Simple Sequence Repeats (SSR) marker-based analysis on the genetic variation and population structure of local and exotic sorghum germplasm collection conserved ex-situ in Sri Lanka

D. V. S. Kaluthanthri
University of Sri Jayewardenepura
S. A.C.N. Perera
University of Peradeniya
P. N. Dasanayaka
nilanthiedas@sjp.ac.lk

University of Sri Jayewardenepura

Research Article

Keywords: Ex-situ conservation, Genetic variation, Molecular characterization, Population structure, Sorghum bicolor

Posted Date: May 3rd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4332824/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on August 22nd, 2024. See the published version at https://doi.org/10.1007/s10722-024-02128-7.

Abstract

Sorghum (*Sorghum bicolor* (L.) Moench) is one of the most important cereal crops occupying the fifth position based on the cultivated extent among the cereal crops in the world. Characterization of genetic resources is a pre-requisite for utilization of conserved genetic resources in breeding programmes and cultivation. The present study was carried out to reveal the genetic variation and population structure of local and exotic sorghum germplasm collection conserved in *ex-situ* seed gene bank at the Plant Genetic Resources Centre, Sri Lanka. Total genomic DNA was extracted from 60 germplasm accessions using CTAB miniprep DNA extraction protocol. A two-step PCR amplification was performed at 16 Simple Sequence Repeat (SSR) loci. Four differentially labeled PCR products were multiplexed and size-fractioned using capillary electrophoresis. Data analyses were performed using GeneMapper 4.0, OSIRIS, PowerMarker 3.25, Structure 2.2 and STRUCTURE HARVESTER. The 16 SSR loci recorded polymorphism and the dendrogram revealed four distinct clusters. The optimum number of subpopulations was three in addition to two admixture subpopulations. The revealed population structure did not depict the geographical origin of the germplasm accessions. The present study confirmed that the majority of local sorghum germplasm accessions tested were genetically distinct. Varying degrees of outcrossing selfing in subsequent generations may have led to the creation of novel sorghum genotypes at global level.

Introduction

Sorghum bicolor is an annual, monoecious cereal belonging to family *Poaceae*. It is a predominantly self-pollinated diploid (2n = 20), displaying varying degrees of natural cross-pollination. (Doggett, 1970). Eastern Africa, centering Ethiopia is considered to be the origin of sorghum with several domestication events leading to current cultivars. Sorghum is widely cultivated globally for food and feed and being a nutritious staple for the rural poor in certain tropical areas, it is reported to be the fifth most important cereal crop in the world (Statista, 2023). The C_4 photosynthetic pathway and the potential for drought and heat tolerance make sorghum an ideal crop for marginal lands and arid areas in the tropics playing a crucial role as a resilient crop in facing the challenges of climate change.

Diverse genetic forms such as wild and related species, breeding stocks, mutant lines etc. are vital sources of desirable alleles, providing material for crop breeding. However, genetically improved novel cultivars used in modern intensive agriculture are genetically highly uniform and consist of a narrow genetic base which make them vulnerable to natural catastrophes, diseases and pests (Liu et al., 2003). Understanding the distribution of genetic diversity among gene pools is crucial in breeding for desirable traits, efficient management of germplasm collections and trans-boundary germplasm exchanges. Accordingly, for ever-changing breeding goals, different genes/alleles need to be reserved in cultivated and cultivable crop species and such genetic resources should be extensively characterized for effective utilization. Unravelling genetic diversity at molecular level facilitates allelic level differentiation and evaluation of genetic stocks. Simple sequence repeats (SSR) are quite popular for the study of genetic diversity of crops due to numerous advantages that they offer (Ghebru et al., 2002).

Population structure analysis has been employed to categorize sorghum accessions into distinct subpopulations or groups based on genetic relatedness in various studies (Mamo et al., 2023: Yahaya et al., 2023). This helps in understanding the underlying genetic structure within the species. Principal Component Analysis (PCA) and model-based clustering algorithms such as STRUCTURE are commonly used to infer population structure.

Diverse domestication events, exchange of seeds via formal and informal channels for extensive worldwide cultivation and the considerable levels of cross-pollination followed by fixing of alleles via self-pollination has led to wide variability among sorghum landraces world over (Mutegi et al., 2010). In Sri Lanka, the *ex-situ* gene bank of

sorghum is rich with a considerable number of local accessions with different vernacular names and accessions of exotic origin. Yet, the effective use of material from this collection is hindered due to lack of information on their genetic diversity. Accordingly, the aim of the current study was to identify the genetic diversity and to elucidate the population structure of *ex-situ* conserved local and exotic sorghum germplasm collection using SSR markers and to make recommendations for conservation and breeding.

Materials and methods

Germplasm materials and experimental design

The study was based on 60 sorghum germplasm accessions including both local (Table 1) and exotic (Table 2) germplasm that are conserved at the seed gene bank of Plant Genetic Resources Center (PGRC), Gannoruwa, Sri Lanka. Seeds of germplasm accessions were sown in small plastic pots filled with soil and the plants were grown in a plant house under controlled environment.

DNA extraction and Capillary electrophoresis

For each of the germplasm accessions, the total genomic DNA was extracted using leaf materials harvested from twenty-day-old plants. Fresh, immature leaves were surface sterilized by washing with sterilized distilled water and 70% ethanol, air dried and cut into small pieces. An approximate weight of 5 g was obtained and put into a 1.5 mL microcentrifuge tube and stored at -80 0C for 24 hours. The frozen pieces of immature sorghum leaves were crushed and total genomic DNA was extracted from ground leaf material using CTAB miniprep DNA extraction protocol (Taylor and Powell, 1982) followed by quantification and dilution to a final concentration of 25 ng μL^{-1}.

Extracted DNA was amplified at 16 SSR loci (Table 3) via a two-step PCR protocol using a modified M13 tagged (5'TGTAAAACGACGGCCAGT3') forward primer and a pig tailed (5'GTTTCTT3') reverse primer with M13 oligo labeled with one of four fluorescent dyes; 6-HEX, FAM, TAMN or PET in a T100™ Thermal cycler (Bio-Rad laboratories Inc.). Four differentially labeled PCR products were multiplexed and size-fractioned via capillary electrophoresis on an ABI3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Data analysis

Electrophoretic data derived from capillary electrophoresis were imported to GeneMapper software version 4.0 for allele characterization based on molecular weights of SSR products. OSIRIS (Open Source, Independent Review and Interpretation System) software version 2.13 was applied for further validation of size peak patterns, using the ABI-LIZ-600-80-TO-400 internal lane standard for allele calling. The peaks from polymorphic SSR markers on capillary electrophoresis were scored as present (1) or absent (0) or missing data (-9) for each germplasm accession by manual inspection. Data were tabulated in a matrix using Excel version 2010 followed by analyzing using PowerMarker V3.25 software (Liu and Muse, 2005). Determination of total number of alleles and polymorphism information were estimated from the genotyping data, considering that each SSR primer pair represents a unique locus. Nei's (1972) standard genetic distances between the germplasm accessions were obtained.

