Genome-wide association studies of *Striga* resistance in extra-early maturing quality protein maize inbred lines

Gbemisola Okunlola,[1,2] Baffour Badu-Apraku [ID] ,[1,*] Omolayo Ariyo,[2] Paterne Agre [ID] ,[1] Queen Offernedo [ID] ,[1] Moninuola Ayo-Vaughan[2]

[1]Maize Improvement Programme, International Institute of Tropical Agriculture, IITA, Oyo Road, Ibadan 200001, Oyo ,5320, Nigeria.
[2]Department of Plant Breeding and Seed Technology, Federal University of Agriculture, Abeokuta 110124, Ogun, 2240, Nigeria

*Corresponding author: International Institute of Tropical Agriculture (IITA), PMB 5320, Oyo Road, Ibadan, Nigeria. Email: b.baduapraku@cgiar.org

Abstract

Identification of genes associated with *Striga* resistance is invaluable for accelerating genetic gains in breeding for *Striga* resistance in maize. We conducted a genome-wide association study to identify genomic regions associated with grain yield and other agronomic traits under artificial *Striga* field infestation. One hundred and forty-one extra-early quality protein maize inbred lines were phenotyped for key agronomic traits. The inbred lines were also genotyped using 49,185 DArTseq markers from which 8,143 were retained for population structure analysis and genome wide-association study. Cluster analysis and population structure revealed the presence of 3 well-defined genetic groups. Using the mixed linear model, 22 SNP markers were identified to be significantly associated with grain yield, *Striga* damage at 10 weeks after planting, number of emerged *Striga* plants at 8 and 10 weeks after planting and ear aspect. The identified SNP markers would be useful for breeders for marker-assisted selection to accelerate the genetic enhancement of maize for *Striga* resistance in sub-Saharan Africa after validation.

Keywords: SNP markers; *Striga hermonthica*; marker-assisted selection; quality protein maize

Introduction

Striga hermonthica parasitism is fast becoming an endemic in West Central Africa because of its increased dispersal mechanisms which include wind, animals, farm implements, and surface water (Ejeta 2007). The geographical distribution and the level of infestation of this parasitic weed are steadily increasing and more increase is expected because of the adverse effects of climate change (Mohamed *et al.* 2007). *Striga* lacks its own root system, and therefore depends completely on maize for nutrients and water for survival for 6–8 weeks while still under the ground once attached to the host plant (Bebawi and Mutwali 1991; Badu-Apraku and Fakorede 2017). The seedling develops haustoria, which penetrates the roots of the maize plant to syphon nutrients and photosynthates (Badu-Apraku and Fakorede 2017). As a result, *Striga* parasitism affects the crop growth, biomass partitioning, and the nutrient status of the maize plant. *Striga* infestation in maize causes chlorotic blotches, scorching, or "firing" of leaves particularly around the margins, wilting of foliage, stunting, spindly stems caused by the preferential allocation of biomass to the roots, and poor grain filling (Menkir *et al.* 2012; Badu-Apraku and Fakorede 2017). These deleterious effects of *Striga* are observed on the host plant even before its emergence from the soil (Parker 2013; Badu-Apraku and Fakorede 2017).

The deleterious impact of *Striga* has been reported in 25 countries in Africa, the most severely affected being sub-Saharan Africa (SSA) countries (Parker 2012). A single *S. hermonthica* plant can inflict an approximately 5% loss in yield on a host plant (Parker and Riches 1993), and high infestation can lead to total crop failure (Badu-Apraku and Fakorede 2017). *Striga* has thus become a major threat to food security, worsening hunger and poverty, especially SSA countries (Pennisi 2010; Khan *et al.* 2014). This impact of *Striga* is further complicated in SSA due to moisture and nutrient stress caused by increased population pressure, short land fallow periods, and minimal use of inorganic fertilizer (De Groote *et al.* 2005). The severity of *Striga* attack increases with the extent of the soil seed bank, existence of strain, variants and races with different virulence, the reaction of the host cultivar, and the environment (Babiker 2007).

