Plant Breeding, **136**, 338–343 (2017)
doi:10.1111/pbr.12479

Genetic diversity of tropical maize inbred lines combining resistance to *Striga hermonthica* with drought tolerance using SNP markers

Wende Abera Mengesha, Abebe Menkir*, Nnanna Unakchukwu, Silvestro Meseka, Adetutu Farinola, Gezahegn Girma and Melaku Gedil

International Institute of Tropical Agriculture (IITA), Ibadan200001, Nigeria; *Corresponding author, E-mail: A.Menkir@cgiar.org

With 2 figures and 5 tables

Received December 28, 2016 / Accepted March 4, 2017
Communicated by T. Lübberstedt

Abstract

Striga hermonthica and drought are the major stresses limiting maize yields in sub-Saharan Africa. The search for diverse maize lines' tolerance to drought and resistance to *S. hermonthica* (DTSTHR) is very crucial for yield improvement in areas affected by the two stresses. Understanding the genetic diversity among the lines is important to develop cultivars resistant to *S. hermonthica* and tolerant to drought. The lines were developed from biparental crosses of drought-tolerant and *Striga*-resistant lines. A total of 128 DTSTHR maize lines were characterized using single-nucleotide polymorphism (SNP) markers. Results of the cluster analysis based on 3297 SNP markers showed four distinct groups consistent with the pedigrees of the lines. Furthermore, model-based analysis also formed the same groups of the DTSTHR lines. Integrating the pedigree information with combining ability and the SNP analyses may provide defined heterotic groups for maize improvement work in West and Central Africa. These results also help breeders to utilize DTSTHR lines present at IITA for developing biparental crosses without disrupting the heterotic groups they have established in their breeding programmes.

Key words: *Striga hermonthica* — genetic diversity — inbred lines — SNP markers

Many African countries produce less maize than what they consume, making them net importers of maize grain for food as this crop provides the bulk of the calories in their diet (Vivek et al. 2009). This is due to the low average maize yield of 1.5 t/ha (FAOSTAT 2012), resulting from pests and disease pressure, low input use and devastating effects of a parasitic weed known as *S. hermonthica*. Most of the maize production zones, particularly in sub-Saharan Africa (SSA), are infested with *Striga*, which causes yield losses of 40–60% in smallholder farmers' fields (Kanampiu and Friesen 2003). Increasing maize production in *Striga*-affected areas can be achieved through development of resistant varieties and hybrids from diverse *Striga*-resistant maize inbred lines developed using conventional breeding methods.

Breeding maize for resistance to *S. hermonthica* and tolerance to drought has been the target for maize breeders at the International Institute of Tropical Agriculture (IITA). Several maize inbred lines with varying levels of resistance to *S. hermonthica* have been developed from different broad-based source populations improved using different recurrent selection schemes (Menkir et al. 2007). Menkir et al. (2005) studied the diversity of inbred lines with different reactions to *S. hermonthica* derived from these broad-based source populations using AFLP and SSR markers and found that the inbred lines were separated into groups based on the source population from which they were derived. Several of these inbred lines were crossed with drought-tolerant lines of diverse origin to create biparental crosses, which were sources of new inbred lines that combine resistance to *S. hermonthica* with tolerance to drought stress. As the heterotic affinities of the diverse parental lines that formed these biparental crosses are unknown, diversity assessment of the resulting drought-tolerant and *S. hermonthica*-resistant inbred lines will be useful to classify them into well-defined heterotic groups.

A heterotic group is a collection of closely related inbred lines which tend to result in vigorous hybrids when crossed with lines from a different heterotic group, but not when crossed to other lines of the same heterotic group (Lee 1995). Molecular markers have been extensively used to define the heterotic patterns of maize inbred lines (Yongsheng et al. 2016) and their use for assessment of the genetic diversity of the new drought-tolerant and *S. hermonthica*-resistant maize inbred lines can play an important role in designing biparental crosses to develop better inbred lines (Mohammadi and Prasanna 2003, Collard and Mackill 2008, Naresh et al. 2009, Semagn et al. 2012, Abdalla et al. 2014). Furthermore, molecular markers can be used in facilitating the selection of potential parents with diverse genetic backgrounds for the development of productive hybrids and open pollinated varieties for areas that are affected by *S. hermonthica* and drought. Hybrids involving parents of diverse origin produce higher grain yields than hybrids formed from lines with the same genetic background (Menkir et al. 2010).