Cluster analysis was performed based on the unweighted pair group mean algorithm (UPGMA) to draw the dendrogram. Genotypic data were analyzed using Structure 2.2 (Pritchard et al., 2000), to determine structured groups or subpopulations (K clusters) in the collection. Analysis was performed using settings for admixture ancestry model and correlated allele frequencies of the programme. The K values were set at 1 to 10, with a burn in value of 10,000 and 30,000 of Markov Chain Monte Carlo replications (MCMC reps) after burning followed by 3

iterations. Optimal K clusters were identified using the online based software STRUCTURE HARVESTER (Earl and VonHoldt, 2012).

Results

The studied sorghum population recorded polymorphic loci at all 16 SSR marker loci. Frequency of major SSR alleles at different SSR loci recorded a mean of 0.257 ranging from 0.17 to 0.33. A total of 170 alleles were recorded across the 16 loci with number of genotypes ranging from 4 to 29 per locus with a mean of 16.8 (Table 4). The mean number of alleles per locus was recorded as 10.6. The results provide ample evidence for high allelic richness in this specific germplasm collection in Sri Lanka. Availability, which is an indication of the amplification efficiency of the marker was more than 50% at all marker loci with an average of 80.63%. Three markers (Xtxp32, Xtxp266 and Xtxp265) showed a high amplification efficiency exceeding 90% of availability. Gene diversity, which is a measure of genetic diversity, varied within the range of 0.694 to 0.887 recording a mean of 0.821. More than one allele was amplified in some accessions at certain loci indicating the residual heterogeneity within these accessions. The observed heterozygosity at each locus over all accessions ranged from zero at marker locus Xtxp295 to 0.672 at Xtxp265 with an average of 0.399 per locus. All the primer pairs used in the study were highly informative with PIC values ranging from 0.632 to 0.877 at marker loci Xtxp295 and Xtxp265 respectively, recording an average of 0.797.

Frequency based standard genetic distances and dendrogram

The expected value of the standard genetic distance developed by Nei in 1972, is proportional to evolutionary time based on both effects of mutation and genetic drift. The genetic distance of 1.0 was observed between a large numbers of sorghum accessions indicating the high genetic divergence between them. The lowest genetic distance of 0.0 was observed between the accessions UNKNOWN 8917 – ETHIOPIA 6014 and UNKNOWN 8915 – UNKNOWN 8943 indicating probable genetic duplication.

The dendrogram of the 60 accessions (Fig. 1) constructed based on Nei's (1972) standard genetic distance matrix using unweighted pair group mean algorithm (UPGMA) revealed four distinct clusters, I-IV. Cluster I was the most homogenous including 15 local sorghum accessions. The remaining clusters comprised of local, origin unknown and exotic germplasm accessions. Two of local and thirteen of exotic accessions were identified in cluster II whereas four local accessions and eleven accessions with unknown origin grouped in cluster III. The remaining fifteen-germplasm accessions including local, exotic (European) and two accessions of unknown origin clustered in cluster IV.

Summary of Nei's standard genetic distances with respect to each cluster is given in Table 5. The lowest means and standard deviations (SD) of the genetic distances were observed in cluster I whereas they were the highest in cluster II. Considerably low mean standard genetic distance and SD values of cluster I revealed the homogenous nature and comparatively low genetic divergence of the 15 local germplasm accessions that grouped in cluster I. On the other hand, cluster II consisted of genetically more divergent germplasm accessions by recording considerably high mean standard genetic distance and SD values. Meanwhile, the difference of mean standard genetic distance was the highest between clusters I and II.

Determination of population structure of sorghum accessions

The admixture ancestry model-based approach with correlated allele frequencies was performed using STRUCTURE software to examine the relatedness among 60 sorghum germplasm accessions, using genotypic data at 16 SSR loci. The K values were set at 1 to 10, with a burn in value of 10,000 and 30,000 of Markov Chain Monte Carlo

replications (MCMC reps) after burning followed by three iterations. Optimal K clusters were identified following the method described by Evanno et al. (2005) based on the improvements in the estimated natural log of likelihood [Ln P(D)] in three independent runs (Casa et al., 2008) through the software STRUCTURE HARVESTER (Earl and VonHoldt, 2012). The estimated LnP(D) values and the highest value of the ad hoc quantity showed that the optimum number of subpopulation K which best explained the population structure was 3. In order to better understand the geographical differentiation based on the defined geographical regions, K = 4 which had the second highest ad hoc quantity Δk was also considered. Graphical representations of K = 3 and K = 4 are presented in Figs. 2 and 3 respectively. Germplasm accessions with probabilities greater than 0.95 and less than 0.95 were considered as pure lines and admixtures respectively.

Discussion

The mean number of alleles per locus recorded in this study was greater than the values recorded by Ali et al. (2008) and Schloss et al. (2002) on 72 and 25 sorghum accessions which recorded 3.22 and 3.4 mean number of alleles respectively. Comparatively low mean number of alleles per locus were also reported by Agrama and Tuinstra (2003), Menz et al. (2004), Egbadzor (2007) and Pei et al. (2010). However, Burow et al. (2012) reported a considerably high number of alleles per locus in their study on Chinese sorghum landraces.

Sorghum genetic diversity values were reported to be 0.6 across 29 SSR loci (Schloss et al. 2002) and ranged from 0.44 to 0.79 at five SSR loci in Somalian landraces (Manzelli et al. 2005). Gene diversity values for sorghum were recorded as 0.29 (Dje et al., 2000) and 0.32 (Barnaud et al., 2007) in Morocco and 0.65, 0.61, 0.66 and 0.66 respectively in Eritrea (Ghebru et al., 2002), Niger (Deu et al., 2008), Kenya (Ngugi and Onyango, 2012) and Ethiopia (Adugna, 2014) respectively. All the above studies evaluated the genetic diversity of sorghum in Africa which is the center of origin for sorghum and despite being in South Asia, the Sri Lankan collection is noted to be harboring a higher diversity.

The observed heterozygosity was comparatively higher to what were observed by Dje et al. (2000), Barnaud et al. (2007) and Deu et al. (2008) recording 0.134, 0.11 and 0.042 respectively in their studies. In contrast to the mean values for heterozygosity and gene diversity recorded in this study, Ng'uni et al. (2011) and Motlhaodi et al. (2014) reported considerably lower values in their studies. However, the observed heterozygosity in the current study was not significantly high and it indicated that the majority of genotypes carried homozygous loci which led to a high level of genetic stability. High levels of allelic variability but low levels of heterozygosity at SSR loci were reported on Eritrean sorghum landraces (Ghebru et al., (2002). The predominantly inbreeding nature of sorghum may be the cause for the low level of heterozygosity observed, yet with the gene pool as a whole maintaining a high level of allelic variation. Besides, the high gene diversity (expected heterozygosity) values compared to the observed heterozygosity values across the 16 loci indicated low cross-pollination rates among the accessions as expected. In congruence with this study, most of the genetic diversity studies in sorghum using SSRs (e.g., Ghebru et al., 2002; Barro-Kondombo et al., 2010; Deu et al., 2010; Ngugi and Onyango, 2012) reported high gene diversity (expected heterozygosity) values compared to the observed heterozygosity.

