livestock farming research (Rodríguez-Ortega *et al*, 2014; Alexandre *et al*, 2014; Ryschawy *et al*, 2017; Dumont *et al*, 2019). Steger *et al* (2018) argue that the diversity of definitions and approaches in ecosystem services research has prevented it from being structured by a single discipline, and maintained it as a boundary object. Choosing the ecosystem services notion as an entry point of our analysis did not provide us with a turnkey approach, but rather gave us room to include different visions (and notions) of animal scientists, social scientists and economists in our analysis. Some of the concepts associated with ecosystem services in the literature appeared to be relevant to our empirical findings in case studies (e.g. the bundle of services) or fed our reflection on the processes underlying the provision of services (e.g. the beneficiaries). However, we also considered other concepts, not related to the ecosystem services literature, which were useful in our analysis, like the concept of animal abilities used in animal sciences. Choosing the notion of ecosystem services as an entry point also favoured a systemic analysis. Indeed, such a view is central to our communities as the farming system is a key notion, particularly for the livestock farming systems community (Dedieu *et al*, 2008). The meeting point of systemic views in several scientific communities is the agroecosystem. However, the view on agroecosystem might favour complementary foci depending on the communities, which could enrich each other. To go further, it would indeed be interesting to broaden our already multidisciplinary vision by including the points of view of ecological scientists in our framework.

Other notions or theoretical frameworks could also be adapted to tackle the question of the multiple benefits obtained by farming systems which use domestic animal biodiversity and are complementary to the ecosystem services approach. The notion of multifunctionality, for instance, is equally constructed and situated (Barnaud and Couix, 2020), and makes it possible to consider the different functions of farming activities. Huang *et al* (2015) reported the proximity of the scientific communities that use multifunctionality and ecosystem services concepts when dealing with agriculture. However, the entry point of the approach is agricultural activity whereas the entry point of the ecosystem services notion is the ecosystem. Huang *et al* (2015) pointed out that the two approaches would benefit from being integrated, and Barnaud and Couix (2020) associated them by using an ecosystem services lens to answer a multifunctionality question. The framework proposed by Díaz et al. (2018) is another relevant framework for our purpose. Díaz et al. argued that the “Nature’s Contribution to People” (NCP), the core concept of this framework, is a concept which makes it possible to go beyond some of the limits of the ecosystem services concept, in particular to respond to criticisms regarding the lack of social and humanities sciences involved. This framework indeed emphasizes the importance of the cultural context in understanding NCP, which our case studies highlight. Díaz and her colleagues’ proposal led to a broad debate on the real novelty of the concept compared to ecosystem services (Braat, 2018; Faith, 2018; Peterson *et al*, 2018). Without going further in this debate, we note that the use of the word ‘nature’ (instead of ‘ecosystem’) as the subject of the contribution, and the people only as the beneficiaries, can question the importance of human action in the production of several of these services. As noted by Peterson *et al* (2018): “a focus on ‘nature,’ therefore de-emphasizes the ecosystems that are home to and provide the necessities of life to most of the world’s population”. In our case, the services are clearly co-produced by humans and animals in farming systems and talking about them as NCP may seem counterproductive.

The framework we propose in this paper is the first step in identifying and understanding the services provided by farming systems using livestock biodiversity. This framework could however be put to the test and enriched by applying it to other cases. Steps to enrich it could be to include the points of view, values and practices of the different stakeholders involved. Answers to the questions proposed in the framework can consequently include a diversity of points of view. The interest of the framework is not to provide an irrevocable single answer to each question, but rather contribute to better understanding the processes underlying services production, in their complexity. However, this framework cannot be used for a quantitative assessment of the dynamics of livestock biodiversity or the values of services. The framework can contribute to better understanding the link between livestock biodiversity dynamics and the provision of services in the farming systems concerned. Indeed, from a perspective of livestock biodiversity management and conservation, one of the scientific challenges is to understand the issues underlying the management dynamics of each breed, the corresponding management objectives, and their translation into practices, at both individual and collective levels. This will be a key step in understanding whether or not the provision of services production is at stake for the people concerned. In our framework, we gathered important items to be taken into consideration to analyze the dynamics of services produced as social constructs and to understand the underlying processes. Such an approach implies including a diversity of biotechnical and sociotechnical dimensions, and accounting for their interactions and their dynamics in multiple services over time. The notion of ‘breed attributes’ actually helps explain how livestock biodiversity contributes to the provision of ecosystem services. We propose this notion to describe not only the biological but also the sociotechnical characteristics of breeds. Our aim is not to objectify the role of a given breed in providing a given ecosystem service (Martin-Collado *et al*, 2019). As ecosystem services are seen as constructs and breeds are seen as dynamics (and not categories with static properties), our aim is rather to question whether accounting for the wide range of productions (including the different services) of farming systems using local breeds would enable a better understanding of the overall dynamics of livestock biodiversity conservation and management .

To conclude, if the ecosystem services notion can help include different visions and develop a multidisciplinary approach, it is not the only notion that can be used to tackle the question of the multiple benefits of farming systems using local breeds. Understanding the processes underlying service production requires going even further in developing interdisciplinary approaches.

Acknowledgments

The authors have a special thought for their late colleague Claire Gaillard, with whom they had the pleasure to work in this study, and who left us too soon. This work was conducted in the frame of the SECOYA project, funded by the INRA Metaprogram ECOSERV.

Conflict of interest statement

The authors declare no conflict of interest.

Author contributions

Anne Lauvie: Study conception and design, data collection, analysis and interpretation of results, draft manuscript preparation and manuscript revision. Claire Gaillard: Study conception and design, data collection, analysis and interpretation of results. Gisèle Alexandre, Valérie Angeon, Nathalie Couix, Olivia Fontaine, Michel Meuret, Catherine Mougenot, Charles-Henri Moulin, Michel Naves, Marie-Odile Nozières-Petit, Jean-Christophe Paoli, Lola Perucho, Jean Michel Sorba, Emmanuel Tillard, Etienne Verrier: Study conception and design, data collection, analysis and interpretation of results and manuscript revision.

References

Ajmone-Marsan, P. and Consortium GLOBALDIV (2010). A global view of livestock biodiversity and conservation - GLOBALDIV. *Animal Genetics* 41(s1), 1–5. doi: https://doi.org/10.1111/j.1365-2052.2010.02036.x

Alexandre, G., De Beauville, S. A., Bienville, Y., and Shitalou, E. (2002). La Chèvre multifonctionnelle dans la société antillaise. *Ethnozootechnie* 70, 35–52.

Alexandre, G., Fanchone, A., Ozier-Lafontaine, H., and Diman, J. L. (2014). Sustainable Agriculture Reviews 14: Agroecology and Global Change, ed. H, O.-L. and M, L.-J. (Cham: Springer International Publishing), 83-115.

Amills, M., Capote, J., and Tosser-Klopp, G. (2017). Goat domestication and breeding: a jigsaw of historical, biological and molecular data with missing pieces. *Animal Genetics* 48(6), 631–644. doi: https://doi.org/10.1111/age.12598

Audiot, A. (1995). Races d'hier pour l'élevage de demain (Paris, France: INRA Editions).

Aznar, O., Guérin, M., and Perrier-Cornet, P. (2007). Agriculture de services, services environnementaux et politiques publiques: éléments d'analyse économique. *Revue d'Économie Régionale & Urbaine* 4, 573–587. doi: https://doi.org/10.3917/reru.074.0573

Barnaud, C., Antona, M., and Marzin, J. (2011). Vers une mise en débat des incertitudes associées à la notion de service écosystémique. *Vertigo* 11. doi: https://doi.org/10.4000/vertigo.10905

Barnaud, C., Corbera, E., Muradian, R., Salliou, N., Sirami, C., Vialatte, A., Choisis, J. P., Dendoncker, N., Mathevet, R., Moreau, C., Reyes-García, V., Boada, M., Deconchat, M., Cibien, C., Garnier, S., Maneja, R., and Antona, M. (2018). Ecosystem services, social interdependencies, and collective action: a conceptual framework. *Ecology and Society* 23(1). doi: https://doi.org/10.5751/ES-09848-230115

Barnaud, C. and Couix, N. (2020). The multifunctionality of mountain farming: Social constructions and local negotiations behind an apparent consensus. *Journal of Rural Studies* 73, 34–45. doi: https://doi.org/10.1016/j.jrurstud.2019.11.012

Berland, F., Signoret, F., and Roche, B. (2006). Conserver et valoriser la race bovine Maraîchine et les prairies naturelles de marais. *Les Actes du BRG* 6, 485–494.

Beudou, J., Martin, G., and Ryschawy, J. (2017). Cultural and territorial vitality services play a key role in livestock agroecological transition in France. *Agron. Sustain. Dev* 37(36). doi: https://doi.org/10.1007/s13593-017-0436-8

Boval, M., Coppry, O., Naves, M., and Alexandre, G. (2012). L'élevage traditionnel, une source et un support pour l'innovation agro-écologique : la pratique du piquet aux Antilles. *Courrier de l'Environnement de l'INRA* 62, 87–97.

Braat, L. C. (2018). Five reasons why the Science publication "Assessing nature's contributions to people" (Diaz et al. 2018) would not have been accepted in Ecosystem Services. *Ecosystem Services* 30, A1–A2. doi: https://doi.org/10.1016/j.ecoser.2018.02.002

Cord, A. F., Bartkowski, B., Beckmann, M., Dittrich, A., Hermans-Neumann, K., Kaim, A., Lienhoop, N., Locher-Krause, K., Priess, J., Schröter-Schlaack, C., Schwarz, N., Seppelt, R., Strauch, M., Václavík, T., and Volk, M. (2017). Towards systematic analyses of ecosystem service trade-offs and synergies: Main concepts, methods and the road ahead. *Ecosystem Services* 28(C), 264–272. doi: https://doi.org/10.1016/j.ecoser.2017.07.012

Couix, N., Gaillard, C., Lauvie, A., Mugnier, S., and Verrier, E. (2016). Des races localement adaptées et adoptées, une condition de la durabilité des activités d'élevage. *Cahiers d'Agriculture* 25(6), 650009. doi: http://dx.doi.org/650010.651051/cagri/2016052

Dedieu, B., Faverdin, P., Dourmad, J. Y., and Gibon, A. (2008). Systèmes d'élevage, un concept pour raisonner les transformations de l'élevage. *INRA Prod Anim* 21(1), 45–58. doi: https://doi.org/10.20870/productions-animales.2008.21.1.3374

Dendoncker, N., Boeraeve, F., Crouzat, E., Dufrêne, M., König, A., and Barnaud, C. (2018). How can integrated valuation of ecosystem services help understanding and steering agroecological transitions? *Ecology and Society* 23(1). doi: https://doi.org/10.5751/ES-09843-230112

Díaz, S., Pascual, U., Stenseke, M., Martín-López, B., Watson, R. T., Molnár, Z., Hill, R., Chan, K. M. A., Baste, I. A., Brauman, K. A., Polasky, S., Church, A., Lonsdale, M., Larigauderie, A., Leadley, P. W., Van Oudenhoven, A. P. E., Van Der Plaat, F., Schröter, M., Lavorel, S., Aumeeruddy-Thomas, Y., Bukvareva, E., Davies, K., Demissew, S., Erpul, G., Failler, P., Guerra, C. A., Hewitt, C. L., Keune, H., Lindley, S., and Shirayama, Y. (2018). Assessing nature's contributions to people. *Science* 359(6373), 270–272. doi: https://doi.org/10.1126/science.aap8826

Droste, N., Amato, D., and Goddard, J. J. (2018). Where communities intermingle, diversity grows - The evolution of topics in ecosystem service research. *PLOS ONE* 13, (9):e0204749. doi: https://doi.org/10.1371/journal.pone.0204749

Dumont, B., Ryschawy, J., Duru, M., Benoit, M., Chatellier, V., Delaby, L., Donnars, C., Dupraz, P., Lemauviel-Lavenant, S., Méda, B., Vollet, D., and Sabatier, R. (2019). Review: Associations among goods, impacts and ecosystem services provided by livestock farming. *Animal* 13(8), 1773–1784. doi: https://doi.org/10.1017/S1751731118002586

Faith, D. P. (2018). Avoiding paradigm drifts in IPBES: reconciling "nature's contributions to people",biodiversity, and ecosystem services. *Ecology and Society* 23(2). doi: https://doi.org/10.5751/ES-10195-230240

FAO (2015). The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture, ed. Scherf, B. D. and Pilling, D. (Rome: FAO Commission on Genetic Resources for Food and Agriculture Assessments). url: https://www.fao.org/3/i4787e/i4787e.pdf.

Fisher, B., Turner, R. K., and Morling, P. (2009). Defining and classifying ecosystem services for decision making. *Ecological Economics* 68(3), 643–653. doi: https://doi.org/10.1016/j.ecolecon.2008.09.014

Flori, L., Gonzatti, M. I., Thevenon, S., Chantal, I., Pinto, J., Berthier, D., Aso, P. M., and Gautier, M. (2012). A Quasi-Exclusive European Ancestry in the Senepol Tropical Cattle Breed Highlights the Importance of the slick Locus in Tropical Adaptation. *PLOS ONE* 7(5), e36133. doi: https://doi.org/10.1371/journal.pone.0036133

Fontaine, O., Niobé, D., Shitalou, E., Fontaine, D., and Choisis, J. P. (2008). Hindouisme et sacrifice de boucs à l'Île de la Réunion. *Ethnozootechnie* 85, 101–110.

Gaillard, C., Mougenot, C., and Petit, S. (2018). Le fromage de Comté dans l'ère du temps. Succès et tensions. . *Temporalités, Revue de sciences sociales et humaines* 28. doi: https://doi.org/10.4000/temporalites.5377

Gandini, G. C. and Villa, E. (2003). Analysis of the cultural value of local livestock breeds: a methodology. *J. Anim. Breed. Genet* 120, 1–11. doi: https://doi.org/10.1046/j.1439-0388.2003.00365.x

Gautier, M. and Naves, M. (2011). Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Molecular Ecology* 20, 3128–3143. doi: https://doi.org/10.1111/j.1365-294X.2011.05163.x

Gourdine, J. L., Fourcot, A., Lefloch, C., Naves, M., and Alexandre, G. (2021). Assessment of ecosystem services provided by livestock agroecosystems in the tropics: a case study of tropical island environment of Guadeloupe. *Tropical Animal Health and Production* 53, 435–435. doi: https://doi.org/10.1007/s11250-021-02880-3

Haines-Young, R. and Potschin, M. B. (2018). Common International Classification of Ecosystem Services (CICES) V5.1 and Guidance on the Application of the Revised Structure. url: https://cices.eu/resources/.

Hall, S. J. G. (2019). Livestock biodiversity as interface between people, landscapes and nature. *People and Nature* 1(3), 284–290. doi: https://doi.org/10.1002/pan3.23

Hall, S. J. G. and Bradley, D. G. (1995). Conserving livestock breed biodiversity. *Trends in Ecology & Evolution* 10(7), 267–270. doi: https://doi.org/10.1016/0169-5347(95)90005-5

Hoffmann, I., From, T., and Boerma, D. (2014). Ecosystem services provided by livestock species and breeds, with special consideration to the contributions of small-scale livestock keepers and pastoralists. url: https://www.fao.org/3/at598e/at598e.pdf.

Huang, J., Tichit, M., Poulot, M., Darly, S., Li, S., Petit, C., and Aubry, C. (2015). Comparative review of multifunctionality and ecosystem services in sustainable agriculture. *Journal of Environmental Management* 149, 138–147. doi: http://dx.doi.org/10.1016/j.jenvman.2014.10.020

Hubert, B. (2011). La rusticité: l'animal, la race, le système d'élevage? Pastum Hors Série (Association Franc
Cardère éditeur).

INRA (2014). Races animales françaises menacées d'abandon pour l'agriculture. url: https://agriculture.gouv.fr/sites/default/files/races_menacees_rapport_methodologique.pdf.

Kull, C., De Sartre, X. A., and Castro, M. (2015). The political ecology of ecosystem services. *Geoforum* 61. doi: https://doi.org/10.1016/j.geoforum.2015.03.004

Lauvie, A., Alexandre, G., Couix, N., Markey, L., Meuret, M., Nozières-Petit, M. O., Perucho, L., and Sorba, J. M. (2017). Comment les diverses formes de valorisation des races locales interagissent avec leur conservation. *Ethnozootechnie* 103, 7–12.

Leroy, G., Baumung, R., Boettcher, P., Besbes, B., From, T., and Hoffmann, I. (2018). Animal genetic resources diversity and ecosystem services. *Global Food Security* 17, 84–91. doi: https://doi.org/10.1016/j.gfs.2018.04.003

Lescourret, F., Magda, D., Richard, G., Adam-Blondon, A. F., Bardy, M., Baudry, J., Doussan, I., Dumont, B., Lefèvre, F., Litrico, I., Martin-Clouaire, R., Montuelle, B., Pellerin, S., Plantegenest, M., Tancoigne, E., Thomas, A., Guyomard, H., and Soussana, J. F. (2015). A social-ecological approach to managing multiple agro-ecosystem services. *Current Opinion in*

Environmental Sustainability 14(0), 68–75. doi: http://dx.doi.org/10.1016/j.cosust.2015.04.001

Ligda, C. and Casabianca, F. (2013). Adding value to local breeds: challenges, strategies and key factors. *Animal Genetic Resources* 53, 107–116.

Mace, G. M., Norris, K., and Fitter, A. H. (2012). Biodiversity and ecosystem services: a multilayered relationship. *Trends in Ecology & Evolution* 27(1), 19–26. doi: https://doi.org/10.1016/j.tree.2011.08.006

Marshall, K., Mtimet, N., Wanyoike, F., Ndiwa, N., Ghebremariam, H., Mugunieri, L., and Costagli, R. (2016). Traditional livestock breeding practices of men and women Somali pastoralists: trait preferences and selection of breeding animals. *Journal of Animal Breeding and Genetics* 133(6), 534–547. doi: https://doi.org/10.1111/jbg.12223

Marsoner, T., Vigl, L. E., Manck, F., Jaritz, G., Tappeiner, U., and Tasser, E. (2018). Indigenous livestock breeds as indicators for cultural ecosystem services: A spatial analysis within the Alpine Space. *Ecological Indicators* 94, 55–63. doi: https://doi.org/10.1016/j.ecolind.2017.06.046

Martin-Collado, D., Boettcher, P., and Bernués, A. (2019). Opinion paper: livestock agroecosystems provide ecosystem services but not their components - the case of species and breeds. *Animal* 13, 1–3. doi: https://doi.org/10.1017/S1751731119001277

Mathias, E., Mundy, P., and Koehler-Rollefson, I. (2010). Marketing products from local livestock breeds: an analysis of eight cases. *Animal Genetic Resources* 47, 59–71.

Millennium Ecosystem Assessment (2005). Ecosystems and human well-being: synthesis (Washington, DC: Island Press). url: https://www.millenniumassessment.org/documents/document.356.aspx.pdf.

Naves, M., Alexandre, G., Mahieu, M., Gourdine, J. L., and Mandonnet, N. (2011). Les races animales locales: bases du développement innovant et durable de l'élevage aux Antilles. *Innovations Agronomiques* 16, 193–205.

Norgaard, R. B. (2010). Ecosystem services: From eye-opening metaphor to complexity blinder. *Ecological Economics* 69, 1219–1227. doi: https://doi.org/10.1016/j.ecolecon.2009.11.009

Nozieres-Petit, M. O. and Lauvie, A. (2018). Diversité des contributions des systèmes d'élevage de races locales. Les points de vue des éleveurs de trois races ovines méditerranéennes. *Cah Agric* 27, 65003. doi: https://doi.org/10.1051/cagri/2018039

Ovaska, U. and Soini, K. (2016). Native Breeds as Providers of Ecosystem Services:The Stakeholders' Perspective. *TRACE finnish journal for human-animal studies* 2, 28–51.

Perucho, L., Paoli, J. C., Ligda, C., Moulin, C. H., Hadjigeorgiou, I., and Lauvie, A. (2020). Diversity of breeding practices is linked to the use of collective tools for the genetic management of the Corsican sheep breed. *Italian Journal of Animal Science* 19, 158–172. doi: https://doi.org/10.1080/1828051X.2020.1713027

Peterson, G. D., Ková, Z. V. H., Meacham, M., Queiroz, C., Jiménez-Aceituno, A., Kuiper, J. J., Malmborg, K., Sitas, N., and Bennett, E. M. (2018). Welcoming different perspectives in IPBES: Nature's contributions to people and Ecosystem services. *Ecology and Society* 23. doi: https://doi.org/10.5751/ES-10134-230139

Poussard, A., Ribere, C., Bonnelle, R., Couix, N., and Labatut, J. (2016). Rusticité des races: Définition. Dictionnaire d'Agroecologie. url: https://dicoagroecologie.fr/encyclopedie/rusticite-des-races/.

Rege, J. E. O. and Gibson, J. P. (2003). Animal genetic resources and economic development: issues in relation to economic valuation. *Ecological economics* 45, 319–330. doi: https://doi.org/10.1016/S0921-8009(03)00087-9

Rodríguez-Ortega, T., Oteros-Rozas, E., Ripoll-Bosch, R., Tichit, M., Martín-López, B., and Bernués, A. (2014). Applying the ecosystem services framework to pasture-based livestock farming systems in Europe. *Animal* 8, 1361–1372. doi: https://doi.org/10.1017/S1751731114000421

Ryschawy, J., Disenhaus, C., Bertrand, S., Allaire, G., Aznar, O., Plantureux, S., Josien, E., Guinot, C., Lasseur, J., Perrot, C., Tchakerian, E., Aubert, C., and Tichit, M. (2017). Assessing multiple goods and services derived from livestock farming on a nation-wide gradient. *Animal* 1-12. doi: https://doi.org/10.1017/s1751731117000829

Steger, C., Hirsch, S., Evers, C., Branoff, B., Petrova, M., Nielsen-Pincus, M., Wardropper, C., and Van Riper, C. J. (2018). Ecosystem Services as Boundary Objects for Transdisciplinary Collaboration. *Ecological Economics* 143, 153–160. doi: https://doi.org/10.1016/j.ecolecon.2017.07.016

Teixeira, H. M., Vermue, A. J., Cardoso, I. M., Claros, M. P., and Bianchi, F. J. J. A. (2018). Farmers show complex and contrasting perceptions on ecosystem services and their management. *Ecosystem Services* 33, 44–58. doi: https://doi.org/10.1016/j.ecoser.2018.08.006

Velado-Alonso, E., Gómez-Sal, A., Bernués, A., and Martín-Collado, D. (2021). Disentangling the Multidimensional Relationship between Livestock Breeds and Ecosystem Services. *Animals* 11, (9):2548. doi: https://doi.org/10.3390/ani11092548

Verrier, E., Tixier-Boichard, M., Bernigaud, R., and Naves, M. (2005). Conservation and value of local livestock breeds: usefulness of niche products and/or adaptation to specific environments. *Animal Genetic Resources* 36, 21–31. doi: https://doi.org/10.1017/S1014233900005538

Woo, S. E., O'Boyle, E. H., and Spector, P. E. (2017). Best practices in developing, conducting, and evaluating inductive research. *Human Resource Management Review* 27(2), 255–264. doi: https://doi.org/10.1016/j.hrmr.2016.08.004

Zhang, W., Ricketts, T. H., Kremen, C., Carney, K., and Swinton, S. M. (2007). Ecosystem services and dis-services to agriculture. *Ecological Economics* 64(2), 253–260. doi: https://doi.org/10.1016/j.ecolecon.2007.02.024

Zoom-Guadeloupe (2012). Bois Jolan. url: http://zoom-guadeloupe.fr/component/content/article?d4dad6935f632ac35975e3001dc7bbe8=a9597d0cde95aa2b7dedeee2f2dd7eca&showall=1&id=42.

Genetic Resources

SHORT COMMUNICATION

Genetic Resources (2023), 4 (8), 29–36
DOI: 10.46265/genresj.VOOZ8371
https://www.genresj.org
ISSN: 2708-3764

The genetic composition of the Traditional Irish Horse – towards the development of a DNA-ancestry test for the preservation of traditionally bred Irish Sport Horses

Beatrice A McGivney[a], **Deirdre Harty**[b], **Alison Corbally**[b] and **Emmeline W Hill** [*,a,c]

[a] *Plusvital Ltd, Dun Laoghaire Industrial Estate, Pottery Road, Co. Dublin, Dun Laoghaire, Ireland*
[b] *Horse Sport Ireland, Beech House, Millennium Park, Co. Kildare, Osberstown, Naas, Ireland*
[c] *School of Agriculture and Food Science, University College Dublin, Belfield, Dublin 4, Ireland*

Abstract: The traditionally bred Irish Sport Horse, known as the Traditional Irish Horse, is an important cultural asset to horse genetic resources in Ireland. We tested the hypothesis that the Irish Sport Horse, which was originally developed from the Irish Hunter, may contain a genetic background distinct from European Warmblood horse populations that would be valuable to preserve. Using genome-wide single nucleotide polymorphism (SNP) data, the results show that Traditional Irish Horses (with confirmed pedigrees) have lower levels of European Warmblood ancestry components than other Irish Sport Horses. These results indicate that measurement of the levels of European Warmblood ancestry components in the Irish Sport Horse may assist in the preservation of traditional Irish lineages.

Keywords: Equine, Genomics, SNP variation, Admixture, Irish Sport Horse, Conservation, Population Genetics

Citation: McGivney, B. A., Harty, D., Corbally, A., Hill, E. W. (2023). **The genetic composition of the Traditional Irish Horse – towards the development of a DNA-ancestry test for the preservation of traditionally bred Irish Sport Horses**. *Genetic Resources* 4 (8), 29–36. doi: 10.46265/genresj.VOOZ8371.

Introduction

The horse has a long history in Ireland, developing through millennia into several distinct native breeds, from small riding ponies to larger work horses (McCormick, 2007). The earliest archaeological evidence for domesticated horses in Ireland dates to 2,400 BCE in the Early Bronze Age contexts of Newgrange (Bendrey *et al*, 2013). The introduction of the Celtic horse into Ireland can be traced to the 7th century, with images of ridden horses contained within the Book of Kells (Unattributed, 800). The medieval Irish horse was known as the Irish Hobby, resulting from crossing horses from northern Spain with local animals (Hendricks, 1995). The Hobby was noted for speed and agility, and was mentioned in 1296 in documents relating to an Irish contingent of horses brought to Scotland by John de Wogan, Judicier of Ireland (Lydon, 1954). The Irish Hobby was likely a founder breed for the extant native breeds (McCormick, 2007), the Connemara Pony and Irish Draught, and also likely contributed to the Thoroughbred (Hendricks, 1995).

The Irish Draught was bred as a light, versatile farm, carriage, riding and hunting horse while the Connemara Pony was developed as a multipurpose animal for the harsher conditions of the western Irish seaboard (Hendricks, 1995). The Thoroughbred was developed in the 17th and 18th centuries when stallions from the Middle East were brought to England and crossed with the best racing mares in Britain, many of which were Irish or had been bred from Irish stock (Bower *et al*, 2011).

---

[*] Corresponding author: Emmeline W Hill (emmeline.hill@ucd.ie)

Received: 18.02.2023 | Accepted: 07.08.2023 | Published online: 12.09.2023

Prior to the development of the Thoroughbred, a centuries-old tradition of hunting in Ireland led to the breeding of specialist hunters originating from the Irish Draught horse (Hendricks, 1995). When the Thoroughbred became established in Ireland, it was used extensively for further improvement of the Irish Hunter (Hendricks, 1995). The Irish Hunter was referred to by William Youatt who noted that the Irish bred "a large, long blood-horse of considerable value" (Youatt, 1831). Until 1922 the Irish Hunter was registered in the Hunter Stud Book, but it was later redefined as the Irish Sport Horse and has been registered in the Irish Sport Horse Studbook since 1970. The Irish Sport Horse became the world leader in eventing and has topped the World Breeding Federation for Sport Horses eventing studbook rankings for all but two years since its inception in 1994.

In the 1980s, the crossing of European Warmblood stallions with Irish Sport Horse mares gained popularity (Doyle *et al*, 2022). In 1990, the percentage of foals registered in the Irish Sport Horse Studbook by foreign-bred sires was 1.1% and by 2014 this had increased to 37.5%. This dramatic change in the breed profile of the Irish Sport Horse population has led to a growing concern among breeders that its traditional traits are being displaced by a pan-European sport horse. Demand has developed amongst breeders to identify traditionally bred Irish Sport Horses, referred to as the Traditional Irish Horse, and to establish conservation measures for the lineages descended from the original Irish Hunter.

To date, there is no published study on genetic diversity in the Irish Sport Horse population, and there are only limited published studies on genetic diversity and relatedness among the other Irish horse breeds. Mitochondrial DNA (mtDNA) analyses showed no evidence of clustering of Irish Draught and Connemara Pony mtDNA haplotypes to the exclusion of other breeds, with their distribution among the most common European horse sequences. On the other hand, the Kerry Bog Pony, a rare, recently recharacterized breed, contained an uncommon mtDNA haplotype (McGahern *et al*, 2006). In a study comparing British and Irish native ponies, the Connemara Pony showed no greater divergence among breeds than between British native breeds, which may be due to recent gene flow resulting from historic migration (Reilly *et al*, 1998).

In consideration of the concerns from breeders about the dilution of Irish Sport Horse genetics with European Warmblood genetics, the aim of this study was to establish whether there are genetic differences between the Traditional Irish Horse and European Warmblood breeds by testing the hypothesis that Irish Sport Horses with Warmblood ancestry in their pedigree can be identified using admixture analysis to test for Warmblood introgression. The purpose of a molecular genetics approach is to provide novel information to augment pedigree-based classifications of the Traditional Irish Horse for the long-term preservation of traditional Irish lineages.

Methods

Samples

A total of 133 samples from 8 horse populations (Table 1) were used in the initial analyses, including 2 native Irish breeds (21 Irish Draught and 24 Connemara Pony), 23 Irish Sport Horse (comprised of 8 certified Traditional Irish Horse, 7 Irish Sport Horse with known Warmblood influence in their pedigree, and 8 Irish Sport Horse with some missing pedigree), as well as 3 breeds with known ancestral influence on the Irish Sport Horse (25 Irish Thoroughbred, 14 Hanoverian, and 14 Swiss Warmblood). The results were validated using a set of 12 pedigree-certified Traditional Irish Horses.

Genotyping

SNP genotypes, derived from the equine Illumina SNP50 genotyping array (SNP50), were publicly available for Swiss Warmblood and Hanoverian (Petersen *et al*, 2013). For the other samples, genomic DNA was isolated from blood or hair samples and genotyped on the Affymetrix Axiom Equine genotyping array (SNP670). Quality control (QC) procedures were carried out in PLINK 1.9 (Purcell *et al*, 2007; Chang *et al*, 2015). Only individuals and SNPs with a genotyping rate $> 95\%$ were included. A minor allele frequency threshold of > 0.05 was applied. QC criteria were performed separately on the data from the SNP50 and SNP670 genotyping platforms. Only SNPs that passed QC on both platforms were included in the SNP dataset. SNPs were then pruned on the basis of linkage disequilibrium (LD) using the parameter –indep 50 5 2 in PLINK. Pruning of SNPs that are in high LD have been shown to counter the effect of ascertainment bias and to generate meaningful comparisons among breeds (Molomane *et al*, 2018). Following QC and pruning, a set of 9,015 SNPs was used for the analyses.

Population genetics analyses

Pairwise genetic distance (D) between all individuals was calculated using PLINK 1.9 (–distance 1-IBS) and was represented by constructing a neighbour-joining tree. Principal component analysis (PCA) was conducted using PLINK 1.9. PCA plots were generated within the R environment (R Core Team, 2014). The PopGenome package in R was used to calculate the pairwise population differentiation index F_{ST}. Population structure was visualized using an unrooted neighbour-joining tree based on the between-population F_{ST}.

For the analysis of population substructure, model-based clustering was performed using the software package ADMIXTURE (Alexander *et al*, 2009). The model assigns ancestry based on a predefined number of K ancestral populations. Individuals are assigned to K clusters based on allele frequencies and the proportion of ancestry from each population is estimated. The analysis was performed for K ranging from 2 to 8.

Table 1. Breeds, abbreviations, sample sizes (N) and study cohort for horses used in the study.

Population	Abbreviation	N	Cohort
Connemara Pony	CP	24	Ancestral
Hanoverian	HAN	14	Ancestral
Irish Draught	ID	21	Ancestral
Irish Thoroughbred	ITB	25	Ancestral
Irish Sport Horse with Warmblood	ISH WB	8	Test
Irish Sport Horse	ISH	7	Test
Swiss Warmblood	SZW	14	Ancestral
Traditional Irish Horse	TIH	8	Test
Traditional Irish Horse	TIH	12	Validation
Total		**133**	

Results

Genetic distance within and among breeds

The greatest genetic distance between individuals was between a Connemara Pony and a Traditional Irish Horse (D = 0.29) and the closest genetic distance between individuals was between two Irish Thoroughbreds (D = 0.19). The average genetic distances among individuals within breeds ranged from 0.22 (Irish Thoroughbred) to 0.26 (Irish Draught).

The pairwise population differentiation index, F_{ST}, was calculated among breeds (Table 2) and used to create a neighbour-joining tree (Figure 1) to illustrate the topological relationship among breeds. The Traditional Irish Horse was identified as a distinct subset of the Irish Sport Horse population and was more similar to the Irish Sport Horse that did not contain foreign blood (F_{ST} = 0.006) than the Irish Sport Horse with Warmblood (F_{ST} = 0.011). The pairwise genetic difference between the Traditional Irish Horse and Irish Sport Horse with Warmblood (F_{ST} = 0.011) was greater than between Hanoverian and Swiss Warmblood (F_{ST} = 0.008), which are distinct breeds.

To visualize the branching structure among the Irish Sport Horse populations and Warmbloods, genetic distances between individuals were used to create a neighbour-joining tree (Figure 2). Traditional Irish Horse individuals were on branches separate to the two main Warmblood branches with the exception of a single individual that was observed among a predominantly Swiss Warmblood cluster. Five of the eight Irish Sport Horse with Warmblood samples were contained in the two main Hanoverian/Swiss Warmblood branches.

To further visualize the overall population structure, a PCA analysis was performed. In the PCA, PC1 explained 7% of the variance with clear separation of Connemara Pony from the other breeds as well as divergence between Irish Thoroughbred and Irish Draught (Figure 3A). In the plot created from PC2 and PC3 (Figure 3B), there was considerable overlap among the Hanoverian and Swiss Warmblood with a general European Warmblood cluster observed. There was a broad distribution of the Irish Sport Horse population centred among the recorded ancestral breed populations reflecting the varying contributions from the other breeds.

Population substructure

In the admixture analysis, $K = 4$ had the lowest estimate of the standard error of the cross-validation. Based on this result as well as the observations in the PCA plot and knowledge of breed history, $K = 4$ was chosen as the most appropriate number of clusters to use to establish ancestry and quantify admixture within the Irish Sport Horse population (Puechmaille, 2016; Lawson *et al*, 2018).

Unsupervised modelling was used to predict allele frequencies in four ancestral genetic lineages and each individual's genome was partitioned and proportionally assigned to one of the lineages (Figure 4). There was clear evidence of four distinct lineages – Irish Thoroughbred, Connemara Pony, Irish Draught and Warmblood – among the five distinct breed populations. The proportion of Warmblood ancestry in individual Irish Sport Horses ranged from 0.04 to 0.88.

To focus on the Irish Sport Horse alone, the proportion of Warmblood ancestry assigned to each of the 23 Irish Sport Horse samples was used to categorize the samples (Figure 5). Using the proportion of Warmblood admixture as an indicator, the seven Irish Sport Horse samples with the lowest levels of Warmblood lineage were classified as Traditional Irish Horse (Warmblood lineage $\leq$ 0.15 (15%)), the eight with the highest levels were classified as Irish Sport Horse with Warmblood influence and the remainder were assigned as Irish Sport Horse of unknown lineage. Six of the eight Traditional Irish Horse were correctly assigned. Two Traditional Irish Horse were assigned as Irish Sport Horse of unknown lineage, one of which was the Traditional Irish Horse that clustered with the predominantly Swiss Warmblood horses in the neighbour-joining tree. One Irish Sport Horse with missing pedigree was assigned as a Traditional Irish Horse.