The advances in high-throughput genotyping, including genotyping by sequencing (GBS), coupled with the declining cost of genotyping has prompted routine use of molecular assay for quality control and diversity assessment studies of gene pools in crop plants (Melchinger 1999, Moose and Mumm 2008, Jianbing et al. 2009, Semagn et al. 2012). Single-nucleotide polymorphism (SNP) markers have become the marker system of choice as they are highly polymorphic and have the potential for high-throughput low-cost genotyping. Single-nucleotide polymorphisms have been used for a variety of functions in crop improvement, including genetic diversity analysis, linkage map construction, marker–trait association analysis and marker-assisted selection (MAS). This study was, therefore, conducted to (i) assess the extent of genetic diversity among *S. hermonthica*-resistant and drought-tolerant maize inbred lines from crosses with mixed genetic backgrounds and (ii) classify the lines into heterotic groups using SNP markers.

Materials and Methods

Plant material: A total of 128 maize inbred lines were included in this diversity study. The inbred lines were derived from biparental crosses of inbred lines from different broad-based populations, composites and synthetics selected for resistance to *S. hermonthica* and tolerance to

Table 1: Sources of the *Striga hermonthica*-resistant and drought-tolerant inbred lines used in this study

Line designations	Source population/biparental cross
L001-L010	(1393/Z.Diplo.BC4-19-4-1
L011-L019	ZDiploBC4
L020-L026	(ACR97SYN-Y-S1-24/ACR97TZLComp1-YS155)
L027-L034	(ACR97SYN-Y-S1-79/ACR97TZLComp1-YS155)
L035-L045	(ACR97TZLComp1-YS155/ACR97SYN-Y-S1-76)
L046-L056	(ACRSYN-W-S2-173/TZLComp1C4S1-37-1
L057-L070	(ACRSYN-W-S2-173/TZLComp1C4S1-37-5)
L071	ACRSYN-W
L072-L078	TZLComp1C4
L079-L096	(ZDiploBC4-472-2-2-1-2-3/ZDiploBC4-19-4-1)
L097-L124	IWD-SYN-STR-C3
L125-L126	STR Low Emerg Pool
L127	(1368/B73LPA/1368)
L128	(9450/CM 116/9450)

drought stress (Table 1). Details on the genetic composition of the broad-based source populations of the inbred lines have been described in different publications (Kim 1996, Kling et al. 2000, Menkir et al. 2005, 2007).

DNA extraction: Young leaves were sampled from 3-week-old maize plants in the field. Leaves from four to 15 plants were collected from each line and bulked for genomic DNA isolation. DNA extraction was made using a CTAB-based extraction protocol modified from Saghai-Maroof et al. (1984). In brief, leaf samples were freeze-dried and approximately 1.5 g of leaves for each genotype were cut into tubes and ground into powder using a GENOGRINDER 2000® instrument (BT and C Inc., New Jersey, USA). The ground leaves were homogenized in 600 μl 2% CTAB extraction buffer comprising 1 M pH 7.5 Tris–HCl, 5 M NaCl and 0.5 M pH 8.0 EDTA. The homogenized samples were subsequently incubated for 30 min at 65°C in a water bath. A mixture of chloroform and isoamyl alcohol in the ratio 24 : 1 was added into each tube and the solutions mixed by repeated inversion for 5 min and centrifuged at 2058 g for 15 min. This process was repeated once to ensure that all impurities are removed from the aqueous solution.

The aqueous supernatant was transferred to new tubes and iso-propanol was added to precipitate DNA. The precipitated DNA was pelleted by centrifuging at 3500 rpm for 15 min, and the iso-propanol supernatant was discarded. The DNA pellet was purified by adding 70% ethanol to the pellet and centrifuging at 3500 rpm for 15 min to clean the DNA pellet. The dried pellets were dissolved in 100 μl ultra-pure water. The quality of DNA was checked on 1% agarose gel and quantified using a NANODROP® spectrophotometer (Thermo Fisher Scientific Inc., Denver, CO, USA). The quality of the DNA for GBS (genotyping by sequencing) was further ascertained by digesting the DNA with restriction enzyme Hind III before shipping it for genotyping.