The discriminatory power of SSR markers can be assessed with polymorphic information content (PIC) as calculated by the number of alleles detected and their relative frequencies (Smith et al., 2000). In the present study, the tested markers are highly recommended for allelic studies in sorghum. Burrow *et al.* (2012), Sapey et al. (2012) and Madhusudhana et al. (2012) reported less average values for both number of alleles per locus and PIC compared to the results of the current study. A total of 23 unique alleles which may serve as diagnostic tools for particular regions

of the genome of respective genotypes, were observed in certain accessions at eleven SSR loci; Xtxp266, Xtxp258, Xtxp141, Xtxp210, Xtxp230, Xtxp265, Xtxp21, Xtxp12, Xtxp297, Xtxp46 and Xtxp302.

The lowest genetic distance of 0.0 was observed between two pairs of germplasm accessions indicating probable genetic duplication. Such duplicates might be an indication of germplasm exchange within and between regions. It is important to note that, clusters I and IV depicted the influence of geographic origin on genetic diversity by solely comprising of local, Italian, and French accessions respectively. Exchange of seeds may be the reason for the genetic proximity and clustering of such accessions together. Seed exchange has been reported in the baseline survey conducted in eastern Kenya on sorghum seed systems, which revealed a tradition among farmers in selection and exchange of seeds (Muui et al., 2013). Furthermore, Ochieng et al. (2011) studied sorghum production systems in Kenya and reported the important role of farmers in seed selection, exchange and movement. Research conducted by, Ayana et al. (2000), Ghebru et al. (2002), Nkongolo and Nsapato (2003) and Deu et al. (2008), revealed similarities in genotypes as a result of region proximity in Africa for cultivated sorghum and its wild relatives. Mutegi et al. (2010) reported the clustering of sorghum in Kenya into a distinct and unique genetic group. Similarly, the current study revealed the inclusion of a majority of Sri Lankan sorghum accessions in a unique cluster. This is attributed to the fact that Sri Lanka is an island, which is relatively geographically remote from other sorghum growing regions. The majority of genetic material from local sorghum accessions tested in this study appear to be genetically distinct.

Clustering of local and exotic accessions together in cluster IV of dendrogram revealed the existing genetic similarity irrespective of the geographical origin. The existence of accessions with wider genetic distances (e.g., MONARAGALA 9002 and MONARAGALA 2971 in clusters III and IV respectively) within the same locality indicates the introduction of seeds of unique landraces from distant localities. Furthermore, some of local accessions collected from Kurunegala district in Sri Lanka with vernacular name *Edal Eringu*, grouped into different two clusters I and III. Generally, farmers save seeds from previous harvest and in cases of insufficiency they tend to obtain seeds from neighboring farmers or markets. As a consequence, there would be a lot of variation derived from generations of selection even within the varieties having the same vernacular name. This also indicates the high stepwise mutation rates of SSR (Dje et al., 1999). Despite being a predominantly inbreeding species, outcrossing rates as high as 0.10% – 0.15% (Doggett, 1988) and spontaneous hybridization are reported in sorghum. This might possibly explain the molecular level differences that exist between other studied sorghum germplasm accessions as well.

Under K = 3 in population structure analysis, SP1comprised of 28 germplasm accessions including six local, 21 of unknown origin and one exotic (ETHIOPIA 6014) accession. Purelines of 15 local accessions assigned into SP2 while purelines of five local and eight exotic accessions of European origin (Italian and French) assigned into SP3 along with two accessions of unknown origin. The results thus provide further evidence for the genetic similarity of certain accessions irrespective of geographical origin spanning across three continents. The formation of SP2 solely with local accessions indicates their genetic uniqueness which makes them even more exclusive. The admixtures of all three subpopulations were observed in two germplasm accessions with unknown origins (UNKNOWN 8918 and UNKNOWN 8942). This observation could be due to the heterogonous nature of the two accessions, sorghum being a cross-pollinating crop. In parallel to the present study, Nemera et al. (2022) revealed two sub-populations irrespective to the geographical origin of 80 Ethiopian sorghum accessions studied at 15 SSR loci. Similarly, population structure of 100 genotypes of Ethiopian sorghum profiled at 15 SSR loci, revealed three sub-populations with no clear geographical origin-based structure Mamo et al. (2023). Furthermore, a total of evenly distributed 108,107 high-quality single-nucleotide polymorphism (SNPs) markers differentiated six sub-populations within 304 Ethiopian sorghum accessions (Wondimu et al. (2021), Enyew et al. (2022) identified two sub-populations within 24 Ethiopian sorghum accessions using 5,000 SNPs. Interestingly, Chakrabarty et al. (2022) recently inferred five sub-populations within 3333 Ugandan sorghum accessions using 12,742 SNP markers. However, Menamo et al. (2021) reported that

agro-ecology is more important than the administrative region in defining the genetic variation in sorghum based on redundancy analysis.

Results revealed the inclusion of certain morphologically different local accessions with different vernacular names such as '*Waguru*' (MONARAGALA 106), '*Edal Iringu*' (MATALE 285) and '*Karal Iringu*' (HAMABANTOTA 971) in the same subpopulation (SP2). Similarly, SP3 contained morphologically different exotic germplasm accessions collected from Italy ('Roce' [ITALY 6006]; '*Vespa*' [ITALY 6008]; 'Wonder Dwarf' [ITALY 6013]) and from France ('*Arval*' [FRANCE 6004]; '*Argence*' [FRANCE 6005]). The occurrence of morphologically distinct accessions having different vernacular names in the same sub-population could be due to a multitude of reasons; narrow genetic base despite the morphological differences, farmers' directional selection for different morphological traits and non-detectable nature of observed morphological differences with neutral genetic markers being the most probable reasons. This observation is consistent with the observations made by Sagnard et al. (2011) and Labeyrie et al. (2014).

At K = 4 of population structure analysis, SP1 and SP3 remained unchanged from that of K = 3 confirming their genetic sub-populations. At K = 4, SP2 contained 15 accessions including only two accessions (RATNAPURA 8638 and KEGALLA 8867) with local origin while the rest had exotic and unknown origins. SP4 included purelines of two local germplasm accessions (MONARAGALA 9002 and KEGALLA 9012) and nine accessions of unknown origin.

Population structure of common self-pollinated field crops such as wheat (Vanzetti et al., 2013) and barley (Naeem et al., 2011) have been shown to be strongly influenced by geographic adaptation. Previous studies have reported sorghum population structure based on botanical races and geography (Deu et al., 2010; Maina et al., 2018). It was not possible to relate the observed genetic structure with racial category because the germplasm accessions used in the current study have not been characterized for racial groups. Of the different sorghum growing agroecological regions in Sri Lanka, the country's 'wet zone', the western face of the mountain range, and the southwest windy slopes receive an average of 2500 mm of rainfall each year whereas the 'dry zone' receives a rainfall between 1200 to 1900 mm, mainly through the northeast monsoon (Alahacoon and Edirisinghe, 2021). This spatial analysis separated SP1 which consisted of purelines of local germplasm accessions inhabiting contrasting environments, at least in terms of rainfall. In contrast, two local germplasm accessions (MATALE 310 and KURUNEGALA 380) and two accessions of unknown origin (UNKNOWN 8953 and UNKNOWN 8956) were observed with admixtures of all four subpopulations suggesting that the population structure may also be affected by other factors such as human activities including seed exchange and food preferences. Deu et al. (2010) reported similar observation on spatio-temporal dynamics of genetic diversity in *Sorghum bicolor* in Niger.