Striga hermonthica is one of the most difficult parasitic weeds to control (Nickrent and Musselman 2004). The approaches for *Striga* control have been grouped into 4 independent options, which includes cultural, chemical, genetic, that is the use of resistant cultivars and biological control methods (Babiker 2007; Sibhatu 2016). The control measures, such as hand pulling, irrigation, crop rotation, herbicides, fallowing, high level of N application, have proven ineffective in small holder farms (Gressel *et al.* 2004). This is primarily because of the parasite's highly specialized cycle, which is synchronized with the host's growth, ability of the parasite to parasitize a broad host range, and longevity of seed of the parasite in the soil (Ejeta 2007; Herne 2009). An integrated management method that involves the use of a number of individual technologies combined together to act at different stages of the parasite's life is sometimes employed for an

effective *Striga* control (Kim 1996). However, host plant resistance is the most effective approach against *Striga* damage in maize production and reduces *Striga* seed bank in the soil (Badu-Apraku 2005). This method is the most economical, suitable, and environmentally friendly approach for *Striga* control (Hearne 2009; Sibhatu 2016; Mandumbu *et al.* 2019). Considerable progress in breeding for *Striga* resistance/tolerance have been made by scientists at International Institute of Tropical Agriculture (Badu-Apraku and Lum 2007; Menkir *et al.* 2012; Badu-Apraku, Fakorede *et al.* 2016; Badu-Apraku *et al.*, 2016; Yallou *et al.* 2016; Akogu *et al.* 2019; Menkir and Meseka, 2019).

The advancement in molecular breeding through rapid genotyping and next-generation sequencing technologies has enabled the use of genome-wide association study (GWAS) to be used as a tool for revealing genotype–phenotype associations in crop species (Liu and Yan 2019). GWAS has been found to be a powerful approach for identifying functional genes and alleles that are associated with complex traits in certain environments (Li *et al.* 2016; Zhu *et al.* 2018). This association is based on linkage disequilibrium which is a result of association of a particular trait with a neighboring genetic variation of another trait (Liu and Yan 2019). Unlike quantitative trait locus (QTL) mapping which results in a relatively low-resolution map, genome-wide sequence association mapping gives a relatively high-resolution mapping for identifying genes or regions associated with a particular trait (Xue *et al.* 2013; Liu and Yan 2019). Several studies have employed GWAS to detect QTLs and genomic regions associated with biotic and abiotic stresses (Shikha *et al.* 2021). Wang *et al.* (2012) identified 18 novel candidate genes associated with head smut resistance in maize, 22 QTLs were revealed for gray leaf spot among biparental populations and association mapping panel of 410 tropical/subtropical inbred lines (Kibe *et al.* 2020). Adewale *et al.* (2020) conducted GWAS on *Striga* resistance traits with 132 early maturing inbred lines. Two putative genes (ZmCCD1 and amt5) located on chromosome 9 and 10 were found to

be linked to plant defense mechanism against *Striga*. In another study involving 380 diverse tropical inbred lines, Gowda *et al.* (2021) identified a set of 32 candidate genes physically near the significant SNPs with varying roles in plant defense against biotic stresses. Although some GWAS studies have been conducted to detect candidate genes for *Striga* resistance, none of the QTLs detected have been employed in *Striga* resistance breeding. Therefore, there is a need to conduct additional studies using different genotypes to detect more QTL so as to facilitate the introgression of novel *Striga* resistant genes into maize breeding programs in SSA. The objectives of this research were to (1) determine the genetic structure of a panel of 141 diverse extra-early maturing white quality protein maize (QPM) inbred lines with varying levels of resistance to *S. hermonthica* parasitism and (2) identify significant SNPs and putative genes associated with grain yield and other *Striga* adaptive traits under *Striga*-infested conditions.

Materials and methods
Genetic materials
One hundred and sixty-nine extra-early QPM inbred lines from the International Institute of Tropical Agriculture (IITA) Maize improvement program (MIP) were used for this study. The inbred lines comprised 163 S_8 inbred lines, 4 standard IITA testers (TZEEQI 294, TZEEQI 321, TZEEQI 7, and TZEEQI 134), and 2 inbred checks (TZEEQI 11 and TZEEQI 60) with combined resistance

to *Striga*, tolerance to drought and low soil N. The 163 inbred lines were extracted from the F_1 maize hybrids of 9 biparental crosses involving crosses among extra-early white QPM inbred testers and early maturing white QPM inbred testers. The testers and the checks were extracted from *Striga*-resistant populations. The F1 hybrids were taken through a cycle of backcrossing to the extra-early inbred testers to recover the earliness. The BC_1F_1 with desirable agronomic characteristics were selected using the pedigree selection method from each backcrossed population, and advanced through repeated inbreeding to the S_8 generation.