Genotyping-by-sequencing Procedure (GBS): Genotyping by sequencing was used to generate genotype data at the Genomic Diversity Facility (GDF), Cornell University. GBS libraries were prepared and analysed as described by Elshire et al. (2011), using the enzyme ApeKI for digestion and creating a library with unique barcodes for each genotype. 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) and the filtered sequences were aligned to the maize reference genome B73 REFGEN v1 (Schnable et al. 2009) using the Burrows-Wheeler alignment tool (BWA).

Genomic DNA sample of 10 ng/μl concentration and 10 μl quantity was dispensed into 96-well plates separate wells already treated with barcoded adapter and common adapter. DNA samples were digested with *Ape*KI restriction enzyme, a type II restriction endonuclease that recognizes a degenerate 5-bp sequence GCWGC, where W can be A or T and creates a 3-bp overhang. It also has relatively few recognition sites in the major classes of maize retro-transposons and is partially methylation sensitive. The digested DNA is then ligated to the barcoded adapter and common adapter that have been designed to ligate to the sticky ends of the *Ape*KI cut site. The barcoded adapters are unique 4- to 8-bp sequences that are used to identify individual samples.

The adapter-modified DNA samples are then pooled together in a single tube by taking 5 μl of each DNA sample and then purified for PCR using a Qiagen PCR cleanup kit. PCR is performed using primers that are complementary to the adapter sequences. PCR-amplified sample pools constitute the sequencing library. Only fragments having a combination of a barcoded and common adapter at either ends of the fragments are amplified. PCR products are purified with a QIAquick PCR Purification Kit and sequenced using Illumina next-generation sequencing technique. The sequence data are then subjected to the GBS reference genome pipeline where raw sequencing reads are called and filtered. Filtered SNPs are then aligned to a maize reference genome, which allows the data to be interpreted to provide the SNP positions and their corresponding polymorphisms.

The GBS generated 143 415 SNPs covering all the ten chromosomes of maize genome. The genotype data generated by GBS were filtered using TASSEL version: 4.1.12 based on two criteria: missing data and allele frequency. SNP loci, which had missing values and having minimum allele frequency <0.05 for any of the individuals, were excluded from genetic diversity analysis. After filtering all the SNPs, only 3297 SNP loci with no missing values were used in the current genetic diversity assessment study.

Table 2: SNPs used per chromosome

Chromosome No.	SNPs from GBS data	SNPs used for genetic diversity assessment
1	23 180	459
2	16 897	353
3	15 822	369
4	15 489	327
5	15 971	315
6	11 653	298
7	11 916	324
8	12 513	309
9	10 382	274
10	9592	269
Total	143 415	3297

GBS, genotyping by sequencing.

Statistical analysis: Summary statistics, including gene diversity, observed heterozygosity and polymorphism information content (PIC), were calculated using the Power Marker software (Liu and Muse 2005). The PIC value, described by Botstein et al. (1980), was used to refer to the relative value of each marker with respect to the extent of polymorphism it revealed. Heterozygosity was calculated to quantify the genetic variation in the maize lines sampled. Genetic distance estimates were computed using the Roger's genetic distance (Rogers 1972) with the Power Marker software. A dendrogram was constructed from the Roger's genetic distance matrix using the neighbour-joining algorithm (Saitou and Nei 1987) with the Power Marker and the resulting trees were visualized using MEGA version 5.2.2 (Tamura et al. 2011). An admixture model-based clustering method was also used to confirm the population structure of the 128 DTSTHR maize inbred lines using the software package STRUCTURE, version 2.3.3 (Pritchard et al. 2000). The inbred lines with membership probabilities of >0.5 were assigned to the corresponding clusters.

Results

SNP and genetic characteristics

A total of 128 intermediate to late maturing DTSTHR maize inbred lines were genotyped by GBS at Cornell University. The initial GBS data of 143 415 SNPs for all maize inbred lines

Table 3: Gene diversity, heterozygosity and polymorphic information content (PIC) within each group of DTSTHR maize lines