Sorghum is a globally widely cultivated crop which has been distributed through transboundary seed exchanges across continents and within different areas within countries. In addition, the crop is cultivated in diverse agro-climatological areas and the cultivars may have genetic adaptations specific to diverse localities. Sorghum displays varying degrees of outcrossing leading to the creation of novel genotypes and getting them genetically fixed upon selfing in subsequent generations. Accordingly, the wide scale cultivation across continents coupled with the creation of novel genotypes have led to a huge genetic variation of the global sorghum germplasm. Sorghum being an ideal crop candidate in marginal areas especially in facing the adverse effects of global climate change the vast diversity unveiled in the current study as well as in other parts of the world should be appropriately utilized in breeding programs.

Conclusions

The current study attempted to elucidate the genetic diversity and population structure of a collection of 60, Sri Lankan *ex-situ* conserved sorghum germplasm accessions comprising of Sri Lankan, South Asian, African and European collections. The collection was observed to be highly polymorphic at 16 SSR loci that were tested giving evidence for high allelic richness and the availability of three to four genetic subpopulations within the collection. The revealed genetic diversity and population structure did not exactly depict the geographical origin of the germplasm accessions. Furthermore, the majority of local sorghum germplasm accessions tested in this study were genetically distinct. The study gave evidence for the importance of systematic studies of sorghum genetic resources for better decision making for sustainable cultivation, further collection and conservation and in utilization of conserved material in genetic improvement of the sorghum crop as a candidate for marginal areas and as a climate resilient crop. The findings will be of high importance with regard to decision making on sustainability of cultivation, further collection and conservation of genetic resources and genetic improvement of sorghum at a global level. The findings of the study reiterate the importance of systematic evaluation of the genetic diversity and genetic structures of populations for better decision making in selection for cultivation and more importantly in genetic improvement programmes.

Declarations

Acknowledgments:

Authors gratefully acknowledge Plant Genetic Resources Center, Gannoruwa, Sri Lanka for providing seeds and the Institute of Biochemistry, Molecular Biology and Biotechnology (IBMBB), Sri Lanka for providing the laboratory facility.

Author contribution

D. V. S. Kaluthanthri contributed to the Conceptualization, Data curation, Formal analysis, Software, Writing – original draft. S. A. C. N. Perera contributed to the Conceptualization, Methodology, Supervision, Writing – review and editing. P. N. Dasanayaka contributed to the Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review and editing.

Financial support

This research was supported by University of Sri Jayewardenepura, Sri Lanka. (Grant No. ASP/01/RE/SCI/2017/07, ASP/01/RE/SCI/2021/08). University of Sri Jayewardenepura, Sri Lanka had no role in the design, analysis or writing of this article.

Competing interests: The authors declare there are no competing interests.

Ethical standards: Not applicable

References

1. **Adugna, A.** (2014). Analysis of *in-situ* diversity and population structure in Ethiopian cultivated *Sorghum bicolor* (L.) landraces using phenotypic traits and SSR markers. *Springerplus* **3**, 212.
2. **Agrama, H. A.** and **Tuinstra, M. R**. (2003). Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology*, **2**(10), 334-340.
3. **Alahacoon, N.** and **Edirisinghe, M**. (2021). Spatial Variability of Rainfall Trends in Sri Lanka from 1989 to 2019 as an Indication of Climate Change. *International Journal of Geo-Information*, **10**, 84-102.

4. **Ali, M. L., Rajewski, J. F., Baenzieger, P. S., Gill, K. S.** and **Dweikat, E. I.** (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Molecular Breeding,* **21**, 497-509.
5. **Ayana, A., Bryngelsson, T.** and **Bekele, E**. (2000). Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resource Crop Evolution*, **47**, 471-482
6. **Barnaud, A., Deu, M., Garine, E., McKey, D**. and **Joly, H.I.** (2007). Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical Applied Genetics,* **114**(2), 237-248.
7. **Barro-Kondombo, C., Sagnard, F., Chantereau, J., Deu, M., Brocke, K.V., Durand, P., Goze, E.** and **Zongo, J.D.** (2010). Genetic structure among sorghum landraces as revealed by morphological variation and microsatellite markers in three agro climatic regions of Burkina Faso. *Theoretical and Applied Genetics,* **120**, 1511-1523.
8. **Burow, G., Franks, C.D., Xin, Z.** and **Burke, J.J.** (2012). Genetic Diversity in a Collection of Chinese Sorghum Landraces Assessed by Microsatellites. *American Journal of Plant* Sciences, **3**, 1722-1729.
9. **Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., Franks, C. D.** and **Kresovich, S.** (2008). Community Resources and Strategies for Association Mapping in Sorghum. *Crop Science*, **48**(1), 30-40.
10. **Chakrabarty, S., Mufumbo, R., Windpassinger, S., Jordan, D., Mace, Snowdon, R. J.** and **Hathorn, A.** (2022). Genetic and genomic diversity in the sorghum gene bank collection of Uganda. *BMC Plant Biology.* **22**:378.
11. **Deu, M., Sagnard, F., Chantereau, J., Calatayud, C.** and **Vigouroux, Y.** (2010). Spatio-temporal dynamics of genetic diversity in Sorghum bicolor in Niger. *Theoretical and Applied Genetics*, ***120***, 1301-1313. doi:10.1007/s00122-009-1257-1.
12. Deu, M., Sagnard, F., Chantereau, J., Calatayud, C., Herault, D., Mariac, C., Pham, J. L., Vigouroux, Y., Kapran, I., Traore, P. S., Mamadou, A., Gerard, B., Ndjeunga, J. and **Bezanaon, G.** (2008). Niger-wide assessment of *in-situ* sorghum genetic diversity with microsatellite markers. *Theoretical and Applied Genetics*, **116**, 903-913.
13. **Dje, Y., Heuertz, M., Lefebvre, C.** and **Vekemans, X.** (2000). Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. *Theoretical and Applied Genetics,* **100**, 918-925.
14. **Dje, Y., Forcioli, D.** and **Ater, M.** (1999). Assessing population genetic structure of sorghum landraces from northwestern Morocco using allozyme and microsatellite markers. *Theoretical and Applied Genetics,* **99**, 157–163.
15. **Doggett, H.** (1970). *Sorghum*. Longmans Ltd. London, UK.
16. **Doggett, H.** (1988) *Sorghum.* 2nd Edition, Longmans Scientific and Technical, UK.
17. **Earl, D.A.** and **VonHoldt, B.M**. (2012). Structure HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, **4**, 359-361.
18. **Egbadzor, F. K.** (2007). *Characterization of Fifty-four sorghum accessions using morphological and molecular markers.* MPhil. Thesis University of Ghana, Ghana.
19. **Enyew, M., Feyissa, T., Carlsson, A.S., Tesfaye, K., Hammenhag, C.** and **Geleta, M.** (2022) Genetic Diversity and Population Structure of Sorghum [Sorghum Bicolor (L.) Moench] Accessions as Revealed by Single Nucleotide Polymorphism Markers. *Frontiers in Plant Science*. **12**:799482. doi: 10.3389/fpls.2021.799482
20. **Evanno, G., Regnaut, S.** and **Goudet, J.** (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, **14**, 2611-2620.