All eight Irish Sport Horse with Warmblood were correctly assigned and overall, 20 of the 23 horses were correctly assigned using genotyping data from 9,015 SNPs. These results demonstrate a sensitivity of

Table 2. Pairwise population differentiation index (F_{ST}) among horse breeds: Connemara pony (CP), Hanoverian (HAN), Irish Draught (ID), Irish Sport Horse with Warmblood (ISH WB), Irish Sport Horse (ISH), Swiss Warmblood (SZW), Irish Thoroughbred (ITB), Traditional Irish Horse (TIH).

	CP	HAN	ID	ITB	ISH WB	ISH	SZW	TIH
CP	-							
HAN	0.049	-						
ID	0.041	0.031	-					
ITB	0.096	0.043	0.069	-				
ISH WB	0.049	0.012	0.027	0.040	-			
ISH	0.044	0.014	0.012	0.033	0.012	-		
SZW	0.048	0.008	0.029	0.038	0.011	0.014	-	
TIH	0.052	0.018	0.022	0.020	0.011	0.006	0.014	-

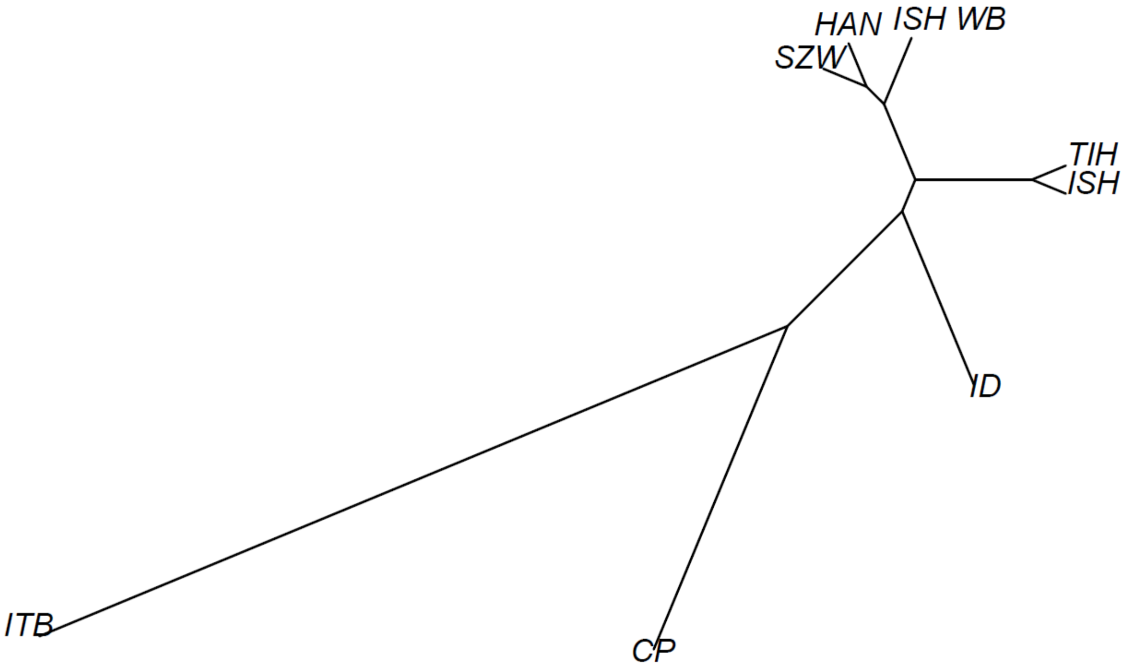


Figure 1. Neighbour-joining tree illustrating the genetic distance between breeds: Connemara pony (CP), Hanoverian (HAN), Irish Draught (ID), Irish Sport Horse with Warmblood (ISH WB), Irish Sport Horse (ISH), Swiss Warmblood (SZW), Irish Thoroughbred (ITB), Traditional Irish Horse (TIH).

75% and specificity of 93% for assigning Irish Sport Horse of unknown ancestry as Traditional Irish Horse, with an accuracy of predicting Traditional Irish Horse from genetic data of 87.5%. Admixture analysis of a validation set of 12 pedigree-certified Traditional Irish Horse revealed similar results with Warmblood lineage contributions in individuals ranging from 0.01 to 0.15.

Discussion

There is a growing trend in equestrian breeding to move away from the traditional approach of selection within breeds, towards developing hybrid sport horses (Gilbert and Gillet, 2011). Originally the Irish Sport Horse was a breed that was established by crossing the Irish Draught with Thoroughbred horses (Reilly *et al*, 1998; Alexander *et al*, 2009). More recently, crossbreeding the Irish Sport Horse with European Warmblood breeds has become popular with the intention of improving the quality of the horses for sport (Doyle *et al*, 2022). This practice threatens the maintenance of genetic variation within the traditional lineages of the Irish Sport Horse that may be beneficial to preserve.

In our phylogenetic analyses, we found that the Irish Sport Horse is genetically distinct from European Warmblood populations. By contrast, there was no clear genetic distinction between the two European Warmblood breeds (Hanoverian and Swiss Warmblood) included in this study. This agrees with a microsatel-

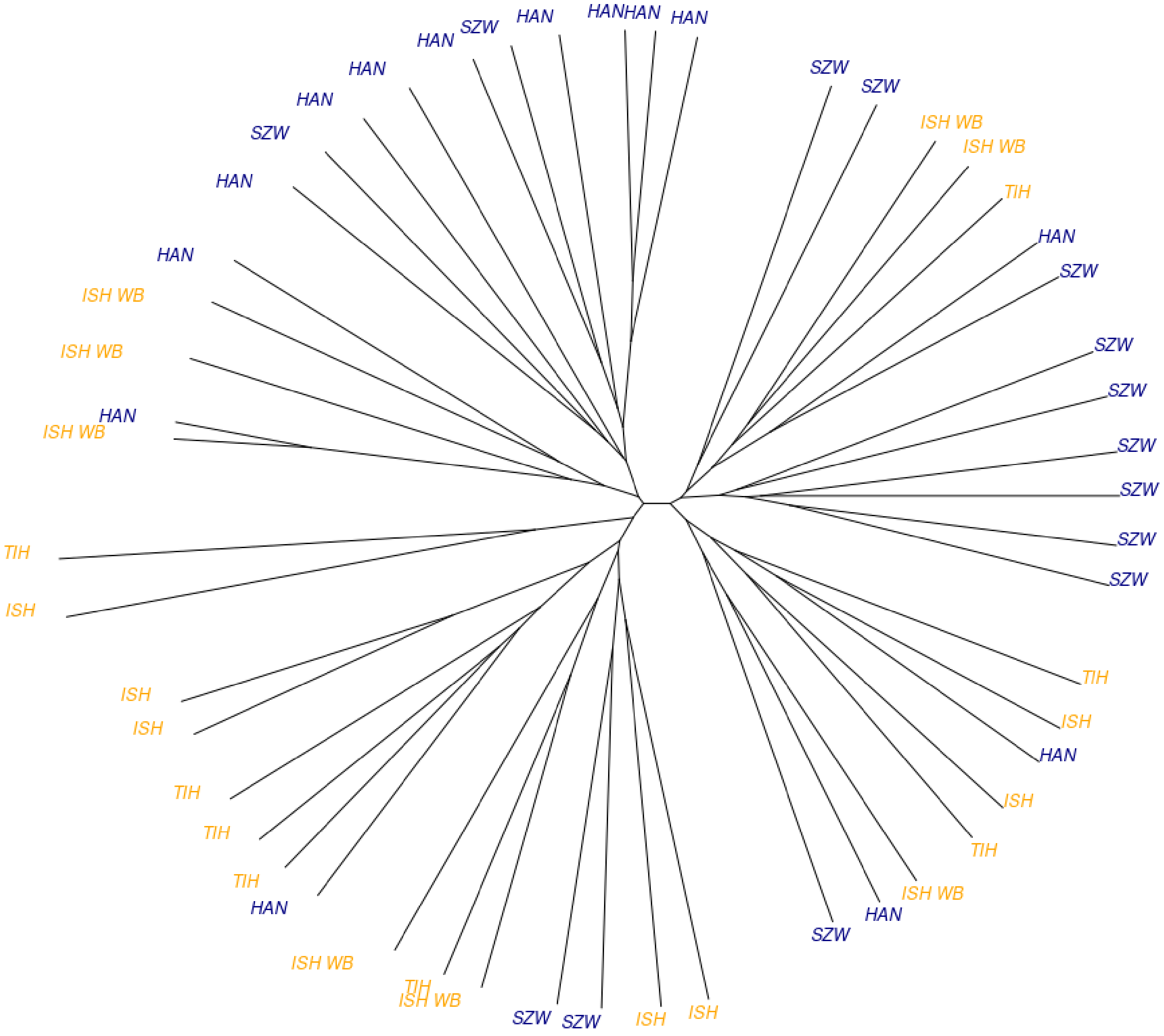


Figure 2. Neighbour-joining tree illustrating individual relationships among Irish Sport Horses (yellow) and Warmbloods (blue; HAN, SZW); including ISH with no foreign blood (ISH), ISH with foreign blood (ISH WB) and Traditional Irish Horse (TIH).

lite genotype-based phylogeny in which Hanoverian clustered closest to other European warmblood breeds (including the Italian Maremmano horse breed) on a branch most distant from other geographically proximal German coldblood breeds (Felicetti *et al*, 2010). The Irish Sport Horse has not previously been included in population genetic diversity studies and there are limited studies including the Irish Draught and Connemara Pony. In a microsatellite, protein and blood group marker study, Hanoverian clustered with Holsteiner, Thoroughbred, Quarter Horse and Irish Draught, suggested to be due to the influence of Thoroughbred in these breeds (Luis *et al*, 2007), and reflected in a more recent genome-wide SNP analysis of genetic diversity in which the European warmbloods (Hanoverian, Swiss Warmblood, Maremmano) clustered with recently admixed Thoroughbred breeds (Petersen *et al*, 2013).

Between breeds, the pairwise genetic difference (F_{ST}) indicated that the Traditional Irish Horse was more different from the Irish Sport Horse with Warmblood than the two European warmblood breeds (Hanoverian, Swiss Warmblood) were from each other. These genetic results indicate that there may be grounds to identify the Traditional Irish Horse as a distinct breed for conservation purposes, should this be desired by breeders. The term 'breed' is generally used to describe a population of animals with common phenotypically distinct traits, and it has been suggested that a breed

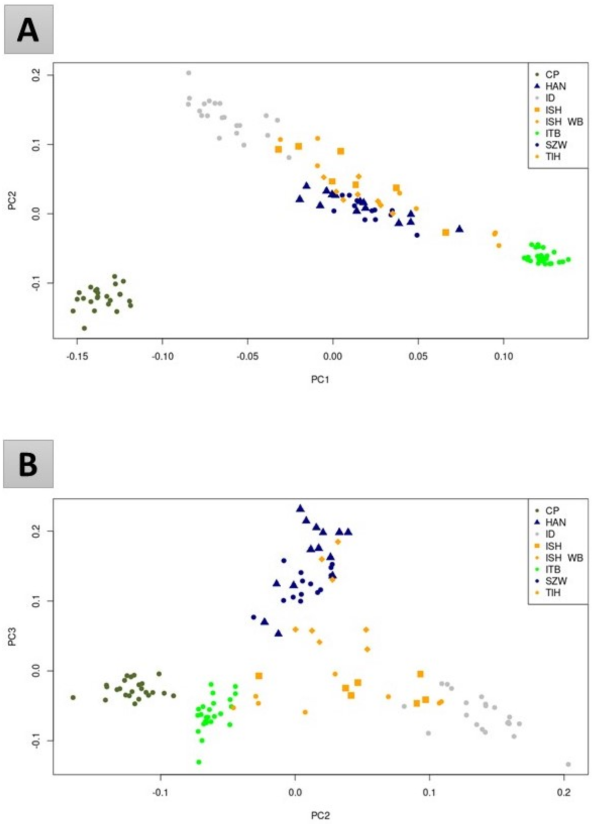


Figure 3. Principal component analysis plots for PC1 and PC2 (A), and PC2 and PC3 (B) for n = 121 individuals coloured according to breed: Connemara Pony (CP), Hanoverian (HAN), Irish Draught (ID), Irish Sport Horse with Warmblood (ISH WB), Irish Sport Horse (ISH), Swiss Warmblood (SZW), Irish Thoroughbred (ITB), Traditional Irish Horse (TIH). PC1, PC2 and PC3 explained 7%, 4% and 3% of the genetic variance, respectively.

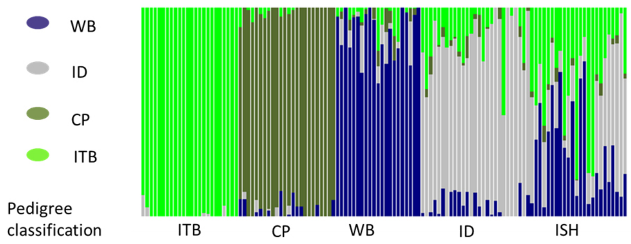


Figure 4. Proportion of genetic ancestry assigned to each breed where a column indicates a single horse and colours represent the four ancestry clusters. The colour of the primary cluster representing each breed ancestral component is depicted in the legend. Breed identifiers are noted at the bottom: Irish Thoroughbred (ITB), Connemara Pony (CP), Warmblood breeds (WB), Irish Draught (ID), Irish Sport Horse (ISH).

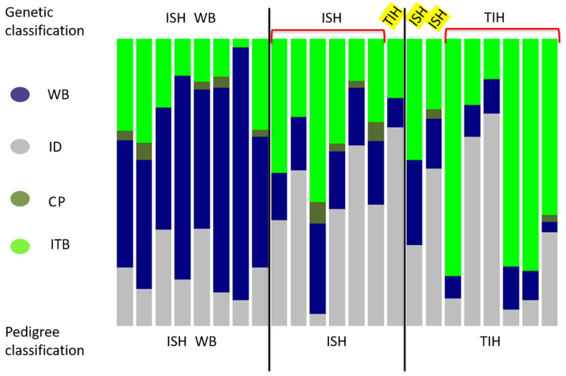


Figure 5. Proportion of genetic ancestry of 23 Irish Sport Horses assigned to each cluster where a column indicates a single horse and colours represent the four clusters. The colour of the primary cluster representing each breed ancestral component is depicted in the legend. The Irish Sport Horses based on genetic classification (top) and pedigree classification (bottom) are indicated as Warmblood (WB), Irish Draught (ID), Connemara Pony (CP), Irish Thoroughbred (ITB), Irish Sport Horse with Warmblood (ISH WB), Irish Sport Horse (ISH), and Traditional Irish Horse (TIH). The vertical black lines separate the horses on the basis of pedigree classification. All except for three horses (yellow highlights) were classified correctly on the basis of genetics; two horses with Traditional Irish Horse pedigrees were genetically assigned as Irish Sport Horse, one horse with Irish Sport Horse pedigree was genetically assigned as Traditional Irish Horse.

may also be a cultural concept defined by breeders (FAO, 2023). However, genetic analyses now enable the examination of distinct and/or common genotypes to provide scientifically informed categorization of breeds (FAO, 2023).

In order to genetically differentiate the Traditional Irish Horse from phenotypically similar Irish Sport Horse with Warmblood, here we evaluated the ancestral genetic contributions to various cohorts of the Irish Sport Horse, defined on the basis of pedigree, to delineate the maximal ancestry component of Warmblood genetics required to determine Traditional Irish Horse status. We observed four of the eight Irish Sport Horse with Warmblood to have comparable Warmblood ancestry to horses in the Hanoverian and Swiss Warmblood cohorts; however, a distinguishing feature of the Irish Sport Horse with Warmblood was a greater Irish Draught influence compared to the Hanoverian and Swiss Warmblood cohorts. In some of the Irish Sport Horse with Warmblood, there was more Warmblood ancestry component than in some of the European Warmblood animals.

Since by 2014 more than 37% of registered Irish Sport Horses had Warmblood ancestry recorded in their pedigrees, our results indicate that concerns regarding European Warmblood introgression and the creation of a homogeneous pan-European sport horse population are valid. Without a dedicated breeding programme for the Traditional Irish Horse, the unique genetic heritage of the Irish Sport Horse may be lost. Current efforts to recognize and assign Traditional Irish Horse status to traditionally bred horses are hindered by the fact that many Irish Sport Horses may have an unregistered dam or sire in their pedigree. This is a situation where ancestral profiling using genomics-based technologies could provide a basis for identifying the Irish Sport Horse with traditional lineages. The FAO practical guide for genomic characterization of animal genetic resources notes that "Genetic tools now allow for errors in recording of parentage to be identified and remedied; this is especially useful in the case of unknown paternity" (FAO, 2023). This has immediate practical implications, since in this study there was one Irish Sport Horse with missing pedigree that had Warmblood ancestry within the range of the Traditional Irish Horse (10%), indicating that this individual has a traditional genetic heritage. Applying a cut-off of $\leq$ 15%, Warmblood ancestry would have an accuracy of prediction of 87.5% for the Traditional Irish Horse.

Conclusion

Considering the unique position of the Irish Sport Horse within the international sport horse world, with traits for elite eventing abilities and versatility as a valued leisure horse, the Traditional Irish Horse should be conserved and the genetic legacy of the Irish Hunter protected. The FAO practical guide for genomic characterization of animal genetic resources notes that the results of genomic characterization studies "should also be used to improve the management of the AnGR [Animal Genetic Resources] involved" (FAO, 2023). The application of genetic ancestry testing to identify the Traditional Irish Horse, particularly in the absence of complete pedigree information for a horse, would provide a novel route for conservation efforts and the promotion of the Irish horse in equestrian sport.

Acknowledgments

We thank the Traditional Irish Horse Association for assistance with the study. The project was funded by the Department of Agriculture, Food and the Marine under Equine Technical Support & Equine Breeding Schemes and by Horse Sport Ireland with in-kind contribution from Plusvital Ltd.

Data availability

Data will be publicly available on the European Variation Archive http://www.ebi.ac.uk/eva/

Authors' contribution

BAM contributed to the study design, performed the data analysis and interpretation, drafted the manuscript and approved the submitted manuscript. DH conceived the project, contributed to the study design and sample acquisition. AC conceived the project, contributed to the study design and sample acquisition. EWH conceived the project, contributed to the study design, performed the interpretation, drafted the manuscript, revised the manuscript and produced the final submitted manuscript.

Conflict of interest

BAM is employed by and EWH is a shareholder in Plusvital Ltd. Plusvital Ltd funded the project in-kind with salaries and technical resources.

References

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

Bendrey, R., Thorpe, N., Outram, A., and Van Wijngaarden-Bakker, L. (2013). The Origins of Domestic Horses: New Direct Dates on the Horses of Newgrange, Ireland. *Proceedings of the Prehistoric Society* 79, 91–103. doi: https://doi.org/10.1017/ppr.2013.3

Bower, M. A., Campana, M. G., Whitten, M., Edwards, C. J., Jones, H., Barrett, E., Cassidy, R., Nisbet, R. E., Hill, E. W., Howe, C. J., and Binns, M. (2011). The cosmopolitan maternal heritage of the Thoroughbred racehorse breed shows a significant contribution from British and Irish native mares. *Biol Lett* 7. doi: https://doi.org/10.1098/rsbl.2010.0800

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4. doi: https://doi.org/10.1186/s13742-015-0047-8

Doyle, J. L., Carroll, C. J., Corbally, A. F., and Fahey, A. G. (2022). An overview of international genetic evaluations of show jumping in sport horses. *Transl Anim Sci* 6. doi: https://doi.org/10.1093/tas/txac038

FAO (2023). Genomic characterization of animal genetic resources, ed. Ajmone-Marsan, P., Boettcher, P. J., Colli, L., Ginja, C., Kantanen, J., and Lenstra, J. A. (Rome, Italy: FAO Animal Production and Health Guidelines No. 32). doi: https://doi.org/10.4060/cc3079en

Felicetti, M., Lopes, M. S., Verini-Supplizi, A., Ada, C. M., Silvestrelli, M., Mendonca, D., and Distl, O. (2010). Genetic diversity in the Maremmano horse and its relationship with other European horse breeds. *Anim Genet* 41(2). doi: https://doi.org/10.1111/j.1365-2052.2010.02102.x

Gilbert, M. and Gillet, J. (2011). Equine athletes and interspecies sport. *International Review for the Sociology of Sport* 47. doi: https://doi.org/10.1177/1012690211416726

Hendricks, B. L. (1995). International Encyclopedia of Horse Breeds (University of Oklahoma Press).

Lawson, D. J., Van Dorp, L., and Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun* 9. doi: https://doi.org/10.1038/s41467-018-05257-7

Luis, C., Juras, R., Oom, M. M., and Cothran, E. G. (2007). Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Anim Genet* 38, 20–27. doi: https://doi.org/10.1111/j.1365-2052.2006.01545.x

Lydon, J. F. (1954). The Hobelar: An Irish Contribution to Medieval Warfare .

McCormick, F. (2007). The horse in early Ireland. *Anthropozoologica* 42, 85–104.

McGahern, A. M., Edwards, C. J., Bower, M. A., Heffernan, A., Park, S. D., Brophy, P. O., Bradley, D. G., MacHugh, D. E., and Hill, E. W. (2006). Mitochondrial DNA sequence diversity in extant Irish horse populations and in ancient horses. *Anim Genet* 37. doi: https://doi.org/10.1111/j.1365-2052.2006.01506.x

Molomane, D. K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A. R., and Simianer, H. (2018). Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics* 19(1). doi: https://doi.org/10.1186/s12864-017-4416-9

Petersen, J. L., Mickelson, J. R., Cothran, E. G., Andersson, L. S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M. M., Borges, A. S., Brama, P., Machado, A. D. C., Distl, O., Felicetti, M., Fox-Clipsham, L., Graves, K. T., Guerin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E. W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M. S., Mcgivney, B. A., Mikko, S., Orr, N., Penedo, M. C., Piercy, R. J., Raekallio, M., Rieder, S., Roed, K. H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Mccue, M. W., and E, M. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One* 8. doi: https://doi.org/10.1371/journal.pone.0054997

Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol Ecol Resour* 16, 608–635. doi: https://doi.org/10.1111/1755-0998.12512

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Manuel, A. R. F., Bender, D., Maller, J., Sklar, P., and De Bakker (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics* 81. doi: https://doi.org/10.1086/519795

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing (Vienna, Austria) .

Reilly, M., Foran, M., Kelleher, D., Flanagan, M., and Brophy, P. (1998). Estimation of genetic value of showjumping horses from the ranking of all performances in all competitions. *Journal of Animal Breeding and Genetics* 115, 17–25. doi: https://doi.org/10.1111/j.1439-0388.1998.tb00324.x

Unattributed (c. 800). The Book of Kells (Trinity College Library Dublin) . doi: https://doi.org/10.48495/hm50tr726

Youatt, W. (1831). The horse; with a treatise on draught (London, UK: Baldwin and Cradock).

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2023), 4 (8), 37–54
DOI: 10.46265/genresj.HFWZ5263
https://www.genresj.org
ISSN: 2708-3764

Evaluating agromorphological traits of Greek wheat landraces and exploring their potential for bread and pasta making based on seed physical properties

Styliani V Protonotariou [*,a], **Ricos Thanopoulos** [b], **Anastasios Katsileros** [c], **Penelope J Bebeli** [*,c] **and Ioanna Mandala** [a]

[a] *Laboratory of Food Process Engineering, Department of Food Science and Human Nutrition, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece*

[b] *Farm Unit, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece*

[c] *Laboratory of Plant Breeding and Biometry, Department of Crop Science, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece*

Abstract: Physical and morphological characteristics of Greek wheat landrace and cultivar seeds were evaluated, aiming to assess their bread- and pasta-making quality. Furthermore, the plant agromorphological traits of wheat landraces and cultivars were measured and correlated with seed physical properties. Fifteen *Triticum* spp. accessions, out of which four cultivars (two *Triticum aestivum* and two *T. durum*), nine landraces (two *T. durum*, five *T. aestivum*, one *T. dicoccum* and one *T. polonicum*) and two *T. durum* of unknown status were studied. Seventeen morphological plant and seed traits were measured based on UPOV descriptors. Ear emergence was earlier in cultivars than in landraces. The plant weight of the landraces was, in some cases, almost twice the cultivars'. Seed firmness (62.96–194.85N) was positively correlated with thousand kernel weight and volume (0.840 and 0.791, $P < 0.05$, respectively). Based on the agromorphological traits, cluster analysis clearly separated the species and the cultivars from the landraces, and the unknown status accessions were grouped with the cultivars. Cluster analysis with all measured traits set the wheat accessions in the following distinct groups: (1) *T. durum* cultivars including the unknown status accessions, (2) *T. aestivum* cultivars, and (3) *T. aestivum* and *T. durum* landraces. *T. polonicum* and *T. dicoccum* each formed a separate group. Seed physical properties of the analyzed Greek landraces indicated their suitability for bread and/or pasta making.

Keywords: Agromorphological characteristics, image analysis, seed firmness, wheat landraces, seed physical properties, pasta and bread making, bread wheat, durum wheat

Citation: Protonotariou, S. V., Thanopoulos, R., Katsileros, A., Bebeli, P. J., Mandala, I. (2023). **Evaluating agromorphological traits of Greek wheat landraces and exploring their potential for bread and pasta making based on seed physical properties**. *Genetic Resources* 4 (8), 37–54. doi: 10.46265/genresj.HFWZ5263.

Introduction

Many researchers in the past have tried to define the meaning of a landrace, which is also known as "a local, farmer's, traditional variety or a local population" (Zeven, 1998). Recently, a more general definition for landrace was suggested: "a cultivated, genetically heterogeneous variety that has evolved in a certain ecogeographical area and is therefore adapted to the edaphic and climatic conditions and to its traditional management and uses" (Casañas *et al*, 2017).

Landraces have been the main crop genetic resources used since the emergence of agriculture (Purugganan, 2019). In the last century, with agricultural intensification, several landraces have been replaced with high-yielding cultivars (hereafter the term cultivar means scientifically improved cultivated variety). The main differences between landraces and cultivars are that the latter are the results of formal breeding research (when compared to farmers' mass selection for landraces) and they are provided to farmers and growers through formal seed systems (when compared to traditionally infor-

---

[*] Corresponding authors: Styliani V Protonotariou (s.protonotariou@yahoo.gr), Penelope J Bebeli (bebeli@aua.gr)

Received: 14.02.2023 | Accepted: 24.08.2023 | Published online: 11.10.2023

mal channels, such as exchanges and local markets for landraces). Landrace cultivation has been almost abandoned and many landraces have been lost, leading to severe genetic erosion (Robbana *et al*, 2021), which is the loss of crop variation due to the modernization of agriculture (Van De Wouw *et al*, 2010).

In the case of wheat, a study conducted by Haudry *et al* (2007), concluded that the estimated initial biodiversity decreased by 84% for durum wheat and by 69% for bread wheat, although such estimations could be controversial (Khoury *et al*, 2022). Genetic erosion of *Triticum aestivum* landraces in Ethiopia was found to be 66.7% in 2016 (Gelelcha *et al*, 2023).

In Greece, only 3% of the total land surface cultivated with durum wheat was planted with landraces in 1984, and 1.3% in 1991 (Kyzeridis *et al*, 1995). On an expedition encompassing 36 villages on Lefkada island, from 2010 to 2012, researchers were able to collect only three wheat accessions, two of which were landraces (Thomas *et al*, 2013). Although the genetic erosion of wheat landraces was shown by collecting expeditions carried out in Greece, they can still be found cultivated on farms (Thanopoulos *et al*, 2021). The study of landrace properties is crucial both for local communities, which would benefit from the added value of their products, and their use in breeding programmes. Thanopoulos *et al* (2021), who collected landraces in Arcadia, Greece, estimated wheat genetic erosion between 1942 and 2021 to be 45.7%. Moreover, recent research revealed a 97% genetic erosion of winter wheat for the years 1955–2015 in Central Europe (Cseh *et al*, 2021).

Cereal cultivars are high-yielding, but the majority of them is less resilient to abiotic and biotic stresses compared to landraces (Newton *et al*, 2010). Reduced adaptation to organic, low-input production farming systems and poorer genetic diversity were observed in Tunisian improved durum wheat cultivars, in contrast to landraces (Ayadi *et al*, 2020). Although further research is required, there is also evidence that ancient wheat types, such as emmer wheat, provide health benefits compared to modern wheat cultivars (Shewry, 2018) and could be a potential source of antioxidants, carotenoids and phenolics (Newton *et al*, 2010). For such ancient species, like *T. dicoccum*, and neglected and underutilized species, like *T. polonicum*, (Bieńkowska *et al*, 2020; Carvalho *et al*, 2019) more research to investigate their properties is needed.

Currently, climate change, extensive loss of genetic diversity, increased vulnerability of cultivars to abiotic and biotic stresses and consumer awareness of the value of local products, have raised interest in wheat landraces (Brush, 1995; Moghaddam *et al*, 1997; Saleh, 2020; Tomás *et al*, 2020). Wheat landraces could play a crucial role in mitigating the negative effects of climate change due to their ability to perform sufficiently under high temperatures (Ulukan, 2021) and exhibit higher yield stability under water stress (Karamanos *et al*, 2017). Recent research focused on the collection of landraces, their traits and cultivation revival as well as the promotion of sustainable agricultural production (Newton *et al*, 2010). It also highlighted the characteristics associated with their utilization in the production of high-quality products for consumers (López-Fernández *et al*, 2021; Ruisi *et al*, 2021).

The analysis of the agronomic performance of wheat landraces revealed significant differences in plant height, seed colour and biological cycle (Gharib *et al*, 2021). A study of morphological characteristics of old Sicilian landraces showed high variability among durum wheat accessions (Sciacca *et al*, 2014). Similar results were also obtained for days to emergence, tillering, booting, flowering and physiological maturity, plant height, and thousand kernel weight (TKW) in 23 Moroccan durum wheat accessions (Zarkti *et al*, 2010). In a study of 63 durum wheat landraces, grain weight was the most significant criterion for improving their adaptation to the northern parts of the Mediterranean basin (Moragues *et al*, 2006).

Several studies compared grain characteristics of landraces and other variety types. In one study, ten obsolete varieties (varieties that were developed by systematic breeding efforts in the past and now have been replaced by newer and better-performing cultivars) (Maxted *et al*, 2011) and two wheat landraces from Bosnia and Herzegovina were evaluated for grain length, width, thickness, volume, surface area, TKW and protein content, and most of the traits presented significant interactions among them. Obsolete wheat genotypes 'Brkulja' and 'Šianka' showed the most promising grain characteristics (Kondić *et al*, 2020). Kernel morphology is important for manufacturing final food products, which requires specific grain characteristics (Dholakia *et al*, 2008). Machine-vision techniques (image analysis) were used as a useful tool for the classification of varieties based on morphology, colour and seed texture of different species (Smykalova *et al*, 2011, 2013; Chaugule and Mali, 2016) and wheat seed quality assessment (Venora *et al*, 2009). Other studies evaluated quality parameters, such as starch and protein content of wheat flour obtained from landraces in comparison to cultivars. Boukid *et al* (2018) detected higher values of protein and gliadin in Tunisian durum wheat landraces, while modern genotypes had significantly higher values of total starch. Zheng *et al* (2012) studied 477 Chinese bread wheat landraces (*T. aestivum* L.) and concluded that there was a significant correlation between starch and protein content. Mediterranean durum wheat landraces presented a higher amount of protein content but lower gluten strength than cultivars (Roselló *et al*, 2018). In addition, a lower content of water-extractable arabinoxylans, the main non-starch polysaccharides in cereals, was found in Turkish landraces (*T. monococcum*, *T. durum* and *T. aestivum*) than in cultivars (*T. aestivum*) (Cetiner *et al*, 2020). The comparison of different aromatic profiles and the characteristics of the processed products showed high diversity in Sicilian durum wheat landraces in contrast to the uniformity of cultivars (Ruisi *et al*, 2021). Recently, some studies focused on the physical or functional characteristics of final products made from wheat landraces. For example, some soft wheat landraces could be used for biscuit production, since their grain hardness did not significantly reduce biscuit quality (Igrejas *et al*, 2002). Spanish landraces showed lower bread-making quality properties because of their low-quality protein compared to cultivars (Gómez *et al*, 2009). High variability of Mediterranean landraces could be used for improving gluten quality (Nazco *et al*, 2014).

Even though many papers separately examine either plant agromorphological traits or seed physico-mechanical properties, there is limited published work that studies holistically the above traits with both agronomic and food technology approaches. For instance, in a comparative study of five cultivars and two landraces, the latter were significantly taller, flowered later, and had lower yields and gluten index but higher protein content in comparison to wheat cultivars (Preiti *et al*, 2022). This field needs more data to critically evaluate this issue. The purpose of the present study was to characterize wheat landraces and cultivars according to morphological, agronomical, seed quality and technological traits and investigate their potential for bread and/or pasta making. Specifically, the research aimed to: (1) study the agromorphological characteristics and seed physical properties, (2) compare *T. durum*, *T. aestivum*, *T. dicoccum*, and *T. polonicum* wheat landraces with cultivars, (3) classify all the accessions (cultivars and landraces) based on agromorphological traits, separately or on the whole set of traits (both agromorphological and seed physical properties) and (4) identify the wheat landraces more suitable for bread- and pasta-making purposes.

Materials and methods

Plant material

The plant material consisted of 15 *Triticum* spp. accessions, namely two cultivars and five landraces of *T. aestivum*, two cultivars, two landraces and two accessions of unknown status of *T. durum*, one landrace of *T. polonicum* and one landrace of *T. dicoccum*. The latter was collected in the early 20th century in Greece, it was conserved in USDA-GRIN (Code PI 94682) and repatriated recently. Landraces were collected from different regions in Greece and donated to the Plant Breeding and Biometry Laboratory of the Agricultural University of Athens (AUA) (Table 1 and Figure 1). The Institute of Plant Breeding and Genetic Resources (IPGRB) donated four cultivars, recommended for bread making ('Yekora' and 'Elisavet' belonging to *T. aestivum*) and pasta making ('Mexicali 81' and 'Elpida', of *T. durum*) which were used as controls.

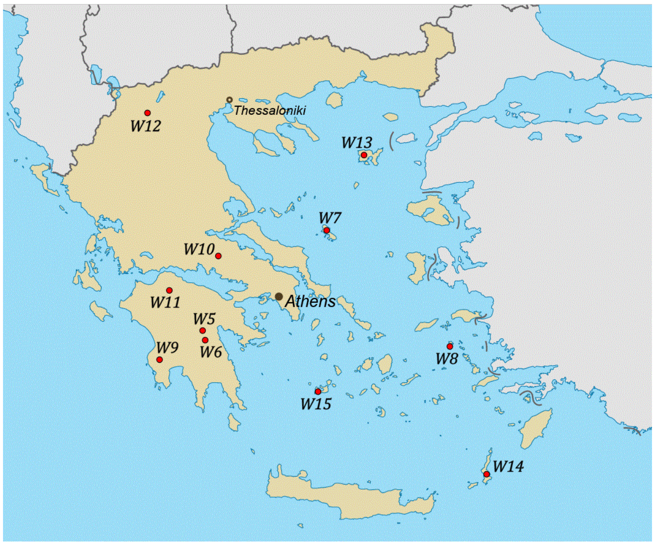


Figure 1. Map of Greece with collection sites of wheat landraces characterized in this study (for code explanations see Table 1).

Experimental design

The experiment was conducted at an experimental field of the Agricultural University of Athens (location: 37° 59' 06.8" N, 23° 42' 24.7" E, altitude 24m a.s.l.) during the winter growing season of 2019–2020. The meteorological data are presented in Figure 2. The values of mean air temperature and precipitation for the Athens region during the years 2015–2020 were provided by the Hellenic National Meteorological Service (HNMS, 2023). A randomized complete block design with three replications was used. Each replication consisted of 20 plants of each accession sown in rows spaced 0.2m apart. Sowing distances among plants in each row was 20cm. The sowing date was mid-November, a typical period for sowing in central Greece.

