Genetic diversity assessment of sorghum (*Sorghum bicolor* (L.) Moench) accessions using single nucleotide polymorphism markers

G. Afolayan[1,2]* [iD], S. P. Deshpande[3]*, S. E. Aladele[1], A. O. Kolawole[2,4], I. Angarawai[5], D. J. Nwosu[1], C. Michael[1], E. T. Blay[2] and E. Y. Danquah[2]

[1]*National Center for Genetic Resources and Biotechnology (NACGRAB), Ibadan, Nigeria, [2]West African Center for Crop Improvement, University of Ghana, Legon, Ghana, [3]International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, [4]Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria and [5]International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Kano, Nigeria*

Abstract

Sorghum (*Sorghum bicolor* (L.) Moench) is an important resource to the national economy and it is essential to assess the genetic diversity in existing sorghum germplasm for better conservation, utilization and crop improvement. The aim of this study was to evaluate the level of genetic diversity within and among sorghum germplasms collected from diverse institutes in Nigeria and Mali using Single Nucleotide Polymorphic markers. Genetic diversity among the germplasm was low with an average polymorphism information content value of 0.24. Analysis of Molecular Variation revealed 6% variation among germplasm and 94% within germplasms. Dendrogram revealed three groups of clustering which indicate variations within the germplasms. Private alleles identified in the sorghum accessions from National Center for Genetic Resources and Biotechnology, Ibadan, Nigeria and International Crop Research Institute for the Semi-Arid Tropics, Kano, Nigeria shows their prospect for sorghum improvement and discovery of new agronomic traits. The presence of private alleles and genetic variation within the germplasms indicates that the accessions are valuable resources for future breeding programs.

Keywords: accessions, genetic diversity, single nucleotide polymorphism, sorghum

Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) is an important cereal crop in dryland regions of the world (National Research Council, 1996) and the main staple crop that is adapted to various agro-climatic regions of the world (Smith and Frederiksen, 2000). It is cultivated under several biotic and abiotic production constraints (National Research Council, 1996; Smith and Frederiksen, 2000; Ejeta, 2007; AATF, 2011). Sorghum is adapted to a wide range of climatic gradients and has high variation for flowering time, plant and inflorescence morphology (Morris *et al*., 2013; Lasky *et al*., 2015; Zhang *et al*., 2015). Africa is considered as the continent of domestication and a centre of diversity for cultivated sorghum (Harlan and De Wet, 1971; Li *et al*., 2010), with two centres overlapping with the borders of Nigeria (Harlan, 1971, 1992). However, Nigeria is the largest producer of sorghum grain (Nzeka and Akhidenor, 2018).

In order to obtain an efficient germplasm management and utilization, accurate genetic fingerprinting and genotype selection in crop improvement, the information on genetic diversity is very important and has been successfully employed (Engles *et al*., 2002; Bucheyekei *et al*., 2009).

*Corresponding authors. E-mail: ogo246@yahoo.com; s.deshpande@cgiar.org

The extent of genetic diversity within a species is mostly used to measure the level of species adaptability because genetic diversity is advantageous for survival in an erratic environmental condition (Rao and Hodgkin, 2002; Govindaraj *et al*., 2015). Genetic variation within a population is the foundation for selection in plant breeding and is also fundamental to plant improvement (Mohammadi and Prasanna, 2003). Therefore the conservation of plant genetic diversity is pivotal to human food security (Rao and Hodgkin, 2002).

Genetic diversity among crop species can be estimated by using morphological, biochemical and molecular markers (Geleta and Labuschagne, 2005; Mehmood *et al*., 2008). Although with morphological characters, the measurement of genetic diversity is much easier, however, these traits are highly affected by the environment (Tuinstra *et al*., 1996; Van Beuningen and Busch, 1997; Abdi *et al*., 2002). Previous researches have reported diversity assessment in sorghum using Random Amplified Polymorphism DNA (Agrama and Tuinstra, 2003; Kanbar *et al*., 2019), Restricted Fragment Length Polymorphism (Deu *et al*., 1994; Smith *et al*., 2000), Simple Sequence Repeat (Smith *et al*., 2000; Agrama and Tuinstra, 2003; Casa *et al*., 2005; Deu *et al*., 2006; Wang *et al*., 2006) and Single Nucleotide Polymorphism (Murray *et al*., 2009; Hu *et al*., 2019)

Despite all work done on genetic diversity studies in sorghum, the genetic diversity of Nigerian sorghum has not been well fully characterized (Olatoye *et al*., 2018). Therefore additional research needs to be done on Nigerians sorghum genetic diversity for conservation and sustainable utilization. Thus, a highly informative, codominant, more effective and very potent tool like the Single Nucleotide Polymorphism (SNP) marker (Varshney *et al*., 2007) would be useful. In this study, SNPs markers were used to estimate the genetic diversity and relationships among gene bank sorghum accessions and other sorghum lines obtained from institutions within and outside Nigeria, for the efficient and effective conservation and utilization of the available genetic diversity.