21. **Ghebru, B., Schmidt, R. J.** and **Bennetzen, J. L.** (2002). Genetic diversity of Eritrea sorghum landraces assessed with Simple sequence repeats (SSRs) markers. *Theoretical and Applied Genetics*, **105**, 229-236.
22. **Labeyrie, V., Deu, M., Barnaud, A., Calatayud, C., Buiron, M., Wambugu, P., Manel, S., Glaszmann, J.C.** and **Leclerc, C.** (2014). Influence of ethnolinguistic diversity on the sorghum genetic patterns in subsistence farming systems in eastern Kenya. *PLoS One,* **9**(3), 92178.
23. **Liu, K.** and **Muse, S.V**. (2005). PowerMarker; Integrated analysis environment for genetic marker data. *Bioinformatics*, **219**, 2128-2129.
24. **Liu, K., Goodman, M., Muse, S., Smith, J. S., Buckler, E.** and **Doebley, J.** (2003). Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics,* **165**, 2117-2128.
25. **Madhusudhana, R., Balakrishna, D., Rajendrakumar, P., Seetharama, N.** and **Patil, J. V.** (2012). Molecular characterization and assessment of genetic diversity of sorghum inbred lines. *African Journal of Biotechnology*, **11**(90), 15626-15635.
26. **Maina, F., Bouche, T. S., Marla, S. R., Hu, Z.** and **Wang, J.** (2018). Population genomics of sorghum across diverse agro-climatic zones of Niger. *Genome*, **61**, 223-232.
27. **Mamo, W., Enyew, M., Mekonnen, T., Tesfaye, K.** and **Feyissa, T.** (2023). Genetic diversity and population structure of sorghum [*Sorghum bicolor* (L.) Moench] genotypes in Ethiopia as revealed by microsatellite markers. *Heliyon.* **9**.
28. **Manzelli, M., Benedettelli, S.** and **Vecchio, V.** (2005). Agricultural biodiversity in Northwest Somalia – an assessment among selected Somali sorghum (*Sorghum bicolor* (L.) Moench) germplasm. *Biodiversity Conservation*, **14**, 3381-3392.
29. **Menamo, T., Kassahun, B., Borrell, A., Jordan, D., Tao, Y.** and **Hunt, C.** (2021). Genetic diversity of Ethiopian sorghum reveals signatures of climatic adaptation. *Theoretical and Applied Genetics*. **134**, 731–742. doi: 10.1007/s00122-020-03727-5
30. **Menz, M., Klein, R., Unruh, N., Rooney, W., Klein, P.** and **Mullet, J.** (2004) Genetic diversity of public inbred of sorghum determined by mapped AFLP and SSR markers. *Crop Science*, **44**, 1236-1244.
31. **Motlhaodi, T., Geleta, M., Bryngelsson, T., Fatih, M., Chite, S.** and **Ortiz, R.** (2014). Genetic diversity in ex-situ conserved sorghum accessions of Botswana as estimated by microsatellite markers. *Australian Journal of Crop Science*, **8**(1), 35-43.
32. **Mutegi, E., Sagnard, F., Semagn, K., Deu, M., Muraya, M., Kanyenji, B., de Villiers, S., Kiambi, D., Herselman, L.** and **Labuschagne, M. (**2010). Genetic structure and relationships within and between cultivated and wild sorghum (Sorghum bicolor (L.) Moench) in Kenya as revealed by microsatellite markers. *Theoretical and Applied Genetics*, **122**, 989-1004.
33. **Muui, C. W., Muasya R. M**. and **Kirubi, D. T.** (2013). Baseline survey on factors affecting sorghum production and use in eastern Kenya. *African Journal of Food, Agriculture, Nutrition and Development*, **13**(1), 7339 -7342.
34. **Naeem, M., Khan, M.M.A., Moinuddin, Idrees, M.** and **Aftab, T.** (2011). Triacontanol-mediated regulation of growth and other physiological attributes, active constituents and yield of *Mentha arvensis* L. *Plant Growth Regulation*, **65**(1), 195-206.
35. **Nei, M.** (1972). Genetic Distance between Populations. *American Naturalist*, **106,** 283-292.
36. **Nemera, B., Kebede, M., Enyew, M.** and **Tileye Feyissa** (2022). Genetic diversity and population structure of sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia as revealed by microsatellite markers. *Acta Agriculturae Scandinavica, Section B — Soil and Plant Science*. **72** (1), 873–884.

37. **Ng'uni, D., Geleta, M.** and **Bryngelsson, T.** (2011). Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) accessions of Zambia as revealed by simple sequence repeats (SSR). *Hereditas*, **148**(2), 52-62.
38. **Ngugi, K.** and **Onyango, C. M.** (2012). Analysis of the molecular diversity of Kenyan Sorghum Germplasm using microsatellites. *Journal of Crop Science and Biotechnology,* **15**(3),189-194.
39. **Nkongolo, K. K.** and **Nsapato, L.** (2003). Genetic diversity in *Sorghum bicolor* (L.) Moench accessions from different ecogeographical regions in Malawi assessed with RAPDs. *Genetic Resource Crop Evolution*, **50**, 149-156.
40. **Ochieng, L. A., Mathenge, P. W.** and **Muasya, R. M.** (2011). A survey of on farm seed production practices of sorghum (*Sorghum bicolor* L. Moench) in Bomet district of Kenya, *African Journal of Food, Agriculture, Nutrition and Development*, **11**(5), 5232-5253.
41. **Pei, Z., Gao, J., Chen, Q., Wei, J., Li, Z., Luo, F., Shi, L., Ding, B.** and **Sun, S.** (2010). Genetic diversity of elite sweet sorghum genotypes assessed by SSR markers. *Biologia planetarium,* **54**(4), 653-658.
42. **Pritchard, J. K., Stephens, M.** and **Donnelly, P.** (2000). Inference of Population Structure Using Multi locus Genotype Data. *Genetics*, **155**(2), 945-959.
43. **Sagnard, F., Deu, M., Dembele, D., Leblois, R., Toure, L., Diakite, M., Calatayud, C., Vaksmann, M., Bouchet, S., Malle, Y., Togola, S.** and **Traore, P. C.** (2011). Genetic diversity, structure, gene flow and evolutionary relationships within the *Sorghum bicolor* wild-weedy-crop complex in a western African region. *Theoretical Applied Genetics,* **123**(7), 1231-1246.
44. **Sapey, E., Offei, S. K., Danquah, E. Y** and **Asante, I. K** (2012). Assessment of genetic diversity among 60 sorghum accessions in Ghana using microsatellites. *Elixir Bio Technology*, **44**, 7171-7177.
45. **Schloss, S. J., Mitchell, S. E., White, G. M., Kukatla, R., Bowers, J. E., Paterson, A. H.** and **Kresovich, S.** (2002). Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* L. Moench. *Theoretical and Applied Genetics,* **105**, 912-920.
46. **Smith, J. S. C., Kresovich, S., Hopkins, M. S., Mitchell, S. E., Dean, R. E., Woodman, W. L., Lee, M.** and **Porter, K.** (2000). Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Science Journal,* **40**, 226-232.
47. **Statista** (2023). *Worldwide production of grain in 2022/23, by type (in million metric tons)*| Statista*. https://www.statista.com/statistics/263977/world-grain-production-by-type/
48. **Taylor, B.** and **Powell, A.** (1982). Isolation of Plant DNA and RNA. *Focus*, **4**, 4-6
49. **Vanzetti, L.S., Yerkovich, N., Chialvo, E., Lombardo, L., Vaschetto, L.** and **Helguer, M.** (2013). Genetic structure of Argentinean hexaploid wheat germplasm. *Genetics and Molecular Biology*, **36**(3), 391-399.
50. **Wondimu, Z., Dong, H., Paterson, A. H., Worku, W.,** and **Bantte, K.** (2021). Genetic diversity, population structure and selection signature in Ethiopian Sorghum (*Sorghum bicolor* L.[Moench]) germplasm. *Genes, Genomes, Genetics*. **11**(6), 87-97.
51. **Yahaya ,M.A., Shimelis, H., Nebie, B., Ojiewo, C.O., Rathore, A.** and **Das, R.** 2023. Genetic Diversity and Population Structure of African Sorghum (*Sorghum bicolor* L. Moench) Accessions Assessed through Single Nucleotide Polymorphisms Markers. *Genes*.**14**(7):1480. https://doi.org/10.3390/genes14071480