Group	Sample size	Gene diversity	Heterozygosity	PIC
ACRSYN-W-S2	33	0.15 ± 0.0027	0.14 ± 0.0032	0.13 ± 0.0024
IWD-SYN-STR-C3	28	0.21 ± 0.0030	0.11 ± 0.0025	0.17 ± 0.0023
ACR97TZLComp1	26	0.20 ± 0.0030	0.17 ± 0.0032	0.16 ± 0.024
ZDiploBC4	34	0.15 ± 0.0027	0.14 ± 0.0034	0.12 ± 0.0025
STR Low Emerg Pool	2	0.07 ± 0.0026	0.11 ± 0.0042	0.05 ± 0.0019
Miscellaneous	2	0.12 ± 0.0036	0.11 ± 0.0045	0.09 ± 0.0025

were distributed across the 10 chromosomes. The number of SNPs on each chromosome ranged from 23 180 on chromosome 1 to 9592 on chromosome 10. Of the 143 415 SNPs, only 3297 SNPs with minor allele frequency of 0.05 were selected and used in the present diversity assessment study (Table 2).

Maize inbred lines included in this study had a broader genetic base as revealed by their range of PIC and gene diversity values. The PIC values varied from 0.10 to 0.40, with an average of 0.25. More than half (56%) of the inbred lines showed PIC values of ≥0.20. The gene diversity for 55% of the inbred lines ranged from 0.20 to 0.40, indicating the presence of high level of genetic diversity among the DTSTHR maize inbred lines developed at IITA. About 80% of the inbred lines showed heterozygosity values varying from 0.0 to 0.3, with only 5% of the inbred lines showing heterozygosity of 0.6–1.0.

Cluster analysis

The genetic distance-based neighbour-joining cluster analysis separated the DTSTHR maize inbred lines into four major groups (Fig. 1). The resulting marker-based grouping of the inbred lines was consistent with grouping based on the pedigree information of the lines, indicating that inbred lines from the same source populations were close to each other than those derived from different sources. Cluster I consisted of 35 DTSTHR maize inbred lines derived from biparental populations of 1393/ZDiploBC4-19-4-1, (ZDiploBC4-472-2-2-1-2-3/ZDiploBC4-19-4-1) and ZDiploBC4, all having *Zea diploperenses* as a common parent. Cluster II comprised 36 inbred lines, derived from biparental crosses of ACRSYN-W/TZLCOMP1C4 and a composite TZLCOMP1C4. Cluster III consisted of 27 DTSTHR inbred lines derived from biparental populations of ACR97SYN-Y/ACRTZLComp1 (9450/CM116/9450) and a composite TZLcomp1C4. All inbred lines grouped under cluster IV were derived from IWD-SYN-STR-C3, except an inbred line (L029) that originated from a biparental population (ACR97SYN-Y/ACR97TZLComp1-Y).

Population structure analysis

An admixture model-based clustering method in STRUCTURE was implemented to infer population structure for all the DTSTHR inbred lines included in this study (Fig. 2) and confirm consistency of results obtained from cluster analysis. The ADMIXTURE is used to include a cross-validation procedure that allows the user to identify the value of K for which the model has best predictive accuracy. A good value of K will exhibit a low cross-validation error compared to other K values (Phase and Stephens 2006). As a result, a good value of K with low cross-validation error as compared to other K values was when $K = 4$, indicating that the 128 lines were divided into four groups. When the results from different K values were compared