Manipulation of sowing time is one of the most effective techniques for the production of high-quality wheat kernels (Butkovskaya and Kozulina, 2021). In the present study, for comparison reasons, plant sowing was carried out simultaneously for all accessions, without considering the best sowing time for each accession.

Measurement of agromorphological traits

During vegetative and reproductive stages, time of ear emergence (days), total number of tillers, ear length excluding awns (cm), ear length including awns (cm), number of spikelets in the spike of the first tiller, plant length (measured from the base of the plant to the tip of the highest awn, cm), stem length (measured from plant base to ear base, cm) and whole plant weight (physically dried by sun, in g/plant) were measured. In addition, the qualitative variables ear colour, ear shape in profile, growth habit and awn colour, were scored. Agromorphological traits were evaluated on all the plants per plot to calculate the mean value for the quantitative traits and the median for the qualitative traits. The measurements of the following traits: plant

Table 1. Accession codes, taxonomic classification, names and origin of wheat accessions.[1]Nomenclature (Hammer *et al*, 2011). [2]About 5% of the sample also contained *T. aestivum*, [3]Institute of Plant Breeding and Genetic Resources (IPGRB)

Accession code	Name	***Triticum*** **species**[1]	Breeding status	Origin	Collection date
W1	Yekora	*T. aestivum* L.	Cultivar	IPGRB[3]	Donated 2019
W2	Elisavet	*T. aestivum* L.	Cultivar	IPGRB	Donated 2019
W3	Elpida	*T. durum* Desf.	Cultivar	IPGRB	Donated 2019
W4	Mexicali 81	*T. durum* Desf.	Cultivar	IPGRB	Donated 2019
W5	Zoulitsa	*T. aestivum* L.	Landrace	Arcadia-Peloponnese	2018
W6	Ntopio	*T. aestivum* L.	Landrace	Arcadia-Peloponnese	2018
W7	Mavragani-Skyros	*T. durum* Desf.	Unknown status	Skyros Island	2012
W8	Grilos	*T. aestivum* L.	Landrace	Patmos Island	2013
W9	Leventis	*T. polonicum* L.	Landrace	Messinia-Peloponnese	2013
W10	Kopaida	*T. dicoccum* Schrank	Landrace	USDA-GRIN, Central Greece	Repatriation 2013
W11	Asprositi -Kalavrita	*T. aestivum* L.	Landrace	Kalavrita-Peloponnese	2019
W12	Asprositi - Kozani	*T. durum* Desf.	Landrace	Kozani-West Makedonia	2005
W13	Mavragani-Lemnos	*T. durum* Desf.	Unknown status	Lemnos Island	2019
W14	Aspratheri	*T. durum* Desf.[2]	Landrace	Karpathos Island	2019
W15	Kokkinositaro	*T. aestivum* L.	Landrace	Milos Island	2019

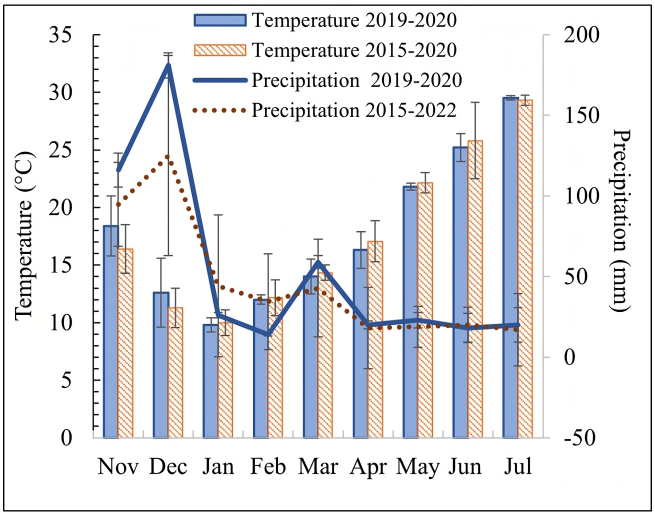


Figure 2. Mean air temperature (°C) and total monthly precipitation (mm) for Athens during the growing period (November 2019 to July 2020) and the average of the last five years (2015–2020). Data provided by HNMS (2023).

length, growth habit, ear emergence, ear length, ear colour, ear shape, and awn colour were carried out using UPOV descriptors (UPOV, 2012, 2017).

Physical properties of seeds

Before measurements, shrunken, broken kernels, and other impurities were removed. The weight of 1,000 randomly selected kernels of wheat was measured for thousand kernel weight (TKW). The same 1,000 selected kernels were then filled into a volumetric cylinder to determine their volume. Tap volume was measured after gently tapping the cylinder 100 times or until there was no further decrease in the sample level. Bulk density (ρ_b) and tap density (ρ_t) were calculated as weight of sample per volume and per tap volume of sample, respectively (g/ml) (Chaloulos *et al*, 2021).

Determination of flow properties is essential in processes such as transportation, mixing and storage (Fitzpatrick *et al*, 2004). The Carr index (CI) was used to calculate the flowability of the seeds (Carr, 1965):

$$CI\ (\%) = (\rho_t - \rho_b)\ /\rho_t \times 100$$

The firmness of seed (SF), defined as the force required to crush the seed, was estimated using an Instron (Universal Testing Machine, Model 3343, Nordwood, MA, USA) equipped with a 1kN load cell. Fifty intact seeds of each accession were compressed to 50% of their initial width with a 4cm diameter probe and a speed of 1mm/min.

Seed colour was measured using the colourimeters 3nh High-Quality Spectrophotometer NS800S (Shenzhen 3nh Technology, China), according to CIE-L*a*b* uniform colour space, where L* indicates lightness, a* indicates hue on a green (−) to red (+) axis, and b* indicates hue on a blue (−) to yellow (+) axis. White index (WI) was also calculated according to the following equation:

$$WI = 100 - \sqrt{(100 - L^{*2}) + a^{*2} + a^{*2}}$$

Image analysis-shape factor measurements

One hundred seeds of each accession were scanned twice using a flatbed scanner (HP scan jet 4370, Hewlett Packard, USA). Images of the seeds were captured for further analysis using an image analysis software (Image Pro Plus 7, Media Cybernetics, USA). An indicative image of the scanned seeds is presented in Figure 3A. Area (mm^2), aspect, average optical density (density/intensity mean), mean, maximum and minimum diameter (mm), perimeter (mm), roundness, size length (mm) and width (mm) were determined. Detailed descriptions and definitions of shape factors are presented in Figure 3B. The measurements were done in triplicates.

Statistical analysis

Statistical analysis of the data was performed with Statgraphics Centurion XV (Statgraphics, Rockville, MD, USA). For qualitative variables (plant growth habit, colour of ear, ear shape in profile and awn colour) frequency percentages (Supplemental Figure 1) and median values were calculated (Supplemental Table 1). Friedman Test was used to test the null hypothesis that the mean ranks of the samples groups are the same (Supplemental Table 1).

Quantitative trait data were subjected to analysis of variance after testing the assumptions of normality and homoscedasticity with Levene's and Bartlett's tests respectively (Supplemental Table 2). The comparisons of the means were performed using the Tukey HSD criteria with a level of significance of $\alpha = 0.05$.

The correlation between quantitative parameters was assessed by Pearson's correlation test (significance level at $\alpha = 0.05$). The correlation between qualitative parameters (nonparametric measures) was assessed by Spearman rank correlation test (significance level at $\alpha = 0.05$). To correlate quantitative and qualitative variables the Spearman test was also used (Supplemental Table 3). Cluster analysis was used to group the quantitative observations with similar characteristics using nearest neighbour (single linkage) and distance metric of squared Euclidean on standardized data.

Results and discussion

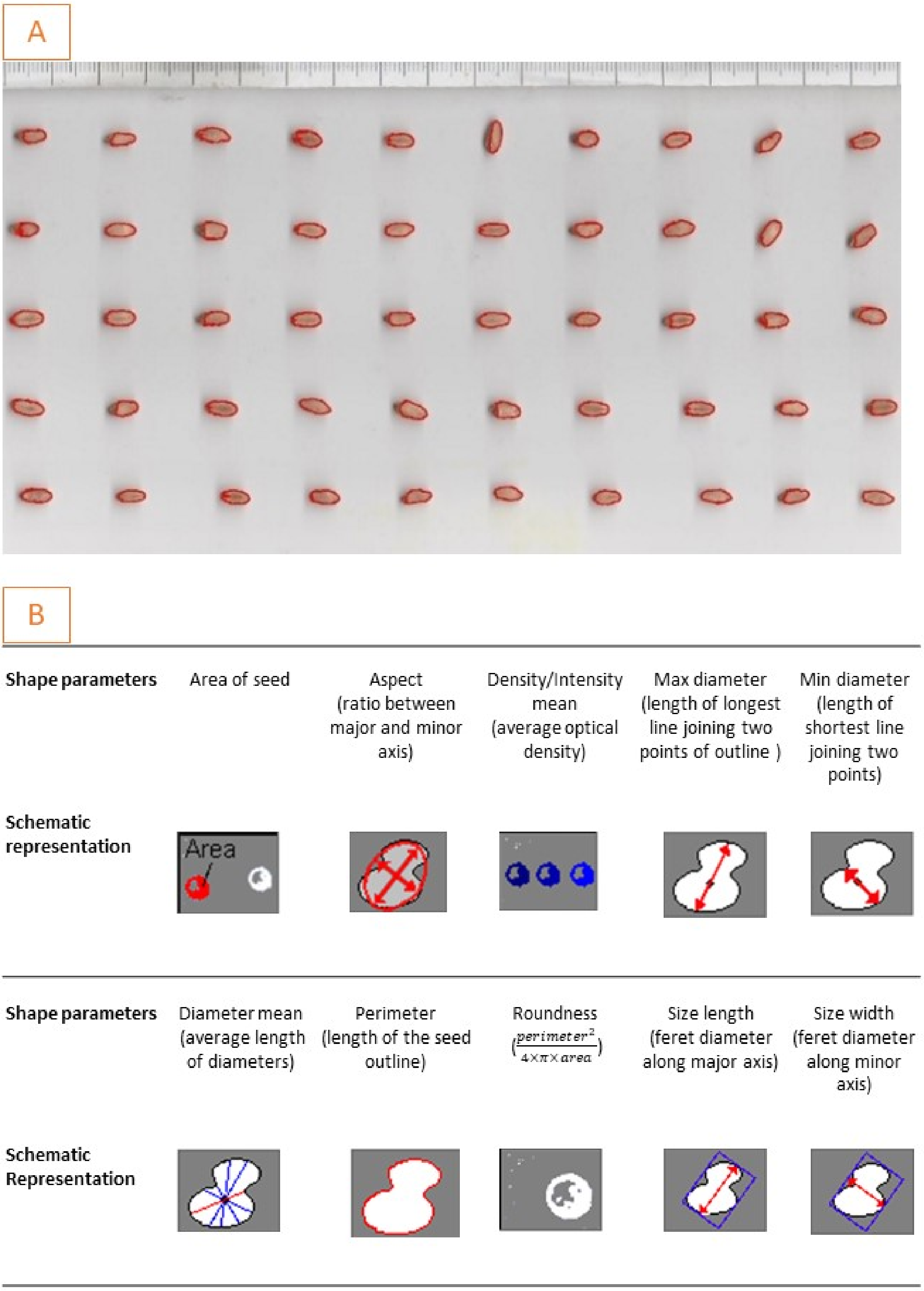
Agromorphological traits

Mean values of quantitative agromorphological traits for all studied wheat accessions are presented in Table 2. Time of ear emergence ranged from 107 to 138 days (Yekora - *T. aestivum*, cultivar, and Zoulitsa - *T. aestivum*, landrace, respectively). All cultivars, except Elisavet - *T. aestivum*, were earlier than landraces. Similarly, Frankin *et al* (2021) reported cultivars to be on average 12% earlier than landraces. On the other hand, the landraces Zoulitsa (*T. aestivum*), Kopaida (*T. dicoccum*) and Leventis (*T. polonicum*) had the latest heading time of all accessions. The *T. aestivum* landraces, Asprositi Kozani (W12), Ntopio (W6) and Asprositi Kalavrita (W11) exhibited early maturation compared to other landraces. Consequently, these landraces have the potential to be incorporated into breeding programmes focused on selecting for early maturation traits (Javaid *et al*, 2005). Time of ear emergence was positively correlated with total number of tillers, plant length (from the base of the plant to the tip of the highest awn) and stem length (0.584, 0.674 and 0.710, $P < 0.05$, respectively) (Table 3). A positive correlation between days of heading and number of fertile tillers was also recorded in a study with 64 bread wheat genotypes in West Shewa and found that days to heading were positively correlated with the number of productive tillers (Mecha *et al*, 2017). In the study of Bilgrami *et al* (2020), where bread wheats were used, the total tiller number was also significantly correlated with days from sowing to heading (0.22, $P < 0.05$). Although landraces had more tillers than cultivars, the difference between landraces and specific cultivars was not statistically significant and the number of tillers ranged from 4.18 (Mavragani Lemnos - *T. durum*, unknown status, W13) to 10.38 (Grilos - *T. aestivum*, landrace, W8).

In most cases, the landraces had higher total plant weight than the cultivars. The weight of the landraces ranged from 28.97g/plant (Asprositi Kalavrita - *T. aestivum*, landrace, W11) to 73.01g/plant (Asprositi Kozani - *T. durum*, landrace, W12), while cultivars' weight ranged from 17.55g/plant (Yekora - *T. aestivum*, cultivar, W1) to 25.33g/plant (Elpida - *T. durum*, cultivar, W3). This could be attributed to the higher biomass produced by landraces, a characteristic desirable by farmers who need straw for their animals (Thanopoulos *et al*, 2021). Plant weight was positively and moderately correlated with the total number of tillers and ear length including awns (0.570 and 0.526, $P < 0.05$, respectively) and highly correlated with the number of spikelets, plant length and stem length (0.644, 0.796 and 0.759, $P < 0.01$, respectively). The number of spikelets per spike of the first tiller did not differ significantly between accessions.

Cultivars were significantly shorter than the landraces, which is due to the introduction of Rht (Reduced height) genes in modern wheat varieties. These dwarfing genes have been used in wheat breeding to develop cultivars with short stature (Acquaah, 2012). Similar results were recorded in another study, where Italian landraces were found to have taller culms compared to cultivars (Preiti *et al*, 2022). The plant length for the cultivars was between 83.88cm (Yekora - *T. aestivum*, cultivar, W1) and 103.87cm (Mexicali 81 - *T. durum*, cultivar, W4), while for the landraces the plant length was between 118cm (Kopaida - *T. dicoccum*, landrace, W10) and 159.04cm (Aspratheri - *T. durum* landrace, W14). Mavragani accessions, W7 and W13, of unknown breeding status, were short, with 87.07cm and 93.18cm, respectively. These values did not differ significantly from cultivar plant length (W1 - W4) (Table 2). Low plant height is a desirable trait for commercial varieties, as it is correlated with good standing ability and resistance to lodging (Würschum *et al*, 2017). The weak point of wheat landraces is the lodging, but under climatic change and high temperatures, taller plants could demonstrate higher yield than short ones (Jatayev *et al*, 2020).

Plant growth habit was significantly correlated with weight per plant, total number of tillers, plant length and stem length (base to ear) (0.525, 0.623, 0.538,



Shape parameters	Area of seed	Aspect (ratio between major and minor axis)	Density/Intensity mean (average optical density)	Max diameter (length of longest line joining two points of outline)	Min diameter (length of shortest line joining two points)
Schematic representation					
Shape parameters	Diameter mean (average length of diameters)	Perimeter (length of the seed outline)	Roundness $\left(\frac{perimeter^2}{4\times\pi\times area}\right)$	Size length (feret diameter along major axis)	Size width (feret diameter along minor axis)
Schematic Representation					

Figure 3. A) Indicative image from scanned seeds processed with Image Pro Plus 7 analysis software and B) shape factors definitions and schematic representations.

Table 2. Mean values of quantitative agromorphological traits of 15 Greek wheat accessions. For each trait, different letters following the means indicate significant differences at a 95% level among accessions, using Tukey's honestly significant difference (HSD) method. W1 and W2, *T. aestivum*, cultivar; W3 and W4 *T. durum*, cultivar; W5, W6, W8, W11 and W15, *T.aestivum*, landrace; W7 and W13, *T. durum*, unknown status; W9, *T.polonicum*, landrace; W10, *T. dicoccum* Schrank, landrace; W12 and W14, *T. durum*, landrace.

Accession code and name	**Time of ear emergence (days)**	**Plant weigth (g/plant)**	**Total no. of tillers**	**Ear length excluding awns (cm)**	**Ear length including awns (cm)**	**No. of spikelets in the first tiller**	**Stem length (cm)**	**Plant length (cm)**
W1, Yekora	107.00^{a}	17.55^{a}	7.05^{abcde}	10.24^{cde}	15.53^{abc}	17.56^{abcd}	68.35^{a}	83.88^{a}
W2, Elisavet	115.67^{bcd}	21.11^{a}	5.91^{abc}	11.18^{de}	14.94_{ab}	20.18^{de}	73.62^{a}	88.56^{a}
W3, Elpida	107.00^{a}	25.33^{ab}	6.96^{abcde}	7.33^{ab}	17.35_{abc}	14.90^{a}	88.68^{b}	106.03^{abc}
W4, Mexicali	108.00^{a}	25.02^{ab}	6.49^{abcd}	7.14^{ab}	17.97^{abc}	15.53^{ab}	85.90^{b}	103.87^{abc}
W5, Zoulitsa	138.00^{h}	35.70^{abcd}	8.64^{cde}	10.97^{cde}	18.54^{abc}	15.57^{abc}	122.06^{ef}	140.60^{defg}
W6, Ntopio	121.50^{ef}	47.14^{bcd}	8.28^{bcde}	12.27^{e}	20.16^{abc}	17.88^{abcd}	137.17	158.25^{fg}
W7, Mavragani Skyros	110.00^{ab}	24.42^{ab}	5.18^{ab}	6.53^{a}	18.04^{abc}	15.80^{abc}	75.14^{a}	93.18^{ab}
W8, Grilos	121.33^{def}	39.17^{abcd}	10.38^{e}	11.44^{de}	18.01^{abc}	16.57^{abc}	110.41^{cd}	127.26^{cde}
W9, Leventis	128.67^{g}	58.30^{de}	7.11^{abcde}	10.69^{cde}	23.81^{cd}	19.26^{cde}	130.12^{fg}	153.97^{efg}
W10, Kopaida	135.00^{h}	31.22^{abc}	10.28^{e}	7.57^{ab}	18.20^{abc}	19.00^{bcde}	101.62^{c}	118.00^{bcd}
W11, Asprositi Kalavrita	123.33^{fg}	28.97^{abc}	7.01^{abcde}	12.11^{e}	12.67^{a}	16.31^{abc}	118.17^{de}	125.15^{cde}
W12, Asprositi Kozani	117.50^{cde}	73.01^{e}	9.74^{de}	9.40^{bcd}	22.16^{bcd}	21.81^{e}	118.33^{de}	144.36^{cdef}
W13, Mavragani Lemnos	112.67^{abc}	20.42^{a}	4.18^{a}	6.52^{a}	18.36^{abc}	15.97^{abc}	71.12^{a}	87.07^{a}
W14, Aspratheri	123.67^{fg}	50.82^{cde}	8.73^{cde}	8.41^{abc}	21.77^{bcd}	20.22^{de}	137.27^{g}	159.04^{g}
W15, Kokkinositaro	123.33^{fg}	32.68^{abc}	8.06^{bcde}	11.58^{de}	29.17^{d}	17.65^{abcd}	123.11^{ef}	152.28^{efg}

Table 3. Correlation coefficients between quantitative agromorphological traits of the studied wheat accessions.The P-values are in brackets. Values in bold denote significance level at P < 0.05.

	Time of ear emergence (days)	**Plant weight (g)**	**Total no. of tillers**	**No. of spikelets (first tiller)**	**Plant length (cm)**	**Stem length (base to ear base) (cm)**	**Ear length excluding awns (cm)**
Plant weight (g)	0.404 (0.135)						
Total no. of tillers	**0.584 (0.022)**	**0.570 (0.026)**					
No. of spikelets (first tiller)	0.256 (0.358)	**0.644 (0.010)**	0.392 (0.148)				
Plant length (cm)	**0.674 (0.006)**	**0.796 (0.000)**	**0.608 (0.016)**	0.375 (0.168)			
Stem length (base to ear base) (cm)	**0.710 (0.003)**	**0.759 (0.001)**	**0.611 (0.016)**	0.330 (0.230)	**0.988 (0.000)**		
Ear length excluding awns (cm)	0.409 (0.131)	0.242 (0.385)	0.344 (0.209)	0.207 (0.459)	0.456 (0.088)	0.499 (0.058)	
Ear length including awns (cm)	0.267 (0.337)	**0.526 (0.044)**	0.267 (0.337)	0.325 (0.238)	**0.643 (0.010)**	**0.523 (0.045)**	0.042 (0.883)

0.529, respectively $P < 0.05$) (Supplemental Table 3). A higher number of tillers leads to an increased inclination of the plant.

Cluster analysis based on the quantitative agromorphological traits grouped all cultivars with two Mavragani accessions of unknown status (W7 and W13) in cluster 1, while all landraces were included in a distinct cluster (cluster 2) (Figure 4). The similarity of Mavragani durum wheat accessions from Skyros and Lemnos, with wheat cultivars indicates that these accessions could be old cultivars grown and reproduced for years by farmers, who considered them of local origin (Thomas *et al*, 2013). Overall, the quantitative agromorphological traits based on UPOV descriptors were sufficient to separate wheat accessions by their breeding status, namely landraces or cultivars.

Physical properties of seeds

TKW is of great interest (Javaid *et al*, 2005) as it is an important index for the prediction of flour extraction rate (Posner, 2009). Higher values of TKW indicate a higher percentage of endosperm leading to higher flour yield (Wiersma *et al*, 2001). TKW values of the studied accessions are presented in Figure 5. TKW of W9 (*T. polonicum*) was significantly higher (80.45g) than all other accessions. W10 (*T. dicoccum*) presented the lowest value of TKW (26.92g), consistent with previous research, where the TKW of 38 emmer wheat accessions, collected in several European countries, ranged from 22.9g to 42.6g (Mondini *et al*, 2013). *T. dicoccum* presented significantly lower values of TKW compared to *T. durum*, 32.02–36.12g and 45.17–46.31g, respectively (Shoormij *et al*, 2022). TKW in the rest of the accessions, either cultivars or landraces, ranged from 34.08g (Elisavet - *T. aestivum*, cultivar, W2) to 67.16g (Mavragani Lemnos - *T. durum*, unknown status, W13) and were in accordance with Travlos *et al* (2012) who studied five Greek bread wheat landraces (30–50.7g). Specifically, in the present study, W5 (Zoulitsa) had an average TKW of 43.66g, while in Travlos *et al* (2012), TKW in Zoulitsa ranged from 36g to 44g, indicating that climatic conditions can affect TKW but also the origin of the genetic material (Papadakis, 1929). Moreover, it is worth noting that in the present study, *T. aestivum* in most cases presented significantly lower values of TKW compared to *T. durum* accessions. Landraces had similar or higher values of TKW compared to cultivars of the same species, indicating that landraces could give high flour yield. However, results from previous research are contrasting as no significant difference between landraces' and cultivars' TKW was reported for Turkish (Cetiner *et al*, 2020) and Sicilian wheats (Ruisi *et al*, 2021). On the contrary, Spanish bread wheat landraces showed higher TKW compared to a reference set (López-Fernández *et al*, 2021). In another research, similar TKW values for *T. aestivum* landraces and cultivars were detected, while *T. durum* cultivars, compared to landraces, had significantly higher TKW values (Frankin *et al*, 2021).

Due to their high heritability grain perimeter, area and volume are important indirect indices for grain yield improvement (Abdipour *et al*, 2016). Seed volume ranged from 48ml to 113ml (Elisavet - *T. aestivum*, cultivar, W2 and Leventis - *T. polonicum*, landrace, W9, respectively) (Table 4). It is worth noting that W9 (*T. polonicum*) had significantly the highest values of TKW, volume and seed firmness (SF). Although Papadakis (1929) mentioned that *T. polonicum* did not provide any agronomic value, our findings show that this species has bread-making value.

There are many methods to determine the required force to evaluate grain mechanical properties. In the present study, a rapid measurement of the mechanical properties of a single grain was used, which provided valuable results (Shewry and Hey, 2016). SF ranged from 62.96N to 194.85N (W10, Kopaida - *T. dicoccum*, landrace and W9, *T. polonicum*, landrace, respectively) and was correlated positively with seed weight and volume (0.840 and 0.791, $P < 0.05$, respectively) (Table 5). High values of weight and hardness indicate healthy seeds. Seed texture determines, to a great extent, milling properties and end uses. Softer seeds require more energy to mill and produce finer flours suitable for biscuit production, while flour from harder seeds is more appropriate for bread making (Liu, 2008). As W10 (Kopaida - *T. dicoccum*, landrace) had the lowest SF values, it could be suitable for cake and biscuit production. A recent research concluded that in general old cultivars of *T. aestivum* were softer than modern cultivars (Cetiner *et al*, 2020). In the present research, this statement was not fully confirmed, as there was no clear trend. However, it should be mentioned that fewer old and modern varieties were used in the present work compared to the earlier study.

Bulk density could be used as a criterion for yield improvement, as it was correlated positively with grain yield (Karimizadeh *et al*, 2012). In the present study, density did not differ significantly between accessions except for W10 (Kopaida - *T. dicoccum*, landrace), having the lowest value (0.55g/ml).

Seed flowability for all accessions was very good as indicated by Carr indices (CI) with values lower than 15% (Carr, 1965). W11 (Asprositi Kalavrita - T. *aestivum*, landrace) had the lowest CI value, 1.79%, and W13 (Mavragani Lemnos - *T. durum*, unknown status) had the highest value, 6.53%.

In general, colour parameters did not differ significantly among accessions and were not correlated with the other physical properties (Table 5). The exceptions were the green/red hue (a*), which correlated negatively with volume (-0.516, $P < 0.05$), indicating that seeds were redder as their volume increased, and lightness (L*), blue/yellow hue (b*) and white index (WI), which were correlated positively with density (0.649, 0.719 and 0.565, $P < 0.05$, respectively). Increased density led to increased lightness, yellowness and whiteness of the seeds. W10 (Kopaida - *T. dicoccum*, landrace) was the only accession with a significantly lower value of b*

Table 4. Mean values of seed physical properties of 15 Greek wheat accessions. For each trait, different letters following the means indicate significant differences at 95% level, using Tukey's honestly significant difference (HSD) method. W1 and W2, *T. aestivum*, cultivar; W3 and W4, *T. durum*, cultivar; W5, W6, W8, W11 and W15, *T. aestivum*, landrace; W7 and W13, *T. durum*, unknown status; W9, *T. polonicum*, landrace; W10, *T. dicoccum*, landrace; W12 and W14, *T. durum*, landrace. CI: Carr Index, L* indicates lightness, a* indicates hue on a green (–) to red (+) axis, b* indicates hue on a blue (–) to yellow (+) axis and WI: White Index.

Accession code and name	**Seed firmness (N)**	**Volume (ml)**	**Tap Volume (ml)**	**Bulk density (g/ml)**	**Tap density (g/ml)**	**CI (%)**	**L***	**a***	**b***	**WI**
W1, Yekora	88.82^{de}	61.50^{b}	60.50^{b}	0.70^{bc}	0.72^{bcde}	1.63^{a}	56.06^{de}	7.51^{ab}	24.47^{ab}	49.13^{cdef}
W2, Elisavet	62.96^{a}	48.00^{a}	47.00^{a}	0.71^{bcd}	0.73^{bcde}	2.08^{a}	51.79^{bc}	10.21^{d}	23.78^{b}	45.26^{abc}
W3, Elpida	101.00^{ef}	88.50^{h}	84.50^{g}	0.70^{bc}	0.74^{def}	4.52^{a}	58.00^{e}	6.80^{a}	23.98^{ab}	51.14^{f}
W4, Mexicali	104.94^{fg}	81.50^{fg}	79.50^{f}	0.71^{bc}	0.72^{bcde}	2.45^{a}	53.61^{bcd}	7.61^{ab}	24.47^{b}	46.97^{bcde}
W5, Zoulitsa	70.30^{ab}	65.00^{bc}	63.50^{bc}	0.67^{b}	0.69^{b}	2.31^{a}	50.97^{b}	9.05^{bcd}	21.34^{ab}	45.77^{abcd}
W6, Ntopio	86.39^{cd}	68.50^{cd}	65.50^{c}	0.71^{bcd}	0.74^{ef}	4.38^{a}	50.03^{ab}	8.43^{bc}	21.37^{ab}	44.99^{ab}
W7, Mavragani Skyros	105.43^{fg}	84.00^{gh}	80.50^{f}	0.67^{b}	0.70^{bc}	4.17^{a}	53.10^{bcd}	7.55^{ab}	24.29^{b}	46.64^{bcde}
W8, Grilos	74.64^{abc}	74.00^{de}	72.00^{d}	0.68^{b}	0.70^{bcd}	2.70^{a}	51.76^{bc}	8.59^{bc}	22.86^{ab}	45.91^{abcd}
W9, Leventis	194.85^{i}	113.00^{j}	110.00^{i}	0.70^{bc}	0.73^{cde}	4.35^{a}	49.42^{ab}	7.51^{ab}	21.35^{ab}	44.56^{ab}
W10, Kopaida	59.29^{a}	49.00^{a}	47.00^{a}	0.55^{a}	0.57^{a}	4.08^{a}	46.14^{a}	8.07^{abc}	18.50^{a}	42.47^{a}
W11, Asprositi Kalavrita	83.06^{bcd}	63.50^{bc}	63.00^{bc}	0.70^{bc}	0.71^{bcde}	0.79^{a}	57.12^{de}	7.77^{ab}	25.49^{b}	49.49^{def}
W12, Asprositi Kozani	139.75^{h}	80.00^{efg}	78.33^{ef}	0.77^{d}	0.78^{g}	2.08^{a}	59.27^{e}	7.56^{ab}	25.40^{b}	51.40^{ef}
W13, Mavragani Lemnos	93.09^{def}	99.50^{i}	93.00^{h}	0.67^{b}	0.72^{bcde}	6.53^{a}	55.68^{cde}	7.51^{ab}	24.97^{b}	48.55^{bcdef}
W14, Aspratheri	117.94^{g}	75.00^{e}	72.00^{d}	0.74^{cd}	0.77^{fg}	4.00^{a}	53.93^{bcde}	9.97^{cd}	26.42^{b}	45.96^{abcde}
W15, Kokkinositaro	83.07^{bcd}	77.00^{ef}	75.00^{de}	0.69^{b}	0.70^{bcd}	2.60^{a}	52.56^{bcd}	7.58^{ab}	22.31^{ab}	47.02^{abcdef}

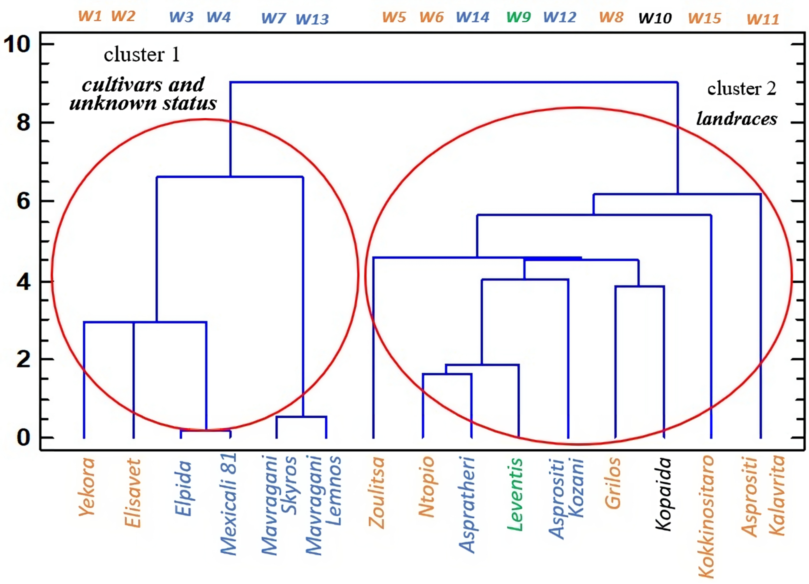


Figure 4. Cluster analysis of 15 Greek wheat accessions based on their agromorphological traits. Different colours indicate different *Triticum* species: *aestivum* (orange); *durum* (blue); *polonicum* (green); *dicoccum* (black).

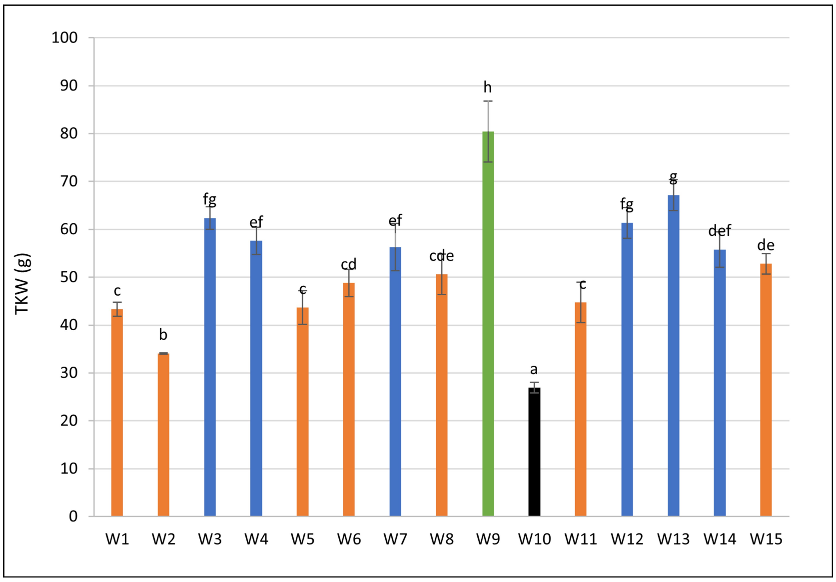


Figure 5. Thousand kernel weight (TKW) of wheat accessions. Different colours indicate different *Triticum* species: *aestivum* (orange); *durum* (blue); *polonicum* (green); *dicoccum* (black). Error bars represent standard deviation. Different small letters state significantly different values of TKW ($P < 0.05$). W1, Yekora - *T. aestivum*, cultivar; W2, Elisavet - *T. aestivum*, cultivar; W3, Elpida - *T. durum*, cultivar; W4, Mexicali 81- *T. durum*, cultivar; W5, Zoulitsa - *T. aestivum*, landrace; W6, Ntopio - *T. aestivum*, landrace; W7, Mavragani Skyros - *T. durum*, unknown status; W8, Grilos - *T. aestivum*, landrace; W9, Leventis - *T. polonicum*, landrace; W10, Kopaida - *T. dicoccum* Schrank, landrace; W11, Asprositi Kalavrita - *T. aestivum*, landrace; W12, Asprositi Kozani - *T. durum*, landrace; W13, Mavragani Lemnos - *T. durum*, unknown status; W14, Aspratheri - *T. durum* landrace and W15, Kokkinositaro - *T.aestivum*, landrace.

among all other accessions. Colour determines to a great extent the overall consumer acceptance of food (Mandala and Protonotariou, 2021). Wheat kernel colour is related to different milling and baking attributes of the final products (Dowell, 1998). Bright yellow or amber colour of durum wheat results in high-quality pasta products (Cole *et al*, 1991), thus sample W10 (Kopaida - *T. dicoccum*, landrace) probably is the least suitable for pasta making among the studied accessions. Probably, W9 (Leventis - *T. polonicum*, landrace) and W13 (Mavragani Lemnos - *T. durum*, unknown status) could be the most suitable for pasta making considering their high values of b*, TKW and seed volume. However, further investigation is needed. Mechanical properties and sensory evaluation of produced pasta could verify our claim.