Tables

Table 1: Details of studied sorghum germplasm accessions of local origin

Accession number	Accession Name	Origin	Gene bank
000004	IDAL IRINGU	ANURADHAPURA	PGRC
000093	EDAL ERINGU	MONARAGALA	PGRC
000101	EDAL ERINGU	MONARAGALA	PGRC
000104	POTH IRINGU	MONARAGALA	PGRC
000106	WAGURU	MONARAGALA	PGRC
000110	KARAL IRINGU	MONARAGALA	PGRC
000111	RATHU THIRINGU	MONARAGALA	PGRC
000112	SUDU THIRINGU	MONARAGALA	PGRC
000117	WAGURU	MONARAGALA	PGRC
000130	EDAL ERINGU	BADULLA	PGRC
000207	IDAL IRINGU	MONARAGALA	PGRC
000253	EDAL ERINGU	HAMBANTOTA	PGRC
000285	EDAL ERINGU	MATALE	PGRC
000310	SORGHUM	MATALE	PGRC
000318	KARAL IRINGU	MATALE	PGRC
000380	EDAL ERINGU	KURUNEGALA	PGRC
000403	EDAL ERINGU	KURUNEGALA	PGRC
000421	EDAL ERINGU	KURUNEGALA	PGRC
000430	KARAL IRINGU	KURUNEGALA	PGRC
000736	RATA KURAKKAN	MATALE	PGRC
000744	KARAL IRINGU	MATALE	PGRC
000844	KARALLIYA	GALLE	PGRC
000845	SORGHUM	GALLE	PGRC
000971	KARAL IRINGU	HAMBANTOTA	RSRS/AK
001532	RICE SORGHUM	ANURADHAPURA	RARS/MI
001546	EDAL ERINGU	NUWARA ELIYA	PGRC
001701	SORGHUM	SRI LANKA	ICRISAT
001824	EDAL ERINGU	KANDY	PGRC
002971	IDAL IRINGU	MONARAGALA	PGRC
003015	BATH IRINGU	HAMBANTOTA	PGRC
007969	SHAN JATHAKA	KANDY	NYES

008638	UNKNOWN	RATNAPURA	PGRC
008867	IDAL IRINGU	KEGALLA	PGRC
009002	RED SORGHUM	MONARAGALA	PGRC
009012	Not Given	KEGALLA	PGRC

Table 2: Details of sorghum germplasm accessions of known and unknown exotic origins

Acc. No.	Accession Name	Origin	Gene bank
001700	RICE SORGHUM	UNKNOWN	ICRISAT
005363		UNKNOWN	RARS/AR
005364		UNKNOWN	RARS/AR
005365		UNKNOWN	RARS/AR
006004	ARVAL	FRANCE	SEIC
006005	ARGENCE	FRANCE	SEIC
006006	ROCE	ITALY	SEIC
006007	SOAVE	ITALY	SEIC
006008	VESPA	ITALY	SEIC
006010	MN 150	ITALY	SEIC
006012	SOAVE	ITALY	SEIC
006013	WONDER DWARF	ITALY	SEIC
006014	GAMBELLA 107	ETHIOPIA	SEIC
008911	IS 1159	UNKNOWN	RARS/MI
008912	IS 2311	UNKNOWN	RARS/MI
008913	IS 2341	UNKNOWN	RARS/MI
008914	IS 2432	UNKNOWN	RARS/MI
008915	IS 3525	UNKNOWN	RARS/MI
008916	IS 4010	UNKNOWN	RARS/MI
008917	IS 8344	UNKNOWN	RARS/MI
008918	IS 9371	UNKNOWN	RARS/MI
008940	OOO234	UNKNOWN	RARS/MI
008941	PJH 53	UNKNOWN	RARS/MI
008942	PJH 62	UNKNOWN	RARS/MI
008943	93064	UNKNOWN	RARS/MI
008944	93050	UNKNOWN	RARS/MI
008945	94002	UNKNOWN	RARS/MI
008946	94004	UNKNOWN	RARS/MI
008947	PSLL 8320	UNKNOWN	RARS/MI
008948	OOO745	UNKNOWN	RARS/MI
008949	IS 2941	UNKNOWN	RARS/MI

008952	ICSV - 112	UNKNOWN	RARS/MI
008953	ICSV - 745	UNKNOWN	RARS/MI
008954	ICSV 239	UNKNOWN	RARS/MI
008955	ICSV 93050	UNKNOWN	RARS/MI
008956	ICSV 758	UNKNOWN	RARS/MI
008957	ICSV 93028	UNKNOWN	RARS/MI
008958	ICSV 93046	UNKNOWN	RARS/MI
008959	ICSV 96007	UNKNOWN	RARS/MI
008960	ICSV 95076	UNKNOWN	RARS/MI
008961	A-2267-2	UNKNOWN	RARS/MI

Table 3: Details of SSR primers used to determine the genetic diversity and population structure of sorghum.