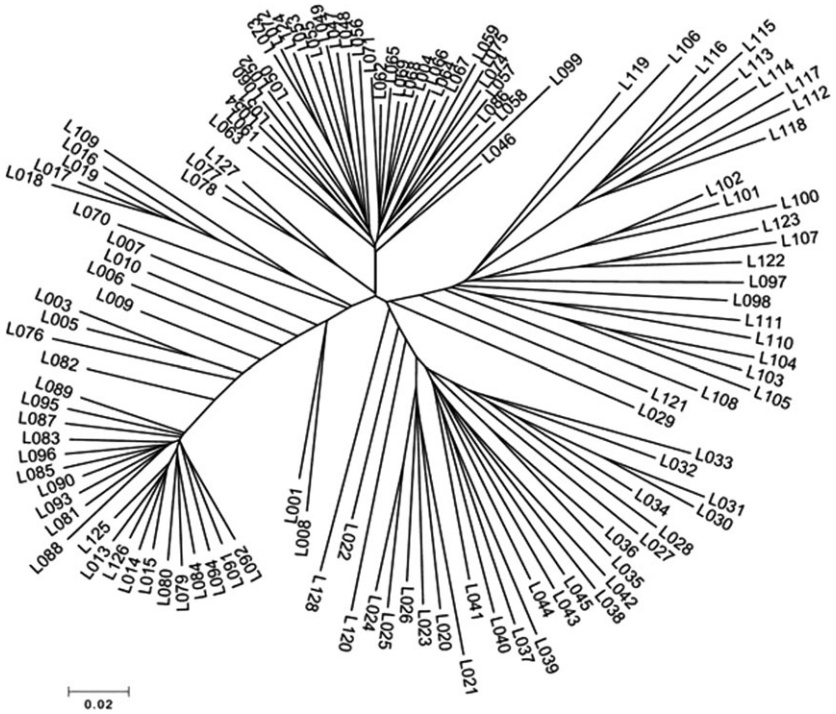

Fig 1: Neighbour-joining (NJ) tree for the 128 DTSTHR maize inbred lines based on the Roger's genetic distance

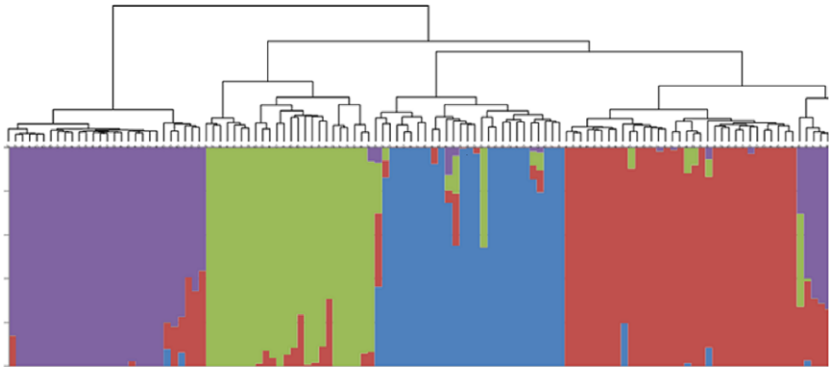

Fig 2: Population structure of the 128 DTSTHR maize inbred lines. Each vertical bar represents one maize line, which is partitioned into K coloured segments [Color figure can be viewed at wileyonlinelibrary.com]

Table 4: Genetic distance within groups and the entire panel of DTSTHR inbred lines

Group	Sample size	Minimum	Maximum	Mean ± SE
Entire panel	125	0.0545	0.2870	0.1949 ± 0.0039
ACRSYN-W-S2	33	0.0291	0.1545	0.0746 ± 0.0036
IWD-SYN-STR-C3	28	0.0477	0.2158	0.1681 ± 0.0049
ACR97TZLComp1	26	0.0419	0.1918	0.1273 ± 0.0041
ZDiploBC4	34	0.0338	0.1647	0.0875 ± 0.0057

with the pedigree information of the DTSTHR inbred lines, the grouping was in agreement with groupings based on the pedigree information of the lines. Groups 1 included DTSTHR inbred lines originating from a biparental population 1393/ZDiploBC4-19-4-1 and ZDiploBC4. Groups 2, 3 and 4 comprised of inbred lines originating from the biparental crosses of ACR97SYN-Y/ACR97TZLComp1-Y, ACRSYN-W/TZLComp1C4 and a *S. hermonthica*-resistant synthetic IWD-SYN-STR-C3, respectively. The majority of the inbred lines (45) were assigned to group 1, whereas the second, the third and the fourth group consisted of 27, 24 and 28 inbred lines, respectively. In general, the groups identified based on both cluster and the populations STRUCTURE analysis were in agreement with the pedigree information of the inbred lines (Fig. 2).

Genetic diversity within and among groups

The genetic diversity and PIC values of the lines originating from synthetic IWD-SYN-STR-C3 and the biparental crosses ACR97SYN-Y/ACR97TZLComp1-Y were greater than those of lines originating from other source populations (Table 3). Also, heterozygosity values of the lines derived from the biparental population of ACR97SYN-Y/ACR97TZLComp1-Y were the highest. The maximum, minimum and mean genetic distance within the DTSTHR inbred lines were calculated based on their source populations and are presented in Table 4. Lines derived from IWD-SYN-STR-C3 had higher genetic distance in comparison with other groups. The highest genetic distance of 0.0825 was observed between inbred lines derived from a backcross population ZDiploBC4 and a biparental population ACR97SYN-Y/ACR97TZLCOMP1-Y, whereas the lowest genetic distance of 0.0516 was found between ZDiploBC4 and ACRSYN-W-S2 (Table 5).