Field trials
The experiments were conducted under *Striga*-infested conditions at Mokwa (9° 18′N and 5° 04′E, 457 m asl, 1,100 mm annual rainfall) in 2019 and 2020 and Abuja (9° 15′N and 7° 20′E, 300 m asl, 1,700 mm annual rainfall) in 2020. At all locations, the experiments were laid out using a 13 × 13 lattice design with 2 replicates and single row plots each 3-m long, spaced 0.75 m apart with 0.4 m between plants in each row. Three seeds were sown per hill, and later thinned to 2 plants per hill at 2 weeks after planting (WAP) to obtain a plant population density of 66,000-plants ha^{-1}.

Each plot was artificially infested with about 5,000 germinable *S. hermonthica* seeds/hill. The *Striga* infestation method developed by IITA-MIP was adopted to ensure uniform *Striga* infestation with no escapes (Kim 1991; Kim and Winslow 1991). The amount of fertilizer applied was about 30 kg ha^{-1} and was split applied. The time of first application was delayed to 21 days after planting so as to subject the maize plants to stress to stimulate the production of strigolactones in an effort to enhance good germination of *Striga* seeds, and the attachment of the *Striga* plants to the roots of the maize plants. Top dressing was done at about 35 days after planting. Weeds other than *Striga* were constantly removed by hand to ensure good weed control.

Data collection
Data were collected on the number of emerged *Striga* plants at 8 and 10 WAP and host plant damage syndrome rating at 10 WAP. The host plant damage syndrome rating was recorded on a scale of 1–9 (1 = normal plant growth, no visible symptoms, and 9 = complete scorching of all leaves, causing premature death or collapse of host plant and no ear formation; Kim 1991). Data were also collected on ear aspect, number of ears per plant, and grain yield (Badu-Apraku *et al.* 2011).

Data analysis
Analysis of variance was performed for the inbreds evaluated in *Striga*-infested environments using the PROC GLM in SAS 2014. The entry means were adjusted for block effects, according to the lattice design. Each year–location combination was considered as a test environment. The environments, replications, and blocks were treated as random factors. Data on the number of emerged *Striga* plants were transformed as [log (counts + 1)] to reduce the heterogeneity of variance for *Striga* counts. Restricted maximum likelihood estimates of the genetic and phenotypic variances of the inbreds were obtained with SAS PROC Varcomp and used to compute the broad-sense heritability for each trait. Correlation analysis was done using the performance analytics package in R. The phenotypic data across environments were collapsed to a single best linear unbiased estimate (BLUE) value using the linear mixed models in META—R (Bates *et al.* 2015; Alvarado *et al.* 2020) as follows:

$$Y_{IJKL} = \mu + B(E)_{J(i)} + G_k + GE_{ij} + e_{ijkl},$$

where Y_{ijkl} = phenotypic observation for a trait, μ = grand mean, E = environmental effect (location), $B(E)$ = replication effects nested in location, G = genotypic effect, GE = genotype by environment interaction, e = random residual error. Broad sense heritability (H^2) estimates were calculated from the phenotypic variance ($\sigma 2p$) and the genotypic variance ($\sigma 2g$) (Hallauer *et al.* 2010).

Genotyping and genotypic data analysis

One leaf per plant was collected from 15 representative plants to form a bulk of each of 141 inbred lines 2 WAP in the IITA maize breeding nursery in Ibadan. The leaf tissues were placed in jute bags and freeze-dried using FreeZone Freeze Dryer (Labconco, USA) following the recommendations of the manufacturer's manual. Genomic DNA was isolated from freeze-dried leaf tissues of each inbred line following the modified Cetyl-trimethyl ammonium bromide (CTAB) protocol as described by Azmach *et al.* (2013). The DNA quality and quantity analysis were performed using the UV/Vis Absorbance protocol in the FlUOstar Omega microplate reader (BMG LABTECH) following the manufacturer's manual. Genotyping analysis of the inbred lines was performed using the high-density whole-genome profiling of Diversity Arrays Technology sequencing (DArTseq). The extracted genomic DNA samples were sent to DArT Pty Ltd, Australia (https://www.diversityarrays.com) for DArTseq analysis following the protocol described by Jaccoud *et al.* (2001).