Materials and methods

Plant materials

A total of 214 sorghum accessions consisting of 92 accessions from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Kano, Nigeria, seven breeding lines from Institute for Agricultural Research (IAR), Zaria, Nigeria, 11 breeding lines from ICRISAT, Bamako, Mali and 104 accessions for National Center for Genetic Resources and Biotechnology (NACGRAB), Ibadan, Nigeria (online Supplementary Fig. S1) were collected for this study.

Genomic DNA extraction and SNP genotyping

Seeds of all the accessions were shipped to ICRISAT-India and sown on the plant quarantine field. Genomic DNA was isolated from 2 weeks old plants using the Qiagen Inc., Valencia, CA, USA, DNeasy Plant Mini Kit and digested using the thermostable ApeKI restriction enzyme. The accessions were then genotyped using the ICRISAT-CEG (https://www.icrisat.org/ceg/) Genotyping-by-Sequence (GBS) Platform. Raw reads from the sequenced GBS library were called in the GBS analysis pipeline (Tassel version: 3.0.147), an extension to the Java program TASSEL (Bradbury *et al*., 2007), was used to call SNPs from the sequenced library by aligning tags to the BTx623 sorghum genome [*Sorghum bicolor* v.1.0; (Paterson *et al*., 2009)]; This procedure provided 317,279 SNPs covering all the 10 chromosomes of the sorghum genome. Out of these 4980 SNP loci having a minimum of 0.05 allele frequency and no missing data for 157 sorghum accessions (online Supplementary Table S1), were selected using TASSEL version 4.1.12 and used for analysing the genetic diversity of sorghum accessions in the current study. Map distribution of the SNP loci on the 10 sorghum chromosomes (ch) was 693 in ch1, 630 in ch2, 656 in ch3, 559 in ch4, 580 in ch5, 437 in ch6, 302 in ch7, 367 in ch8, 377 in ch9 and 379 in ch10.

Data analysis

To quantify genetic variation among the 157 sorghum accessions, genetic diversity estimates which include minor allele frequency (MAF), gene diversity, observed and expected heterozygosity, frequency of private alleles, percentage of polymorphic loci and polymorphic information content (PIC) which refers to the relative significance of each marker with respect to the amount of polymorphism according to Botstein *et al*. (1980) were calculated using the Genetic Analysis in Excel (GenAIEx) version 6.4 (Peakall and Smouse, 2012) and PowerMarker (Liu and Muse, 2005) softwares. The Nei's genetic distance (Nei and Li 1987) between accessions was generated by using the PowerMarker software.

Identification and partitioning of the sorghum accessions into genetic groups was performed using two complementary clustering approaches: (i) Dendrogram was constructed based on the Neighbour-Joining (Saitou and Nei, 1987) algorithm according to shared-allele distance and visualized using MEGA version 5 (Tamura *et al*., 2011). (ii) Structure version 2.3.3 software (Pritchard *et al*., 2000; Falush *et al*., 2003) was used for the Bayesian clustering of the 157 Sorghum accessions. A burn-in period of 50,000 and a Markov chain Monte Carlo (MCMC) replications set at 50,000 was used. Three runs were performed

by setting the number of clusters (*K*) from 1 to 10. The optimal subpopulation model was investigated in several ways: (a) by applying the informal pointers (i.e. geographical origin) proposed by Pritchard *et al.* (2000) (b) by considering Δ*K*, a second order rate change with respect to *K*, as defined in Evanno *et al.* (2005), and (c) by plotting the log probability *L*(*K*) and Δ*K* of the data overruns, as implemented in STRUCTURE HARVESTER (Earl, 2012) and thus the most likely value of *K* determined. Analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992) and the fixation index (F_{ST}) used for population differentiation and level of significance for variance component estimates were computed using the GenAlEx software.

Results

Genetic diversity

From the 4980 informative SNPs markers, 0.08% had a MAF < 0.05, 38.88% had a MAF ⩽ 0.2, approximately 50% of the markers had a MAF > 0.2, and 8.78% showed almost equal allele frequencies (with MAF of 0.5) (Fig. 1).

Accessions from NACGRAB-Ibadan, Nigeria showed the highest values for Nei's gene diversity and PIC, followed by ICRISAT-Kano, Nigeria and ICRISAT-Bamako, Mali (Table 1).