Table 5: Genetic distance (Nei 1983) between groups

Groups	1	2	3	4
1. ACRSYN-W-S2				
2. IWD-SYN-STR-C3	0.0534			
3. ACR97TZLComp1	0.0527	0.0643		
4. ZDiploBC4	0.0516	0.0800	0.0825	

Discussion

The present study investigated the genetic diversity among 128 DTSTHR maize inbred lines using SNP markers. We identified 3297 SNP markers after excluding rare allele and SNPs with missing values. In the current study, a higher mean allele frequency of 5% was used for SNP filtering, which was used as key criteria to reduce ascertainment bias and select more reliable SNPs. The genetic diversities of the DTSTHR maize inbred lines were assessed using neighbour-joining cluster analysis and model-based population structure analysis. The results of these analyses revealed the presence of four distinct groups, with a good match of the inbred lines that grouped together based on the two methods of analysis. The results also showed clear population structure and genetic diversity between different groups of DTSTHR inbred lines, which was in agreement with previous studies (Wen et al. 2012, Wu et al. 2016).

The observed 0.0–2.2 heterozygosity value for most of the inbred lines included in the present study indicated that these lines were fairly homozygous. The observed PIC, heterozygosity and gene diversity values in the present study were in agreement with that of Semagn et al. (2012) who reported the PIC and gene diversity values of approximately 0.2. The low level of heterozygosity present in the majority of the inbred lines confirms the purity of the inbred lines that is important for breeders to maintain the inbred lines, ease in quality control during seed production, high level of heterosis in hybrid formation and ease in assigning inbred lines into different heterotic groups.

Molecular marker analyses provide an additional approach to combining ability studies for separation of a large number of inbred lines into distinct heterotic groups (Flint-Garcia et al. 2009, Reif et al. 2003, Wu et al. 2016). Semagn et al. (2012) used multivariate analysis of SNP markers to separate 220

CIMMYT lines into heterotic groups, but did not find a clearly defined population structure and genetic differentiation of the inbred lines into the heterotic groups defined by CIMMYT breeders. In contrast, the results of cluster and population structure analysis in our study clearly separated the inbred lines into four groups in line with their genetic backgrounds. Lu et al. (2009) observed that lines within a group or subgroup have high level of genetic similarity and that crosses between genetically distinct divergent lines selected from different groups or subgroups generally produce better-performing hybrids than the closely related parents. Crossing lines selected from the four groups defined in the present study can generate promising new maize hybrids resistant to *S. hermonthica* and tolerant to drought.

A clear separation of the lines was observed in line with their source populations (TZL COMP1C4, ZDiploBC4, ACR97SYN-Y/ACR97TZLComp1-Y and IWD-SYN-STR-C3). The first group was composed of 35 inbred lines, derived mainly from a broad-based backcross population ZDiploBC4 that contained a wild relative of maize *Z. diploperenses* as sources of genes for resistance to *S. hermonthica* and tolerance to drought in its genetic background. The other major group of inbred lines was derived from one of the adapted *S. hermonthica*-resistant synthetic variety, IWD-SYN-STR-C3. The synthetic was formed from diverse germplasm unrelated to other Broad-based source populations that were sources of the majority of the inbred lines included in the present study. IWD-SYN-C3 is also resistant to lodging, tolerant to drought, has good ear and plant aspect as well as excellent husk cover (IITA 2008). It has thus been an important source of *S. hermonthica*-resistant and drought-tolerant inbred lines that are currently being used as parents for developing elite hybrids. In the current study, the inbred lines derived from this synthetic had higher gene diversity and genetic distance estimates compared to lines originating from other source populations, confirming potential value of the synthetic as sources of lines with distinct heterotic pattern that can be used as parents for developing conventional maize hybrids.