The determination of kernel morphology is of great importance for bread making, as kernel shape and uniformity may influence the milling quality (Campbell *et al*, 1999). Image analysis could play a key role in distinguishing different wheat varieties by defining the

Table 5. Correlation coefficients between seed physical properties of all the studied wheat accessions. The P-values are in brackets. a*, hue on a green (−) to red (+) axis; b*, hue on a blue (−) to yellow (+) axis; CI, Carr Index; SF, Seed Firmness; L*, lightness; TKW, Thousand Kernel Weight; and WI, White Index. Values in bold denote significance level at $P < 0.05$.

	SF	**Volume**	**Volume tap**	**Bulk density**	**Tap density**	**TKW**	**CI**	**L***	**a***	**b***
Volume	**0.791** **(0.000)**									
Volume tap	**0.806** **(0.000)**	**0.998** **(0.000)**								
Bulk density	0.462 (0.083)	0.283 (0.308)	0.311 (0.259)							
Tap density	0.500 (0.058)	0.401 (0.138)	0.418 (0.121)	**0.969** **(0.000)**						
TKW	**0.840** **(0.000)**	**0.978** **(0.000)**	**0.982** **(0.000)**	0.473 (0.075)	**0.574** **(0.025)**					
CI	0.223 (0.425)	**0.546** **(0.035)**	0.496 (0.060)	-0.271 (0.328)	-0.053 (0.851)	0.446 (0.096)				
L*	0.170 (0.545)	0.214 (0.444)	0.233 (0.403)	**0.649** **(0.009)**	0.651 **(0.009)**	0.337 (0.220)	-0.215 (0.442)			
a*	-0.317 (0.249)	**-0.516** **(0.049)**	**-0.524** **(0.045)**	0.100 (0.724)	0.076 (0.787)	-0.456 (0.088)	-0.174 (0.535)	-0.352 (0.198)		
b*	0.197 (0.481)	0.192 (0.492)	0.207 (0.460)	**0.719** **(0.003)**	**0.731** **(0.002)**	0.326 (0.236)	-0.166 (0.554)	**0.840** **(0.000)**	0.014 (0.962)	
WI	0.160 (0.569)	0.240 (0.389)	0.260 (0.349)	**0.565** **(0.028)**	**0.565** **(0.028)**	0.343 (0.210)	-0.202 (0.472)	**0.974** **(0.000)**	-0.501 (0.057)	**0.701** **(0.004)**

product traceability for wheat landraces (Khoshroo *et al*, 2014; Grillo *et al*, 2017). In the present research, mean seed diameter ranged between 4.17mm (W2, Elisavet - *T. aestivum*, cultivar) and 5.66mm (W9, Leventis - *T. polonicum*, landrace). Particle size distribution curves differed between accessions, indicating that some landraces had a more uneven seed population than cultivars (Figure 6). Curve of W12 (Asprositi Kozani - *T. durum*, landrace) presented an almost bimodal distribution, indicating the presence of two distinguished populations of seeds, with different sizes, which could be explained by the fact that sometimes landraces consisted of two or more populations with different morphological characteristics (Papadakis, 1929). This could be linked with seed heteromorphism, characteristic of wild populations. A similar curve was presented by W13 (Mavragani Lemnos - *T. durum*, unknown status), where the bimodal distribution was less evident, but also presented a shoulder on the left (Figure 6). In contrast, W3 (Elpida - *T. durum*, cultivar) presented the narrowest unimodal distribution curve, with the peak around 5.5mm. All the other accessions presented frequency curves similar to W2 (Elisavet - *T. aestivum*, cultivar) or W9 (Leventis - *T. polonicum*, landrace). The main characteristic of these curves is that distribution is broad, due to the uneven seed populations. W10 (Kopaida - *T. dicoccum*, landrace) presented a similar curve to W2 (Elisavet - *T. aestivum*, cultivar), despite the fact that they belonged to different species.

Bread wheat cultivars (W1, Yekora - *T. aestivum* and W2, Elisavet - *T. aestivum*) had significantly more roundness and shorter seeds compared to pasta wheat cultivars (W3, Elpida - *T. durum*, and W4, Mexicali 81 - *T. durum*), as referred to in the literature (Campbell *et al*, 1999). Perimeter was significantly lower for W2, Elisavet - *T. aestivum*, cultivar (15.71mm) and significantly higher for W9, Leventis - *T. polonicum*, landrace (23.62mm), compared to all other accessions (Supplemental Table 4). Similar results were observed for seed length, with W9 having the longest seeds (9.3mm) and W2 the shortest (5.96mm). *T. polonicum* (represented by W9) had characteristically long seed (Percival, 1921). Regarding the width of the seeds, the differences were not significant in all the cases and the values ranged from 2.72mm to 3.94mm (W10 and W13, respectively). Grain width could be used as a predictive index for mean grain weight determination (Haghshenas *et al*, 2022). However, in the present study, the most valuable parameter was the mean diameter.

The average grain area was 18.54mm^2 (from 13.85mm^2 to 26.08mm^2, for W2, Elisavet - *T. aestivum*, cultivar and W9, Leventis - *T. polonicum*, respectively), in accordance with other studies (Gegas *et al*, 2010; Okamoto *et al*, 2013; Abdipour *et al*, 2016). Only Kondić *et al* (2020) recorded significantly higher values of seed area, 53.5mm^{2}and 40.17mm^2 in two experimental years in Bosnia and Herzegovina for seeds of *T. aestivum*.

Correlations among all parameters

In the present study, TKW was significantly positively correlated with area of seed, mean diameter, SF, perimeter, length, width and volume of the seeds (0.871, 0.887, 0840, 0.786, 0.677, 0.829, and 0.975, $P <$

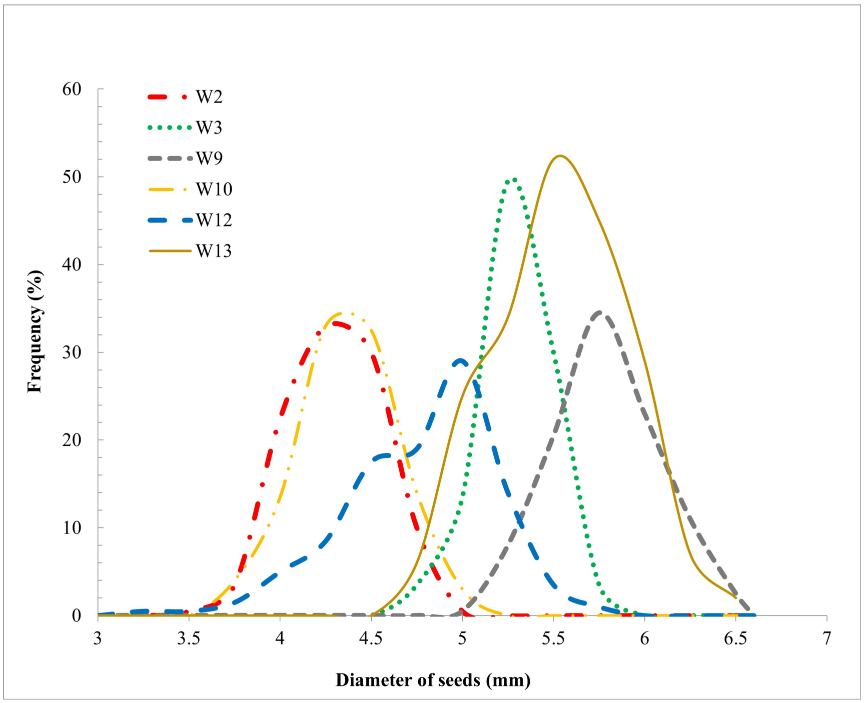


Figure 6. Indicative particle size distribution curves by frequency based on mean diameters of different wheat seeds (W2 -*T. aestivum*, Elisavet; W3 - *T. durum*, Elpida; W9 - *T. polonicum*; W10 - *T. dicoccum*; W12 - *T. durum*, Asprositi Kozani and W13 - *T. durum*, Mavragani Lemnos).

0.05, respectively, Supplemental Table 5). There is previous research which also associated kernel weight with its width and length (Campbell *et al*, 1999; Kondić *et al*, 2020), but some of them could not associate these parameters when they studied only bread wheats (Schuler *et al*, 1995). Moreover, Javaid *et al* (2005) found that TKW was positively correlated with plant height in Pakistani bread wheat landraces, and Moghaddam *et al* (1997) found a significant correlation between TKW and number of tillers in Irani bread wheat landraces, which was not observed in the present study, suggesting possibly different patterns.

There are several studies evaluating wheat kernels by compression tests (Ponce-García *et al*, 2016). However, there is a research gap on the correlation between wheat seed mechanical properties with plant-agromorphological and seed-morphological traits. In the present study, SF was significantly correlated with plant weight, and seed area, mean diameter, perimeter, length, width, volume and TKW (0.626, 0.711, 0.709, 0.688, 0.615, 0.570, and 0.840, $P < 0.05$, respectively). Thus, SF could be used as an additional tool to predict all the above parameters. Fitting a simple linear regression model to describe the relationship between SF and TKW explained 70.58% of firmness variability ($P < 0.05$) (Supplemental Figure 2).

Grouping of accessions

Cluster analysis, using all quantitative data, allowed us to group the wheat accessions in accordance with their species and breeding status (Figure 7). W9 (Leventis - *T. polonicum*, landrace) and W10 (Kopaida - *T. dicoccum*, landrace) created a distinct clade each, as anticipated, considering that they were different species. The control cultivars of *T. durum* species (W3 and W4) created a distinct cluster, along with the *T. durum* accessions of unknown breeding status W7 and W13 (Mavragani - *T. durum*, unknown status from Skyros and Lemnos, respectively). Interestingly, except for W11 (Asprositi Kalavrita) all landraces, whether *T. aestivum* or *T. durum*, formed a separate cluster, related to the *T. aestivum* cultivars.

It is worth noting that with the addition of physical properties in the cluster analysis, the grouping of *Triticum* spp. accessions changed compared to Figure 4, where only observations according to agromorphological traits were used. The cluster of landraces was closer to *T. aestivum* cultivars, indicating that landraces – regardless of the species they belonged to – could be more suitable for bread making, considering their physical properties, but also the main preference and use (mainly for bread) by farmers in Greece (Douma *et al*, 2016).

Conclusion

The agromorphological traits recorded in this study proved to be sufficient to highlight the differences between wheat species, landraces and cultivars, grouping them into two different clusters. TKW did not differ significantly between landraces and cultivars, indicating that landraces could also give a high flour yield. Increased values of TKW and seed volume indicated higher values of seed firmness. Cluster analysis of all data separated the wheat accessions into groups accord-

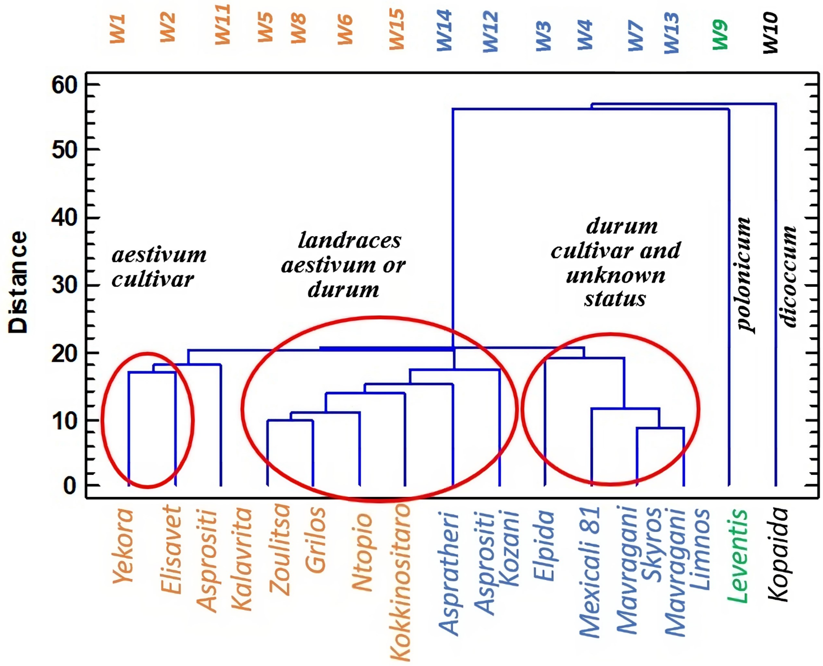


Figure 7. Cluster analysis of 15 Greek accessions based on all quantitative variables (nearest-neighbour method, square Euclidean distance). Different colours indicate different *Triticum* species: *aestivum* (orange); *durum* (blue); *polonicum* (green); *dicoccum* (black).

ing to their attributes, indicating that Greek wheat landraces, either *T. durum* or *T. aestivum* could be probably used for bread making. Regarding the W10 (*T. dicoccum*, landrace) and W9 (*T. polonicum*, landrace) wheat accessions, their incorporation in both bread and pasta making would be of great interest, as they presented distinct characteristics. The evaluation of both agromorphological and physical properties of wheat seeds provided valuable information that could be used to distinguish wheat varieties for pasta and bread making based on quality characteristics.

Acknowledgements

The authors are grateful to Dr Michiel Van Slageren (Royal Botanical Gardens Kew, UK) who helped us with wheat taxonomy, to the student T. Sarros for helping with sowing, to the PhD candidates P. Chaloulos and P. Loukopoulos and Mr V. Protonotarios for helping with harvesting. Special acknowledgements go to the Institute of Plant Breeding and Genetic Resources and all the local producers for providing seeds of wheat cultivars and landraces, respectively. We thank the Hellenic National Meteorological Service for their prompt response to our request for meteorological data. We would also like to thank cordially the anonymous reviewers for their valuable comments and suggestions that improved the manuscript.

This research was co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme ‘Human Resources Development, Education and Lifelong Learning’ in the context of the project ‘Reinforcement of Postdoctoral Researchers – 2nd Cycle’ (MIS-5033021), implemented by the State Scholarships Foundation (*ΙΚΥ*).

Author contributions

SVP contributed to study conception and design, established the experimental trials, made field and laboratory measurements, performed data analysis and interpretation, and wrote the initial draft. RT contributed to study conception and design, conducted experiments on the agronomic research trial, provided the know-how in this field, worked on the draft and contributed to the writing of the final manuscript. AK contributed to statistical analysis, visualization and interpretation of the results. PJB contributed to the study conception and design, interpretation of the results, writing of the final manuscript, resource provision and supervision of the research regarding agronomic traits. IM contributed to study conception and design, interpretation of the results, final manuscript revision, resource provision and supervision of the research on the physical properties of seeds. All authors discussed the results and commented on the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Supplemental data

Supplemental Figure 1: The frequency percentage for the four qualitative morphological traits observed in the 15 wheat accessions studied.

Supplemental Figure 2: Simple linear regression model to describe the relationship between seed firmness (SF) and thousand kernel weight (TKW) of wheat seeds.

Supplemental Table 1: Summary table of Friedman test results and medians of qualitative variables; plant growth habit; colour of ear; ear shape in profile and awn colour, for 15 *Triticum* spp accessions.

Supplemental Table 2: Analysis of Variance for quantitative agromorphological traits (ANOVA tables).

Supplemental Table 3: Spearman rank correlations between each pair of variables (qualitative: plant growth habit, colour of ear, ear shape in profile and awn colour versus all the variables).

Supplemental Table 4: Mean values of Image Analysis-Shape factor measurements of 15 Greek wheat accessions.

Supplemental Table 5: Pearson correlation test between quantitative parameters.

References

Abdipour, M., Ebrahimi, M., Izadi-Darbandi, A., Mastrangelo, A. M., Najafian, G., Arshad, Y., and Mirniyam, G. (2016). Association between Grain Size and Shape and Quality Traits, and Path Analysis of Thousand Grain Weight in Iranian Bread Wheat Landraces from Different Geographic Regions. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 44(1), 228–236. doi: https://doi.org/10.15835/nbha44110256

Acquaah, G. (2012). Principles of Plant Genetics and Breeding (Hoboken, NJ, Wiley) . doi: https://doi.org/10.1002/9781118313718

Ayadi, S., Jallouli, S., Landi, S., Capasso, G., Chamekh, Z., Cardi, M., Paradisone, V., Lentini, M., Karmous, C., Trifa, Y., and Esposito, S. (2020). Nitrogen assimilation under different nitrate nutrition in Tunisian durum wheat landraces and improved genotypes. *Plant Biosystems* 154(6), 924–934. doi: https://doi.org/10.1080/11263504.2020.1722274

Bieńkowska, T., Suchowilska, E., and Wiwart, M. (2020). *Triticum polonicum* L. as promising source material for breeding new wheat cultivars. *Journal of Elementology* 25(1), 237–248. doi: https://doi.org/10.5601/JELEM.2019.24.2.1873

Bilgrami, S. S., Ramandi, H. D., Shariati, V., Razavi, K., Tavakol, E., Fakheri, B. A., Nezhad, N. M., and Ghaderian, M. (2020). Detection of genomic regions associated with tiller number in Iranian bread wheat under different water regimes using genome-wide association study. *Scientific Reports* 10(1). doi: https://doi.org/10.1038/s41598-020-69442-9

Boukid, F., Vittadini, E., Prandi, B., Mattarozzi, M., Marchini, M., Sforza, S., Sayar, R., Seo, Y. W., Yacoubi, I., and Mejri, M. (2018). Insights into a century of breeding of durum wheat in Tunisia: The properties of flours and starches isolated from landraces, old and modern genotypes. *LWT* 97, 743–751. doi: https://doi.org/10.1016/J.LWT.2018.08.014

Brush, S. B. (1995). In Situ Conservation of Landraces in Centers of Crop Diversity. *Crop Science* 35(2), 346–346. doi: https://doi.org/10.2135/cropsci1995.0011183X003500020009x

Butkovskaya, L. K. and Kozulina, N. S. (2021). Sowing time and seeding rate in the new wheat varieties cultivation for seeds. *IOP Conference Series: Earth and Environmental Science* 839(4), 042012. doi: https://doi.org/10.1088/1755-1315/839/4/042012

Campbell, K. G., Bergman, C. J., Gualberto, D. G., Anderson, J. A., Giroux, M. J., Hareland, G., Fulcher, R. G., Sorrells, M. E., and Finney, P. L. (1999). Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. *Crop Science* 39(4), 1184–1195. doi: https://doi.org/10.2135/cropsci1999.0011183X003900040039x

Carr, R. L. (1965). Evaluating flow properties of solids. *Chemical Engineering* 18, 163–168. doi: https://ci.nii.ac.jp/naid/10016048504

Carvalho, M. A. A. P., De, Ragonezi, C., Macedo, F. L., De, Antunes, G. N., Freitas, G., and Nóbrega, H. (2019). Agrodiversidade e a Agricultura na Madeira. *Revista de Ciências Agrárias* 42(3), 575–605. doi: https://doi.org/10.19084/rca.16658

Casañas, F., Simó, J., Casals, J., and Prohens, J. (2017). Toward an Evolved Concept of Landrace. *Frontiers in Plant Science* 145-145. doi: https://doi.org/10.3389/fpls.2017.00145

Cetiner, B., Tömösközi, S., Török, K., Salantur, A., and Koksel, H. (2020). Comparison of the arabinoxylan composition and physical properties of old and modern bread wheat (*Triticum aestivum* L.) and landraces genotypes. *Cereal Chemistry* 97(2), 505–514. doi: https://doi.org/10.1002/CCHE.10265

Chaloulos, P., Bazanis, A. E., Georgiadou, M., Protonotariou, S., and Mandala, I. (2021). Effect of drying and grinding or micro-grinding process on physical and rheological properties of whole cladode (*Opuntia ficus-indica*) flour. *LWT* 151, 112171. doi: https://doi.org/10.1016/j.lwt.2021.112171

Chaugule, A. A. and Mali, S. N. (2016). Identification of paddy varieties based on novel seed angle features. *Computers and Electronics in Agriculture* 123, 415–422. doi: https://doi.org/10.1016/J.COMPAG.2016.03.012

Cole, M. E., Johnson, D. E., Cole, R. W., and Stone, M. B. (1991). Color of Pregelatinized Pasta as Influenced by Wheat Type and Selected Additives. *Journal of Food Science* 56(2), 488–493. doi: https://doi.org/10.1111/j.1365-2621.1991.tb05310.x

Cseh, A., Poczai, P., Kiss, T., Balla, K., Berki, Z., Horváth, A., Kuti, C., and Karsai, I. (2021). Exploring the legacy of Central European historical winter wheat landraces. *Scientific Reports* 11(1), 1–14. doi: https://doi.org/10.1038/s41598-021-03261-4

Dholakia, B. B., Ammiraju, J. S. S., Singh, H., Lagu, M. D., Röder, M. S., Rao, V. S., Dhaliwal, H. S., Ranjekar, P. K., Gupta, V. S., and Weber, W. E. (2008). Molecular marker analysis of kernel size and shape in bread wheat. *Plant Breeding* 122(5), 392–395. doi: https://doi.org/10.1046/j.1439-0523.2003.00896.x

Douma, C., Koutis, K., Thanopoulos, R., Tsigou, R., Galanidis, A., and Bebeli, P. J. (2016). Diversity of agricultural plants on Lesvos Island (Northeast Aegean, Greece) with emphasis on fruit trees. *Scientia Horticulturae* 210, 65–84. doi: https://doi.org/10.1016/J.SCIENTA.2016.07.009

Dowell, F. E. (1998). Automated Color Classification of Single Wheat Kernels Using Visible and Near-Infrared Reflectance. *Cereal Chemistry Journal* 75(1), 142–144. doi: https://doi.org/10.1094/CCHEM.1998.75.1.142

Fitzpatrick, J. J., Barringer, S. A., and Iqbal, T. (2004). Flow property measurement of food powders and sensitivity of Jenike's hopper design methodology to the measured values. *Journal of Food Engineering* 61(3), 399–405. doi: https://doi.org/10.1016/S0260-8774(03)00147-X

Frankin, S., Roychowdhury, R., Nashef, K., Abbo, S., Bonfil, D. J., and Ben-David, R. (2021). Infield comparative study of landraces vs. Modern wheat genotypes under a mediterranean climate. *Plants* (12), 10–10. doi: https://doi.org/10.3390/plants10122612

Gegas, V. C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J. H., and Snape, J. W. (2010). A genetic framework for grain size and shape variation in wheat. *Plant Cell* 22(4), 1046–1056. doi: https://doi.org/10.1105/tpc.110.074153

Gelelcha, F., Kumsa, F., and Kuma, T. (2023). On-farm genetic diversity of wheat (*Triticum aestivum* spp) Digalu Tijo District, Arsi zone, Ethiopia. *Heliyon* 9-9. doi: https://doi.org/10.1016/J.HELIYON.2022.E12775

Gharib, M. A. A. H., Qabil, N., Salem, A. H., Ali, M. M. A., Awaad, H. A., and Mansour, E. (2021). Characterization of wheat landraces and commercial cultivars based on morpho-phenological and agronomic traits. *Cereal Research Communications* 49(1), 149–159. doi: https://doi.org/10.1007/s42976-020-00077-2

Gómez, M., Aparicio, N., Ruiz-París, E., Oliete, B., and Caballero, P. A. (2009). Evolución de la calidad panadera de variedades de trigo cultivadas en España. *Spanish Journal of Agricultural Research* 7(3), 585–595. doi: https://doi.org/10.5424/sjar/2009073-443

Grillo, O., Blangiforti, S., and Venora, G. (2017). Wheat landraces identification through glumes image analysis. *Computers and Electronics in Agriculture* 141, 223–231. doi: https://doi.org/10.1016/j.compag.2017.07.024

Haghshenas, A., Emam, Y., and Jafarizadeh, S. (2022). Wheat grain width: a clue for re-exploring visual indicators of grain weight. *Plant Methods* 18(1), 1–16. doi: https://doi.org/10.1186/S13007-022-00891-1

Hammer, K., Filatenko, A. A., and Pistrick, K. (2011). Taxonomic remarks on *Triticum* L. and ×*Triticosecale* Wittm. *Genetic Resources and Crop Evolution* 58(1), 3–10. doi: https://doi.org/10.1007/s10722-010-9590-4

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S., and David, J. (2007). Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication. *Molecular Biology and Evolution* 24(7), 1506–1517. doi: https://doi.org/10.1093/MOLBEV/MSM077

HNMS (2023). Hellenic National Meteorological Service. url: http://www.emy.gr/emy/en/climatology/climatology.

Igrejas, G., Guedes-Pinto, H., Carnide, V., Clement, J., and Branlard, G. (2002). Genetical, biochemical and technological parameters associated with biscuit quality. II. Prediction using storage proteins and quality characteristics in a soft wheat population. *Journal of Cereal Science* 36(2), 187–197. doi: https://doi.org/10.1006/jcrs.2002.0452

Jatayev, S., Sukhikh, I., Vavilova, V., Smolenskaya, S. E., Goncharov, N. P., Kurishbayev, A., Zotova, L., Absattarova, A., Serikbay, D., Hu, Y. G., Borisjuk, N., Gupta, N. K., Jacobs, B., De Groot, S., Koekemoer, F., Alharthi, B., Lethola, K., Cu, D. T., Schramm, C., and Langridge, P. (2020). Green revolution 'stumbles' in a dry environment: Dwarf wheat with Rht genes fails to produce higher grain yield than taller plants under drought. *Plant Cell and Environment* 43(10), 2355–2364. doi: https://doi.org/10.1111/pce.13819

Javaid, M. M., Rabbani, M. A., and Rastina, A. (2005). Phenotypic diversity and trait association in bread wheat (*Triticum aestivum* L.) landraces from Baluchistan. *Pakistan Journal of Botany* 37(4), 949–957. url: https://www.researchgate.net/publication/287626437_Phenotypic_diversity_and_trait_association_in_bread_wheat_Triticum_aestivum_L_landraces_from_Baluchistan_Pakistan.

Karamanos, A. J., Economou, G., Sotirakoglou, K., Lyra, D., and Papastavrou, A. (2017). Assessing Greek Bread Wheat Landraces for their Drought Resistance Strategies. *Crop Science* 57(1), 416–426. doi: https://doi.org/10.2135/CROPSCI2016.06.0524

Karimizadeh, R., Sharifi, P., and Mohammadi, M. (2012). Correlation and path coefficient analysis of grain yield and yield components in durum wheat under two irrigated and rainfed condition. *International Journal of Agriculture: Research and Review* 2(3), 277–283.

Khoshroo, A., Arefi, A., Masoumiasl, A., and Jowkar, G. H. (2014). Classification of Wheat Cultivars Using Image Processing and Artificial Neural Networks. *Agricultural Communications* 2(1), 17–22.

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

Kondić, D., Žuljević, S. O., Hajder, and Selimbegović, E. (2020). Evaluation of grain characteristics of domestic wheat (*Triticum aestivum* L.) obsolete cultivars and landraces. *Italian Journal of Agronomy* 15(1), 3–9. doi: https://doi.org/10.4081/ija.2020.1345

Kyzeridis, N., Biesantz, A., and Limberg, P. (1995). Vergleichende Feldversuche mit Land- und Zuchtsorten von Durumweizen (*Triticum turgidum* var. durum) unter verschiedenen ökologischen Bedingungen im Mittelmeerraum. *Journal of Agronomy and Crop Science* 174(2), 133–144. doi: https://doi.org/10.1111/J.1439-037X.1995.TB00204.X

Liu, K. (2008). Measurement of Wheat Hardness by Seed Scarifier and Barley Pearlerand Comparison with Single-Kernel Characterization System. *Cereal Chemistry* 85(2), 165–173. doi: https://doi.org/10.1094/CCHEM-85-2-0165

López-Fernández, M., Pascual, L., Faci, I., Fernández, M., Ruiz, M., Benavente, E., and Giraldo, P. (2021). Exploring the End-Use Quality Potential of a Collection of Spanish Bread Wheat Landraces. *Plants* 10(4). doi: https://doi.org/10.3390/plants10040620

Mandala, I. G. and Protonotariou, S. V. (2021). Physical properties of food materials. *Engineering Principles of Unit Operations in Food Processing* 45-64. doi: https://doi.org/10.1016/b978-0-12-818473-8.00015-3

Maxted, N., Kell, S., and Brehm, J. M. (2011). Options to Promote Food Security: on-farm Management and in situ Conservation of Plant Resources for Food and Agriculture. url: https://www.fao.org/3/am489e/am489e.pdf.

Mecha, B., Alamerew, S., Assefa, A., Assefa, E., and Dutamo, D. (2017). Correlation and Path Coefficient Studies of Yield and Yield Associated Traits in Bread Wheat (*Triticum aestivum* L.) Genotypes. *Advances in Plants & Agriculture Research* 6(5), 128–136. doi: https://doi.org/10.15406/apar.2017.06.00226

Moghaddam, M., Ehdaie, B., and Waines, J. G. (1997). Genetic variation and interrelationships of agronomic characters in landraces of bread wheat from southeastern Iran. *Euphytica* 95(3), 361–369. doi: https://doi.org/10.1023/A:1003045616631

Mondini, L., Grausgruber, H., and Pagnotta, M. A. (2013). Evaluation of European emmer wheat germplasm for agro-morphological, grain quality traits and molecular traits. *Genetic Resources and Crop Evolution* (1), 69–87. doi: https://doi.org/10.1007/S10722-013-0016-Y

Moragues, M., Moral, L. F. G., Del, Moralejo, M., and Royo, C. (2006). Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin I: Yield components. *Field Crops Research* 95(2-3), 194–205. doi: https://doi.org/10.1016/j.fcr.2005.02.009

Nazco, R., Peña, R. J., Ammar, K., Villegas, D., Crossa, J., and Royo, C. (2014). Durum wheat (*Triticum durum* Desf.) Mediterranean landraces as sources of variability for allelic combinations at Glu-1/Glu-3 loci affecting gluten strength and pasta cooking quality. *Genetic Resources and Crop Evolution* (6), 1219–1236. doi: https://doi.org/10.1007/s10722-014-0104-7

Newton, A. C., Akar, T., Baresel, J. P., Bebeli, P. J., Bettencourt, E., Bladenopoulos, K. V., Czembor, J. H., Fasoula, D. A., Katsiotis, A., Koutis, K., Koutsika-Sotiriou, M., Kovacs, G., Larsson, H., De Carvalho, M. A. A. P., Rubiales, D., Russell, J., Santos, T. M. M. D., and Patto, M. C. V. (2010). Cereal landraces for sustainable agriculture. A review. *Agronomy for Sustainable Development* 30, 237–269. doi: https://doi.org/10.1051/agro/2009032

Okamoto, Y., Nguyen, A. T., Yoshioka, M., Iehisa, J. C. M., and Takumi, S. (2013). Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. *Breeding Science* 63(4), 423–429. doi: https://doi.org/10.1270/jsbbs.63.423

Papadakis, I. (1929). Scientific report I. Plant Breeding Station Thessaloniki.

Percival, J. (1921). The wheat plant; a monograph . url: https://doi.org/10.5962/bhl.title.17154. doi: https://doi.org/10.5962/bhl.title.17154

Ponce-García, N., Ramírez-Wong, B., Escalante-Aburto, A., Torres-Chávez, P. I., and Figueroa-Cárdenas, J. D. (2016). Mechanical Properties in Wheat (*Triticum aestivum*) Kernels Evaluated by Compression Tests: A Review. *Viscoelastic and Viscoplastic Materials*. doi: https://doi.org/10.5772/64171

Posner, E. S. (2009). Wheat Flour Milling. In *Wheat: Chemistry and Technology*, ed. Khan, K. 119-152. 4th edition.

Preiti, G., Calvi, A., Giuffrè, A. M., Badagliacca, G., Virzì, N., and Bacchi, M. (2022). A Comparative Assessment of Agronomic and Baking Qualities of Modern/Old Varieties and Landraces of Wheat Grown in Calabria (Italy). *Foods* 11(15), 2359. doi: https://doi.org/10.3390/foods11152359

Purugganan, M. D. (2019). Evolutionary Insights into the Nature of Plant Domestication. *Current Biology* 29(14), 705–714. doi: https://doi.org/10.1016/J.CUB.2019.05.053

Robbana, C., Kehel, Z., Ammar, K., Guzmán, C., Naceur, M., Ben, A., and Amri, A. (2021). Unlocking the Patterns of the Tunisian Durum Wheat Landraces Genetic Structure Based on Phenotypic Characterization in Relation to Farmer's Vernacular Name. *Agronomy* 11(4), 634. doi: https://doi.org/10.3390/agronomy11040634

Roselló, M., Royo, C., Álvaro, F., Villegas, D., Nazco, R., and Soriano, J. M. (2018). Pasta-making quality QTLome from mediterranean durum wheat landraces. *Frontiers in Plant Science* 871. doi: https://doi.org/10.3389/fpls.2018.01512

Ruisi, P., Ingraffia, R., Urso, V., Giambalvo, D., Alfonzo, A., Corona, O., Settanni, L., and Frenda, A. S. (2021). Influence of grain quality, semolinas and baker's yeast on bread made from old landraces and modern genotypes of Sicilian durum wheat. *Food Research International* 140, 110029–110029. doi: https://doi.org/10.1016/J.FOODRES.2020.110029

Saleh, M. M. (2020). Stress breeding of neglected tetraploid primitive wheat (*Triticum dicoccum*, *Triticum carthlicum* and *Triticum polonicum*). *Current Botany* 99-110. doi: https://doi.org/10.25081/CB.2020.V11.6100

Schuler, S. F., Bacon, R. K., Finney, P. L., and Gbur, E. E. (1995). Relationship of Test Weight and Kernel Properties to Milling and Baking Quality in Soft Red Winter Wheat. *Crop Science* 35(4), 949–953. doi: https://doi.org/10.2135/cropsci1995.0011183X003500040001x

Sciacca, F., Cambrea, M., Licciardello, S., Pesce, A., Romano, E., Virzì, N., and Palumbo, M. (2014). Evolution of durum wheat from Sicilian landraces to improved varieties. In Porceddu, E., Damania, A. B., and Qualset, C. O., Proceedings of the International Symposium on Genetics and Breeding of Durum Wheat, Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 110, CIHEAM, Bari, 139-145.

Shewry, P. R. (2018). Do ancient types of wheat have health benefits compared with modern bread wheat? *Journal of Cereal Science* 79, 469–476. doi: https://doi.org/10.1016/j.jcs.2017.11.010

Shewry, P. R. and Hey, S. J. (2016). Do we need to worry about eating wheat? *Nutrition Bulletin* 41(1), 6–13. doi: https://doi.org/10.1111/nbu.12186

Shoormij, F., Mirlohi, A., Saeidi, G., and Mohammadi, M. (2022). Concurrent improvement of durum and emmer wheat through reciprocal backcrossing and hybridization. *Agronomy Journal* 114(6), 3074–3082. doi: https://doi.org/10.1002/AGJ2.21183

Smykalova, I., Grillo, O., Bjelkova, M., Hybl, M., and Venora, G. (2011). Morpho-colorimetric traits of *Pisum* seeds measured by an image analysis system. *Seed Science and Technology* 39(3), 612–626. doi: https://doi.org/10.15258/SST.2011.39.3.08

Smykalova, I., Grillo, O., Bjelkova, M., Pavelek, M., and Venora, G. (2013). Phenotypic evaluation of flax seeds by image analysis. *Industrial Crops and Products* 47, 232–238. doi: https://doi.org/10.1016/J.INDCROP.2013.03.001

Thanopoulos, R., Chatzigeorgiou, T., Argyropoulou, K., Kostouros, N. M., and Bebeli, P. J. (2021). State of crop landraces in Arcadia (Greece) and in-situ conservation potential. *Diversity* 13(11). doi: https://doi.org/10.3390/D13110558

Thomas, K., Thanopoulos, R., Knüpffer, H., and Bebeli, P. J. (2013). Plant genetic resources in a touristic island: the case of Lefkada. *Genetic Resources and Crop Evolution* 60(8), 2431–2455. doi: https://doi.org/10.1007/S10722-013-0011-3

Tomás, D., Coelho, L. P., Rodrigues, J. C., Viegas, W., and Silva, M. (2020). Assessment of Four Portuguese Wheat Landrace Diversity to Cope With Global Warming. *Frontiers in Plant Science* 11, 1803–1803. doi: https://doi.org/10.3389/fpls.2020.594977

Travlos, I. S., Karamanos, A. J., Economou, G., and Papastavrou, A. (2012). Screening of Greek wheat landraces for their yield responses under arid conditions. *International Journal of Plant Production* 6(2), 225–238. doi: https://doi.org/10.22069/ijpp.2012.777

Ulukan, H. (2021). Climate Change and Global Warming Effect(s) on Wheat Landraces: A General Approach. *Wheat Landraces* 169-191. doi: https://doi.org/10.1007/978-3-030-77388-5_9

UPOV (2012). International union for the protection of new varieties of plants durum wheat UPOV Code: TRITI_TUR_DUR *Triticum turgidum* L. subsp. *durum* (Desf.) Husn. url: https://www.upov.int/edocs/tgdocs/en/tg120.pdf.