High-throughput genotyping was carried out in 96 plex following the DArTseq protocol. The 49,184 DArTseq markers obtained as raw SNPs were filtered to eliminate SNPs with missing rate greater than 10%, heterozygosity greater than 20%, and minor allele frequency (MAF) less than 5%. SNPs with unknown or multiple chromosome locations were also eliminated. After quality filtering, a total of 8,144 DArTseq markers distributed across the 10 maize chromosomes were retained for the population structure and for GWAS analyses.

Population structure and kinship analysis

Population structure and kinship analyses were conducted to determine the extent of genetic diversity among the inbred lines. Structure software version 2.3.3 (Pritchard *et al.* 2000) was used to cluster the 141 inbred lines into populations. Structure simulations were carried out using an admixture model with a burning period of 10,000 iterations, followed by Markov chain Monte Carlo set at 10,000. The assumed number of subpopulations was simulated from $k = 1$ to $k = 10$ for an initial assessment of the most likely number of subpopulations, each K was run 10 times. The ideal number of sub-populations (K) was found by examining the optimal ΔK value (Evanno *et al.* 2005) in STRUCTURE Harvester (Earl and von Holdt 2012). Structure population was then plotted using barplot function implemented in R. The phylogeny tree was constructed using ape version 5.0 implemented in R (Paradis and Schliep 2019). The marker-based kinship matrix K was calculated with the same genotypes using the VanRaden method, and then used to create a clustering heat map of the association mapping panel in the GAPIT (Lipka *et al.* 2012).

Association analysis

The association between SNP genotypes and the phenotypes was determined using a compressed linear model implemented in GAPIT (Genome Association and Prediction Integrated Tool)—R package (Lipka *et al.* 2012). Mixed linear method (MLM) and

SUPER (Tang *et al.* 2016) were tested for association analysis. The MLM adopted was proposed by Yu *et al.* (2006) with each molecular marker considered a fixed effect and evaluated individually: $Y = X\beta + W\alpha + Qv + Zu + \varepsilon$; where Y is the observed vector of means; β is the fixed effect vector ($p \times 1$) other than molecular marker effects and population structure; α is the fixed-effect vector of the molecular markers; ν is the fixed-effect vector from the population structure; u is the random effect vector from the polygenic background effect; X, W, and Z are the incidence matrice from the associated β, α, ν, and u parameters; ε is the residual effect vector. Quantile–quantile (Q–Q) plots were generated by plotting the negative logarithms (−log10) of the P-values against their expected P-values to fit the appropriateness of the GWAS model with the null hypothesis of no association and to determine how well the models accounted for the population structure. The Manhattan plot was generated for visualizing GWAS on the entire genome and zoom mapping was performed on a particular chromosome after identifying a significant SNP marker. The marker effect or SNP contribution was estimated for the significant SNPs using multiple regression analysis using lme4 function implemented in R where the trait was considered as a response variable while the SNP markers above the Bonferroni threshold for the trait was the independent variable. A threshold of –log (p) = 3 was used to declare significant marker–trait associations, which were determined based on the Q–Q plots and distribution of P-values for all the measured traits (Gao *et al.* 2016; Sukumaran *et al.* 2018; Mogga *et al.* 2018).

Results
Evaluation of phenotypic traits

The combined analysis of variance of the 169 inbred lines (including the 2 checks) across *Striga*-infested environments is presented in Table 1. The results revealed significant ($P < 0.05$) environment (E) and genotype mean squares for measured traits. Genotype × environment interactions were not significant ($P < 0.05$) for the number of emerged *Striga* plants at 8WAP and 10 WAP. Broad sense heritability (h^2) estimates on plot mean basis ranged from 46% for *Striga* damage at 10 WAP to 69% for ear aspect. Moderately high broad sense heritability was observed for the measured traits.