Primer	Sequence of forward primer (5'to 3')	Sequence of reverse primer (5' to 3')	Annealing Temp. (^{0}C)
Xtxp302	TAGGTTCTGGACCACTTTTCTTTTTGTGTT	GAATCAACTATGTGCTTGCATTGTGCT	55
Xtxp32	AGAAATTCACCATGCTGCAG	ACCTCACAGGCCATGTCG	60
Xtxp46	GGGCAATCTTGATGGCGACAT	AGGTGTGGCTCGGGGAGAAC	55
Xtxp297	GACCCATATGTGGTTTAGTCGCAAAG	GCACAATCTTCGCCTAAATCAACAAT	55
Xtxp266	GTTGTCTAGTATAGCAAGGTGGG	ATAATAGTAGATGCGTGTCAAAAGAA	55
Xtxp12	AGATCTGGCGGCAACG	AGTCACCCATCGATCATC	55
Xtxp21	GAGCTGCCATAGATTTGGTCG	ACCTCGTCCCACCTTTGTTG	60
Xtxp136	GCGAATAGCATCTTACAACA	ACTGATCATTGGCAGGAC	55
Xtxp265	GTCTACAGGCGTGCAAATAAAA	TTACCATGCTACCCCTAAAAGTGG	55
Xtxp17	CGGACCAACGACGATTATC	ACTCGTCTCACTGCAATACTG	55
Xtxp295	AAATCATGCATCCATGTTCGTCTTC	CTCCCGCTACAAGAGTACATTCATAGCTTA	55
Xtxp210	CGCTTTTCTGAAAATATTAAGGAC	GATGAGCGATGGAGGAGAG	55
Xtxp321	TAACCCAAGCCTGAGCATAAGA	CCCATTCACACATGAGACGAG	55
Xtxp258	CACCAAGTGTCGCGAACTGAA	GCTTAGTGTGAGCGCTGACCAG	55
Xtxp230	GCTACCGCTGCTGCTCT	AGGGGGCATCCAAGAAAT	55
Xtxp141	TGTATGGCCTAGCTTATCT	CAACAAGCCAACCTAAA	55

Table 4: Summary statistics of SSR markers used to study sorghum germplasm accessions.

Marker	Major Allele Frequency	No. of Genotypes	No. of Alleles	Availability	Gene Diversity	Heterozygosity	PIC
Xtxp32	0.219	15.000	7.000	0.950	0.819	0.298	0.793
Xtxp46	0.266	12.000	10.000	0.783	0.802	0.511	0.774
Xtxp302	0.233	17.000	11.000	0.750	0.845	0.511	0.826
Xtxp297	0.260	20.000	14.000	0.833	0.861	0.580	0.848
Xtxp266	0.315	22.000	13.000	0.900	0.835	0.426	0.819
Xtxp21	0.265	17.000	12.000	0.817	0.824	0.490	0.802
Xtxp12	0.287	17.000	12.000	0.783	0.816	0.383	0.793
Xtxp136	0.300	11.000	7.000	0.750	0.783	0.400	0.749
Xtxp17	0.207	19.000	9.000	0.683	0.853	0.342	0.836
Xtxp265	0.173	29.000	17.000	0.917	0.887	0.673	0.877
Xtxp295	0.333	4.000	4.000	0.700	0.694	0.000	0.632
Xtxp321	0.245	18.000	11.000	0.817	0.830	0.347	0.809
Xtxp210	0.271	14.000	10.000	0.800	0.811	0.250	0.786
Xtxp258	0.225	19.000	13.000	0.817	0.838	0.469	0.818
Xtxp230	0.322	13.000	10.000	0.750	0.777	0.222	0.744
Xtxp141	0.186	22.000	10.000	0.850	0.864	0.490	0.849
Mean	**0.257**	**16.800**	**10.600**	**0.806**	**0.821**	**0.400**	**0.797**

Abbreviation: PIC – polymorphic information content

Table 5: Summary of Nei's standard genetic distances with respect to clusters resulted in cluster analysis of sorghum accessions (SD=Standard deviation).

Standard Genetic distance	Cluster I	Cluster II	Cluster III	Cluster IV
Degrees of freedom	14	14	14	14
Mean	0.3	0.6	0.5	0.4
Standard Deviation	0.08	0.27	0.19	0.09
Standard Error	0.02	0.07	0.05	0.02
Minimum	0.073	0.000	0.184	0.125
Maximum	0.506	1.000	1.000	0.568

Figures

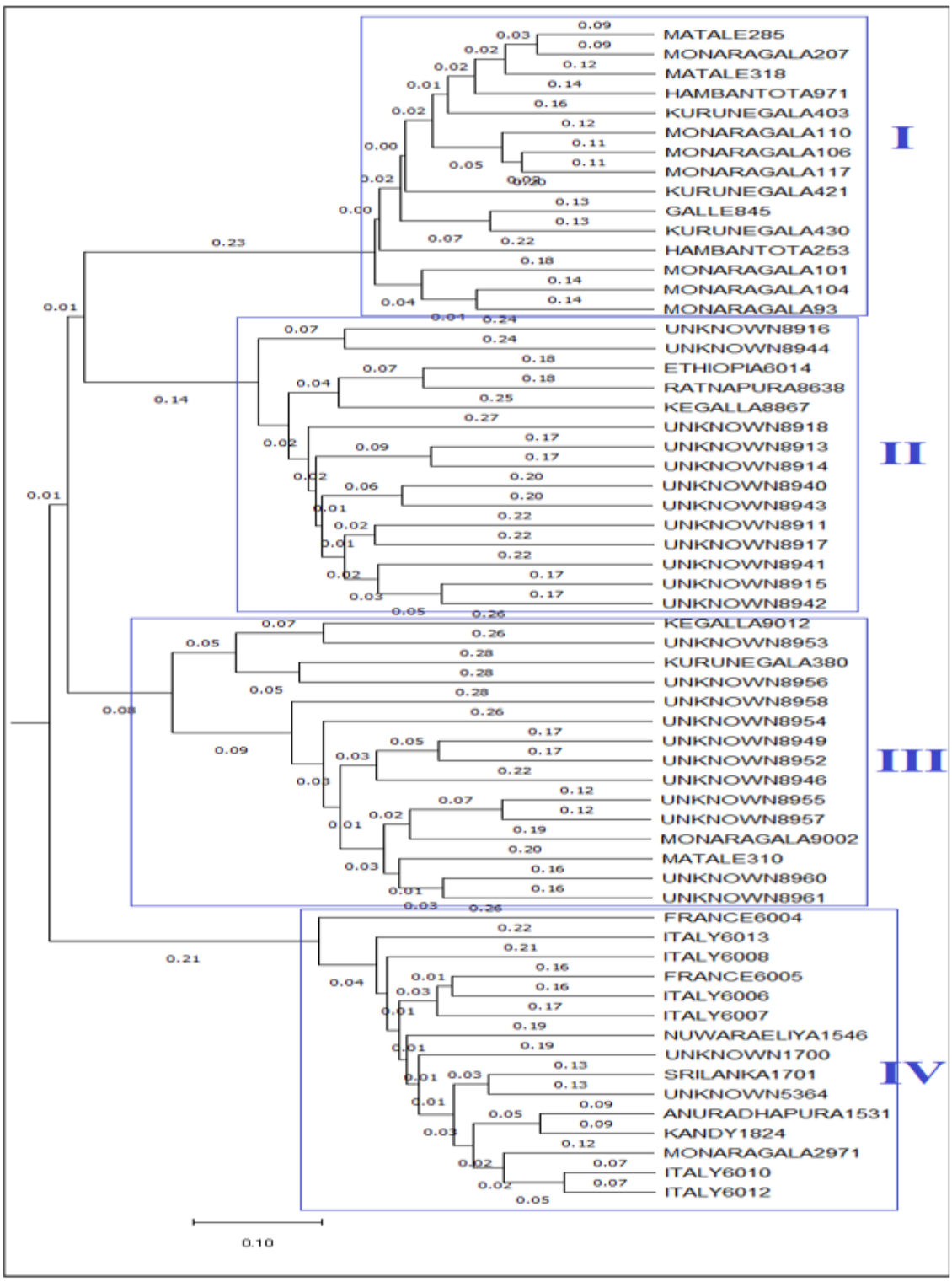

Figure 1

Phylogenetic tree based on the Nei's standard genetic distances of 60 sorghum germplasm accessions.