UPOV (2017). International union for the protection of new varieties of plants wheat UPOV Code(s): TRITI_AES *Triticum aestivum* L. emend. Fiori et Paol. Guidelines for the conduct of tests for distinctness, uniformity and stability. url: https://www.upov.int/edocs/tgdocs/en/tg003.pdf.

Van De Wouw, M., Kik, C., Van Hintum, T., Van Treuren, R., and Visser, B. (2010). Genetic erosion in crops: Concept, research results and challenges. *Plant Genetic Resources: Characterisation and Utilisation* 8(1), 1–15. doi: https://doi.org/10.1017/S1479262109990062

Venora, G., Grillo, O., and Saccone, R. (2009). Quality assessment of durum wheat storage centres in Sicily: Evaluation of vitreous, starchy and shrunken kernels using an image analysis system. *Journal of Cereal Science* 49(3), 429–440. doi: https://doi.org/10.1016/J.JCS.2008.12.006

Wiersma, J. J., Busch, R. H., Fulcher, G. G., and Hareland, G. A. (2001). Recurrent Selection for Kernel Weight in Spring Wheat. *Crop Science* 41(4), 999–1005. doi: https://doi.org/10.2135/cropsci2001.414999x

Würschum, T., Langer, S. M., Longin, C. F. H., Tucker, M. R., and Leiser, W. L. (2017). A modern Green Revolution gene for reduced height in wheat. *The Plant Journal* 92(5), 892–903. doi: https://doi.org/10.1111/TPJ.13726

Zarkti, H., Ouabbou, H., Hilali, A., and Udupa, S. M. (2010). Detection of genetic diversity in Moroccan durum wheat accessions using agro-morphological traits and microsatellite markers. *African Journal of Agricultural Research* 5(14), 1837–1844. url: https://hdl.handle.net/20.500.11766/13403.

Zeven, A. C. (1998). Landraces: A review of definitions and classifications. *Euphytica* 104(2), 127–139. doi: https://doi.org/10.1023/A:1018683119237

Zheng, W., Sun, D., and Pan, F. (2012). Analysis on quality characters diversity of wheat landraces from Yangtze River valley. *Advances in Intelligent and Soft Computing, 134 AISC* 385-396. doi: https://doi.org/10.1007/978-3-642-27537-1_48

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2023), 4 (8), 55–63
DOI: 10.46265/genresj.EMOR6509
https://www.genresj.org
ISSN: 2708-3764

A public mid-density genotyping platform for alfalfa (*Medicago sativa* L.)

Dongyan Zhao[a], **Katherine Mejia-Guerra**[a,b], **Marcelo Mollinari**[c], **Deborah A Samac**[d], **Brian M Irish**[e], **Kasia Heller-Uszynska**[f], **Craig T Beil**[a] and **Moira J Sheehan**[*,a]

[a] *Breeding Insight, Cornell University, Ithaca, 14853, NY, USA*
[b] *Sarepta Therapeutics, Cambridge, 02142, MA, USA*
[c] *Campus Box 7609, North Carolina University, NC, Raleigh, 27695, USA*
[d] *Plant Science Research Unit, USDA-ARS, St. Paul, 55108, MN, USA*
[e] *Plant Germplasm Introduction and Testing Research Unit, USDA-ARS, Prosser, 99350, WA, USA*
[f] *Diversity Arrays Technology, ACT 2617, Bruce, Australia*

Abstract: Small public breeding programmes have many barriers to adopting technology, particularly creating and using genetic marker panels for genomic-based decisions in selection. Here we report the creation of a DArTag panel of 3,000 loci distributed across the alfalfa (*Medicago sativa* L.) genome for use in molecular breeding and genomic insight. The creation of this marker panel brings cost-effective and rapid genotyping capabilities to alfalfa breeding programmes. The open access provided by this platform will allow genetic data sets generated on the marker panel to be compared and joined across projects, institutions and countries. This genotyping resource has the power to make routine genotyping a reality for any breeder of alfalfa.

Keywords: Alfalfa, amplicon-sequencing, plant breeding, DArTag genotyping, microhaplotype

Citation: Zhao, D., Mejia-Guerra, K., Mollinari, M., Samac, D. A., Irish, B. M., Heller-Uszynska, K., Beil, C. T., Sheehan, M. J. (2023). A public mid-density genotyping platform for alfalfa (*Medicago sativa* L.). *Genetic Resources* 4 (8), 55–63. doi: 10.46265/genresj.EMOR6509.

Introduction

Molecular breeding techniques have been used for nearly four decades to enhance and speed breeding efforts for major staple food crops like tomato, maize and barley (Tanksley, 1983; Helentjaris *et al*, 1985; Feuerstein *et al*, 1990; Hasan *et al*, 2021). Over time, these techniques have been augmented with high-quality phenotypic data to perform genome-wide association studies (GWAS) and genomic selection and prediction, further fueling breeding for quantitative or complex traits (Eathington *et al*, 2007; Lorenzana and Bernardo, 2009; Heffner *et al*, 2009). While these achievements are significant, many crop species grown for human consumption and livestock feed are still unable to apply these techniques in breeding efforts. Many breeders would like to adopt molecular breeding tools and techniques, but sometimes doing so is hampered by large barriers-to-entry challenges. The range of barriers and how surmountable they are, varies from species to species and is impacted by species-specific challenges in logistics, technical know-how, biology and the growing environment.

Alfalfa (*Medicago sativa* L.) is the most widely grown perennial forage crop worldwide (Undersander, 2021). In the United States, it was the fourth most cultivated crop in 2021 with an estimated direct value of US$11.6 billion (Putnam and Meccage, 2022) and ranked first among forage crops planting area with a total of 14.9 million acres in 2022 (https://www.nass.usda.gov/). Alfalfa is a key nutritional component for dairy and beef production because it contains a high amount of crude protein, provides dietary fibre needed to

---

[*] Corresponding author: Moira J Sheehan (moirasheehan@cornell.edu)

Received: 20.07.2023 | Accepted: 26.09.2023 | Published online: 13.10.2023

maintain rumen health, and is an excellent source of vitamins and minerals. In addition, it is unparalleled as a component of sustainable agricultural systems because of its ability to fix nitrogen, protect water quality, interrupt pest and pathogen cycles in annual crops, and improve soil carbon storage (Fernandez *et al*, 2019). Alfalfa is adapted to different growth environments and depending on location and management, is highly persistent.

As a highly heterozygous outcrossing autotetraploid species, alfalfa features a predominant pattern of random chromosome pairing during meiosis. The four sets of chromosomes add a layer of complexity to genotyping endeavours. Traditional SNP marker systems, primarily designed for diploid species, often fall short when applied to alfalfa due to their inability to identify allelic dosages accurately. Thus, the intricacies of alfalfa's genetic structure call for more sophisticated SNP genotyping systems capable of addressing the unique challenges posed by its autotetraploid nature. The investment cost and reliance upon skilled bioinformatics support for each genotyping run, make this a high-risk technology for breeders to adopt.

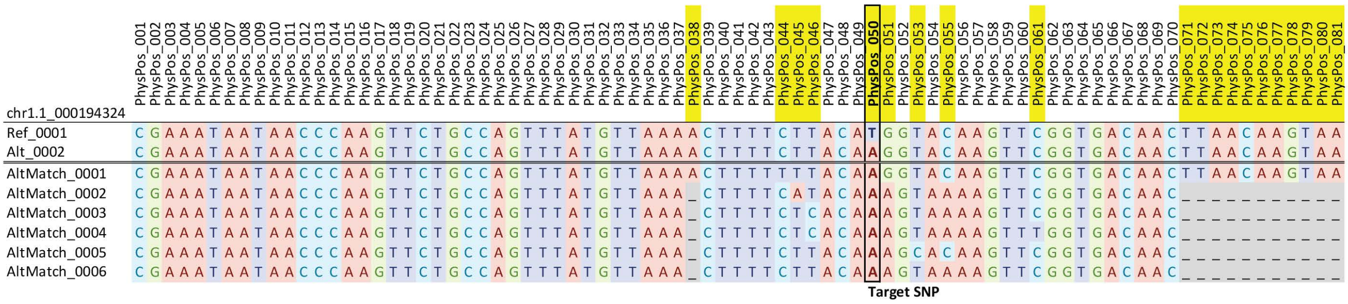
Currently, most alfalfa cultivars are synthetic populations developed by multiple cycles of phenotypic selection for desired traits. Evaluation for biomass yield, winter survival, grow back, disease resistance and forage nutritional quality among other traits is a multiyear process before cultivar registration and commercial seed production. Breeding programmes have operated with half-sib populations originating from polycrosses, where only the maternal parent is known, and the paternal parent can range from a few individuals up to hundreds. Unfortunately, breeding for yield gains in alfalfa using traditional phenotypic evaluation and recurrent selection methods has hit a plateau partly due to its highly heterozygous and heterogeneous population-level breeding. Community genomic tools like those in the 'Tools for Polyploids' project (https://www.polyploids.org/) or developed in other polyploid outcrossing crops have shown promise in accelerating breeding and yield gains (Ferrão *et al*, 2021). The creation of genomic tools that account for the biological and logistic challenges of the crop, has the potential to significantly improve yield gains in alfalfa through breeding.

The first and typically most tractable place to build capacity and tools for molecular breeding is to begin by creating a rapid genotyping pipeline that fits within both the breeding cycle and the selection cycle and can deliver on the breeder's objectives (Hawkins and Yu, 2018; Mejia-Guerra *et al*, 2021). As stated here, a pipeline refers to a complete workflow starting with a genetic marker platform, vendors for service and bioinformatic tools to transform returned raw data into a usable format for breeders. There are several factors to consider when choosing a genetic marker platform: cost per data point, vendor services, turnaround times and what genetic analyses can be done with the resulting data. For alfalfa, we determined that a targeted-amplicon sequenced-based approach would be the most beneficial for breeders. Unlike Genotyping-by-Sequencing (GBS), targeted, amplicon-based genotyping technologies such as DArTag (Diversity Array Technology - DArT), and Capture-Seq (LGC Genomics) have low missing data rates and query the same exact loci in all samples across genotyping projects, allowing new data to be easily appended to existing data (Telfer *et al*, 2019; Darrier *et al*, 2019; Wang *et al*, 2020). The amount of data returned is in the tens of thousands or less, rather than the millions of reads from GBS, simplifying downstream bioinformatics processing (Darrier *et al*, 2019; Milner *et al*, 2019). This in turn speeds up the analysis time for marker-assisted selection (MAS), introgression tracking, linkage mapping and genomic prediction (Darrier *et al*, 2019).

Here, we report the creation of a DArTag panel of 3,000 loci distributed across the alfalfa genome for use in molecular breeding and genomic prediction. DArTag is a hybridization/amplicon-based targeted genotyping platform developed by DArT (Blyton *et al* (2023); https://www.diversityarrays.com/services/targeted-genotying/). Oligos are custom designed to target known genetic variants (SNPs and InDels less than 50bp) with its flanking genomic regions, and sequencing products of 54bp (legacy technology) or 81bp (current technology) in length are produced.

The DArTag assay consists of four steps based on principles described in Krishnakumar *et al* (2008). Briefly, the pool of 3,000 alfalfa oligos, each targeting one genetic variant, is hybridized to denatured gDNA in step 1, followed by SNP/INDEL copying into DArTag molecules by DNA polymerase in step 2. After ligation into circular molecule also in step 2, and nucleases treatment to remove unwanted molecules in step 3, DArTag products are subsequently amplified in step 4 with simultaneous addition of sample unique barcode used downstream for demultiplexing. The products of DArTag assay, after purification and quantification, are sequenced on NGS platforms (e.g. NovaSeq 6000, Illumina) to a depth of around 350x per marker per sample, then demultiplexed and genetic variants detected using DArT's proprietary analytical pipeline.

The alfalfa DArTag panel was designed on the legacy technology to produce 54bp reads but works equally well with the current technology (81bp reads) with the caveat that some residual adapter sequences may be included (read-through of the entire fragment into the adapter). After trimming of any residual adapter sequences, the reads can be used to call SNPs, or in the case of complex genomes like alfalfa, used to identify microhaplotypes (Figure 1). Sequencing reads can contain variants beyond the target SNP, which allows for the detection of more than two alleles at each of the 3,000 loci. As the amplicons are very short, variants found within these reads are assumed to be in complete linkage disequilibrium and therefore can be used for phasing genotyping calls for genetic map construction.



chr1.1_000194324	PhysPos_001	PhysPos_002	PhysPos_003	PhysPos_004	PhysPos_005	PhysPos_006	PhysPos_007	PhysPos_008	PhysPos_009	PhysPos_010	PhysPos_011	PhysPos_012	PhysPos_013	PhysPos_014	PhysPos_015	PhysPos_016	PhysPos_017	PhysPos_018	PhysPos_019	PhysPos_020	PhysPos_021	PhysPos_022	PhysPos_023	PhysPos_024	PhysPos_025	PhysPos_026	PhysPos_027
Ref_0001	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
Alt_0002	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0001	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0002	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0003	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0004	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0005	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T
AltMatch_0006	C	G	A	A	A	T	A	A	T	A	A	C	C	C	A	A	G	T	T	C	T	G	C	C	A	G	T

chr1.1_000194324	PhysPos_028	PhysPos_029	PhysPos_030	PhysPos_031	PhysPos_032	PhysPos_033	PhysPos_034	PhysPos_035	PhysPos_036	PhysPos_037	PhysPos_038	PhysPos_039	PhysPos_040	PhysPos_041	PhysPos_042	PhysPos_043	PhysPos_044	PhysPos_045	PhysPos_046	PhysPos_047	PhysPos_048	PhysPos_049	**PhysPos_050**	PhysPos_051	PhysPos_052	PhysPos_053	PhysPos_054
Ref_0001	T	T	A	T	G	T	T	A	A	A	A	C	T	T	T	T	C	T	T	A	C	A	**T**	G	G	T	A
Alt_0002	T	T	A	T	G	T	T	A	A	A	A	C	T	T	T	T	C	T	T	A	C	A	**A**	G	G	T	A
AltMatch_0001	T	T	A	T	G	T	T	A	A	A	A	C	T	T	T	T	T	T	T	A	C	A	**A**	G	G	T	A
AltMatch_0002	T	T	A	T	G	T	T	A	A	A	_	C	T	T	T	T	C	A	T	A	C	A	**A**	A	G	T	A
AltMatch_0003	T	T	A	T	G	T	T	A	A	A	_	C	T	T	T	T	C	T	C	A	C	A	**A**	A	G	T	A
AltMatch_0004	T	T	A	T	G	T	T	A	A	A	_	C	T	T	T	T	C	T	C	A	C	A	**A**	A	G	T	A
AltMatch_0005	T	T	A	T	G	T	T	A	A	A	_	C	T	T	T	T	C	T	T	A	C	A	**A**	A	G	C	A
AltMatch_0006	T	T	A	T	G	T	T	A	A	A	_	C	T	T	T	T	C	T	T	A	C	A	**A**	A	G	T	A

chr1.1_000194324	PhysPos_055	PhysPos_056	PhysPos_057	PhysPos_058	PhysPos_059	PhysPos_060	PhysPos_061	PhysPos_062	PhysPos_063	PhysPos_064	PhysPos_065	PhysPos_066	PhysPos_067	PhysPos_068	PhysPos_069	PhysPos_070	PhysPos_071	PhysPos_072	PhysPos_073	PhysPos_074	PhysPos_075	PhysPos_076	PhysPos_077	PhysPos_078	PhysPos_079	PhysPos_080	PhysPos_081
Ref_0001	C	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	T	T	A	A	C	A	A	G	T	A	A
Alt_0002	C	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	T	T	A	A	C	A	A	G	T	A	A
AltMatch_0001	C	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	T	T	A	A	C	A	A	G	T	A	A
AltMatch_0002	A	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	_	_	_	_	_	_	_	_	_	_	_
AltMatch_0003	A	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	_	_	_	_	_	_	_	_	_	_	_
AltMatch_0004	A	A	A	G	T	T	T	G	G	T	G	A	C	A	A	C	_	_	_	_	_	_	_	_	_	_	_
AltMatch_0005	C	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	_	_	_	_	_	_	_	_	_	_	_
AltMatch_0006	A	A	A	G	T	T	C	G	G	T	G	A	C	A	A	C	_	_	_	_	_	_	_	_	_	_	_

Figure 1. DArTag sequencing reads from locus Chr1.1_000194324. Each sequence is a microhaplotype detected in breeding material tested on the panel. The DArTag assay was designed to detect the target locus (black rectangle) and distinguish the Reference allele from the Alternative allele. Additional variant nucleotide positions (yellow fill) distinguish the individual microhaplotypes. InDels are shown in grey fill. PhysPos refers to the physical nucleotide position within the sequencing read from left to right.

In addition, accurate allele dosage can be determined for both bi-allelic and multi-allelic haplotypes, allowing genetic effect contributions to be determined for each unique haplotype for traits of interest. As DArT had not tested many polyploid species with DArTag when this study was initiated, we agreed to limit the number of probes to 3,000 loci, though the optimal max may differ by species and genome complexity, and read depth required to sufficiently call genotypes (Andrzej Kilian, DArT, personal communication).

Results

This alfalfa 3K DArTag panel was developed from a diversity panel of 40 individual alfalfa clonal genotypes, focusing on elite breeding and stress-resistant genotypes used in North America. This panel consisted of 17 elite parents with various fall dormancy levels, six samples of diploid-cultivated alfalfa, 13 genotypes with abiotic stress resistance, one genotype with Aphanomyces root rot disease resistance, and three other genotypes (Table 1, column 2). Two biological replicates of the diversity panel were processed, where the sequencing libraries were prepared using either Illumina Nextera WGS library prep at Cornell Institute of Biotechnology or NEBNext Ultra DNA Library Prep Kit with an average insert DNA size of 300bp. The whole-genome sequencing (WGS) was done using Illumina NovaSeq 6000 at Novogene (https://en.novogene.com/). Raw FASTQ sequences were processed by removing residual adapter sequences and low-quality bases using Trimmomatic (LEADING:10 TRAILING:10 SLIDINGWINDOW:4:15 MINLEN:30) (Bolger *et al*, 2014). Cleaned reads were then aligned to the haploid set (the first set out of the four homologous chromosomes) of XinJiangDaYe reference genome (Chen *et al*, 2020) using BWA-MEM (Li, 2012, 2013) and structural variants (SNPs and indels) were called using the DNAseq pipeline developed by Sention (https://www.sentieon.com). A total of 28M SNPs present in both replicates were discovered from the whole-genome re-sequencing of the diversity panel, where a high-confidence set of 10K SNPs (Figure 2) were obtained by requiring them: (1) not located within 5bp distance to an indel, (2) QUAL > 30, (3) minimum and maximum read depths of 20 and 1,900, respectively, (4) for each sample, at least one read supporting reference allele and two reads supporting the alternative allele, (5) no missing genotype per SNP position, (6) with a minor allele frequency greater than 0.25, (7) not located in transposable elements and (8) not within 1Kb of chromosome termini. The 10K SNPs were assessed by DArT, and from those that passed QC, a 3K SNP set targeting even genomic distribution was selected to form a 3K DArTag marker panel. Of the 3,000 loci selected for the panel, 85% (2,542) reside in genic regions and only 15% (458) reside in non-genic regions (Supplemental Table 1). Oligo probes were synthesized, and genotyping done at DArT.

The alfalfa 3K DarTag marker panel was validated using a bi-parental F1 population (n = 184), a backcross (BC1) population (n = 94), and a diverse set of elite genotypes (n = 74) and individual plants from other *Medicago* species (n = 20) (Table 1, column 3). It should be noted that all 40 alfalfa lines used in the SNP discovery were also included in this validation sample set. The material selected for validation was to assess (1) the panel's ability to construct genetic (linkage) maps with the data output and (2) to define the usable limit to the panel with extant species (non-*Medicago sativa*) germplasm.

As expected, the missing data (a marker with < 10 reads in a population) is the lowest among the *Medicago sativa* genotypes. The alfalfa lines used in SNP discovery showed the least missing data (an average of ~9% of the 3K markers with no data) and the rest *M. sativa* lines of the validation sample set had comparable missing rates (10%) (Supplemental Figure 2). Other *Medicago* species had an average of 51% markers with missing data, which is approximately five times higher than the *M. sativa* genotypes.

Table 1. Accessions used in the construction and testing of the alfalfa (*Medicago sativa* L.) 3K DArTag panel. Germplasm used for whole genome sequencing and SNP database construction are indicated by 'Y' in the 'SNP discovery' column. Germplasm used to validate the 3K panel is indicated by 'Y' in the 'Validation set' column.

Sample ID	SNP discovery	Validation set	Contributor	Note
S&W FD4	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 4
S&W FD5	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 5
Legacy FD4	Y	Y	Legacy Seeds	Elite parent; fall dormancy 4
Legacy FD5	Y	Y	Legacy Seeds	Elite parent; fall dormancy 5
S&W FD6	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 6
S&W FD7	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 7
S&W FD8	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 8
S&W FD9A	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 9
S&W SFD9B	Y	Y	S&W Seed Co.	Elite parent; fall dormancy 9
CADL-1	Y	Y	N. Young	Cultivated alfalfa at diploid level
CADL-3	Y	Y	N. Young	Cultivated alfalfa at diploid level
CADL-4-5	Y	Y	N. Young	Cultivated alfalfa at diploid level
CADL-5-3	Y	Y	N. Young	Cultivated alfalfa at diploid level
CADL-13	Y	Y	N. Young	Cultivated alfalfa at diploid level
CADL-18	Y	Y	N. Young	Cultivated alfalfa at diploid level
UMN3988-BIP	Y	Y	D. Samac	Biomass type
RegenSY27x	Y	Y	D. Samac	Regenerator; Ref. genome
I195	Y	Y	N. Young	WAPH5; Aphanomyces root rot
UT14-46 SP	Y	Y	M. Peel	Tetraploid *Medicago falcata*
UT27-62	Y	Y	M. Peel	Elite parent; Salt tolerant
FL99	Y	Y	E. Rios	Elite parent; Fall dormancy 9
Bulldog 505	Y	Y	A. Missaoui	Elite parent; fall dormancy 5
GAMS 1403-FSH	Y	Y	A. Missaoui	Elite parent; fall dormancy 7
GAMS 1404-FSH	Y	Y	A. Missaoui	Elite parent; fall dormancy 8
GAMS 1405-FSH	Y	Y	A. Missaoui	Elite parent; fall dormancy 9
3010	Y	Y	A. Missaoui	Elite parent; fall dormancy 3
CW1010	Y	Y	A. Missaoui	Elite parent; fall dormancy 10
CUF 101	Y	Y	D. Samac	Fall dormancy 10 check
BIP1	Y	Y	M. Peel	Salt tolerant; 27-62
BIP2	Y	Y	M. Peel	Salt tolerant; 31-6
BIP3	Y	Y	M. Peel	SemiP; 1-34
BIP4	Y	Y	M. Peel	SemiP; 6-2
BIP5	Y	Y	M. Peel	SemiP; 14-46
BIP6	Y	Y	M. Peel	Drought (Ut7); 17-43
BIP7	Y	Y	M. Peel	Drought (Ut8); 17-44
BIP8	Y	Y	M. Peel	Drought (Ut9); 18-22
BIP9	Y	Y	M. Peel	Drought (Ut10); 21-3
BIP10	Y	Y	M. Peel	Drought (Ut11); 22-30
BIP11	Y	Y	M. Peel	Drought (Ut26); 7-18
BIP12	Y	Y	M. Peel	Drought (Ut30); 13-14
Wilson	N	Y	L.-X. Yu	Elite parent
WA467895	N	Y	L.-X. Yu	Elite parent
Cornell NY1	N	Y	D. Viands	Elite parent
Cornell NY2	N	Y	D. Viands	Elite parent
Cornell NY3	N	Y	D. Viands	Elite parent
Cornell NY4	N	Y	D. Viands	Elite parent
PAF 13 5, 11-1	N	Y	H. Riday	*Medicago falcata*

Continued on next page

Table 1 continued

Sample ID	SNP discovery	Validation set	Contributor	Note
PAF 13 2, 9-4	N	Y	H. Riday	*Medicago falcata*
PAF 13 9, 10-5	N	Y	H. Riday	*Medicago falcata*
PAF 13 7, 21-2	N	Y	H. Riday	*Medicago falcata*
FAL12 1, 11-5	N	Y	H. Riday	*Medicago falcata*
FAL12 4, 12-4	N	Y	H. Riday	*Medicago falcata*
MAV8	N	Y	D. Samac	Elite parent
Aph 2	N	Y	D. Samac	Elite parent
MAV13	N	Y	D. Samac	Elite parent
MAV14	N	Y	D. Samac	Elite parent
MAV15	N	Y	D. Samac	Elite parent
ZG9	N	Y	D. Samac	Elite parent
ZG20	N	Y	D. Samac	Elite parent
ZG21	N	Y	D. Samac	Elite parent
ZG23	N	Y	D. Samac	Elite parent
ZG25	N	Y	D. Samac	Elite parent
Aph 11	N	Y	D. Samac	Elite parent
Aph 47	N	Y	D. Samac	Elite parent
PI 516640	N	Y	B. Irish	*Medicago arabica*
PI 504540	N	Y	B. Irish	*Medicago arborea*
PI 495215	N	Y	B. Irish	*Medicago bonarotiana*
PI 315458	N	Y	B. Irish	*Medicago cancellata*
PI 498767	N	Y	B. Irish	*Medicago ciliaris*
W6 32886	N	Y	B. Irish	*Medicago daghestanica*
PI 538998	N	Y	B. Irish	*Medicago hybrida*
PI 498849	N	Y	B. Irish	*Medicago laciniata*
PI 537186	N	Y	B. Irish	*Medicago littoralis*
PI 516711	N	Y	B. Irish	*Medicago marina*
PI 287999	N	Y	B. Irish	*Medicago monspelliaca*
PI 537259	N	Y	B. Irish	*Medicago murex*
PI 220021	N	Y	B. Irish	*Medicago orbicularis*
PI 464704	N	Y	B. Irish	*Medicago papillosa*
PI 253450	N	Y	B. Irish	*Medicago pironae*
W6 5252	N	Y	B. Irish	*Medicago polymorpha*
PI 150564	N	Y	B. Irish	*Medicago popovii*
PI 577446	N	Y	B. Irish	*Medicago prostrata*
PI 631912	N	Y	B. Irish	*Medicago ruthenica*
PI 631715	N	Y	B. Irish	*Medicago sativa nothosubsp. tunetana*
PI 631714	N	Y	B. Irish	*Medicago sativa nothosubsp. tunetana*
PI 631952	N	Y	B. Irish	*Medicago sativa nothosubsp. varia*
PI 631920	N	Y	B. Irish	*Medicago sativa nothosubsp. varia*
PI 631923	N	Y	B. Irish	*Medicago sativa subsp. caerulea*
PI 631921	N	Y	B. Irish	*Medicago sativa subsp. caerulea*
PI 641405	N	Y	B. Irish	*Medicago sativa subsp. glomerata*
PI 631978	N	Y	B. Irish	*Medicago sativa subsp. glomerata*
PI 631869	N	Y	B. Irish	*Medicago sativa var. viscosa*
PI 631870	N	Y	B. Irish	*Medicago sativa var. viscosa*
PI 197356	N	Y	B. Irish	*Medicago scutellata*
I195 x J432	N	Y	D. Samac	F1 population (184 progeny)
AphBC1	N	Y	D. Samac	BC1 population (94 progeny)

Using the 3K panel genotyping results, we generated linkage maps for two distinct populations, an F1 and a backcross (BC1) that share the parent I195. For the F1 population, individuals were derived from a cross between parents I195 and J432, which are resistant and susceptible to *Aphamomyces euteiches*, respectively. Meanwhile, the BC1 population was obtained through a cross between I195 and a progeny (85-209) from the above F1 population. Initially, we constructed individual genetic maps for the F1 and BC1 populations. Genotype dosages for both were determined using updog software (Gerard *et al*, 2018). Subsequently, updog-generated objects were fed into MAPpoly software (Mollinari and Garcia, 2019; Mollinari *et al*, 2020) to build separate genetic maps for each population. A standard screening was performed based on missing data and Mendelian segregation fit. We calculated the recombination fraction matrix between all retained markers, using this information to cluster markers into linkage groups. According to available genome information, most of these markers corresponded with specific chromosomes. Notably, a few markers that were mapped outside their physical position still presented consistent linkage with the markers in their assigned group. This pattern held true across both F1 and BC1 populations. For each linkage group formed, we used MAPpoly's functions, *mds_mappoly* and *est_rf_hmm_sequential* to carry out de novo ordering and phasing to obtain the final F1 and BC1 maps. From all mapped markers, only 2.55% were assigned to different chromosomes in the F1 map and 0.97% in the BC1 map. After constructing individual maps for the F1 and BC1 populations, we merged them using the genome order (Figure 3A; Supplemental Figure 1A). The mapped markers were all consistent placed in the two maps, but a few markers were assigned to different linkage groups when comparing the linkage and physical assembly in both maps. These markers were retained in filtering because they had reasonable Mendelian segregation behaviour and their association with linkage groups that do not correspond to their physical chromosome assignment could indicate potential errors in the reference assembly (Figure 3B). Markers mapped out of their physical positions were inserted into the genome-based map using the multidimensional scaling (MDS) de novo information. We then reconstructed a joint map by employing the hidden Markov model (HMM) algorithm's extension, as Mollinari and Garcia (2019) detailed. The implementation for this algorithm can be found in the GitHub repository https://github.com/mmollina/highprecHMM. Finally, haplotypes for all individuals across both F1 and BC1 populations were reconstructed using the same algorithm (Supplemental Figure 1B).

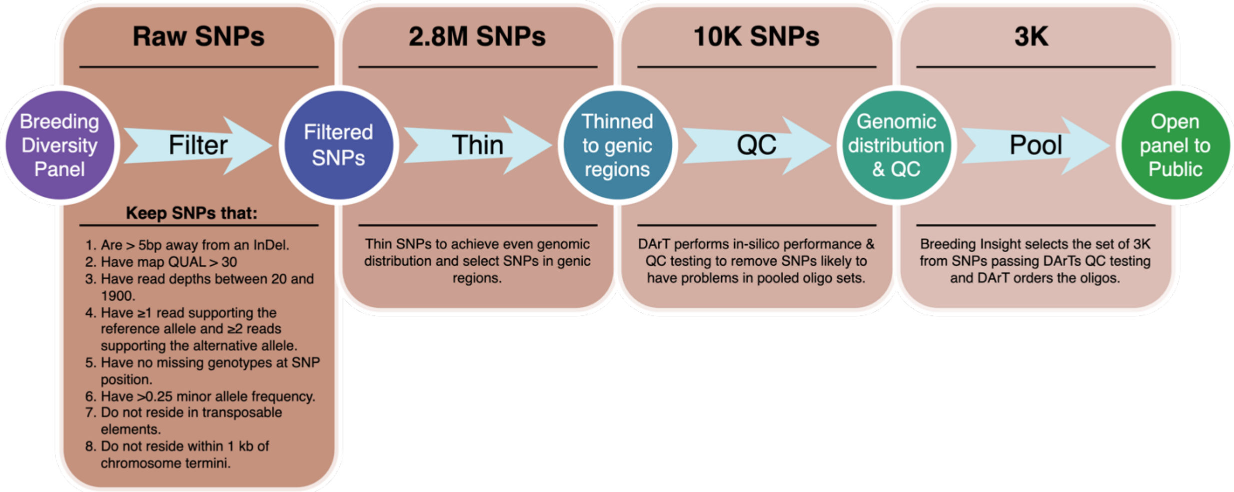
Conclusion

This panel is now publicly available and open for any researcher or breeder to order through DArT (https://www.diversityarrays.com). Researchers interested in using the panel and genotyping services are encouraged to contact DArT directly for pricing details.

Raw data in FASTQ can be requested as can the Missing Allele Discovery File (MADC) that indicates the read depth of each detected haplotype in each sample. The panel and its resulting data are suitable for marker-assisted selection, reconstruction of recombination patterns, allele dosage estimation, and parental confirmation in North American cultivated alfalfa, with some limited application in other *Medicago* species. The efficacy of the panel on breeding materials outside of North America has not been tested, nor has its efficacy in GWAS. Single plant samples were used to create and test the panel. Subsequent testing on samples that are genotyped individually and in tissue or DNA bulks (DNA bulks up to 30 individuals per population) have produced the same allele frequency ratios in both sample types but higher read depth in pools (Esteban Rios, personal communication). More testing is needed to determine the most efficient number of samples to pool to achieve population-level allele frequencies with minimal human labour and monetary costs.

The DArTag assay can be processed from gDNA or from tissue to genotyping data extraction in a three-week turnaround time. The genotyping data report comprises allele dose calls and raw data with custom report formats available upon request. One benefit that DArTag has over fixed array platforms is the ability to update and improve the panel as required over time. The panel is a pool of 3,000 oligos, one per locus, which is used to generate the sequencing libraries from the assayed material. Because the pool is created from individual oligo stocks, the removal of suboptimal loci or the addition of new loci can be easily done by creating a new pool. To determine which loci should be considered for removal, extensive genotyping ($>$ 10,000 samples) is underway to identify those loci that consistently underperform or fail and flag them for removal. Independently, as new significant QTL markers and/or markers specific to other germplasm are detected, they can be targeted for inclusion in the original pool in the next version(s) of the panel. DArT offers re-pooling services once per year at low or no cost, but more frequent requests could result in labour surcharges being applied (Andrzej Kilian, personal communication). Researchers interested in initiating projects with DArT are encouraged to contact DArT directly for consultation.

Another benefit of the deep testing underway is the ability to detect and catalogue all the microhaplotypes into a fixed allele database, which will improve combining data sets across genotyping projects (manuscript in preparation). If after deep testing it is clear that there are too few markers for GWAS for given traits of interest, additional panels can be made to complement this panel. The other option is to add the required loci to the existing panel up to the technical limit of 7K, which is



Keep SNPs that:

1. Are > 5bp away from an InDel.
2. Have map QUAL > 30
3. Have read depths between 20 and 1900.
4. Have ≥1 read supporting the reference allele and ≥2 reads supporting the alternative allele.
5. Have no missing genotypes at SNP position.
6. Have >0.25 minor allele frequency.
7. Do not reside in transposable elements.
8. Do not reside within 1 kb of chromosome termini.

Thin SNPs to achieve even genomic distribution and select SNPs in genic regions.

DArT performs in-silico performance & QC testing to remove SNPs likely to have problems in pooled oligo sets.

Breeding Insight selects the set of 3K from SNPs passing DArTs QC testing and DArT orders the oligos.

Figure 2. Filters and criteria applied to produce the 3K DArTag SNP panel from the whole-genome sequencing (WGS) of the alfalfa diversity panel. M, millions; K, thousands.