Grouping of inbred lines into four groups based on their degree of relatedness will serve as valuable guide for breeders in selecting the inbred lines to develop different types of DTSTHR maize hybrids and synthetics. The diverse inbred lines can then readily be used a source of novel resistant alleles that may be introgressed into locally adapted maize germplasm with desirable agronomic traits (Menkir et al. 2010). This will enhance development of new maize cultivars with durable resistance to *S. hermonthica* and tolerance to abiotic stresses such as drought. The broad range of genetic distance detected with SNP markers highlights the significant reservoir of diversity that exists among the DTSTHR parental lines used in the present study (Menkir et al. 2005). Maize breeders usually cross two inbred lines from the same heterotic group to develop a female parent and choose the third inbred line from another group as a male parent to develop three-way cross hybrids. The observed diversity among the four groups in the present study can be exploited for the development of single-cross, three-way cross and double-cross hybrids with enhanced levels of *S. hermonthica* resistance, drought tolerance and agronomic performance.

In conclusion, our study demonstrated the presence of broad genetic diversity among the DTSTHR inbred lines derived from different source populations which is important for the development of *S. hermonthica*-resistant and drought-tolerant maize hybrids and synthetics. The SNP markers separated the DTSTHR maize inbred lines into four distinct clusters, consistent with their pedigree records. Combining ability studies along with per se performance evaluations of these inbred lines is recommended to have a complete understanding of their usefulness as potential parents of hybrids that can be targeted for areas infested with *S. hermonthica* and affected by drought stress. Integrating the pedigree information with combining ability tests and the genetic diversity information from SNP marker-based analyses may provide well-defined heterotic groups for future maize improvement work in West and Central Africa. These results can also help other breeders to utilize DTSTHR maize inbred lines present at IITA in an effective manner for developing biparental source populations of inbred lines without disrupting the heterotic groups they have established in their breeding programmes.

References

Abdalla, D., S. Jacob, E. M. Sharon, G. Vernon, and Y. D. Eric, 2014: Genetic diversity among INERA maize inbred lines with single nucleotide polymorphism (SNP) markers and their relationship with CIMMYT, IITA, and temperate lines. BMC Genet. **15**, 127.

Botstein, D. R., R. L. White, M. Skolnick, and R. W. Davis, 1980: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32**, 314—333.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler, 2007: TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**, 2633—2635.

Collard, B. C., and D. J. Mackill, 2008: Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. Lond. B Biol. Sci. **363**, 557—572.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, and K. Kawamoto, 2011: A robust, simple Genotype-by-sequence (GBS) approach for high diversity species. PLoS One **6**, e19379. doi:10.1371/journal.pone.0019379.

FAOSTAT, 2012: Statistical Database of the Food and Agriculture of the United Nations. Available at: http://www.fao.org (last accessed on 29 September 2016).

Flint-Garcia, S. A., E. S. Buckler, P. Tiffin, E. Ersoz, and N. M. Springer, 2009: Heterosis is prevalent for multiple traits in diverse maize germplasm. PLoS One **4**, e7433. doi:10.1371.

IITA, 2008: International Institute of Tropical Agriculture. More yield, less crop loss from new Striga-resistant maize. IITA Bull. Available at: www.iita.org/2008-press-releases/

Jianbing, Y., Y. Xiaohong, S. Trushar, S. V. He′ctor, L. Jiansheng, W. Marilyn, Z. Yi, H. C. Jonathan, and X. Yunbi, 2009: High-throughput SNP genotyping with the GoldenGate assay in maize. Mol. Breed. doi:10.1007/s11032-009-9343-2.

Kanampiu, F., and D. Friesen, 2003: A new approach to striga control. R. Soc. Chem. **14**, 51—53.

Kim, S. K., 1996: Horizontal resistance: core to a research break through to combat Striga in Africa. Integr. Pest Manag. Rev. **1**, 229—249.

Kling, J. G., J. M. Fajemisin, B. Badu-Apraku, A. Diallo, A. Menkir, and A. Melake-Berhan, 2000: *Striga* resistance breeding in maize. In: B. I. G. Haussmann, D. E. Hess, M. L. Koyama, L. Grivet, H. F. W. Rattunde, and H. H. Geiger (eds), Breeding for Striga Resistance in Cereals, 103—118. Margraf Verlag, Weikersheim, Germany.

Lee, M., 1995: DNA markers and plant breeding programs. Adv. Agron. **55**, 265—344.

Liu, K., and S. V. Muse, 2005: Power marker: an integrated analysis environment for genetic marker analysis. Bioinformatics **21**, 2128—2129.