We identified 174 private alleles (MAF > 0.01) only from ICRISAT-Kano, Nigeria and NACGRAB-Ibadan, Nigeria sorghum germplasms. Since accessions from IAR-Zaria, Nigeria and ICRISAT-Bamako, Mali had fewer than 10 accessions, we did not estimate the proportion of private alleles. The observed heterozygosity (Ho) within accessions obtained from ICRISAT-Kano, Nigeria, ranged from 0.00 to 0.98 with an overall mean of 0.26. Accessions from IAR-Zaria, Nigeria had observed heterozygosity ranging from 0.00 to 1.00 and the overall mean of 0.19. Within accessions obtained from ICRISAT-Bamako, Mali, observed heterozygosity ranged from 0.00 to 1.00 with an overall mean of 0.20, accessions obtained from the NACGRAB-Ibadan, Nigeria gene bank had observed heterozygosity ranged from 0.00 to 0.96 with a mean of 0.23. The expected heterozygosity (He) within accessions obtained from ICRISAT-Kano, Nigeria, ranged from 0.00 to 0.50 with an overall mean of 0.30. Accessions from IAR-Zaria, Nigeria had the expected heterozygosity ranging from 0.00 to 0.50 and the overall mean of 0.28. Within accessions obtained from ICRISAT-Bamako Mali, expected heterozygosity ranged from 0.00 to 0.50 with an overall mean of 0.30, accessions obtained from NACGRAB-Ibadan, Nigeria gene bank, had expected heterozygosity ranged from 0.00 to 0.50 with a mean of 0.30. The percentage of polymorphic loci obtained from ICRISAT-Kano, Nigeria and NACGRAB-Ibadan, Nigeria also reflects greater genetic variation
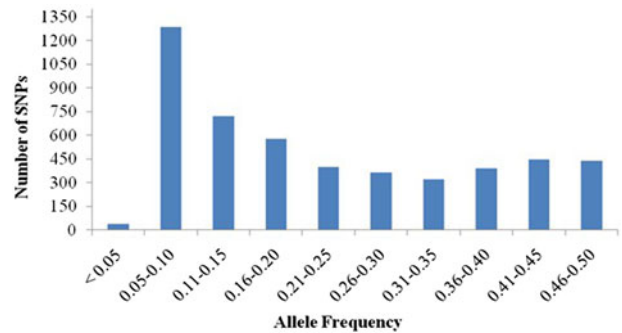

Fig. 1. Frequency distribution of minor allele frequency among 157 sorghum accessions.

compared to accessions from ICRISAT-Bamako, Mali and IAR-Zaria, Nigeria.

Genetic differentiation within and among the four germplasms

Molecular variance component among and within the accessions was compared to serve as a complementary indicator for genetic diversity. Analysis of molecular variance (AMOVA) shows that genetic variation within accessions accounted for 96% of the total variation, while the variation among populations from the four research institutes was 4% (Table 2). In spite of the very low variation among the populations; the fixation index (F_{ST}) of 0.04 was highly significant ($P < 0.001$).

Genetic distance and cluster analysis of 157 sorghum accessions

The pair-wise genetic distances between the germplasms which was calculated using PowerMarker software based on the allele frequency revealed the relatedness of accessions on a germplasm × germplasm basis. From the matrices given in Table 3, accessions from NACGRAB-Ibadan, Nigeria and IAR-Zaria, Nigeria were the most distant followed by ICRISAT-Kano, Nigeria and IAR-Zaria, Nigeria. While accessions from NACGRAB-Ibadan, Nigeria and ICRISAT-Kano, Nigeria were the closest at 0.01.

In order to gain further insight into the genetic diversity among different collection of sorghum, a neighbour joining tree method which was based on shared-allele genetic distances was constructed. The 157 sorghum accessions were classified into three genetically distinct groups (Fig. 2). The first group composed of 7 accessions of which 2 were from NACGRAB-Ibadan, Nigeria and the other 5 from ICRISAT-Kano, Nigeria. The second group was divided into two subgroups (2A and 2B). The subgroup 2A consisted of 17 accessions of which 5 were from NACGRAB-Ibadan, Nigeria and the other 12 were

Table 1. Genetic diversity generated by 4980 SNP markers from 157 sorghum accessions

Germplasm	Sample Size	Frequency of private alleles	Average Diversity	Gene	Average Ho	Average He	Average PIC	% P
ICRISAT-Nigeria	63	0.05	0.34		0.26	0.30	0.24	100
IAR-Zaria	7	0.00	0.29		0.19	0.28	0.22	74
ICRISAT-Mali	9	0.00	0.30		0.20	0.30	0.24	80
NACGRAB-Ibadan	78	0.17	0.37		0.23	0.30	0.25	100

Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphic information content, and % P, percentage of polymorphic loci.