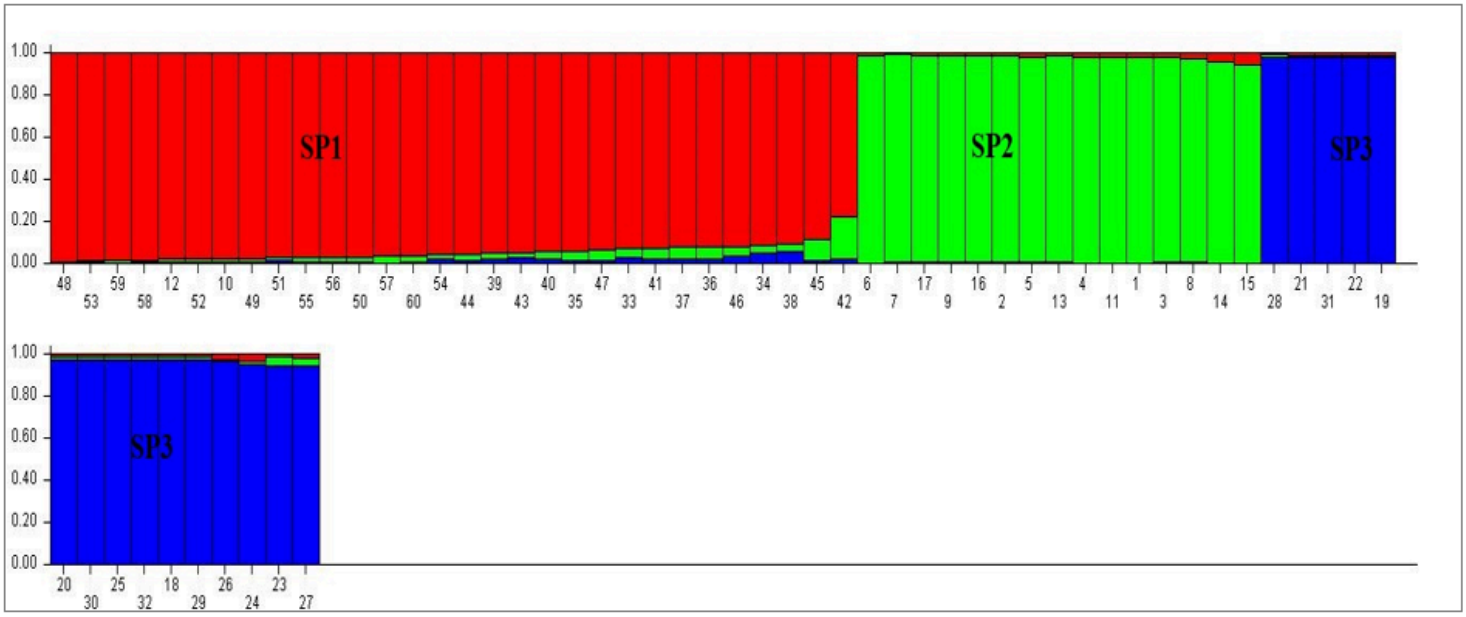

Figure 2

Three subpopulations of 60 sorghum germplasm accessions assessed by STRUCTURE.

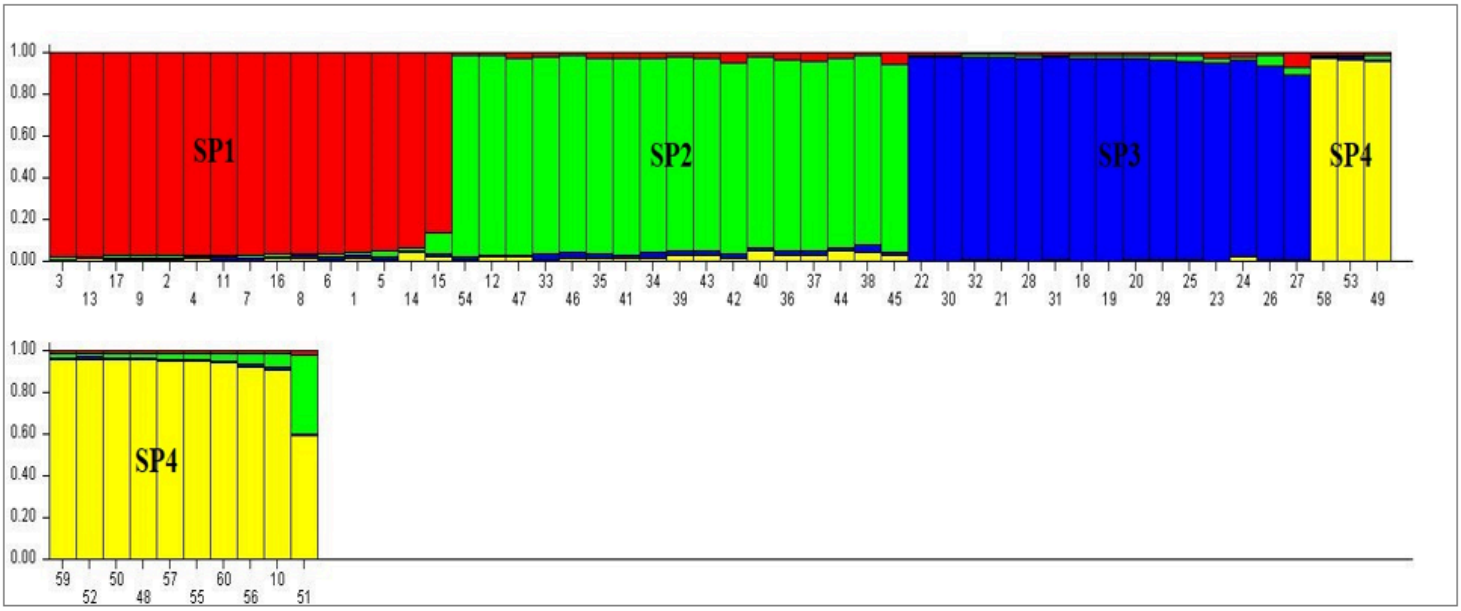

Figure 3

Four subpopulations of 60 sorghum germplasm accessions assessed by STRUCTURE.