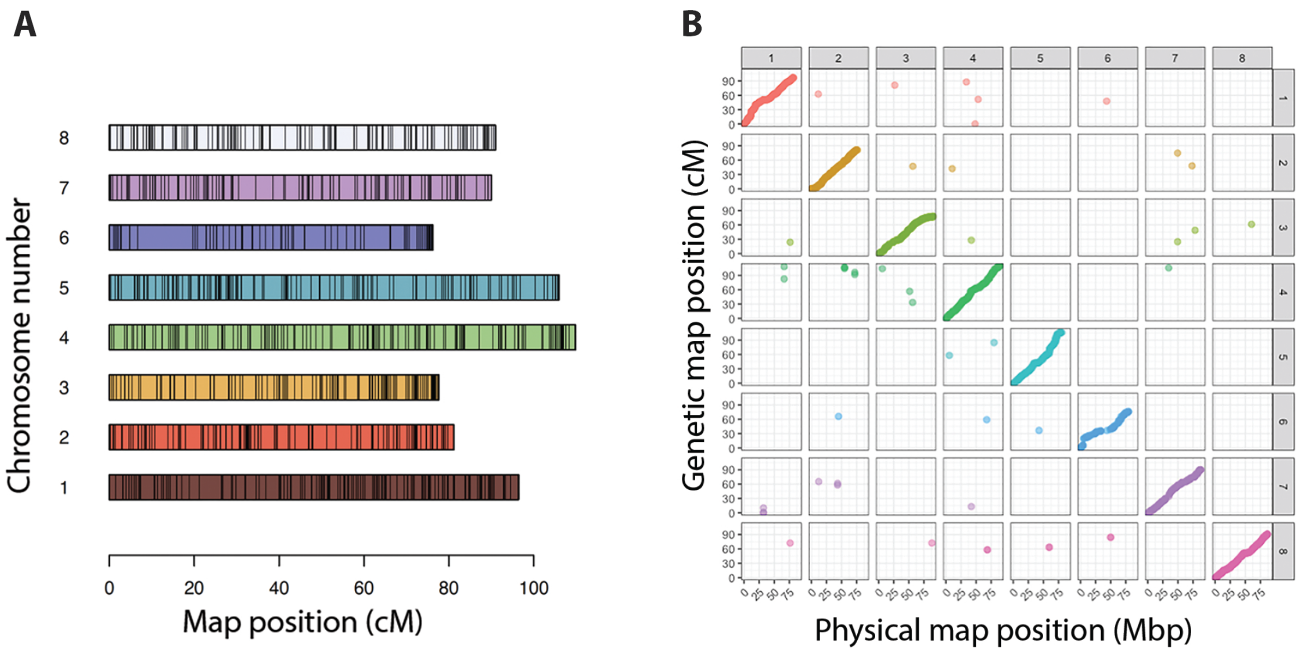


Figure 3. Composite genetic maps of a bi-parental F1 and a backcross (BC1) population. A) Regeneration of the eight linkage groups of alfalfa genome. Scale bar is shown in cM. B) Scatter plots showing the relationship of genetic distance (cM) to physical distance (Mbp) for each of the eight linkage groups.

a more cost-effective option for the routine genotyping service with scalability.

We choose to create a panel of 3,000 loci due to cost and technical reasons, but smaller complementary panels can be made at lower up-front and downstream usage costs. The addition of a complementary 3K panel would nearly double the cost of genotyping per sample but would result in more granular genotyping data.

Data availability statement

The FASTQ files from the whole-genome skim sequencing for the 40 *Medicago sativa* accessions used for identifying the candidate SNP variants are housed in the NCBI Short Read Archive under the BioProject ID PRJNA1014379. The targeted regions used to create the 3K DArTag markers and the haplotypes detected as of 31 May 2023 (v17) are available on DRYAD (https://datadryad.org/stash/share/wJEn32Dfl94EOYMoeM00PJti6MKUliPBTAtsgbWJyOU). The code and data for construction of the F1, BC1 and joint maps in MAPpoly are available in our GitHub repository for those interested in reproducing our analysis (https://github.com/Breeding-Insight/alfalfa_dartag_panel_paper.git).

Acknowledgments

Breeding Insight is acknowledged for project design, marker development, curation and data processing. Diversity Arrays Technology created the oligo array, provided sequencing services and contributed to the manuscript. We thank S&W Seed Company, Legacy Seeds, Debby Samac, Brian Irish, Michael Peel, Long-Xi Yu, Heathcliffe Riday, Nevin Young, Esteban Rios, Ian Ray, Ali Missaoui and Don Viands for providing germplasm and Sharon Mitchell for manuscript review. Breeding Insight was funded for this work through a Cooperative Agreement between USDA-ARS and Cornell (project number: 8062-21000-043-004-A). Marcelo Mollinari was funded by a USDA NIFA-awarded AFRI grant (project number: 2022-67013-36269).

Supplemental Data

Supplemental Figure 1. Alfalfa Genetic map construction for an F1, BC1, and a joint map of the consensus.
Supplemental Figure 2. Missing data rates for different grouped subsets of genetic material.
Supplemental Table 1. Final 3,000 loci selected for the DArTag panel.

Author contributions

DZ, KMMG, DS and MJS contributed to experimental design and planning. DZ, DS and MJS selected the diversity panel for WGS. DS, MP and BI grew and harvested all plant materials used in the study. KMMG performed all the WGS analyses, SNP database creation, filtering pipelines and quality control analyses to create the 3K panel. KHU managed the panel creation at Diversity Arrays Technology. DZ, MM and DS executed the data analyses and genetic mapping. DZ and MJS wrote the initial draft of the manuscript. CB managed experiments and communication among all authors involved. All authors contributed to reviewing the manuscript.

Conflict of interest statement

The authors have no conflicts of interest to report.

References

Blyton, M. D. J., Brice, K. L., Heller-Uszynska, K., Pascoe, J., Jaccoud, D., Leigh, K. A., and Moore, B. D. (2023). A new genetic method for diet determination from faeces that provides species level resolution in the koala. *bioRxiv* 2023.02.12.528172. doi: https://doi.org/10.1101/2023.02.12.528172

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinform* 30(15), 2144–2120. doi: https://doi.org/10.1093/bioinformatics/btu170

Chen, H., Zeng, Y., Yang, Y., Huang, L., Tang, B., Zhang, H., Hao, F., Liu, W., Li, Y., Liu, Y., Zhang, X., Zhang, R., Zhang, Y., Li, Y., Wang, K., He, H., Wang, Z., Fan, G., Yang, H., Bao, A., Shang, Z., Chen, J., Wang, W., and Qiu, Q. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nature Comm* 11, 1–11. doi: https://doi.org/10.1038/s41467-020-16338-x

Darrier, B., Russell, J., Milner, S. G., Hedley, P. E., Shaw, P. D., Macaulay, M., Ramsay, L. D., Halpin, C., Mascher, M., Fleury, D. L., Landridge, P., Stein, N., and Waugh, R. (2019). A Comparison of Mainstream Genotyping Platforms for the Evaluation and Use of Barley Genetic Resources. *Front Plant Sci* 10, 1–14. doi: https://doi.org/10.3389/fpls.2019.00544

Eathington, S. R., Crosbie, T. M., Edwards, M. D., Reiter, R. S., and Bull, J. K. (2007). Molecular markers in a commercial breeding program. *Crop Sci* 47, 154–163. doi: https://doi.org/10.2135/cropsci2007.04.0015IPBS

Fernandez, A. L., Sheaffer, C. C., Tautges, N. E., Putnam, D. H., and Hunter, M. C. (2019). Alfalfa, Wildlife, and the Environment (St. Paul, MN: National Alfalfa and Forage Alliance).

Ferrão, L. F. V., Amadeu, R. R., Benevenuto, J., Oliveira, I. D. B., and Munoz, P. R. (2021). Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons From Blueberry Breeding. *Front Plant Sci* 12, 1–13. doi: https://doi.org/10.3389/fpls.2021.676326

Feurerstein, U., Brown, A. H. D., and Burdon, J. J. (1990). Linkage of Rust Resistance Genes from Wild Barley (*Hordeum spotaneum*) with Isozyme Markers. *Plant Breeding* 104, 318–324. doi: https://doi.org/10.1111/j.1439-0523.1990.tb00442.x

Gerard, D., Ferrão, L. F. V., Garcia, A. A. F., and Stephens, M. (2018). Genotyping Polyploids from Messy Sequencing Data. *Genetics* 210(3), 789–807. doi: https://doi.org/10.1534/genetics.118.301468

Hasan, N., Choudhary, S., Naaz, N., Sharma, N., and Laskar, R. A. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *J Genet Eng Biotech* 19, 1–26. doi: https://doi.org/10.1186/s43141-021-00231-1

Hawkins, C. and Yu, L. X. (2018). Recent progress in alfalfa (*Medicago sativa* L.) genomics and genomic selection. *The Crop Journal* 6, 565–575. doi: https://doi.org/10.1016/j.cj.2018.01.006

Heffner, E. L., Sorrells, M. E., and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci* 49, 1–12. doi: https://doi.org/10.2135/cropsci2008.08.0512

Helentjaris, T., King, G., Slocum, M., Siedenstrang, C., and Wegman, S. (1985). Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant*

Mol Biol 5, 109–118. doi: https://doi.org/10.1007/BF00020093

Krishnakumar, S., Zheng, J., Wilhelmy, J., Faham, M., Mindrinos, M., and Davis, R. (2008). A comprehensive assay for targeted multiplex amplification of human DNA sequences. *PNAS* 105(27), 9296–9301. doi: https://doi.org/10.1073/pnas.0803240105

Li, H. (2012). Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinform* 28(14), 1838–1882. doi: https://doi.org/10.1093/bioinformatics/bts280

Li, H. (2013). Aligning sequence reads, clones sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997. doi: https://doi.org/10.48550/arXiv.1303.3997

Lorenzana, R. and Bernardo, R. (2009). Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor Appl Genet* 120, 151–161. doi: https://doi.org/10.1007/s00122-009-1166-3

Mejia-Guerra, M. K., Zhao, D., and Sheehan, M. J. (2021). Genomic Resources for Breeding in Alfalfa: Availability, Utility, and Adoption. In *The Alfalfa Genome, Compendium of Plant Genomes,* ed. Yu, L. X. and Kole, C., (Cham: Springer), 177-189.

Milner, S. G., Jost, M., Taketa, S., Mazon, E. R., Himmelbach, A., Oppermann, M., Weise, S., Knupffer, H., Basterrechea, M., König, P., Schüler, D., Sharma, R., Pasam, R. K., Rutten, T., Guo, G., Xu, D., Zhang, Z., Herren, G., Müller, T., Krattinger, S. G., Keller, B., Jiang, Y., González, M. Y., Zhao, Y., Habekuß, A., Fäber, S., Ordon, F., Lange, M., Börner, A., Graner, A., Reif, J. C., Scholz, U., Mascher, M., and Stein, N. (2019). Genebank genomics reveals the diversity of a global barley collection. *Nat. Genet* 51, 319–326. doi: https://doi.org/10.1038/s41588-018-0266-x

Mollinari, M. and Garcia, A. A. F. (2019). Linkage Analysis and Haplotype Phasing in Experimental Autopolyploid Populations with High Ploidy Level Using Hidden Markov Models. *Genes, Genomes, Genetics* 3(10), 3297–3314. doi: https://doi.org/10.1534/g3.119.400378

Mollinari, M., Olukolu, B. A., Gds, P., Khan, A., Gemenet, D., Yencho, G. C., and Zeng, Z. (2020). Unraveling the Hexaploid Sweetpotato Inheritance Using Ultra-Dense Multilocus Mapping. *Genes, Genomes, Genetics* 3(1), 281–292. doi: https://doi.org/10.1534/g3.119.400620

Putnam, D. and Meccage, E. (2022). Profitable alfalfa production sustains the environment. In Proceedings, 2022 World Alfalfa Congress, 14-17 November 2022, San Diego, CA.

Tanksley, S. D. (1983). Molecular markers in plant breeding. *Plant Mol Biol Rep* 1, 3–8. doi: https://doi.org/10.1101/2023.02.12.528172

Telfer, E., Graham, N., Macdonald, L., Li, Y., Klápště, J., Resende, M., Neves, L. G., Dungey, H., and Wilcox, P. (2019). A high-density exome capture genotype-by-sequencing panel for forestry breeding in Pinus radiata. *PLoS One* 14(9), 222640–222640. doi: https://doi.org/10.1371/journal.pone.0222640

Undersander, D. (2021). Economic importance, practical limitations to production, management, and breeding targets of alfalfa. In *The Alfalfa Genome, Compendium of Plant Genomes,* ed. Yu, L. X. and Kole, C., (Cham: Springer), 1-11.

Wang, N., Yuan, Y., Wang, H., Yu, D., Liu, Y., Zhang, A., Gowda, M., Nair, S. K., Hao, Z., Lu, Y., Vincente, F. S., Prasanna, B. M., Li, X., and Zhang, X. (2020). Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding. *Sci Rep* 10. doi: https://doi.org/10.1038/s41598-020-73321-8

Genetic Resources

SHORT COMMUNICATION

Genetic Resources (2023), 4 (8), 64–70
DOI: 10.46265/genresj.IBNU2035
https://www.genresj.org
ISSN: 2708-3764

Management practice of the Sheko cattle breed in Ethiopia: A review

Melkam Aleme [*,a] **and Gezahegn Mengistu** [b]

[a] *Ethiopian Institute of Agricultural Research, Teppi Agricultural Research Center, Teppi, Ethiopia*
[b] *Ethiopian Institute of Agricultural Research, Holeta Agricultural Research Center, Addis Ababa, Ethiopia*

Abstract: The great diversity of agroecological conditions and production systems present in Ethiopia contributes to the country's large livestock population and makes it suitable for various forms of livestock production. The Ethiopian livestock sector provides significant economic and social benefits at household levels and to the national economy. Ethiopia has 32 recognized indigenous cattle breeds, of which the Sheko is a known trypano-tolerant breed found in the southwest of the country. However, in recent times, the Sheko population has been declining mainly due to interbreeding with local Zebu cattle and to a shift in the production system. This paper aims to review and illustrate the current status and management practices of the Sheko cattle breed towards conserving and improving the breed, and the production system. The feed resources in the Sheko home area are natural pasture and crop residues, with limited utilization of cultivated improved forage. Husbandry practices such as feeding, watering, housing and veterinary services are priority areas where improvements are needed. Enhancing the conservation and improvement of the breed would greatly benefit from the active improvement of various stakeholders, including governmental policymakers and non-governmental organizations. Therefore giving special attention to enhancing the management systems in the breed's home area is crucial. This involves the direct engagement of research centres, extension workers; and higher learning institutions in proximity to the area all aimed at the conservation and improvement of the Sheko cattle breed.

Keywords: Conservation, Ethiopia, husbandry practice, livestock, trypanosomiasis

Citation: Aleme, M., Mengistu, G. (2023). **Management practice of the Sheko cattle breed in Ethiopia: A review**. *Genetic Resources* 4 (8), 64–70. doi: 10.46265/genresj.IBNU2035.

Introduction

Ethiopia is a forefront country in Africa with an important livestock sector. The great diversity of agroecological conditions and production systems found in Ethiopia contributes to the country's large livestock population and makes it suitable for various forms of livestock production. The country is home to 32 recognized cattle breeds with an estimated population of 65 million heads (DAGRIS, 2007; Dessie, 2012; CSA, 2020). Although livestock farming in Ethiopia is subsistence-oriented, the sector provides significant economic and social contributions at household levels and to the national economy (Tolera *et al*, 2012; Desta *et al*, 2011). For instance, the livestock sector in the country serves as a source of food, income, employment opportunities, draught power and savings. The sector contributed 20% of the total gross domestic product (GDP), 40% of agricultural GDP, and 20% of national foreign exchange earnings (World Bank, 2007).

The occurrence of serious diseases also limits cattle production through increased mortality rates and their effects on fertility, growth and production. The major disease affecting cattle in the area is trypanosomiasis, locally called 'Gendi', which is caused by flagellate protozoa belonging to the genus *Trypanosoma* and transmitted by tsetse fly (*Glossina* spp). The disease appears at all times of the year but it reaches its peak point after the rainy season (May–October). Both direct and indirect loss is brought on by the disease. Lack of equipment and transport for the field services, a weak tsetse and trypanosomiasis control unit, and a shortage of adequately trained workforce are some

---

[*] Corresponding author: Melkam Aleme (melekamaleme@gmail.com)

Received: 17.07.2023 | Accepted: 15.11.2023 | Published online: 19.12.2023

of the problems associated with cattle health in the area (DAGRIS, 2007).

The direct loss is primarily caused by disease-related costs, death, morbidity and infertility in infected cattle. According to Seyoum *et al* (2013), trypanosomiasis-related mortality, morbidity and control expenses result in a yearly direct loss of US$200 million. More substantial than the direct loss is the indirect loss, which includes the lack of use of cattle and animal power for agricultural production in productive land in tsetse-infested areas. Food security and the reduction of poverty are both significantly hampered by trypanosomiasis (Kitila *et al*, 2017). Blackleg and other diseases including babesiosis and pasteurellosis are also present in the area. There is heavy tick infestation, especially in the lowland areas of southwest Ethiopia (Dawit, 1992). The Sheko cattle breed is known for its trypano-tolerance, which makes it very suitable for areas affected by the disease but populations are facing conservation challenges. This paper aims to review and illustrate the current status and management practices of Sheko cattle towards conserving and improving the breed and its production system.

Location and characteristics of Sheko cattle breed

Indigenous cattle breeds have a multipurpose function and are distributed across the country's diverse topographic and climatic conditions. Sheko is one of the taurine-known cattle breeds (humpless) found in the southwestern part of Ethiopia, particularly in the Bench-Sheko, Sheka and Keffa zones of the Southern Nations Nationalities and Peoples' (SNNP) region and adjacent to the Sudanese border (Hanotte *et al*, 2000) (Figure 1). Sheko cattle are also called 'Mizan' or 'Goda' by the local community. Phenotypically, the breed is characterized by short-horned and/or hornless heads, a small body size, humplesseness, particularly in females, and a diverse range of coat colours such as red, red-brown, patchy red, white, and black (Taye *et al*, 2009; Bayou *et al*, 2014).

The breed is known for its hardiness, trypano-tolerance (ability to survive in trypanosomiasis endemic areas), good mothering ability, better feed conversion efficiency, fast growth rate and comparable milk production to other indigenous cattle breeds in Ethiopia (Taye *et al*, 2009; Stein *et al*, 2011; Desta *et al*, 2012). Farmers in the mid-altitude agroecological zone keep the Sheko cattle breed primarily for draught power followed by milk and income whereas farmers in the lowland agroecological zones often keep the breed for milk and as a source of income (Bayou *et al*, 2014). However, the breed is endangered, estimated at around 4,000 heads due to gradual interbreeding with local zebu, castration of the male at an early age, shift of production systems and shrinkage of grazing land (Takele, 2005). Studies done by Hanotte *et al* (2000) indicated that approximately 90% of the inspected Sheko bulls had their particular taurine allele almost 10% replaced by the Zebu-originated indicine allele, showing an alarming amount of crossbreeding Zebu.

Conservation and maintenance of pure-breed individuals are essential to the long-term survival of any species to withstand the changing environmental conditions and ensure their sustainable use in the future. Conservation of animal genetic resources is undertaken to maintain their diversity, fostering contributions to food and agricultural production, enhancing productivity and safeguarding ecological resources and cultural values for current and future generations. According to Mekuriaw and Kebede (2015), genetic conservation can be carried out using different methods. *In situ* conservation refers to the conservation of livestock through continued use by livestock keepers in the production system in which the livestock evolved or are normally found and bred. It involves the production of animals in their original production environment either on-farm or community-based (Figure 2).

Feeds and feeding

Feed is the main input of livestock production and largely determines its profitability since it accounts for 65–75% of the total cost of livestock operation (Walli *et al*, 2012; Makkar, 2018; Matope *et al*, 2020). In addition, feeding is the foundation of livestock systems as it directly or indirectly affects the entire livestock sector, including animal productivity, health and welfare, and the environment (Garg *et al*, 2014). Feed sourcing and feeding are at the very interface where the 'positive' and 'negative' effects of livestock production are negotiated (Blummel *et al*, 2018). Feed resources in the home area of the Sheko cattle breed are natural pasture, crop residues, and to a small extent the cultivation and utilization of cultivated improved forage. For instance, during cropping season (early June to late August), tinned maize (*Zea mays* L.) and the green stovers remaining after the green cob is harvested are used for animal feeding (Bayou *et al*, 2014). Crop residues, such as teff or barley straw are also major feeds. Improved forage like elephant grass (*Pennisetum purpureum*) to some extent, sesbania (*Sesbania sesban*), cassava (*Manihot esculenta*), taro (*Colocasia esculenta*), sweet potato (*Ipomoea batatas*), leaf and stem parts of enset (*Ensete ventricosum*), Steudner's dragon tree (*Dracaena steudneri*), banana (*Musa sapientum*), sugarcane (*Saccharum officinarum*) and homemade conventional feeds are common and well known. No reports are available regarding the production of roughage, such as hay for feeding livestock during the dry season. Straw treatments to produce feed include cutting and chopping long stalks and utilizing molasses and urea to enhance digestibility (Takele and Workneh, 2011; Hailu, 2020).

There is a general lack of feed availability for livestock production in the area due to the use of grazing lands for crop production. Feed shortage is rampant when most of the farmlands are covered with food crops during the wet season, and grasses are depleted

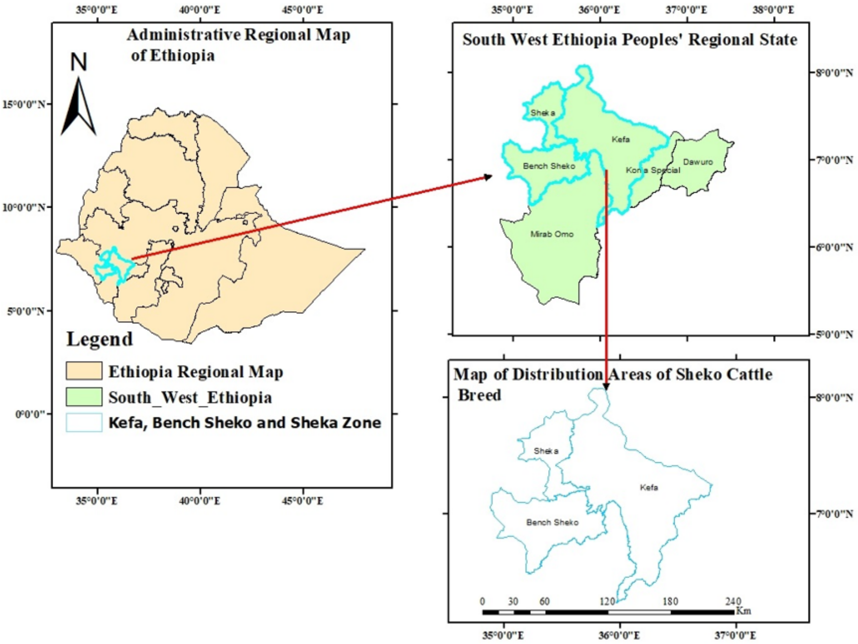


Figure 1. Map of the distribution areas of Sheko cattle breed

Figure 2. Sheko cattle breed maintained at the Teppi Agricultural Research Center

during the dry season (Bayou *et al*, 2014). The lack of improved feed conservation knowledge, shortage of land to plant forage, lack of adequate extension service to inform farmers on improved forage technology and unavailability of forage seed or planting materials are the most common limiting factors in the production and productivity of Sheko cattle breed (Desta *et al*, 2011). Feed resources are declining due to the expansion of crop fields in the face of fast human population growth. Consequently, Sheko cattle keepers increasingly resort to tethered feeding. Therefore, the average cattle herd size is declining, which also discourages the maintenance of the Sheko cattle breed (Desta *et al*, 2011; Mengistu *et al*, 2017). There is a notable lack of awareness and poor feed resource management, utilization, preservation, treatment and development (Hailu, 2020).

Feeding troughs are made of wood clay dishes, and synthetic materials such as plastic bath and plastic pot are used (Desta *et al*, 2011). The challenge lies in creating linkages between innovation sources, such as research centres and agricultural extension services, and the expansion of technology as well as coordination with budgetary administrations. Accordingly, addressing the issues related to the quantity and quality of feeds is essential (MOA, 2013).

Water resources and watering systems

Within the region where the Sheko cattle are grown, agriculture has actively harnessed water resources, predominantly drawing from natural resources such as springs and ponds. However, watering frequency is an additional limiting factor for Sheko cattle rearing (Bunke, 2019). The irrigation system adds further challenges, particularly if access, water quality, timing and frequency of watering are limiting factors (Hailu, 2020).

According to Mengistu *et al* (2017), water sources for cattle in the area are rivers, springs and ponds. The majority of the people give water to their cattle twice a day during the dry season, others offer water to their cattle once a day and some of them water their animals three times a day by tracking the animals to the watering point.

The majority of producers report there is no water shortage in the area, but for the remaining 27%, water is scarce during the dry season. The distance to a water hole is another factor that limits the rearing of cattle. About 86.4% of producers travelled 1–4 km to reach drinking water for their animals (Mengistu *et al*, 2017). Some communities need more accessible watering points and natural mineral licks to fully utilize the genetic potential of the Sheko cattle breed (Desta *et al*, 2011).

Housing system

The majority of cattle keepers in Mejenger (81.2%) and 47% in Bench-Sheko zones keep their animals in separately constructed houses made of wood with a grass-covered roof (Mengistu *et al*, 2017). At night, cattle are housed not far from the family house to protect them from cold, rain, predators and theft (Alemu, 1990). Feeder cattle are well protected from adverse weather conditions, predator attacks and theft. To keep the stall warm, the walls are covered with leaves and grasses. Enough space is provided in the stall to reduce competition for feed among stall-fed cattle (Takele and Habtamu, 2009).

Calves are kept separate from their dams, either within an annex connected to most houses or independently situated in and around the family house, typically within a horse shelter. Often the roof of a cattle house is made of grasses (Mengistu *et al*, 2017; Hailu, 2020).

Husbandry practices

Husbandry practices are carried out with family individuals and sometimes with neighbours. To utilize manure for trim generation, farmers frequently clean cattle faeces and urine and bolster refusals from the barn and the compound (Takele and Habtamu, 2009). Tools such as nose rings are used to tame aggressive Sheko bulls and oxen in the area. Traditionally, farmers use ear rings for the same purpose and this practice needs to be closely examined for its effectiveness (Lund, 2002).

Breaking and training animals at a younger age and alternating herding with tethering can soften the aggressive behaviour of the breed. Larger herd sizes increase the likelihood of selection and maintenance of breeding studs in the villages. Male cattle are usually castrated at an early age. Such males are kept in the herd for a long time adorned with decorations and are a source of pride to their owners. Another reason for the observed high number of castrates in household herds is that the number of castrated males is an indicator of social status ranking (Terefe *et al*, 2012).

Marketing practices

The breed has the potential for high marketability in large parts of south-western Ethiopia, which are challenged with medium to high tsetse and trypanosomiasis presence. Recent government initiatives, aimed at relocating smallholder farmers from densely populated highlands into underutilized fertile and sparsely populated valleys in south-western Ethiopia, have created a demand for adapted breeding cattle with the Sheko breed emerging as the optimal choice. The dairy and draught qualities of the breed may be worth investigating even for other agroecologies. The agriculture extension services and the national research systems need to support this effort until market interests gain momentum. In light of the threat of extinction, implementing an incentive system is essential for the Sheko breed. Owners who successfully rear a Sheko calf could be rewarded, promoting efficiency as proposed in some studies and encouraging breed preservation efforts by discouraging crossbreeding (Zander *et al*, 2009).

Additionally, the creation of a breeders' society for Sheko cattle will ensure farmers' involvement in identifying elite animals, creating market opportunities, recruiting herd registration and taking the lead in conservation activities (Hegde, 2005). Targeted promotion outside the existing market niche should focus on the special merits of the breed, such as its trypano-tolerance and dairy quality potential. The breed management plan should also be widely communicated to relevant stakeholders using appropriate media as part of marketing strategies (Terefe *et al*, 2012).

Production system

The livestock production system dominating in Sheko district and its surrounding areas is a mixed crop–livestock farming system. Permanent crops like coffee and bananas are widely produced. Coffee is the main cash crop. Smaller areas of land are also used for cereal production like maize and sorghum. Fruits such as avocado and mango are also highly produced. Cattle are important as a source of diverse foods such as milk, butter and meat, and also a source of income through sales of live animals and animal products like leather (Fasil, 2004).

The margin for increasing forage production from pasture is limited. Free grazing is the main animal management system in the area, but a few farmers provide supplementary feed at home. Most of the farmers keep their animals outdoors during the night near their homestead. Some farmers have enclosures made for their animals, whereas others keep animals together within the family house in a partially enclosed area (MOA, 1984). Sheko cattle are said to be better milk producers, reproduce faster than other breeds of cattle in tsetse-infested zones, and have good grazing and browsing ability during critical periods of forage scarcity (Workeneh, 2001). The average production and reproduction performances of the Sheko breed are presented in Table 1.

The most important characteristics of Sheko cattle include tolerance to disease especially to trypanosomes, long lactation length, good milking potential, good adaptation to heat stress, good traction power and adaptation to internal parasitic and tick infestations. Other positive traits attributed to Sheko cattle include good mothering ability, less selective feeding behaviour, attractive coat colour appropriate for the local environment, tolerance for biting flies, shorter calving interval and better adaptation to the terrain of the area. On the other hand, Sheko are known for their aggressive behaviour and their relatively high feed requirement of the farm households (Elias, 2008).

Table 1. Production and reproduction performances of the Sheko cattle breed (Takele, 2005)

Parameter and measure	Values
Age at first mating (years)	3
Age at first calving (years)	4
Calving interval (years)	1.5
Gestation length (days)	276.8
Body weight at calving (kg)	20
Body weight at weaning (kg)	102
Daily milk yield (litre)	2.3
Lactation milk yield per calving (litre)	698.3
Lactation length (months)	9.9
Females life span calving (in number)	8

Constraints of production

The overall trend indicates a decline in the population of Sheko cattle in its production region. The reason for this decline could be identified with Sheko cattle's energetic and aggressive behaviour, which decreases their acceptance by the local community (Alemu, 2002). The lower male-to-female ratio in the Sheko population might expose the animals to inbreeding and crossbreeding with bulls of other breeds. This situation is exacerbated by little consideration given to maintaining the endangered breed (Alemu, 2002). Due to their aggressive behaviour, Sheko cattle are difficult to harness and control by old persons, women, children and disabled persons. In addition, the early castration of bulls exercised by some farmers to control their aggressive behaviour seriously limits pure breeding of Sheko cattle (Takele, 2005). The most important constraints affecting Sheko cattle production are their sparse distribution, absence of conservation efforts, declining interest of the community in Sheko cattle due to their aggressive behaviour, inbreeding and hornlessness, which makes them difficult to restrain using rope (Elias, 2008; Bayou *et al*, 2014).

Given that the home area of Sheko cattle is renowned for coffee plantations, nearly all available land except for the swampy areas and hill-tops, is utilized for cultivation. The potential for increased forage production from pasture is severely limited and cattle rely entirely on natural pasture. Consequently, natural pastures are overgrazed causing the proliferation of undesirable plant species. Seasonal feed shortage is common during the dry season. Proper management, along with the allocation of grazing area and conservation of feed, is rarely practised. All stocks are grazed together with no attempt to provide special treatment for different classes of the stock (MOA, 1984).

Conclusion and way forward

In Ethiopia's rural areas, livestock is the most long-lasting feature of the way of life. Sheko cattle are among the three categories of Ethiopian breeds, which include small East Africa zebu and large East Africa zebu. Characterized by short horns and lack of hump, Sheko cattle are predominantly located in the Bench-Maji area, which was formerly part of the Bench-Sheko zone. Additionally, they can be found to some extent in the Sheka and Kefa zones of south-western Ethiopia.

Sheko cattle are raised within mixed farming systems and are sustained through natural pasture and crop residues. The breed's most important traits include resistance to trypanosomiasis, better milk yield compared with other local breeds, drought power and better feed conversion efficiency. Safeguarding the breed is vital to preserving its distinctive characteristics. To counteract the loss of the breed's genetic diversity, a diverse array of conservation and management practices needs to be in place. These include establishing *in situ* breeding stations and breed studs in its breeding tract, promoting niche markets and improving husbandry practices. Crucial measures involve addressing production challenges like feed shortages during the dry season, ensuring proper housing and feeding management, and allowing the breed to thrive in its natural habitat. The creation of incentives can encourage owners to keep breeding stock under good management, contributing to the breed's preservation. Consequently, the establishment of a Sheko breed conservation fund may be necessary.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Melkam Aleme provided contributions to the current versions of the article through collection, review, writing and interpretation, while Gezahegn Mengistu provided a review of the work before its final submission and overall assistance.

References

Alemu, R. (2002). A summarized field report on the Status of Sheko cattle Traditionally Managed in the Bench-Maji zone, South-Western Ethiopia (Addis Ababa, Ethiopia: ILRI). url: https://cgspace.cgiar.org/bitstream/handle/10568/2894/ILRI_2002_AnnualReport.pdf;jsessionid=045C3605E2BF621794F12BAB3ED38990?sequence=1.

Alemu, T. (1990). The unexploited potential of improved forages in the mid-altitude and lowland areas of Ethiopia. PANESA/ARNAB (Pastures Network for Eastern and Southern Africa/African Research Network for Agricultural By-products. Utilization of research results on forage and agriculture.

Bayou, E., Haile, A., Gizaw, S., and Mekasha, Y. (2014). Characterizing husbandry practices and breeding objectives of Sheko cattle owners for designing conservation and improvement strategies in Ethiopia. *Livestock Research for Rural Development* 26. url: http://www.lrrd.org/lrrd26/12/bayo26235.html.

Blummel, M., Teymouri, F., Moore, J., Nielson, C., Videto, J., Kodukula, P., Pothu, S., Devulapalli, R., and Varijakshapanicker, P. (2018). Ammonia Fiber Expansion (AFEX) as spin off technology from 2nd generation biofuel for upgrading cereal straws and stovers for livestock feed. *Animal Feed Science and Technology* 236, 178–186. url: https://hdl.handle.net/10568/89916.

Bunke, Y. (2019). Productive and Reproductive Performances, Husbandry Practices and Associated Problems of Crossbred and Indigenous Dairy Cattle in Gamo Goffa Zone, SNNPR, Ethiopia. Ph.D. thesis.

CSA (2020). Volume II report on livestock and livestock characteristics (private peasant holdings) (Addis Ababa, Ethiopia: Central Statistical Agency (CSA)). url: https://www.statsethiopia.gov.et/wp-content/uploads/2021/05/2013.LIVESTOCK-REPORT.FINAL_.pdf.

DAGRIS (2007). Domestic Animal Genetic Resources Information System. International Livestock Research Institute. url: http://pigtrop.cirad.fr/content/pdf/4199. accessed date: 2016-11-17

Dawit, A. (1992). Prevalence of Bovine Trypanosomiasis in Gimera/Sheko and Zebu Cattle under Natural Condition in Gimera Awraja, Kaffa Administrative Region. Ph.D. thesis, Faculty of Veterinary Medicine, Addis Ababa University, Ethiopia.

Dessie, T. (2012). Biodiversity, resource base, animal breed level characterization, and utility of the information for goat genetic resources in Ethiopia.

Desta, T. T., Ayalew, W., and Hegde, B. P. (2011). Breed and trait preferences of Sheko cattle keepers in southwestern Ethiopia. *Tropical Animal Health and Production* 43(4), 851–856. doi: https://doi.org/10.1007/s11250-010-9772-2

Desta, T. T., Ayalew, W., and Hegde, P. B. (2012). Farmers perceptions on trypanosomosis and trypanotolerance character of the taurine Sheko. *Tropical Animal Health and Production* 44(3), 609–616. doi: https://doi.org/10.1007/s11250-011-9943-9

Elias, B. (2008). Sheko Cattle: Distribution, Management, and Performance in Bench-Maji zone of SNNPRs. url: http://etd.aau.edu.et/handle/123456789/21661.

Fasil, M. (2004). Prevalence of Trypanosomiasis among Domestic Animals in Bench Maji Zone, Mizan Teferi. Ph.D. thesis, Faculty of Veterinary Medicine, Addis Ababa University, Ethiopia.