Table 2. Analysis of molecular variance (AMOVA) and Wright's fixation index (F_{ST}) for 157 sorghum accessions based on by 4980 SNP markers

Level of variation	df	Sum of squares	Mean squares	Variance components	Percentage of variation	F_{ST}
Among accessions	3	391.54	130.51	2.39	4	0.04***
Within accessions	153	8714.60	56.96	56.96	96	
Total	156	9106.13		59.35	100	

***P-value < 0.001.

Table 3. Genetic distance matrices between germplasms calculated according to Nei and Li (1987) for the 157 sorghum accessions

	NACGRAB-Ibadan	ICRISAT-Nigeria	ICRISAT-Mali	IAR-Zaria
NACGRAB-Ibadan	0			
ICRISAT-Nigeria	0.01	0		
ICRISAT-Mali	0.16	0.16	0	
IAR-Zaria	0.28	0.27	0.06	0

from ICRISAT-Kano, Nigeria. Subgroup 2B consisted of the majority of all analysed accession (129) and was separated into smaller groups (2Bi and 2Bii). 2Bi composed of 59 accessions with 42 from NACGRAB-Ibadan, Nigeria and 17 from ICRISAT-Kano, Nigeria. The subgroup 2Bii composed of 70 accessions with all lines from ICRISAT-Bamako, Mali, IAR-Zaria, Nigeria and others from NACGRAB-Ibadan, Nigeria and ICRISAT-Kano, Nigeria. The third major group represented 3 accessions from NACGRAB-Ibadan, Nigeria and one accession from ICRISAT-Kano, Nigeria. The model-based programme STRUCTURE was used to determine genetic relationship among the 157 sorghum accessions. STRUCTURE was run for $K = 1$–10 based on the distribution of alleles at 4980 SNP loci. The number of populations $K = 2$ revealed greater variability of maximum likelihood (Ln $P(D)$) among different tested K values than among repeated runs and was considered as the optimal number of populations (online Supplementary Fig. S2). The analysis of population structure revealed two major

subpopulations (Fig. 3). Majority of the accessions were categorized as having admixed ancestry. We observed 77 accessions with evidence of admixture (41,29,6,1 accessions from ICRISAT-Kano, Nigeria, NACGRAB-Ibadan, Nigeria, ICRISAT-Bamako Mali and IAR-Zaria, Nigeria, respectively). The accessions represented by red colour include 19 from ICRISAT-Kano, Nigeria and 41 from NACGRAB-Ibadan, Nigeria. The remaining 19 accessions from the four research institutes were represented by the green colour.

Discussion

The description of genetic variability in a germplasm is vital for proper application and management of genetic materials (Thudi and Fakrudin, 2011). The average expected heterozygosity among the 157 sorghum accessions in this study is low. Previous research reported expected heterozygosity ranging from 0.6 to 0.8 among reference set of
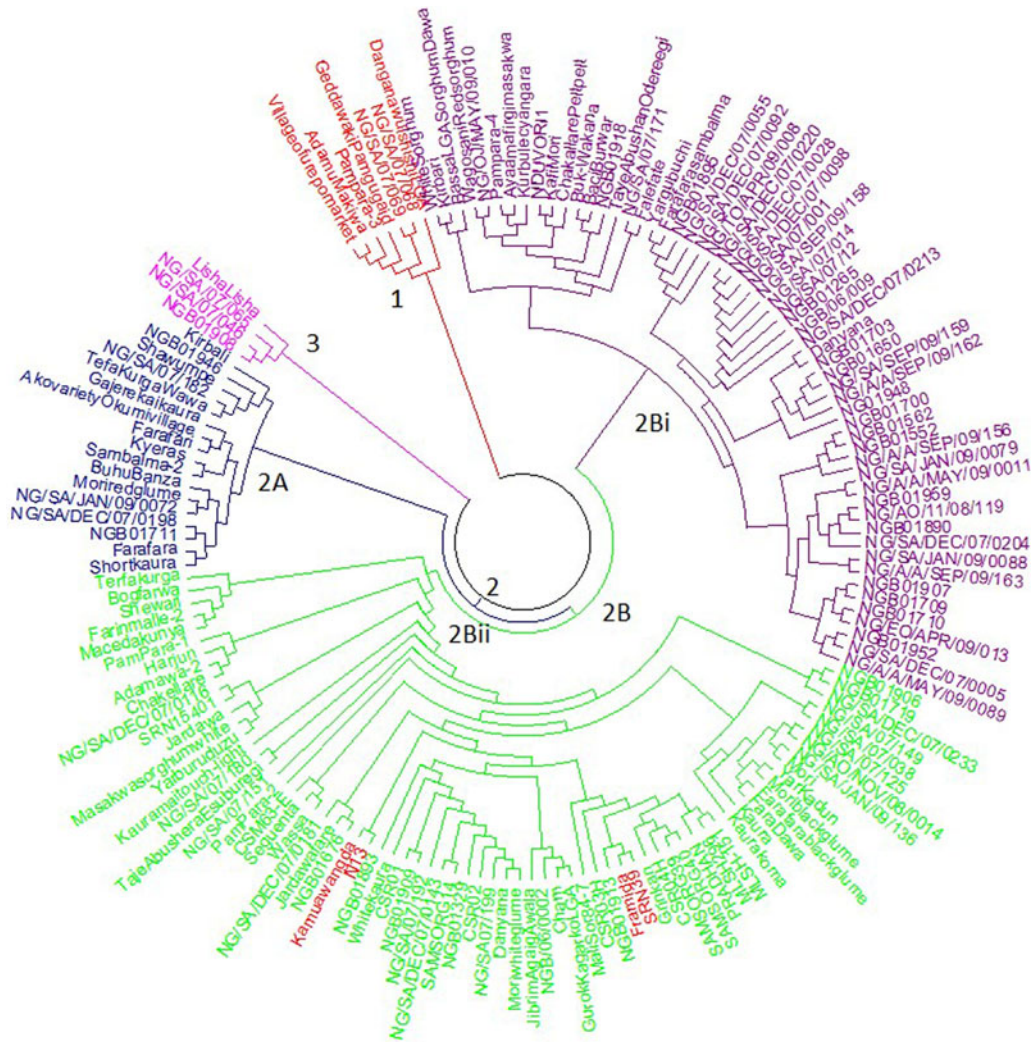
Fig. 2. Dendrogram of 157 sorghum accessions based on the Nei's genetic distance calculated from data of 4981 SNP markers, using the neighbour joining algorithm as the clustering method. Accessions in cluster 1 ▮, Accessions in cluster two with subclusters 2A ▮, 2Bi ▮, and 2Bii ▮. Accessions in Cluster 3 ▮.
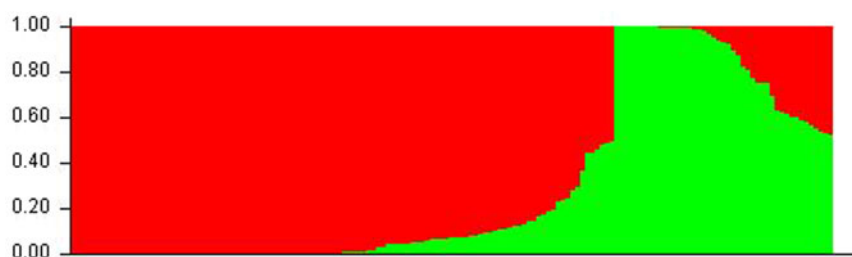