Lu, Y., J. Yan, G. T. Guimaraes, S. Taba, Z. Hao, S. Gao, S. Chen, J. Li, S. Zhang, B. S. Vivek, C. Magorokosho, S. Mugo, D. Makumbi, S. N. Parentoni, T. Shah, T. Rong, J. H. Crouch, and Y. Xu, 2009: Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor. Appl. Genet. **120**, 93—115.

Melchinger, A., 1999: Genetic diversity and heterosis. In: J. Coors, S. Pandey, and J. T. Gerdes (eds), The Genetics and Exploitation of

Heterosis in Crops, 99—118. American Society of Agronomy, Madison, WI.

Menkir, A., J. G. Kling, B. Badu-Apraku, and I. Ingelbrecht, 2005: Molecular marker-based genetic diversity assessment of Striga -resistant maize inbred lines. Theor. Appl. Genet. **110**, 1145—1153.

Menkir, A., B. Badu-Apraku, C. G. Yallou, A. Y. Kamara, and G. Ejeta, 2007: Breeding maize for broad-based resistance to *Striga hermonthica*. In: G. Ejeta, and J. Gressel (eds), Integrating New Technologies for Striga Control, 99—114. World Scientific, Pte Ltd, Singapore.

Menkir, A., V. O. Adetimirin, C. G. Yallou, and M. Gedil, 2010: Relationships of genetic diversity of inbred lines with different reaction to Striga hermonthica (Del.) Benth and the performance of their crosses. Crop Sci. **50**, 602—611.

Mohammadi, S. A., and B. M. Prasanna, 2003: Analysis of genetic diversity in crop plants: statistical tools and considerations. Crop Sci. **43**, 1235—1248.

Moose, S. P., and R. H. Mumm, 2008: Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol. **147**, 969—977.

Naresh, V., K. N. Yamini, P. Rajandrakumar, and V. Kumar, 2009: EST-SSR marker based assay for genetic purity assessment of safflower hybrids. Euphytica **170**, 347—353.

Pritchard, J. K., M. Stephens, and P. Donnelly, 2000: Inference of population structure using multi-locus genotype data. Genetics **155**, 945–959.

Phase, S., and M. Stephens, 2006: A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. **78**, 629—644.

Reif, J. C., A. E. Melchinger, X. C. Xia, M. L. Warburton, D. A. Hoisington, S. K. Vasal, G. Srinivasanb, M. Bohna, and M. Frisch, 2003: Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. Crop Sci. **43**, 1275—1282.

Rogers, J. S., 1972: Measures of genetic similarity and genetic distance. Studies in Genetics, VII. Univ. Texas Publ., Texas. **7213**, 145—153.

Saghai-Maroof, M. A., K. M. Soliman, R. A. Jorgensen, and R. W. Allard, 1984: Ribosomal DNA spacer-length polymorphisms I barley: mendelian inheritance, chromosomal locations, and population dynamics. Proc. Natl Acad. Sci. USA **81**, 8014—8018.

Saitou, N., and M. Nei, 1987: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**, 406—425.

Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, and F. Wei, 2009: The B73 maize genome: complexity, diversity and dynamics. Science **326**, 1112—1115.

Semagn, K., C. Magorokosho, S. V. Bindiganavile, D. Makumbi, Y. Beyene, S. Mugo, B. M. Prasanna, and M. L. Warburton, 2012: Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. BMC Genom. **13**, 113.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, 2011: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**, 2731—2739.

Vivek, B. S., J. Crossa, and G. Alvarado, 2009: Heterosis and combining ability among CIMMYT's mid-altitude early to intermediate maturing maize (*Zea mays L.*) populations. Maydica **54**, 97—107.

Wu, X., Y. Li, Y. Shi, Y. Song, T. Wang, Y. Huang, and Y. Li, 2016: Fine genetic characterization of elite maize germplasm using high-throughput SNP genotyping. Theor Appl Genet. **127**, 621–631.

Wen, W., J. Franco, V. H. Chavez-Tovar, J. Yan, and S. Taba, 2012: Genetic characterization of a core set of a tropical maize race Tuxpeño for further use in maize improvement. PLoS One **7**, e32626.

Yongsheng, W., F. San Vicente, K. Huang, T. Dhliwayo, D. Friesen, K. Semagn, N. Sudha, M. Olsen, M. P. Boddupalli, X. Zhang, and R. Babu, 2016: Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. Theor. Appl. Genet. **129**, 753—765.