Garg, M. R., Sherasia, P. L., Phondba, B. T., and Hossain, S. A. (2014). Effect of feeding a balanced ration on milk production, microbial nitrogen supply and methane emissions in field animals. *Animal Production Science* 54(10), 1657–166. doi: https://krishi.icar.gov.in/jspui/bitstream/123456789/12174/1/493-1-2796-2-1020180712.pdf

Hailu, A. (2020). Evaluation of husbandry practices and constraints for cattle fattening in Bench Maji zone, south west Ethiopia. *International Journal for Research in Agricultural and Food Science* 6(1), 32–48. url: https://gnpublication.org/index.php/afs/article/download/1176/851/.

Hanotte, O., Tawah, C. L., Bradley, D. G., Okomo, M., Verjee, Y., Ochieng, J., and Rege, J. E. O. (2000). Geographic distribution and frequency of a taurine Bos taurus and an indicine Bos indicus Y specific allele amongst sub-Saharan African cattle breeds. *Molecular Ecology* 9(4), 387–396. doi: https://doi.org/10.1046/j.1365-294X.2000.00858.x

Hegde, B. P. (2005). Cattle breeding strategy to improve milk production in Ethiopia. Alemaya University. Unpublished manuscript. url: https://hdl.handle.net/10568/3141.

Kitila, G., Kebede, B., Guta, D., Bekele, F., Wagari, M., Tilahun, B., and Tadesse, A. (2017). Epidemiological investigation of bovine trypanosomosis and its vector apparent densities in Yayo District Illuababora Zone, Western Oromia. *Ethiopia. Austin J. Vet. Sci. Anim. Husb* 4, 1–6. doi: https://doi.org/10.1016/j.parepi.2021.e00218

Lund, V. (2002). Ethics and animal welfare in organic animal husbandry an interdisciplinary approach. Ph.D. thesis, Swedish University of Agricultural Sciences.

Makkar, H. P. S. (2018). Feed demand landscape and implications of food-not feed strategy for food security and climate change. *Animal* 12(8), 1744–1754. doi: https://doi.org/10.1017/S175173111700324X

Matope, A., Zindove, T. J., Dhliwayo, M., and Chimonyo, M. (2020). Mitigating the effects of drought on cattle production in communal rangelands of Zimbabwe. *Tropical Animal Health and Production* 52(1), 321–330. doi: https://doi.org/10.3390/su131810023

Mekuriaw, G. and Kebede, A. (2015). A Review on Indigenous Cattle Genetic Resourses in Ethiopia: Adaptation, Status and Survival. *Online Journal of Animal and Feed Research* 5(5), 125–137. url: http://www.science-line.com/index/.

Mengistu, G., Yadessa, E., Tulu, D., Aleme, M., Bogale, A., and Effa, K. (2017). Survey on livestock production system characterization in Bench-Maji, Sheka and Mejenger Zones. *International Journal of Research in Agricultural Sciences* 4(5), 2348–3997. url: https://ijras.org/administrator/components/com_jresearch/files/publications/IJRAS_594_FINAL.pdf.

MOA (1984). Livestock Sub Sector Reviews. Main Report (Ethiopia: Ministry of Agriculture).

MOA (2013). Major challenges and achievements in Ethiopian livestock production. . Presentation. Ministry of Agriculture (MOA): Addis Ababa, Ethiopia. url: https://slideplayer.com/slide/7321828/.

Seyoum, Z., Terefe, G., and Ashenafi, H. (2013). Farmers' perception of impacts of bovine trypanosomosis and tsetse fly in selected districts in Baro-Akobo and Gojeb river basins, Southwestern Ethiopia. *BMC veterinary research* 9(1), 1–9. doi: https://doi.org/10.1177/11786302221127266

Stein, J., Ayalew, W., Rege, E., Mulatu, W., Lemecha, H., Tadesse, Y., Tekle, T., and Philipsson, J. (2011). Trypanosomosis and phenotypic features of four indigenous cattle breeds in an Ethiopian field study. *Veterinary Parasitology* 178(1-2), 40–47. doi: https://doi.org/10.1016/j.vetpar.2010.12.025

Takele, T. (2005). On-farm Phenotypic Characterization of Sheko Breed of Cattle and Their Habitat in Bench Maji Zone. url: https://www.researchgate.net/publication/269763438.

Takele, T. and Habtamu, L. (2009). Traditional Backyard Cattle Fattening in Wolayta: Systems of Operation and the Routine Husbandry Practices. *Ethiopian Journal of Animal Production* 9(1), 39–56.

Takele, T. and Workneh, A. (2011). Hegde Breed and trait preferences of Sheko cattle keepers in southwestern Ethiopia. *Tropical Animal Health and Production* (43), 851–856. url: https://www.researchgate.net/publication/237100111.

Taye, T., Ayalew, W., and Hegde, B. P. (2009). Status of Ethiopian indigenous Sheko cattle breed and the need for participatory breed management plan. *Ethiopian Journal of Animal Production* 9(1). url: https://www.researchgate.net/publication/237100090.

Terefe, E., Dessie, T., Haile, A., Mulatu, W., and Mwai, O. (2012). Husbandry and breeding practices of cattle in Mursi and Bodi pastoral communities in Southwest Ethiopia. *African Journal of Agricultural Research* 7(45), 5986–5994. doi: https://doi.org/10.5897/AJAR12.1566

Tolera, A., Yami, A., and Alemu, D. (2012). Livestock feed resources in Ethiopia: Challenges, Opportunities and the need for transformation. url: https://www.researchgate.net/publication/312692080.

Walli, T. K., Garg, M. R., and Makkar, H. P. (2012). Crop residue based densified total mixed ration. FAO Animal Production and Health Paper n. 172. url: https://www.fao.org/3/i2728e/i2728e00.pdf.

Workeneh, A. (2001). Revised field report on a survey of a sample of the Sheko cattle maintained at the former Tolley military Training Center in Oromiya Regional State, Ethiopia. Oromiya Agricultural Development Bureau, ILRI, Addis Ababa, Ethiopia.

World Bank (2007). World development report 2008: Agriculture for development. url: https://digitallibrary.un.org/record/1305297?ln=en.

Zander, K. K., Drucker, A. G., and Holm-Müller, K. (2009). Costing the conservation of animal genetic resources: The case of Borana cattle in Ethiopia and Kenya. *Journal of arid environments* 73(4-5), 550–556. doi: https://doi.org/10.1371/journal.pone.0244836

Genetic Resources

ORIGINAL ARTICLE

Genetic Resources (2023), 4 (8), 71–90
DOI: 10.46265/genresj.DVKV8430
https://www.genresj.org
ISSN: 2708-3764

Status and prospects of plant genetic resource conservation in Yemen

Maeen Ali Aljarmouzi [*,a], **Khalil M Alsharjabi** [a] **and Luigi Guarino** [b]

[a] *Agricultural Research and Extension Authority (AREA), Dhamar, Republic of Yemen*
[b] *Global Crop Diversity Trust, Platz der Vereinten Nationen 7, 53113, Bonn, Germany*

Abstract: This study provides a comprehensive examination of the conservation status of plant genetic resources in Yemen, with a particular focus on the national genebank, natural sites and field genebanks. Employing a historical and descriptive approach, the study sheds light on the inception of field inventories and collection surveys, the roles and contributions of various projects and institutions since the late 1980s, and the quantities of seed samples collected from different crop varieties and species. Additionally, the study highlights the challenges faced, especially during the last decade following the eruption of war in early 2015. Specifically, it emphasizes the detrimental impact of the war and sanctions, resulting in the loss of conserved seed samples, and the damage and sabotage of field genebanks. In contrast, there is a need to ensure the enhancement of functioning seed systems and agriculture production even during times of conflict to reduce the impact of food insecurity. To conclude, the study puts forth several proposals, with a strong emphasis on expanding conservation efforts beyond natural sites, enhancing the capabilities for seed conservation in genebanks, and building genetic resources capacity.

Keywords: Agricultural research, Conservation, Genebank, Field genebanks, Natural sites, Yemen

Citation: Aljarmouzi, M. A., Alsharjabi, K. M., Guarino, L. (2023). Status and prospects of plant genetic resource conservation in Yemen. *Genetic Resources* 4 (8), 71–90. doi: 10.46265/genresj.DVKV8430.

Introduction

For thousands of years, farmers in Yemen have inherited and managed their local seeds. In particular, the sorghum crop shows a remarkable level of diversity, serving as a clear testament to their commitment. They have cultivated, conserved, exchanged and managed the seeds of agricultural crops since the early stages of history when dams, irrigation canals and highland terraces were developed across different regions during the eras of ancient Yemeni kingdoms and states dating back to between the tenth and fifth centuries BCE (Alafif Cultural Foundation, 2003). However, the systematic conservation, characterization and use of plant genetic resources in a modern scientific manner can be traced to the establishment of agricultural research work in Yemen during the mid- to late-1940s.

The first agricultural research station was established in the El-Kod district of Abyan governorate, which later evolved into the fully-fledged El-Kod Agricultural Research Center (EARC) in 1955. The first systematic activities related to the conservation and use of plant genetic resources in Yemen took place in this centre in the early 1970s. Since then, various Yemeni government bodies, including the Ministry of Agriculture and Irrigation (MAI), the Ministry of Water and Environment (MWE), as well as institutions like the Agricultural Research and Extension Authority (AREA) and the Environment Protection Authority (EPA), have made diverse contributions to the conservation and sustainable use of genetic resources. These bodies have been actively involved in supporting the implementation of numerous national programmes, projects and activities in collaboration with local, regional and international organizations.

In recent decades, there has been an increasing interest in plant genetic resources due to their recognition as a national asset and a sovereign resource for any country worldwide. These resources play a vital role in achieving food security, sustainable development, resilience, better livelihoods and higher income. Yemen,

---

[*] Corresponding author: Maeen Ali Aljarmouzi (maeen669@gmail.com)

Received: 09.07.2023 | Accepted: 03.11.2023 | Published online: 20.12.2023

characterized by diverse climates and a rich and distinct plant diversity, unfortunately, has not made optimal use of its traditional plant genetic resources. On the contrary, these resources have been diminishing and lost due to various factors, including human interventions such as urban expansion, road construction at the expense of agricultural lands, and the expansion of qat (*Catha edulis*) cultivation, an important cash crop in Yemeni social life but with negative effects on human nutrition and health, as well as the introduction of new crops.

The deterioration and loss of plant genetic resources can in addition be attributed to factors such as the absence of legislation, weak institutional frameworks, limited national programmes and inadequate material and human capacities to leverage modern technologies for conservation and sustainable use.

Since 2015, the ongoing war and conflict have also significantly impacted Yemen's plant genetic resources, resulting in severe economic and social consequences. The collapse of state institutions, the paralysis of public service agencies and facilities, the internal and external displacement of millions of citizens, currency devaluation, limited availability of goods and services, rising prices, declining investment and the loss of income sources for many Yemeni families all exert immense pressure on natural resources, including the country's biodiversity in all its forms.

In recent years, amid worsening conditions and the effects of the war and recurring food crises, official authorities and decision-makers, prompted by cases of tampering with plant genetic resources, have called for more attention to be given to the country's biodiversity. The need for implementing relevant research and raising awareness about these issues has become increasingly urgent. The present study was undertaken in response to these calls. The request made by international bodies and organizations such as the Commission on Genetic Resources for Food and Agriculture (CGRFA) and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), among others, for member states to contribute to monitoring and assessing the state of plant genetic resources, served as further justification for conducting this study.

Materials and methods

To assess the status of plant genetic resource conservation in Yemen, the authors employed a historical descriptive approach. This involved gathering and reviewing national and international documents, reports, studies and surveys related to plant genetic resources. The authors also utilized the available electronic and paper information systems, including databases and records, of the Genebank of the National Genetic Resources Center (NGRC) within the Agricultural Research and Extension Authority (AREA) in Dhamar, Yemen. Furthermore, various institutions contacted during the study provided additional information that was incorporated into the research.

Additionally, the authors relied on their personal knowledge as co-witnesses, drawing from their extensive experience in research and agricultural technologies, as well as their involvement in various activities related to plant genetic resources. They also considered events and developments they have witnessed or been exposed to up until the time of preparing this study.

It is worth noting that the authors followed the methodology established by the Commission on Genetic Resources for Food and Agriculture of the Food and Agriculture Organization of the UN (FAO), as well as the approved indicators, to guide the monitoring of the implementation of the Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture (Second GPA). A total of 55 indicators from this guide were used in the study to assess the state of conservation of plant genetic resources in Yemen from 2012 to 2022 (FAO, 2019). The assessment was conducted based on the themes and activities outlined in the Second GPA, as well as the medium-term plan of AREA and the NGRC's plan.

The execution of this study encountered numerous difficulties and obstacles, with the most significant ones being:

- Limited availability of data: The study faced challenges due to the scarcity of available data. The existing information was dispersed across multiple sources and exhibited a diversity of forms and formats, making it challenging to gather comprehensive and standardized data.
- Low response rate from contacted agencies: Despite efforts to collect relevant and available data, the study encountered a low response rate from most of the local agencies contacted. This lack of cooperation hindered the acquisition of crucial information necessary for the research.
- Impact of war: The study was conducted amidst the ongoing war, which created adverse conditions and posed additional challenges. Coordinating and communicating with relevant entities and individuals became difficult due to the disrupted infrastructure and limited resources. Obtaining the requested information became a daunting task in such a context.

These difficulties and obstacles significantly influenced the implementation of the study, impacting the comprehensiveness and accuracy of the collected data. Nonetheless, the authors made their best efforts to overcome these challenges and ensure the reliability of the findings within the existing constraints.

Table 1. Studies surveying natural vegetation in the different Yemeni regions and numbers of species identified.

Study Title	No. of species	Source
Survey and classification of medicinal and aromatic plants in Sana'a and Dhamar governorates	858	NGRC (2022)
Survey and classification of the forest-pastoral vegetation cover in the Al-Jawf governorates	93	RNRRC (2022)
Inventory of the natural plant species in important areas in the Arabian Peninsula: Bani Omar, Taiz governorate, Yemen	135	Al-Khulaidi *et al* (2000)
Survey of the natural flora of some districts in Ibb, Taiz, Al-Mahweet and Sana'a governorates	1,603	AREA (2015)
Studying the existing natural resources and the climate change dimension in Taiz governorate using remote sensing and geographic information systems techniques (second phase)	389	AREA (2014)
Surveying, collecting and classifying the vegetation cover in Bura'a	240	AREA (2014)
Survey of the natural plants in Al-Makha district	167	NGRC (2013)
Survey of the natural plants in Ibb governorate	316	Al-Khulaidi (2013)
Surveying and collecting plant germplasm of pasture and forest, medicinal and aromatic plants, and the unexploited plants in the highlands region of Ba'adan and Al-Sher'ar districts, Ibb governorate	219	NGRC (2011)
Survey and evaluation of the deterioration of the vegetative cover in the upper stream area of Wadi Rasyan	332	Mufarreh *et al* (2011)
Surveying and collecting unexploited plant germplasm of pastures, forests, and medicinal and aromatic plants in the highland region of Sana'a, Al-Mahweet Amran Hajjah governorates	1,329	AREA (2013)
Surveying and collecting unexploited plant germplasm of pastures, forests, and medicinal and aromatic plants in Mareb governorate	294	NGRC (2009)
Surveying and collecting unexploited plant germplasm of pastures, forests, and medicinal and aromatic plants in 14 districts of the coastal plain of Hadramout and Al-Mahrah governorates	172	NGRC (2008)
Inventory and evaluation of the pastural genetic resources in Al-Kusumah district of Raymah governorate	289	Mufarreh (2004)
Inventory and evaluation of the pastural genetic resources in Belad Al-Ta'am district of Raymah governorate	311	Mufarreh (2003)
Survey of the natural vegetative cover of Madinat Al-Sharq – Wadi Rama'a	78	Al-Khulaidi (1994)

Results and discussion

State of *in situ* conservation and on-farm management

Surveys and inventories of plant genetic resources

Understanding crop diversity, including its distribution and changes over time, is a crucial prerequisite for the development and implementation of effective and efficient conservation strategies. Various studies have been conducted to survey and inventory plant diversity across different regions in Yemen by agencies such as the Agricultural Research and Extension Authority (AREA), colleges of agriculture and sciences in Yemeni universities, and institutions like the Social Fund for Development (SFD).

Al-Khulaidi (2013) reported the identification of 2,838 plant species, as part of the Yemeni flora, of which 2,602 are native, 129 are cultivated and 107 are introduced. Among these species, 608 were found to be endemic or semi-endemic (457 are endemic, among which 307 on Socotra Island alone), confined to Yemen or the Arabian Peninsula. The diverse topography of Yemen's plains, mountains, plateaus, valleys and deserts, with varying altitudes, has contributed to the emergence of this unique plant diversity in the country.

The uniqueness of such diversity extends to the varied crops grown in the country, which includes both tropical and temperate crops. Additionally, there is a diversity within species, with multiple crop varieties adapted to different environments or agroecological zones in the country. For instance, sorghum is characterized by a great diversity of types that mature over a range of durations (after three, four, five or even six months).

Table 1 presents the results of some recent studies and surveys, which involved counting and classifying plant species, compiling lists with their scientific and local names, providing photographs, and mapping their distribution in specific areas. These studies also identified the uses of the plant species, such as food, fodder, medicine and other purposes. Additionally, they focused on identifying threatened, rare, endemic and semi-endemic plants, some of which are introduced plants.

On-farm management and improvement of plant genetic resources for food and agriculture

Plant diversity plays a vital role in food security and adaptation to climate change. Therefore, it is crucial to focus on on-farm management and improvement of local crop varieties, including neglected and underutilized crops. By doing so, the resilience and adaptability of the cropping systems can be enhanced, enabling better withstanding of biotic and abiotic stresses and eventually better production and food security for the farmers' households.

NGRC has undertaken various activities to support the management and improvement of genetic resources on farm through multiple projects (Table 2). These initiatives primarily involved the characterization, evaluation and dissemination of local varieties suitable for different agroclimatic zones. The germplasm collected by various projects has been safety duplicated at the national genebank. The aim was to increase genetic diversity and broaden the range of options available to farmers for selecting varieties that are well-adapted and tolerant to biotic and abiotic stresses prevalent in their specific local conditions.

Providing seeds to farmers and researchers in case of disasters to restore cropping systems

NGRC is ideally positioned to play an important role in preserving plant genetic resources, especially in the face of natural disasters (such as droughts and floods) and man-made disasters (like conflicts and wars). NGRC's mandate includes providing high-quality seeds adapted to farmers' environmental conditions in disaster-affected areas, in coordination with relevant National authorities and organizations like regional agricultural research stations, the General Improved Seed Multiplication Corporation (GISMC), the Public Corporation for Grain Development and Production (PCGDP), and provincial MAI offices. However, due to limited material and human resources, NGRC's efforts are constrained, particularly in the current challenging conditions within the country.

In 2016, the regional station for agricultural research in the western coastal plain (Al-Kadan, Sardoud, Tihama region) was severely damaged by airstrikes. The seed stores in the station were significantly affected, leading to spoilage and loss of different varieties that had been collected and preserved by the station's researchers over many years (Figure 1).

Nonetheless, NGRC has managed to provide services to farmers and researchers in conflict and war-prone zones in some instances. This has been achieved through quality testing of locally distributed seeds and by providing seeds of important local strains to research stations in the respective agricultural regions.

For example, NGRC collaborated with other organizations and international agencies operating in Yemen to provide seeds of cereal, legumes and vegetable crops to affected farmers as part of humanitarian relief activities. Organizations such as FAO, the Red Cross, Acted and others have contributed to seed distribution efforts. However, the distribution mechanism faced drawbacks, including poor coordination with official agencies, poor seed quality and sometimes unsuitable seeds for the targeted agricultural climatic zones. These issues created a negative impression among farmers and decision-makers regarding the effectiveness and safety of the distribution mechanism.

In response, measures were established to mitigate these negative effects. These measures included conducting seed quality tests in laboratories, seed screening and purification by GISMC, and ensuring direct coordination with government authorities and institutions in the targeted areas.

Table 2. Projects implemented by the National Genetic Resources Center (NGRC) that have contributed to enhancing the on-farm conservation of genetic resources in Yemen. AREA, Agricultural Research and Extension Authority; BSF, the Benefit-Sharing Fund; FAO, Food and Agriculture Organization of the UN; GEF, Global Environment Facility; ICARDA, International Center for Agricultural Research in the Dry Areas; IFAD, International Fund for Agricultural Development; MAI, Ministry of Agriculture and Irrigation; WB, World Bank.

Project name	**Funding agency**	**Implementing agencies**	**Implementation period**	**Implementation site**	**Activity**
Participation of the rural community in raising crops (barley and lentils) in the mountain terraces (Ceccarelli, 2002)	ICARDA	AREA/Farmers' groups in the selected sites, Agriculture office	1998–2010	Kuhlan Affar, Hajjah governorate	Evaluation of strains of barley and lentil on farmers' fields. An economic and social study for the farmers of the selected villages; study and analysis of gender (the role of the farm household, male and females).
The rainfed agriculture and livestock Project (YRALP, 2005)	WB	AREA/GRC-Sana'a University, GCISM	2006–2013	Sana'a, Al-Mahweet, Hajjah, Al-Hodeidah and Lahj governorates	Farmer-based seeds improvement and management system. Evaluation of several local genotypes of sorghum, millet and cowpea
The participatory rural development of the Raimah governorate (IFAD, 2010)	IFAD	AREA and Dhamar MAI office	2001–2007	Raimah governorate	Evaluation and dissemination of adapted cereal and legume cultivars
Irrigation improvement project (IIP, 2009)	WB	AREA, Agricultural Services Corporation, Tehama Development Authority and Nasser's Faculty of Agriculture Sciences (Lahj)	2005–2007	Al-Hodeidah, Lahj, and Abyan governorates	Evaluation and dissemination of adapted cultivars of cereals, cotton and sesame
The participatory rural development of Al-Mahrah governorate (IFAD, 2011)	IFAD	AREA, MAI office	2002–2006	Al-Mahrah governorate	Evaluation and dissemination of adapted maize and sesame cultivars
The participatory rural development of Dhamar governorate (IFAD, 2012)	IFAD	AREA, MAI office	2006–2010	Dhamar governorate	Evaluation and dissemination of adapted cereal and legume cultivars
Improvement of the medicinal, and aromatic plants and underutilized crops (Gotor and Cherfas, 2012)	FAO	NGRC, MAI offices	2009	Sana'a, Dhamar and Lahj governorates	Cultivation of neglected and underutilized crops on farmers' fields

Continued on next page

Table 2 continued

Project name	**Funding agency**	**Implementing agencies**	**Implementation period**	**Implementation site**	**Activity**
Enhancing food security in the Arab countries (Yemen) (Communication Team ICARDA, 2018)	ICARDA	AREA and Dhamar MAI office	2012–2019	Dhamar	Dissemination of wheat, lentils and peas varieties on farmers' fields
Agricultural biodiversity and adaptation to climatic changes (YACCAP, 2010)	GEF	AREA, Sana'a College of Agriculture, and the GCISM	2014	Sana'a, Taiz, Ibb and Al-Mahweet governorates	Evaluation of local sorghum strains on farmers' fields
Participatory conservation and sustainable use of landraces to improve farmer's livelihood and their resilience in adapting to climate changes in Yemen (FAO, 2023)	BSF	AREA and MAI offices	2019–2023	Dhamar and Hadramout governorates	Evaluation and characterization of local strains of sorghum, wheat, barley, lentils and peas on farmers' fields

Unfortunately, there have been instances where poor-quality seeds have been distributed. In April 2021, MAI destroyed a large quantity of rotten seeds that were provided by FAO for distribution to farmers in Al-Hodeidah governorate. The spoiled quantity was estimated to be more than 240 tonnes. These seeds, intended as donated assistance to Yemeni farmers, were infected with fungi and deemed unsuitable for cultivation.

This incident highlights the importance of strict quality control measures and coordination among relevant authorities and organizations to ensure the effective and safe distribution of seeds in Yemen in line with the recommendation of the International Plant Protection Convention (IPPC) on "ensuring safe provision of seed during humanitarian assistance disbursement" (IPPC, 2021).

Changes and trends in on-farm genetic resources conservation

Yemen is known as one of the main countries that are growing qat. Qat or khat or khata (*Catha edulis*) is a mild stimulant plant consumed by most Yemenis in a daily afternoon session by chewing its fresh leaves (Zahran *et al*, 2019). The expansion of qat cultivation at the expense of the main cereals and legumes (sorghum, millet, wheat, barley, lentil and pea) poses a significant threat to crop diversity in Yemen. The yearly agricultural statistics book indicates a decline in the cultivated area of cereal crops over the years, while qat cultivation has increased (GDSID/MAI, 2022). However, the actual increase in qat cultivation may be even greater than what is reflected in the official figures (Figure 2). Due to the war and resulting consequences, including the institutional divide and ineffectiveness, no agricultural census has been conducted for the past two decades.

The urban population expansion on agricultural lands is another factor threatening crop diversity. Figure 3 shows the trend in population growth over the past 70 years.

A study conducted by the Renewable Natural Resources Research Center (RNRRC) in 2021 on the urban expansion in Dhamar governorate revealed a dramatic increase over the past 50 years.

Reports from field surveys conducted by NGRC and RNRRC highlight the expansion of urban centres and population at the expense of agricultural lands (Figure 4). This phenomenon is considered one of the most critical factors endangering agricultural biodiversity.

The disappearance of certain plant species and fruit varieties (quince, walnuts, pear, peach, apricot, fig and grapes) from areas like Sana'a governorate, which used to be known for cultivating them, further emphasizes the negative impact of urban expansion on agricultural biodiversity. Parks and orchards (for example of walnuts, peach and grapes) have been replaced by commercial and residential areas, resulting in a significant decrease in green coverage.

On the other hand, there has been an expansion in the cultivation of some vegetable crops (tomato, onion, potato, pepper, carrot, cucumber, zucchini, cabbage), including varieties introduced from abroad. Although the cultivated area of vegetable crops fluctuates, the number of greenhouses in different regions of Yemen has increased. Strawberry cultivation has also grown in response to high demand from fresh juice shops in major cities.

The cultivation of fruit crops, particularly almonds, has expanded due to relatively low water requirements and the profitability of production. The almond cultivation area has increased between 2011 and 2020, with significant expansion in Sana'a governorate, particularly in the districts of Al-Haima and Bani Matar.

Coffee cultivation has also expanded due to increasing demand for Yemeni coffee locally and internationally. The cultivation area has increased over the past few years much more than the estimated figures made in the agricultural census book for 2020, which indicated an increase of 2,000 hectares in the coffee area between 2016 (33,900ha) and 2020 (35,900ha). Also, there has been a focus on improving agricultural processes and post-harvest operations to meet desired quality standards. High government authorities and MAI have shown increased interest in the coffee sector, establishing the National Center for Coffee Research and supporting farmers' associations. A decree was issued in 2022 banning the import of both processed and unprocessed coffee from abroad to encourage domestic cultivation and increase the income of coffee farmers.

Overall, while there is an expansion in certain crops, the threats to crop diversity posed by the expansion of qat cultivation, urbanization and population activities remain significant challenges in Yemen.

Status of *ex situ* conservation of plant genetic resources (genebanks)

The NGRC in Yemen operates under the umbrella of AREA and has its headquarters outside Dhamar city, about 100km from the Yemeni capital Sana'a. In addition, there are two other genetic resource centres working under the faculties of agriculture sciences of Sana'a and Aden universities and some field genebanks as detailed in Table 5.

Supporting the targeted collection of plant genetic resources

The primary drivers for implementing targeted collecting of genetic resources were the risk of loss of on-farm diversity, opportunities for use, and the need to address deficiencies in *ex situ* conservation. Bawazir (2004) conducted a study on cereal diversity in southern Yemen and emphasized the importance of surveying and collecting genetic resources of cereal crops and their wild relatives across different agroclimatic zones in Yemen. The study aimed to document these resources and utilize them in breeding and genetic improvement programmes.

The study's findings revealed that sorghum varieties in these areas exhibited genetic variation within each environmental zone, as evidenced by differences in

Figure 1. Examples of the effects of the air raids of the war coalition against Yemen on the buildings, equipment and seed stores of the Tehama regional agricultural research station in the western coastal plain, Al-Kadan, Sardoud, Tihama zone, Al-Hodeidah governorate (A and C), including damage to the preserved seed samples of various crops (B).

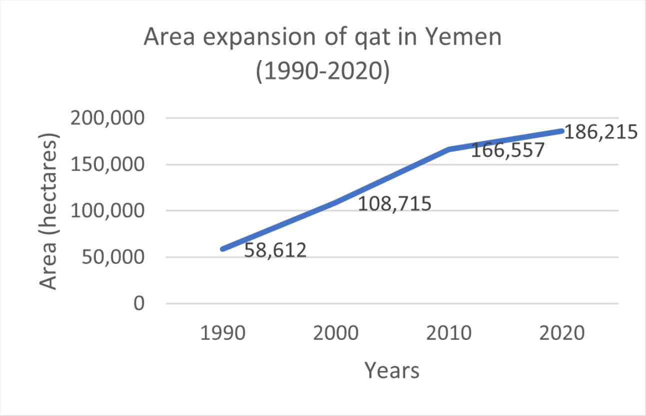


Figure 2. Expansion of qat cultivated area from 1990 to 2020. Source: CSO (2020)

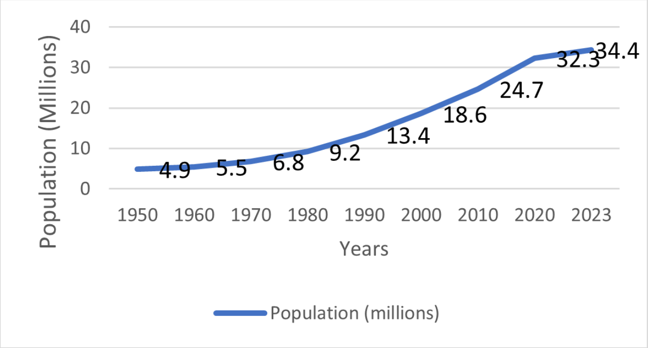


Figure 3. Trend of population growth in Yemen over the past 70 years. Source: Macrotrends (2023)

phenotypic characteristics, anatomical features and resistance to environmental stress conditions. The researcher pointed out that the varieties of sorghum grown in coastal areas showed early maturity (100–120 days), while the varieties grown in the highland areas matured late (160–180 days). According to the study, sorghum varieties in the highland areas were tallest (300cm) while in the coastal areas plant height ranged from 250–300cm.

The study also demonstrated that local wheat varieties possessed drought tolerance. This may be accounted for by their morphological and anatomical traits, such as the smaller number of seminal roots and the small size of seminal root vessels. Most of the local varieties of wheat grown in the southern, medium-altitude regions of Yemen belong to *Triticum aestivum*. In some areas, there may be a mixture of *T. aestivum* and *T. durum*.

The local varieties of millet grown in the southern Yemeni regions belong to *Pennisetum glaucum* (pearl

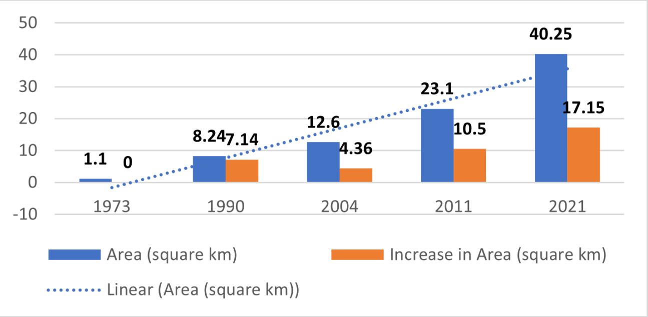


Figure 4. Urban sprawl on fertile agricultural lands in Dhamar City (km^2) (1973–2021)

millet), *Pennisetum setaceum*, *Pennisetum rigidum* (little millet) and *Eleusine coracana* (finger millet). Misibli is the local name for pearl millet throughout Yemen, Kanab is the local name for finger millet, and Heba is the local name for small (short) millet (and an early-maturing type) on Socotra Island.

The efforts made in collecting and conserving plant genetic resources in Yemen can be divided into three phases:

1) First phase (1970s and 1980s): During this period, international organizations collaborated with researchers from the El-Kod Research Station in Abyan governorate and the Southern Highlands Regional Agricultural Research Station (SHRAR) in Osaiferah, Taiz governorate to collect hundreds of samples. Notable collecting missions during this period in the different Yemeni regions are presented in Table 3.

Unfortunately, all the samples collected during this period were lost due to inadequate storage conditions, except for those preserved outside Yemen by supporting organizations, such as FAO, the International Plant Genetic Resources Insitute (IPGRI), the United States Agency for International Development (USAID), the International Maize and Wheat Improvement Center (CIMMYT) and others (Al-Mua'alem *et al*, 1993).

2) Second phase (1990s): During this period, a specialized unit for plant genetic resources was established under RNRRC of AREA in Dhamar. Over 2,000 plant samples were collected during this phase. Notably, a joint team from AREA and the International Center for Crop Research in the Semi-Arid and Arid Tropical Areas (ICRISAT) conducted significant collecting trips. They collected 685 seed samples, mostly of sorghum and millets, from Al-Dhalea, Radfan, Yafe'a, Abyan, Lahj and Tehama. Details on this collection trip and other trips are presented in Table 3.

Despite having a refrigeration room for cold storage (-18°C), frequent power shortages and blackouts led to the loss of many collected samples, despite their data being recorded in NGRC's records.

3) Third phase (2000s to present): This stage witnessed the establishment of NGRC in 2002, an increase in specialized staff and the establishment of various departments within the centre. FAO stated that NGRC in Dhamar held 3,281 samples in its genebank (FAO, 2009). The number of samples collected from different regions of the country has increased to over 6,500 accessions at present. Notable collecting operations, covering most Yemeni governorates, conducted by NGRC during this phase are displayed in Table 3.

The collecting trips conducted by NGRC between 1989 and 2013 to collect seeds from natural and cultivated plants, targeting all geographical regions in Yemen are shown in Figure 5.

Most of the collection activities previously reviewed were carried out with support and funding within bilateral or multilateral cooperation or joint work projects between Yemen and several partners over the previous decades, and copies of samples of genetic resources that were collected found their way abroad and have been preserved by institutions outside Yemen. Early reports indicate that over 8,000 samples were deposited in various genebanks abroad (Al-Ghouri *et al*, 1996). The Second National Report on the State of Plant Genetic Resources for Food and Agriculture stated that international organizations held 8,619 samples from Yemen (FAO, 2009). Currently, data on the Genesys platform show 8,958 samples from Yemen conserved in 19 genebanks, covering 44 plant species, with a focus on grain crops (Genesys, 2021). Data from FAO WIEWS and the Svalbard platform also align closely with these figures (Table 4).

Expanding diversity in genebanks

It is crucial to expand *ex situ* conservation efforts for neglected crops, crop wild relatives and forages to facilitate research and crop improvement. The genebank of NGRC currently conserves seeds from 45 different crops. These crops encompass cereals such as sorghum, maize, millet, wheat and barley; legumes including lentils, beans, kidney beans, cowpeas, peas and fenugreek; vegetables like onions, tomatoes, chilli peppers, cucumbers, zucchini, mallow, radishes, eggplant and okra; oil-producing crops like sesame and peanuts; less commonly utilized crops, such as black seed (*Nigella sativa*), henna (*Lawsonia inermis*), coriander (*Coriandrum sativum*), caladium, arugula (*Eruca sativa*) and mustard (*Brassica juncea* L.).

Additionally, field genebanks situated in various regions hold a diverse range of crops, including date palms, mangoes, citrus fruits, almonds, grapes, bananas and papayas, along with forage and forest crops. There are eight field genebanks in various regions – the northern, central and southern highlands, the Tihama coastal plains region, the southern coast region (Al-Kud, Abyan) and the Eastern Plateau region (Marib and Sayun). More details are provided in Table 5.

It is important to note that the acquisition of new species and samples to expand the genetic resources reserve at NGRC was temporarily halted between 2014 and 2018 due to the prevailing conditions of the war and the resultant socioeconomic deterioration and institutional ineffectiveness in the country.