Fig. 3. Population structure of 157 sorghum accessions.

sorghum using a small set of highly polymorphic Simple Sequence Repeat (SSR) markers (Dje *et al.*, 1999, Uptmoor *et al.*, 2003, Ngugi and Onyango, 2012, Ramu *et al.*, 2013). However, the expected heterozygosity recorded in this study corroborate the reports of Ren *et al.*

(2013) and Dao *et al.* (2014) in which SNPs markers were employed and expected heterozygosity ranged between 0.2 and 0.3. The observed heterozygosity reported in the present study is 0.24 which is much higher than what was reported by Ramu *et al.* (2013). In self-pollinating crops like

sorghum in which the accessions have been kept over several generations by continuous regeneration *via* selfing, it is expected that the observed heterozygosity will be very small. However, due to the bi-allelic nature of SNP markers, the gene diversity or expected heterozygosity and also other values like the PIC cannot be more than 0.50, which is contrary to what is obtainable for multi-allelic markers such as SSRs in which the maximum can approach 1 (Ren *et al.*, 2013). According to Smith *et al.* (2000), the PIC value of markers can give better or more information on variability in a germplasm by considering the number of alleles, then the relative frequencies of the individual allele. The average PIC value (0.24) obtained from this study was similar to other studies where SNPs markers were used (Lu *et al.*, 2009; Hao *et al.*, 2011; Ren *et al.*, 2013; Dao *et al.*, 2014). Thus comparing the four germplasms, with respect to the PIC values and also the heterozygosity values, across the germplasms, all the accessions seem to have a relatively narrow genetic base with accessions from IAR-Zaria, Nigeria showing more similarities and those from NACGRAB-Ibadan, Nigeria being more diverse. We detected private alleles in two germplasm (MAF > 0.01). The distribution of private alleles among accessions suggests these alleles could represent introgression from neighbouring research institutes. These private alleles may be useful in sorghum improvement for the discovery of new agronomic traits.