The phenotypic correlations among grain yield and other measured *Striga* adaptive traits differed under artificial *Striga* infestation (Fig. 1). Grain yield had significant negative correlation with *Striga* damage at 8 WAP ($r = -0.58^{**}$) and 10 WAP ($r = -0.71^{**}$), ear aspect ($r = -0.85^{**}$). Significant and positive correlations were obtained between *Striga* damage at 8 WAP and *Striga* damage at 10 WAP ($r = 0.56^{**}$), number of emerged *Striga* plants at 8 WAP and number of emerged *Striga* plants at 10 WAP ($r = 0.72^{**}$), ear aspect and *Striga* damage at 8 WAP ($r = 0.67^{**}$), and *Striga* damage at 10 WAP ($r = 0.78^{**}$).

Population stratification and genetic diversity

The results revealed that PIC ranged from 0.09 to 0.37 with an average of 0.24 whereas the heterozygosity averaged 0.08 and varied from 0.00 to 0.50 (Supplementary Fig. 1). The mean of the minor allele frequencies of the 8,145 primers was 0.14 with minimum and maximum minor allele frequencies of 0.04 and 0.5, respectively. Gene diversity varied from 0.10 to 0.50 with an average of 0.33.

The population structure analysis of the inbred lines showed that delta K values from the mean log-likelihood probabilities peaked at $k = 3$. At $k = 3$, 84% of the inbred lines were assigned to

Table 1. Mean squares of grain yield and other agronomic traits of 169 extra early maturing QPM inbred lines evaluated across *Striga*-infested conditions in Mokwa 2019 and 2020, and Abuja 2020.

Source	DF	Yield, kg/ha	*Striga* damage ratings at 10WAP	Emerged *Striga* plants at 8 WAP	Emerged *Striga* plants at 10 WAP	Ear aspect	Ears per plants
Environment (E)	2	24,506,574**	324.87**	95.80**	62.48**	182.17**	20.11**
Replicate (ENV)	3	7,491,412**	41.38**	1.77**	1.11**	16.17**	1.21**
Block (ENV*Rep)	74	2,071,579**	2.29**	0.11**	0.13**	1.84**	0.19*
Genotype	168	2,178,729**	2.47**	0.12**	0.15**	4.23**	0.19**
Env*Genotype	336	1,041,579*	1.44**	0.08	0.08	1.39**	0.25**
Error	43	865,017	0.78	0.07	0.07	0.84	0.14
Heritability		65	46	48	47	69	—