Table 3. Details of the main phases of plant genetic resource collection and conservation in Yemen

Main Genus/Species	No. of accessions	Location(s)	Year	Project/programme	Reference
Phase 1: 1970s–1980s					
Sorghum, millet and Sudan grass	4,500	Northern Yemeni Governorates	1975–1977	The American Sorghum Improvement Assistance Project	Hakimi and Ya'ni (2008); Jaradat (1997)
Wheat, barley and some legumes	490		1978–1979	The German Technical Cooperation Agency (GTZ)	
Field crops (Sorghum, millet, wheat, barley and some legumes)	783		1980–1981	The International Board for Plant Genetic Resources (IBPGR)	
30 crops of human and animal food	351	Southern and Eastern Yemeni Governorates	1969–1989	IBPGR	Hakimi and Ya'ni (2008); Guarino (1989)
Phase 2: 1990s					
70 plant species	617	Tehama, central and south highlands, and east regions	1990–1999	AREA programme	Jaradat (1997)
Sorghum and millets	685	Al-Dhalea, Radfan, Yafe'a, Abyan, Lahj, and Tehama.	1992	Sorghum and Millet Improvement Project	
Sorghum and millet	294	Saada, Taiz, Sana'a, Al-Dhalea, Ibb and Socotra	1992	Sorghum and Millet Improvement Project	Amer and Al-Dahmashi (1997)

Continued on next page

Table 3 continued

Main Genus/Species	No. of accessions	Location(s)	Year	Project/programme	Reference
Phase 3: 2000–present time					
Sorghum bicolor, *Pennisetum americanum*, *Zea mays*, *Vigna* spp., *Cajanus cajan*, *Phaseolus vulgaris*, *Trigonella foenum-graecum*, *Sesamum indicum*, *Raphanus* spp. *Corchorus olitorius*, *Abelmoschus esculentus*, *Eruca sativa*, *Cyamopsis tetragonoloba*, *Cucumis* spp, *Nicotiana* spp., *Capsicum annuum*, *Gossypium* spp.	629	Sana'a, Hajjah, Al-Mahweet, Al-Hudiedah and Lahj	2007	The Rainfed Agricultural Project	NGRC (2008)
36 species of human and animal food	390	Western and Eastern Coastal Plain	2008	AREA research programme	NGRC (2009)
Sorghum bicolor, *Zea mays*, *Pennisetum americanum*, *Triticum aestivum*, *Hordeum vulgare*, *Lens culinaris*, *Trigonella foenum-graecum*, *Vicia faba*, *Pisum sativum*, *Vigna* spp., *Lablab vulgaris*, *Phaseolus vulgaris*, *Trifolium* spp., *Abelmoschus esculentus*, *Corchorus olitorius*, *Capsicum annuum*, *Sesamum indicum*, *Brassica napus*, *Linum usitatissimum*, *Coffea arabica*	330	Sana'a, Amran, Ibb, Dhamar, Hajjah, Al-Mahweet and Al-Hodeidah	2009	AREA research programme	NGRC (2010)
Fruit crops (*Vitis vinifera*, *Prunus amygdalus*, *Prunus persica*, *Cydonia oblonga*, *Ficus carica*, *Malus sylvestris*, *Morus alba*, *Pyrus* spp., *Olea* spp., *Ceratonia siliqua*, *Mangifera indica*, *Citrus* spp., *Manilkara achras* (Mill), *Musa* spp., *Psidium guajava* L., *Carica papaya* L., *Passiflora edulis* Sims, *Annona* spp., *Opuntia ficus-indica*, *Phoenix dactylifera*)	150	Sana'a, Amran, Al-Mahweet and Hajjah	2009	AREA research programme	NGRC (2010)
Honeydew pumpkin (*Cucurbita* spp.)	70	Belad Alrous District, Sana'a governorate	2010	AREA research programme	NGRC (2011)
Sorghum bicolor (30), *Pennisetum americanum* (2) *Zea mays* (14), *Triticum durum* (1), *Hordeum vulgare* (2), Faba bean (1), *Vigna* spp. (14), *Sesamum indicum* (1), *Trigonella foenum-graecum* (1), *Phaseolus* spp. (1), *Cajanus cajan* (5)	72	Al-Salafiya Directorate and Bilad Al-Ta'am in Raymah governorate	2012	AREA research programme	NGRC (2013)

Continued on next page

Table 3 continued

Main Genus/Species	No. of accessions	Location(s)	Year	Project/programme	Reference
Sorghum bicolor (33), *Zea mays* (7), *Triticum durum* (26), *Hordeum vulgare* (16), *Lens culenaris* (11), *Vigna* spp. (11), *Pisum sativum* (6), *Trigonella foenum-graecum* (9), *Brassica* spp. (1), *Linum usitatissimum* (3), *Allium* spp. (1), *Cucumis sativum* (1), *Phaseolus* spp. (1)	126	Sabah district in Al-Bayda governorate	2012	AREA research programme	NGRC (2013)
Sorghum bicolor (33), *Zea mays* (7), *Triticum durum* (32), *Hordeum vulgare* (20), *Lens culenaris* (4), *Vigna* spp. (3), *Pisum sativum* (8), *Trigonella toeniccum* (1)	108	Ans, Jahran and Al-Hada'a in Dhamar governorate	2012	AREA research programme	NGRC (2013)
Cereals, vegetables and fruits	282	Sana'a, Al Mahwit, Ibb and Taiz governorates	2014	The Agricultural Biodiversity and Climate Adaptation Project	NGRC (2013)
The project team collected a total of 368 landraces (228 from Dhamar and 140 from Hadramout and Almahrah governorates) and more than 20 species: sorghum, maize, wheat, barley, millet, cowpea, pea, bean, fava bean, lentils, fenugreek (*Trigonella*, spp.), sesame, *Lablab purpureus*, mustard (*Brassica* spp.), flax (*Linum* spp.), pepper (*Capsicum annum*), *Eleusine* spp., *Pennisetum* spp., Roselle (*Hibiscus sabdariffa*), *Plectranthus* spp., black caraway (*Nigella sativa* L.), fennel (*Foeniculum vulgare*)	383	Several districts in the governorates of Dhamar, Hadramout and Al-Mahra governorates	2019	The conservation and sustainable use of local landraces project - BSF	NGRC (2022)

Efforts have also been made to conserve genetic resources of threatened and endemic species *ex situ* in various research stations located in different agricultural regions. Notably, a specialized team from the El-Kod research station in Abyan governorate collected the Socotra wild pomegranate (*Punica protopunica* Balf. f.) in 1989 and 1990. This threatened species, which is endemic to the Island, has been confirmed by international organizations such as the International Union for Conservation of Nature (IUCN) to be one of the species at risk of overexploitation without natural regeneration. The collected seeds were grown, and the trees were nurtured at the El-Kod research farm (Bazara'a, 2000).

Regeneration and multiplication of genebank seed samples

Even under optimal storage conditions, it is necessary to periodically regenerate and multiply seed accessions due to the decline in viability over time and the limited quantity of preserved seeds resulting from their distribution to users such as researchers and farmers.

Hence, NGRC carries out annual processes to regenerate seed accessions that are at risk of viability loss and to multiply accessions with limited quantities. The most recent regeneration and multiplication initiative was conducted by NGRC in 2014. Due to the prevailing war conditions, regeneration operations were suspended from the beginning of 2015 until 2020, resulting in no activity in this regard during that period. Subsequently, with the initiation of the 'Participatory Conservation and Sustainable Use of Local Varieties' project, funded by the Benefit Sharing Fund (BSF) of the ITPGRFA, approximately 200 plant accessions (mainly cereal crops) were regenerated in the governorates of Dhamar and Hadramout (Seiyun) (FAO, 2023).

Table 6 shows the numbers of seed samples that have been regenerated until 2014 and those still requiring regeneration, as reported by NGRC in Dhamar.

Conservation changes and trends in genebanks

The study period witnessed several notable positive changes and trends in the field of genebank conservation, including:

- Increase and expansion of plant samples: The number of seed accessions in the NGRC genebank increased from 3,281 in 2006 to approximately 6,500 accessions in 2021. Similar growth was observed in the Genetic Resources Center (GRC) of Sana'a University, with the number of accessions rising from 1,528 to over 3,000 (Table 5).
- Improved energy security: Noteworthy progress was made in securing electrical energy for cold storage in the NGRC genebank in Dhamar. Solar energy systems were introduced in 2016, funded by the Agricultural and Fisheries Production Encouragement Fund (AFPEF). Additionally, the Public Corporation for Grain Development and Production (PCGDP) provided a solar energy system for the GRC at the Faculty of Agriculture, Sana'a University, in 2019.
- In 2022, with support from the Crop Trust and the ITPGRFA, the NGRC genebank in Dhamar was provided with several solar energy batteries to improve the electricity supply for cooling the long-term seed storage.
- Safety duplication: Copies of seed accessions from cereal and leguminous crops were sent to ICARDA in 2013 and subsequently stored in the Svalbard Seed Vault (Table 4).

However, there were also negative changes and trends observed in germplasm conservation in genebanks.

Negative impact on field genebanks: The conservation of plant genetic resources in field genebanks managed by research stations across the various regions was significantly affected by the war and its associated socioeconomic and institutional consequences. The lack of operational budgets hindered essential agricultural activities such as ploughing, levelling, weeding, pruning, grafting, fertilizing and irrigation. Fuel shortage and high prices further exacerbated the challenges faced by research stations in adequately maintaining the field genebanks. In certain field genebanks in the northern (Al-Errah) and central highlands (Dhamar), some fruit trees were cut down by residents for firewood during periods of gas shortage and high prices resulting from the war and siege (Figure 6). Complete destruction of field genebanks occurred in the Southern Highlands Agricultural Research Station (Osaiferah, Taiz) for mango, guava and coffee varieties, as well as in the Southern Coast Agricultural Research Station in El-Kod, Abyan governorate. Similarly, the field genebank for pomegranate at the College of Agriculture and Veterinary Medicine, University of Dhamar, experienced the same fate, with all preserved trees being cut down due to war and siege conditions.

Limited implementation of biotechnologies: The use of biotechnologies, such as tissue culture, for the conservation of plant genetic resources has not been possible due to a lack of resources, including electrical energy sources, laboratory materials and equipment. Thus, no duplicates of the accessions maintained in field genebanks had been maintained in tissue cultures and thus, as safety duplicates.

Limited multiplication and regeneration of seed samples: Except for a limited number of seed accessions regenerated and multiplied through funded projects such as the 'Participatory Conservation and Sustainable Use of Landraces' (2019–2023) (NGRC, 2022) and activities supported by the General Corporation for Grain Production and Development, the war's economic effects halted vital regeneration programmes and seed multiplication activities. Improper storage conditions, insufficient seed quantities and difficulties in providing seeds to farmers and researchers have led to a decline in seed viability. Hundreds of samples are currently at risk of spoilage and loss, requiring regeneration and multiplication.

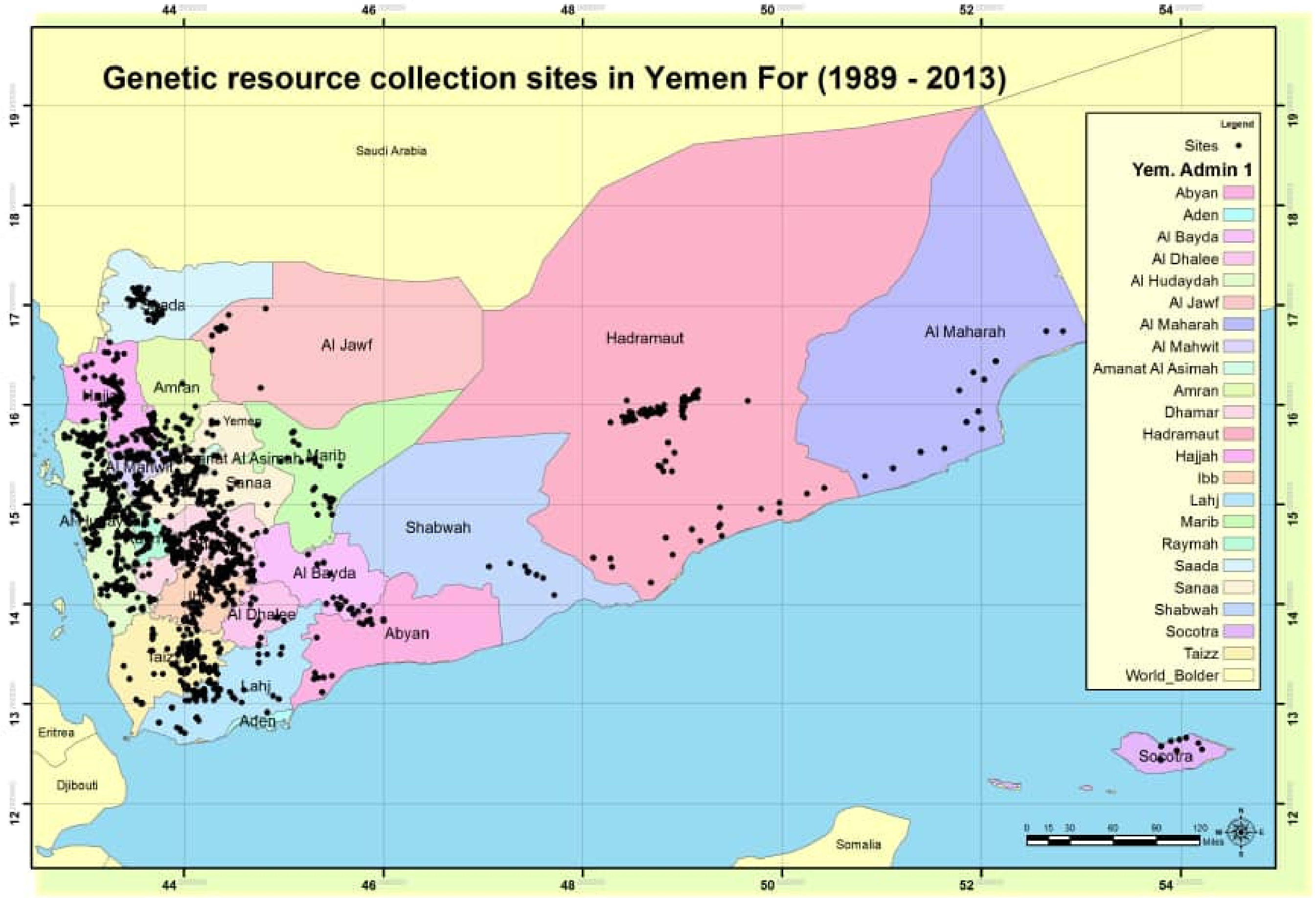


Figure 5. Areas of plant genetic resources collection in various parts of Yemen (1989–2019). Source: NGRC (2021), NGRC database (2021) and FAO (2023)

Table 4. Number of accessions of some types of Yemeni plant genetic resources preserved outside Yemen (SGSV, 2023). ICRISAT, The International Crops Research Institute for the Semi-Arid Tropics; CIMMYT, The International Maize and Wheat Improvement Center; ICARDA, The International Center for Agricultural Research in the Dry Areas; IITA, International Institute of Tropical Agriculture; US NPGS, United States National Plant Germplasm System. *, includes the following genera: *Capsicum, Nigella, Citrullus, Linum, Coriandrum, Medicago, Teramnus* and *Ricinus*.

Crop	ICRISAT	ICARDA	IITA	CIMMYT	US NPGS	Others	Svalbard
Sorghum	2,144				3,376	45	5,565
Maize				2		1	3
Wheat		13		33	12	115	173
Barley		66			25	55	146
Millet	289					49	338
Cowpea			25			12	37
Lentil		38				48	86
Bean		3					3
Peas						50	50
others*	2	7			33	88	130
Total	**2,435**	**127**	**25**	**35**	**3,446**	**463**	**6,531**

Table 5. Changes in the number of species and accessions of genetic material in seed and field genebanks in Yemen (2006–2020). Source: FAO (2009) and AREA (2021).

Centre/Reserve	Institution	Site	No. of species		No. of accessions	
			2006	**2020**	**2006**	**2020**
National Genetic Resource Center (NGRC)	Agricultural Research and Extension Authority	Dhamar	56	47	3,281	6,300
Genetic Resource Center	College of Agriculture, Sana'a University	Sana'a	38	54	1,528	> 3,000
Genebank	Nasser's College of Agricultural Sciences, University of Aden	Lahj	9	0	136	0
Field genebank of date palm and lemon	The Valley and the Desert Agricultural Research Station, Seiyun	Seiyun	1	1	67	48
Field genebank of coffee, guava and mango	The Southern Highlands Agricultural Research Station	Osaifr, Taiz	16	0	36	0
Field genebank of pomegranate	The College of Agriculture and Veterinary Sciences, Dhamar University	Dhamar	1	0	22	0
Field genebank of mango and date palm	Tihama Plateau Agricultural Research Station	Al-Kadan, Surdud, Al-Hudeidah	5	2	64	47
Field genebank of apricot, grapes, almond, pomegranate, fig and olive	The Northern Highlands Agricultural Research Station	Al-Errah, Sana'a	11	6	218	78
Field genebank of apple, olive and almond.	The Central Highlands Agricultural Research Station	Dhamar	–	65	55	144
Field genebank of banana, papaya, mango and date palm	The Southern Coast Agricultural Research Station	Al Kod-Abyan	65	67	230	> 400
Field genebank date palm, coconut and Jasminum sambac	The Eastern Coast Agricultural Research Station	Mukalla, Hadramout	15	24	68	> 200
Total			**217**	**266**	**5,705**	**10,217**

Figure 6. Remains of the field genebank of mango, coffee and guava crops in the field genebank of the Osaifera experimental farm, Taiz

Table 6. Number of seed accessions in NGRC in Dhamar that have been regenerated until 2014 and those that need regeneration. Source: Data were collected during the genetic resources inventory of NGRC in 2014 (NGRC, 2022).

Crop	Scientific name	No. of accessions	No. of regenerated accessions	No. of accessions needing regeneration
Cereals				
Sorghum	*Sorghum bicolor*	2,436	679	1,757
Millet	*Pennisetum glaucum*	589	225	364
Maize	*Zea mays*	505	108	397
Barley	*Hordeum vulgare*	362	164	198
Wheat	*Triticum aestivum*	351	229	122
Finger millet	*Eleusine coracana*	36	0	36
Al-Tahf	*Eragrostis* spp.	11	0	11
Legumes				
Cowpea	*Vigna unguiculata*	451	88	363
Lentils	*Lens culinaris*	180	78	102
Fenugreek	*Trigonella foenum-graecum*	141	141	0
Beans	*Vicia faba*	100	40	60
Kidney beans	*Phaseolus vulgaris*	90	90	0
Peas	*Pisum sativum*	80	80	0
Lablab	*Lablab purpureus*	45	0	45
Vegetables				
Okra	*Abelmoschus esculentus*	95	0	95
Radish	*Raphanus sativus*	47	33	14
Eggplant	*Solanum melongena*	42	0	42
Tomatoes	*Solanum lycopersicum*	41	0	41
Carrot	*Daucus carota*	19	0	19
Cumin	*Cuminum cyminum*	20	0	20
Arugula	*Eruca sativa*	10	7	3
Onion	*Allium cepa*	10	0	10
Leek	*Allium porrum*	3	0	3
Chilli	*Capsicum* spp.	40	0	40
Cucumber	*Cucumis sativus*	8	0	8
Pumpkin	*Cucurbita* spp.	6	0	6
Melon	*Cucumis melo*	60	0	60
Watermelon	*Citrullus lanatus*	10	0	10
Calabash/Squash	*Cucurbita maxima*	53	0	53
Mallow	*Corchorus olitorius*	37	0	37
Oil and cash crops				
Sesame	*Sesamum indicum*	88	71	17
Peanut	*Arachis hypogaea*	9	0	9
Underutilized crops				
Coriander	*Coriandrum sativum*	50	24	26
Black seed	*Nigella sativa*	49	23	26

Gaps and challenges facing the conservation of genetic resources in Yemen

This list of gaps and challenges in the conservation of plant genetic resources in Yemen, in addition to the recommendations and conclusions, is based on government documents, including the first, second and third country reports prepared as contributions to the FAO Reports on the State of the World's Plant Genetic Resources for Food and Agriculture. AREA, as a national public agricultural research institution in Yemen, has asked the national focal point of the ITPGRFA and CGRFA to prepare a working paper to be presented at the planned National Conference on Genetic Resources but postponed it to an undefined future time after preparatory efforts of more than a year. Such a paper was prepared and reviewed by the steering committee and included a set of gaps, conclusions and recommendations. Moreover, the recently prepared documents on a road map of the agricultural research programmes and projects (AREA, 2022), prepared by a specialist team, discussed thoroughly with the AREA management, and endorsed by higher agricultural authorities, are compatible with what is stated in this article. However, the content of this paper and other developments will serve as a background primary paper for any national symposium or conference when the time comes after the resumption of peace in the country.

Based on the above-reported situation, the most significant points can be summarized as follows.

In situ conservation gaps and challenges

- Limited knowledge about rich areas/regions of plant genetic diversity and regions at risk of extinction and loss of diversity
- Inadequate methodologies for monitoring and evaluating genetic diversity, including geographical and plant species gaps
- Lack of a policy for restoring cropping systems after disasters and wars
- Absence of targeted collecting, propagation and reintroduction of rare and threatened species to their natural habitats
- Insufficient sustainable use of economically promising plant species
- Absence of *in situ* conservation and management of crop wild relatives and wild food plants
- Weak technical capabilities, including a shortage of specialized staff in farm systems management, biodiversity, environmental systems management, plant taxonomy, remote sensing, control and monitoring, and evaluation systems. Additionally, laboratories in agricultural research, environmental protection and university science colleges lack certain devices and equipment.
- Lack of support and funding for managing natural sites, and inadequate funding for scientific research activities
- Lack of coordination between relevant authorities, particularly between AREA and EPA, which serve as the national focal points for the ITPGRFA and CBD, respectively.

Ex situ conservation gaps and challenges

- Weak integration between on-farm/*in situ* and *ex situ* conservation
- Absence of a plan for regeneration and multiplication under normal and emergency circumstances.
- Lack of a specialized supportive unit for communication, publication and coordination of genetic resources
- Genetic collection operations are not comprehensive and do not cover most geographical areas, genera or plant species
- Weak human, technical and institutional capabilities in NGRC
- Weakness in the basic infrastructure for managing plant genetic information
- Weak participation of local partners and government institutions in collecting, preserving and using plant genetic resources
- Lack of annual budgets to cover the expenses of collecting and preserving plant genetic resources
- Weak cooperation and participation in regional and international initiatives in the field of conservation and use.

Conclusion and recommendations

Based on the preceding results and discussion, the plant genetic resources system in Yemen faces numerous problems and gaps, which have been exacerbated by the ongoing war since 2015. However, it is possible to identify a set of needs to develop conservation processes for both *in situ* and *ex situ* plant genetic resources in the country. In this regard, the most important recommendations can be summarized as follows:

On farm/*in situ*

- Assess the impact of threats and pressures on crop diversity based on scientific foundations
- Identify areas of diversity and threats and develop an action plan for their conservation
- Conduct targeted collecting of threatened and rare species
- Reintroduce threatened or rare species and varieties into Yemen through a coordinated effort
- Conduct inventory and survey operations for crop wild relatives, in coordination with *ex situ* activities
- Promote *in situ* conservation and management of crop wild relatives and wild food plants
- Develop a national seed policy that includes provisions for seed distribution mechanisms during and after disasters and wars
- Enhance technical expertise in plant taxonomy, agriculture and environmental systems management, and monitoring and evaluation
- Provide adequate funding and financial support for programmes and projects focused on managing natural sites and on-farm conservation
- Develop effective mechanisms to enhance cooperation and coordination among various relevant agencies.

Ex situ (genebanks)

- Improve storage conditions in genebanks by providing electricity, preservation equipment and materials, and conducting seed quality tests
- Establish a centrally organized seed health system by providing resources and expertise
- Enhance the national genetic resources information system, documentation and data exchange, and encourage the use of available information and accumulated knowledge
- Allocate operational budgets to cover the expenses of genebanks
- Implement seed regeneration and multiplication programmes for accessions at risk of loss
- Develop the capabilities of genebank personnel for routine genebank operations including information management systems

- Provide NGRC with young male and female professionals and enhance their scientific and technical capacities in various crucial areas such as laboratory conservation using biotechnologies, genetic data and information management, plant taxonomy and characterization (particularly wild relatives), among others. It is also important to foster NGRC to assume coordinating responsibilities across the different areas and activities.
- Develop an awareness-raising strategy and plan highlighting the importance of the country's genetic resources and biodiversity
- Establish and strengthen communication, cooperation and coordination among all relevant local, regional and international organizations involved in genetic resources and genebanks
- Conduct inventory and survey operations for crop wild relatives, in coordination with on-farm or *in situ* activities.

Acknowledgments

We extend our sincere thanks and appreciation to Dr Muhammad Al-Khasha'a, Director General of the Agricultural Research Station on the southern coast, Al-Kod, Abyan governorate, Mr Amjad Bgigo from the Eastern Coast Agricultural Research Station, Mukalla, Hadramout, Mr Ali Abdul-Mughni Al-Maqtari from the Tihama Plain Agricultural Research Station, and Dr Saeed Saif, Director General of the Agricultural Research Station in the Southern Highlands, Taiz governorate. We would also like to express our gratitude to Dr Muhammad Hizam Al-Mashreqi, Dr Fouad Al-Qadasi and Dr Nadir Al-Absi, from RNRRC, and thanks to all NGRC employees and the leadership of AREA for their support and cooperation in providing the necessary data, information and images for this study. Furthermore, we express our thanks and gratitude to the Benefit Sharing Fund of the ITPGRFA for their support to the NGRC Committee during the period 2019–2023, and in disseminating the project's outputs, including this study.

Author contributions

Maeen Ali Al-Jarmouzi prepared the study proposal, collected, organized, analyzed the data, wrote and improved the manuscript's drafts. Khalil M. Alsharjabi contributed to reviewing and improving the study proposal, collecting the data, writing the manuscript, editing and improving the manuscript's drafts at different stages. Luigi Guarino provided support through guidance, reviewing and improving the manuscript draft, enhancing the translation of the manuscript and providing important references.

Conflict of interest

The authors declare no known conflicts of interest or any financial or personal relationships influencing the work or materials appearing in the article.

References

Al-Ghouri, M., Sailan, A. A., Bahumaish, A., Al-Muallem, A. B., Bazara'a, M., and Al-Khulaidi, A. A. (1996). Country Report Submitted to the FAO International Technical Conference on Plant Genetic Resources (Leipzig, Germany) .

Al-Khulaidi, A. A. (1994). Survey of the Natural Vegetation Cover of the City of Al-Sharq Town-Wadi Rama'a. *The Yemeni Journal of Agricultural Research and Studies* 1, 158–187.

Al-Khulaidi, A. A. (2013). Plant Life in Yemen. The Environment Protection Authority (EPA). The Sustainable Natural Resources Management Project (Sana'a, Republic of Yemen) .

Al-Khulaidi, A. A., Mahyoub, A., and Al-Houshabi, O. S. (2000). Inventory of the Natural Plant Species in Important Plant Areas in the Arabian Peninsula. *Aden University Journal of Basic Applied and Engineering Sciences* 1(3), 135–150.

Al-Mua'alem, A. B., Mansour, M. M., Jahlan, I. M. S., and Al-Ghaz, A. S. (1993). Investigation and Collection of the Local Germplasm of Field Crops in the Southern and Eastern Governorates of the Republic of Yemen. In Proceedings of the First National Workshop on Plant Genetic Resources, 103-155.

Alafif Cultural Foundation (2003). The Encyclopedia of Yemen: The old history of Yemen.

Amer, H. M. and Al-Dahmashi, A. (1997). The Status and Future of the Genetic Resources Activities. In *Proceedings of the First National Workshop on Plant Genetic Resources,* ed. Jaradat, A., (Aleppo, Syria: International Genetic Resources Institute), 33-43.

AREA (2013). The AREA's Annual Report for 2012 (Dhamar, Republic of Yemen: Agricultural Research and Extension Authority (AREA)).

AREA (2014). The AREA's Annual Report for 2013 (Dhamar, Republic of Yemen: Agricultural Research and Extension Authority (AREA)).

AREA (2015). The AREA's Annual Report for 2014 (Dhamar, Republic of Yemen: Agricultural Research and Extension Authority (AREA)).

AREA (2021). The AREA's Annual Report for 2020 (Dhamar, Republic of Yemen: . Agricultural Research and Extension Authority (AREA)).

AREA (2022). A Road Map of the Agricultural Research Programs and Projects (2022-2026) (Sana'a, Yemen: Agricultural Research and Extension Authority (AREA), MAI), 2022-2026.

Bawazir, A. A. (2004). Biological Variation in the Genetic Resources of Cereal Crops Grown in the Southern Region of Yemen. *The Yemeni Journal of Agricultural Research and Studies, issue* 11, 119–132.

Bazara'a, A. (2000). Efforts Made to Preserve the Socotra Pomegranate Plant Outside its Natural Habitat. *The Yemeni Journal of Agricultural Research and Studies* 3, 96–107.

Ceccarelli, S. (2002). Village-Based Participatory Breeding in the Mountain Slopes of Yemen (Aleppo, Syria: ICARDA). url: https://hdl.handle.net/10568/76968.

Communication Team ICARDA (2018). Enhancing Food Security in Arab Countries (Beirut, Lebanon: International Center for Agricultural Research in the Dry Areas). url: https://hdl.handle.net/20.500.11766/8320.

CSO (1990-2020). The Central Statistics Organization Collection of Books and Reports. Ministry of Planning and Development (MIPD) (Sana'a, Yemen) .

FAO (2009). The Second Country Report on the State of the Genetic Resources for Food and Agriculture in Yemen (Rome, Italy: FAO). url: https://www.fao.org/3/i1500e/Yemen.pdf.

FAO (2019). Preparation of Country Reports for the Third Report on the State of the World's Plant Genetic Resources for Food and Agriculture. url: https://www.fao.org/fileadmin/user_upload/wiews/docs/Reporting_Guidelines_2020e.pdf.

FAO (2023). Participatory conservation and sustainable use of local landraces to improve the livelihood and the resilience of farmers to climate change in Yemen (Rome, Italy: Food and Agriculture Organization (FAO) of the UN - International Treaty on Plant Genetic Resources for Food and Agriculture). url: https://www.fao.org/plant-treaty/areas-of-work/benefit-sharing-fund/projects-funded/bsf-details/en/c/1198874/?iso3=YEM.

GDSID/MAI (2022). General Directorate of Statistics and Information Documentation-Ministry of Agriculture and Irrigation Annual Agricultural Statistics Book for the Year 2021 (Sana'a, Republic of Yemen) .

Genesys (2021). Genetic Data. url: https://www.genesys-pgr.org/. accessed date: 2023-02-18

Gotor, E. and Cherfas, J. (2012). Medicinal and aromatic plants improve livelihoods in Yemen. Impact Assessment Brief No. 9. Bioversity International. url: https://hdl.handle.net/10568/74283.

Guarino, L. (1989). Report on the Collection of Crop Germplasm in the P. D. R of Yemen (1988-1989) (Ministry of Agriculture and Agrarian Reform-El-Kod Agricultural Research Center, PDRY-IBPGR), 27p. url: https://www.central-repository.cgiar.org/download.html?path=/Collecting%20Missions%20File%20Repository/Bioversity%20International/CN467/CN467Report5.pdf

Hakimi, A. A. and Ya'ni, A. A. (2008). Women and coping strategies for adaptation to climate change using agrobiodiversity resources in the rainfed highlands of Yemen (English) (Washington, D.C.: World Bank Group), 19-22. url: http://documents.worldbank.org/curated/en/588831468339546090/Women-and-coping-strategies-for-adaptation-to-climate-change-using-agrobiodiversity-resources-in-the-rainfed-highlands-of-Yemen.

IFAD (2010). Rural Development: Raymah Area Development Project. url: https://www.ifad.org/en/-/document/project-completion-report-digest.

IFAD (2011). Rural Development: Al Mahrah Area Development Project. Project Completion Report Digest. url: https://www.ifad.org/en/web/operations/-/project/1100001095?p_l_back_url=%2Fen%2Fsearch%3Fq%3DThe%2Bparticipatory%2Brural%2Bdevelopment%2Bof%2BAl-Mahrah%2Bgovernorate%2BYemen

IFAD (2012). Rural Development: Dhamar Participatory Rural Development Project. Dhamar Supervision Report, December 2012. url: https://www.ifad.org/en/web/operations/-/project/1100001195?p_l_back_url=%2Fen%2Fsearch%3Fq%3DRural%2BDevelopment%253A%2BDhamar%2BParticipatory%2BRural%2BDevelopment%2BProject

IIP (2009). Yemen - Irrigation Improvement Project (Washington, D.C.: World Bank Group). url: http://documents.worldbank.org/curated/en/824371468183283540/Yemen-Irrigation-Improvement-Project.

IPPC (2021). The Commission of Phytosanitary Measures (CPM) Recommendation. url: https://www.ippc.int/en/publications/89786/.

Jaradat, A. (1997). Proceedings of the first workshop on plant genetic resources in Yemen (Aleppo, Syria: International Institute for Genetic Resources).

Macrotrends (2023). Yemen Population. url: https://www.macrotrends.net/countries/YEM/yemen/population. accessed date: 2023-10-28

Mufarreh, M., Mahoub, A. H., and Muthanna, S. (2011). Assessment of the Degradation of the Vegetaion Cover in the Upper Site of Wadi Rasyan. *The Yemeni Journal of Agricultural Research and Studies* 24, 137–158.

Mufarreh, M. M. (2003). Inventory and Evaluation of the Pastoral Resources in the Belad Al-Ta'am District, Raimah Governorate. *The Yemeni Journal of Agricultural Research and Studies, issue* 8, 113–116.

Mufarreh, M. M. (2004). Inventory and Evaluation of the Pastoral Resources in Kusmah District, Raimah Governorate. *The Yemeni Journal of Agricultural Research and Studies* 10, 179–184.

NGRC (2008). The Annual Technical Report of the National Genetic Resources Center for the year 2007 (Dhamar, Republic of Yemen: AREA).

NGRC (2009). The Annual Technical Report of the National Genetic Resources Center for the year 2008 (Dhamar, Republic of Yemen: AREA).

NGRC (2010). The Annual Technical Report of the National Genetic Resources Center for the year 2009 (Dhamar, Republic of Yemen: AREA).

NGRC (2011). The Annual Technical Report of the National Genetic Resources Center for the year 2010 (Dhamar, Republic of Yemen: AREA).

NGRC (2013). The Annual Technical Report of the National Genetic Resources Center for the year 2012 (Dhamar, Republic of Yemen: AREA).

NGRC (2021). The Annual Technical Report of the National Genetic Resources Center for the year 2020 (Dhamar, Republic of Yemen: AREA).

NGRC (2022). The Annual Technical Report of the National Genetic Resources Center for the year 2021. AREA, Dhamar, Republic of Yemen.

RNRRC (2022). The Annual Technical Report of the Renewable Natural Resources Research Center for the year 2021 (Dhamar, Republic of Yemen: AREA).

SGSV (2023). Svalbard Global Seed Vault Data Portal. url: http://www.nordgen.org/sgsv. accessed date: 2023-05-20

YACCAP (2010). Yemen - Agro-biodiversity and Climate Change Adaptation Project. Environmental assessment: Environmental and social impact assessment (English) (Washington, D.C.: World Bank Group). url: http://documents.worldbank.org/curated/en/721331468334834796/Environmental-and-social-impact-assessment.

YRALP (2005). Yemen - Rainfed Agriculture and Livestock Project (Washington, D.C.: World Bank Group). url: http://documents.worldbank.org/curated/en/438861468334798669/Republic-of-Yemen-Rainfed-Agriculture-and-Livestock-Project.

Zahran, M. A., Khedr, A., Dahmash, A., and El-Ameir, Y. A. (2019). Qat farms in Yemen: Ecology, dangerous impacts and future promise. *Pages* 107. doi: https://doi.org/10.1016/j.ejbas.2013.09.002