With sorghum being a self-pollinating crop, it is expected that the percentage variation within the population will be lower than among the population (Hamrick, 1983). Our study revealed a low (6%) percentage of genetic variation among the sorghum accession while 94% genetic variation was found within the germplasms. These results negate those of Manzelli *et al.* (2007) and Motlhaodi *et al.* (2014) who reported higher variation among accessions than within accessions but was similar to those of Tadesse and Feyissa (2013) and Tesfamichael *et al.* (2014). This indicates that the accessions used in this study are not under any selection for crop improvement and there is probably continuous flow of genes among accessions from the four research institutes resulting in similarity among accessions. Also, when there is a high level of seed exchange and open sorghum marketing across localities, there is a high probability of seed mixture which narrows down the overall diversity and in turn a high percentage of the overall genetic variation being present within germplasms than among (Tesfamichael *et al.*, 2014). Additionally, the population size of the germplasm collections from ICRISAT-Bamako, Mali (9) and IAR-Zaria, Nigeria (7), were very low compared to those from NACGRAB-Ibadan, Nigeria and the ICRISAT-Kano, Nigeria. The small sample sizes from those two locations may have also affected the percentage of genetic variation among the accessions (Pimm *et al.*, 1989). It is therefore important for plant breeders to retain

the ideal population size in germplasm conservation for crop improvement. As populations are not genetically the same, in order to sufficiently estimate the genetic diversity, many individuals need to be sampled (Govindaraj *et al.*, 2015).

There was high diversity within each accession, mostly within the NACGRAB-Ibadan, Nigeria and ICRISAT-Kano, Nigeria accessions with the existence of private alleles. This could be very useful to sorghum breeding program since such alleles may be linked to important traits. The ICRISAT-Kano, Nigeria and NACGRAB-Ibadan, Nigeria accessions were generally clustered together. This clustering could be because they were collected from very close regions as most of these accessions were collections obtained as a result of scientist exploration activities. Most of the accessions clustered together in cluster 2, and were further divided into subclusters 2A, 2Bi and 2Bii, indicating that there is still some degree of variability among genotypes of the major cluster. The subclusters, differentiated between accessions from NACGRAB-Ibadan, Nigeria and ICRISAT-Kano, Nigeria with the majority of the accessions from ICRISAT-Bamako, Mali, being grouped together with those from ICRISAT-Kano, Nigeria. Also, the farmers varieties obtained from IAR-Zaria, Nigeria were grouped in the subclusters 2Bii with most of the other released varieties obtained from NACGRAB-Ibadan, Nigeria gene bank and also with the ICRISAT-Kano, Nigeria and ICRISAT-Bamako, Mali accessions. This Indicates genetic relatedness among accessions which may be due to mixtures of seeds from released varieties with farmers' cultivars.

Furthermore, the smaller clustering of accessions within subclusters could be due to different naming systems of the same accessions in the different institutions such as ICRISAT-Kano, Nigeria and NACGRAB-Ibadan, Nigeria. Subcluster 2Bii with the Indian *durra Striga* resistant accession (N13) clustered with most of the NACGRAB-Ibadan, Nigeria accessions and in close proximity with one of the ICRISAT-Kano, Nigeria accession (Kamu awangda). Also, SRN39 and Framida obtained from ICRISAT-Bamako, Mali, which has also been classified as resistant to *Striga* were clustered together with IAR-Zaria, Nigeria accessions (SAMSORG39 and SAMSORG40). These reveal the relatedness of these accessions through common ancestry or sharing of same genetic background. Similar clustering pattern observed in this study has been reported previously by Tesfamichael *et al.* (2014). Clustering of accessions obtained from ICRISAT-Bamako, Mali, in close proximity with accessions from IAR-Zaria, Nigeria also shows the extent of germplasm exchange and common gene pool sharing between these neighbouring countries and geographical zones. Grouping pattern in this research is, therefore, indicating that there could be seed exchange and pollen dispersal which causes similarity between nearby

populations. There was a clear separation of the 157 sorghum accessions into two major groups with the appearance of similarity of the genetic information among accessions. The results from the pair-wise genetic distances and AMOVA corroborate that of the structure analysis, revealing lower variation among the accessions from the four research institutes. We detect concordance between our dendrogram and population structure results. From these results, SNP markers have proven to be very promising for monitoring genetic variability. Understanding the population structure of the Nigerian sorghum accessions is vital to establish new strategies useful for the preservation and utilization its genetic diversity.

Our results suggest that the sorghum accessions from different geographic origins or institutions could come from the same gene pool and may not be useful for continual sorghum improvements. Therefore, when planning for a new germplasm collection and conservation, it is imperative to characterize the available germplasm before commencing on a breeding program. Conclusively, our study revealed high genetic diversity within the sorghum germplasm. Hence, these sorghum accessions will be protected from genetic vulnerability to biotic and abiotic stresses in the growing environments and breeders will be guided on how to utilize and conserve the germplasm.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S1479262119000212

Acknowledgements

References

Abdi A, Bekele E, Asfaw Z and Teshome A (2002) Patterns of morphological variation of sorghum [Sorghum bicolor (L.) Moench] landraces in qualitative characters in North Showa and South Welo, Ethiopia. *Hereditas* 137: 161–172.