*, **Significant at 0.05 and 0.01probability levels, respectively.
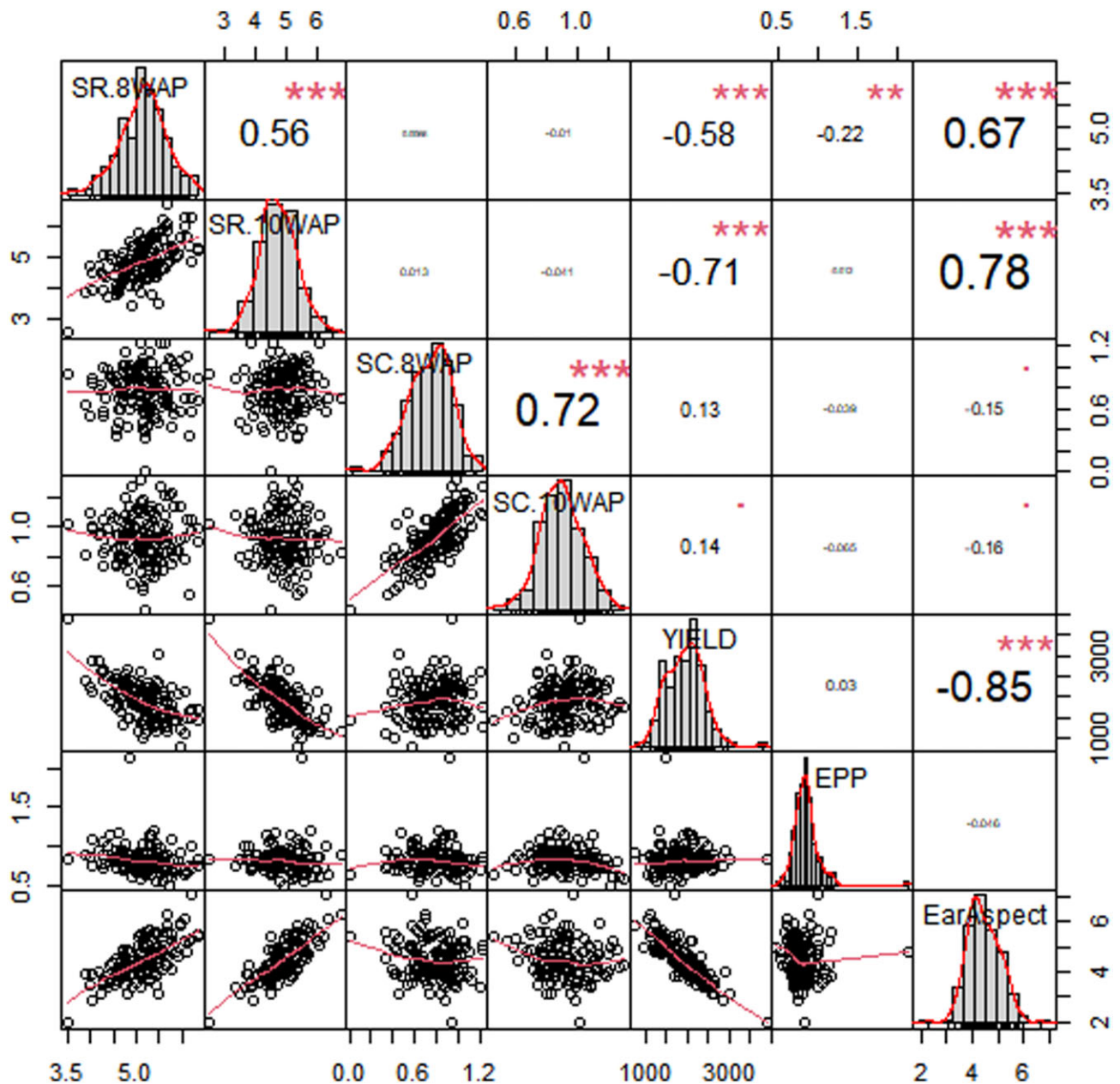
Fig. 1. Correlation coefficients between *Striga* resistance indicator traits and other agronomic traits of early maturing QPM inbred lines under artificial *Striga* infestation at Mokwa and Abuja 2019 and 2020. SR 8WAP = *Striga* damage symptoms rating at 8WAP, SR 10WAP = *Striga* damage symptoms rating at 10 WAP, SC 8WAP= number of emerged *Striga* plants at 8 WAP, SC 10WAP = number of emerged *Striga* plants at 10 WAP, YIELD = grain yield, EPP—number of ears per plant, and Ear Aspect = ear aspect.

3 groups, with only 16% of the lines assigned to the mixed group. A total of 98 inbred lines were placed in group 1, 13 in group 2, 7 in group 3, and 23 in the mixed group. The 3 groups comprised inbred lines derived from 2 or more germplasm sources. The phylogeny tree displayed 3 genetic groups and was aligned with the kinship population structure (Fig. 2). The heat map of the values in the kinship matrix created from the 141 inbred lines also revealed 3 groups which showed relatedness with a few large blocks in the population (Supplementary Fig. 2).

Genome-wide association analysis

Under artificial *Striga* infestation, a total of 22 significant SNPs were associated with 5 different traits at a GWAS threshold of –log (p) = 3 (Table 2). The trait variation explained by individual marker (R^2) varied from 14% to 22%. Five SNP markers were significantly associated with grain yield (Fig. 3). These markers were located on chromosomes 1, 5, 8, and 9, with the phenotypic variation explained by these markers ranging from 20% to 21%. Four of these markers had negative quantitative trait nucleotide (QTN) effects with the MAF ranging from 0.05 to 0.13. *Striga* damage ratings at 10 WAP was associated with 8 markers (Fig. 3). These markers were located on chromosomes 1, 3, 5, 6, 7, 8, 10, and they explained 14–17% phenotypic variation. Four of the associated markers had negative QTN effects with the MAF ranging from 0.05 to 0.12. Six markers located in chromosomes 4, 9, and 10 were detected for emerged *Striga* plants at 8 WAP (Fig. 4).