African Agricultural Technology Foundation (AATF) (2011) *Feasibility Study on Striga Control in Sorghum*. Nairobi: The African Agricultural Technology Foundation, 11–20.

Agrama HA and Tuinstra MR (2003) Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology* 2: 334–340.

Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of a genetic linkage in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314–331.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y and Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics (Oxford, England)* 23: 2633–2635.

Bucheyekei TL, Gwanama C, Mgonja M, Chisi M, Folkertsma R and Mutegi R (2009) Genetic variability characterization of Tanzania sorghum landraces based on simple sequence repeats (SSRs) molecular and morphological markers. *African Crop Science Journal* 17: 71–86.

Casa AM, Mitchell SE, Hamblin MT, Sun H, Bowers JE, Paterson AH, Aquadro CF and Kresovich S (2005) Diversity and selection in Sorghum: simultaneous analyses using simple sequence repeats. *Theoretical and Applied Genetics* 111: 23–30.

Dao A, Sanou J, Mitchell SE, Gracen V and Danquah EY (2014) Genetic diversity among INERA maize inbred lines with single nucleotide polymorphism (SNP) markers and their relationship with CIMMYT, IITA, and temperate lines. *BMC Genetics* 15: 1–14.

Deu M, Gonzalez-de-Leon D, Glaszmann JC, Degremont I, Chantereau J, Lanaud C and Hamon P (1994) RFLP diversity in cultivated sorghum in relation to racial differentiation. *Theoretical and Applied Genetics* 88: 838–844.

Deu M, Rattunde F and Chantereau J (2006) A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* 49: 168–180.

Dje Y, Forcioli D, Ater M, Lefèbvre C and Vekemans X (1999) Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. *Theoretical and Applied Genetics* 99: 157–163.

Earl AE (2012) Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

Ejeta G (2007) Breeding for resistance in sorghum: exploitation of an intricate host–parasite biology. *Crop Science* 47 (Supplement_3): S-216.

Engles JMM, Rao VR, Brown AHD and Jackson MT (2002) *Managing Plant Genetic Diversity*. Wallingford, UK: CABI, pp. 487.

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.

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitocondrial DNA restriction sites. *Genetics* 131: 479–491.

Falush D, Stephens M and Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

Geleta N and Labuschagne MT (2005) Qualitative traits variation in sorghum (Sorghum bicolor (L.) Moench) germplasm from, eastern highlands of Ethiopia. *Biodiversity and Conservation* 14: 3055–3064.

Govindaraj M, Vetriventhan M and Srinivasan M (2015) Importance of genetic diversity assessment in crop plants

and its recent advances: an overview of its analytical perspectives. *Genetics Research International* 2015: 1–14.

Harlan JR (1971) Agricultural origins: centers and noncenters. *Science* 174: 468–474.

Harlan JR (1992) *Crops and Man*, 2nd edn. Madison, WI: American Society of Agronomy, pp. 284.

Harlan JR and de Wet JMJ (1971) Toward a rational classification of cultivated sorghums. *Crop Science* 12: 172–176.

Hamrick JL (1983) The distribution of genetic variation within and among natural plant populations. In: Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas WL (eds) *Genetics and Conservation*. California: Benjamin-Cummings, pp. 335–348.

Hao Z, Li X, Xie C, Weng J, Li M, Zhang D and Zhang S (2011) Identification of functional genetic variations underlying drought tolerance in maize using SNP markers. *Journal of Integrative Plant Biology* 53: 641–652.

Hu Z, Olatoye MO, Marla S and Morris GP (2019) An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *The Plant Genome* 12: 1–15.

Kanbar A, Shakeri E, Alhajturki D, Horn T, Emam Y, Tabatabaei SA and Nick P (2019) Morphological and molecular characterization of sweet, grain and forage sorghum (Sorghum bicolor L.) genotypes grown under temperate climatic conditions. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 1–10. doi: 10.1080/11263504.2019.1569568.

Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J …and Morris GP 2015 Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances* 1: e1400218.

Li R, Zhang H, Zhou X, Guan Y, Yao F, Song G, Wang J and Zhang C (2010) Genetic diversity in Chinese sorghum landraces revealed by chloroplast simple sequence repeats. *Genetic Resources and Crop Evolution* 57: 1–15.

Liu K and Muse SV (2005) Powermarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics (Oxford, England)* 21: 2128–2129.

Lu Y, Yan J, Guimaraes GT, Taba S, Hao Z, Gao S and Xu Y (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. *Theoretical and Applied Genetics* 120: 93–115.

Manzelli M, Pileri L, Lacerenza N, Benedettelli S and Vecchio V (2007) Genetic diversity assessment in Somali sorghum (Sorghum bicolor (L.) Moench) accessions using microsatellite markers. *Biodiversity and Conservation* 16: 1715–1730.