These markers accounted for 14–15% of the phenotypic variation. Four of these associated markers had negative QTN effects with the MAF ranging from 0.06 to 0.48. Two markers located on chromosomes 8 and 10 were associated with the number of emerged *Striga* plants at 10 WAP (Fig. 4) and explained 21% of the phenotypic variation. One of the SNP markers had MAF varying from 0.13 to 0.48. One SNP marker located on chromosome 8 was associated with ear aspect (Fig. 5) and explained 22% of the phenotypic variation. The MAF of this marker was 0.50. Marker S1_163520946 located on chromosome 1 was repeatedly found to be associated with grain yield and *Striga* damage ratings at 10 WAP.

Discussion
Phenotypic variation

The significant genotypic variation observed among the inbred lines for grain yield and other measured traits in our study indicated the existence of genetic variability among the extra-early maturing QPM inbred lines. The significant environmental effects for measured traits indicated that the research conditions were unique and provided distinct information on the QPM inbred lines. The differential response of the genotypes for grain yield and other measured traits suggested that the inbred lines performed differently under the research condition. This differential performance could be attributed to variation in environmental
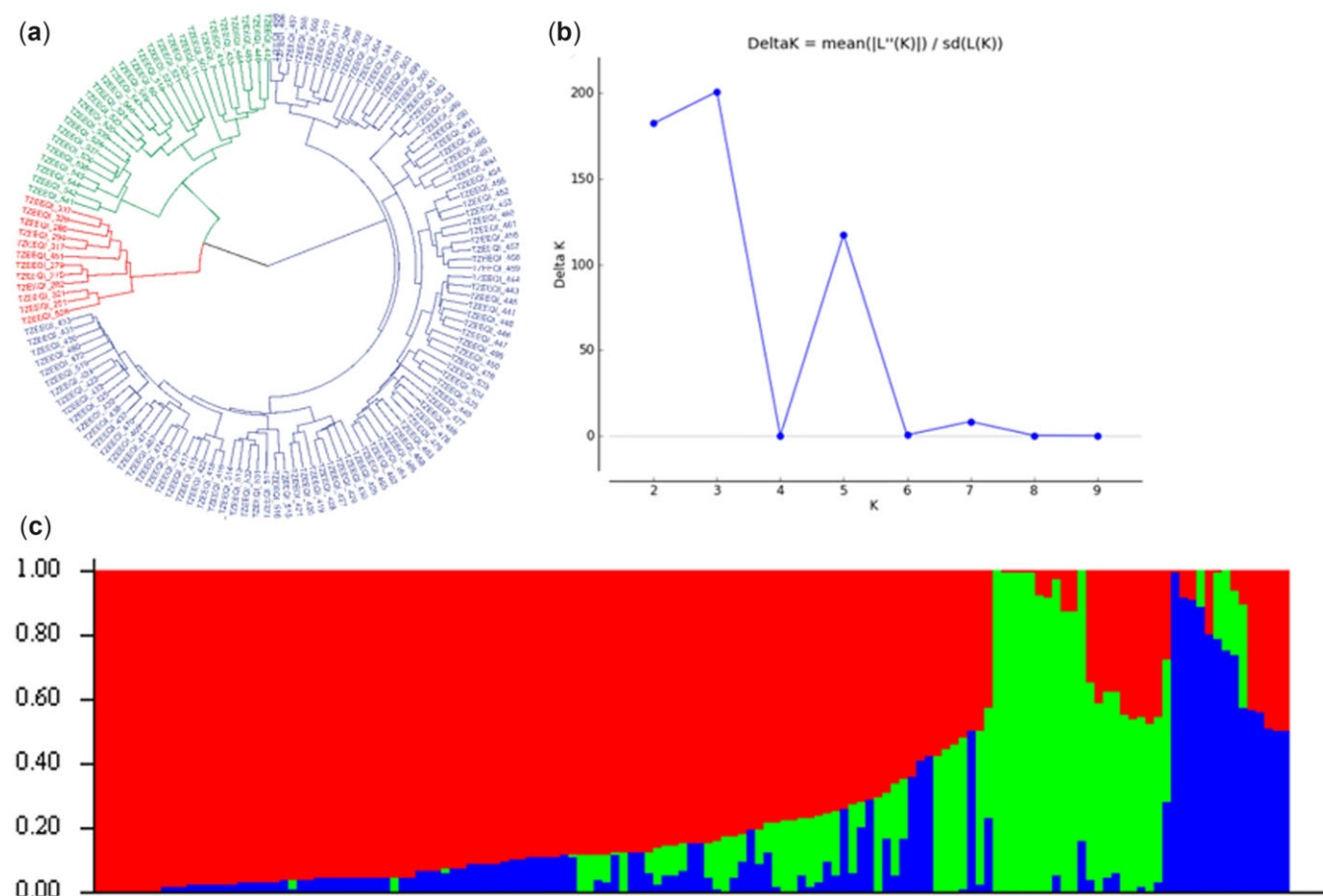