Mehmood S, Bashir A, Ahmad A, Akram Z, Jabeen N and Gulfraz M (2008) Molecular characterization of regional Sorghum bicolor varieties from Pakistan. *Pakistan Journal of Botany* 40: 2015–2021.

Mohammadi SA and Prasanna BM (2003) Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Science* 43: 1235–1248.

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD and Kresovich S (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences* 110: 453–458.

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: 35–43.

Murray SC, Rooney WL, Hamblin MT, Mitchell SE and Kresovich S (2009) Sweet sorghum genetic diversity and association mapping for brix and height. *The Plant Genome* 2: 48–62.

National Research Council (1996) *Lost Crops of Africa: Volume I: Grains*. Washington, DC: National Academy Press, pp. 408.

Nei M and Li WH (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press, pp. 512.

Ngugi K and Onyango CM (2012) Analysis of the molecular diversity of Kenyan sorghum germplasm using microsatellites. *Journal of Crop Science and Biotechnology* 15: 189–194.

Nzeka U and Akhidenor J (2018) Nigeria: Grain and Feed Annual. GAIN Reports, USDA-FAS. https://gain.fas.usda.gov/RecentGAINPublications/GrainandFeedAnnualLagosNigeria.

Olatoye MO, Hu Z, Maina F and Moris GP (2018) Genomic signatures of adaptation to a precipitation gradient in Nigerian sorghum. *G3: Genes Genomes Genetics* 8: 3269–3281.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H and Rokhsar DS (2009) The Sorghum bicolor genome and the diversification of grasses. *Nature* 457: 551–556.

Peakall ROD and Smouse PE (2012) GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Pimm SL, Gittlaman JL, McCracken GF and Gilpin M (1989) Genetic bottlenecks: alternative explanations for low genetic variability. *Trends in Ecology and Evolution* 4: 176–177.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Ramu P, Billot C, Rami JF, Senthilvel S, Upadhyaya HD, Ananda Reddy L and Hash CT (2013) Assessment of genetic diversity in the sorghum reference set using EST-SSR markers. *Theoretical and Applied Genetics* 126: 2051–2064.

Rao VR and Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture* 68: 1–19.

Ren J, Sun D, Chen L, You FM, Wang J, Peng Y and Peng J (2013) Genetic diversity revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. *International Journal of Molecular Sciences* 14: 7061–7088.

Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.

Smith CW and Frederiksen RA (2000) *Sorghum: Origin, History, Technology, and Production*, Vol. 2. New York: John Wiley and Sons, pp. 824.

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M and Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Science* 40: 226–232.

Tadesse H and Feyissa T (2013) Analysis of genetic diversity of sorghum bicolor ssp. bicolor (L.) moench using ISSR markers. *Asian Journal of Plant Sciences* 12: 61–70.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.

Tesfamichael TA, Githiri SM, Kasili RW, Skilton RA, Solomon M and Nyende AB (2014) Genetic diversity analysis of eritrean sorghum (*Sorghum bicolor* (L.) Moench) Germplasm using SSR markers. *Molecular Plant Breeding* 5: 1–12.

Thudi M and Fakrudin B (2011) Identification of unique alleles and assessment of genetic diversity of rabi sorghum accessions using simple sequence repeat markers. *Journal of Plant Biochemistry and Biotechnology* 20: 74–83.

Tuinstra MR, Grote EM, Goldsbrough PB and Ejeta G (1996) Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. *Crop Science* 36: 1337–1344.

Uptmoor R, Wenzel W, Friedt W, Donaldson G, Ayisi K and Ordon F (2003) Comparative analysis on the genetic relatedness of Sorghum bicolor accessions from Southern Africa by RAPDs, AFLPs and SSRs. *Theoretical and Applied Genetics* 106: 1316–1325.

Van Beuningen LT and Busch RH (1997) Genetic diversity among North American spring wheat cultivars: III. Cluster analysis based on quantitative morphological traits. *Crop Science* 37: 981–988.

Varshney RK, Beier U, Khlestkina E, Kota R, Korzun V, Röder M, Graner A and Börner A (2007) Single nucleotide polymorphisms in rye: discovery, frequency and applications for genome mapping and diversity studies. *Theory of Applied Genetics* 114: 1105–1116.

Wang ML, Dean R, Erpelding J and Pederson G (2006) Molecular genetic evaluation of sorghum germplasm differing in response to fungal diseases: rust (*Puccinia purpurea*) and anthracnose (*Collectotrichum graminicola*). *Euphytica* 148: 319–330.

Zhang D, Kong W, Robertson J, Goff VH, Epps E, Kerr A and Paterson AH (2015) Genetic analysis of inflorescence and plant height components in sorghum (Panicoidae) and comparative genetics with rice (*Oryzoidae*). *BMC Plant Biology* 15: 107.