Fig. 2. Graphical representation of the population structure of 141 maize inbred lines. a) Phylogeny tree showing the 3 sub-populations. The colors represent each sub-population. b) Plot of mean likelihood of delta K against the number of K groups. The highest peak observed at $K = 3$ signified the grouping of the accessions into 3 groups. c) Population structure originated from the STRUCTURE based on $K = 3$. Each vertical barplot represented a single maize inbred.

Table 2. Summary of SNP markers associated with *Striga*-adaptive traits evaluated under *Striga*-infested conditions.

Traits	SNP maker	Chr	Position	MAF	-log10(P)'	P-value	r^2 (%)	Effect
Grain yield	chr8_328928	8	328928	0.13	3.55	2.82×10^{-4}	21.38	−390.03
	chr8_135475183	8	135475183	0.07	3.42	3.80×10^{-4}	20.99	−325.99
	chr5_210580022	5	210580022	0.12	3.36	4.41×10^{-4}	20.79	−409.87
	chr9_23155189	9	23155189	0.10	3.35	4.47×10^{-4}	20.77	254.16
	chr1_163520946	1	163520946	0.05	3.24	5.76×10^{-4}	20.44	−334.13
Striga damage ratings at 10 WAP	chr6_79398477	6	79398477	0.09	3.73	1.8×10^{-4}	17.17	0.48
	chr8_17232945	8	17232945	0.48	3.71	1.96×10^{-4}	17.08	−0.24
	chr5_194107033	5	194107033	0.09	3.67	2.14×10^{-4}	16.96	−0.68
	chr10_112661466	10	112661466	0.12	3.50	3.17×10^{-4}	16.41	0.38
	chr1_163520946	1	163520946	0.05	3.47	3.38×10^{-4}	16.32	0.51
	chr3_135777833	3	135777833	0.06	3.31	$4.85\text{-} \times 10^{-4}$	15.81	−0.46
	chr6_79656637	6	79656637	0.07	3.20	6.30×10^{-4}	15.45	−0.44
	chr7_14310701	7	14310701	0.07	3.04	9.07×10^{-4}	14.95	0.42
Emerged *Striga* plants at 8WAP	chr4_18563948	4	18563948	0.20	3.14	7.22×10^{-4}	16.16	−0.08
	chr4_240388223	4	240388223	0.13	3.11	7.80×10^{-4}	16.06	0.10
	chr9_156894729	9	156894729	0.06	3.08	8.31×10^{-4}	15.97	0.19
	chr10_134476659	10	134476659	0.07	3.02	9.57×10^{-4}	15.78	−0.12
	chr4_238158257	4	238158257	0.17	3.01	9.74×10^{-4}	15.75	−0.09
	chr4_183275539	4	183275539	0.10	3.00	9.96×10^{-4}	15.72	−0.15
Emerged *Striga* plants at 10WAP	chr10_85787634	10	85787634	0.48	3.83	1.50×10^{-4}	21.86	1.65
	chr8_180192054	8	180192054	0.13	3.80	1.58×10^{-4}	21.79	−2.99
Ear aspect	chr6_125302216	6	125302216	0.50	3.35	4.49×10^{-4}	22.38	0.23
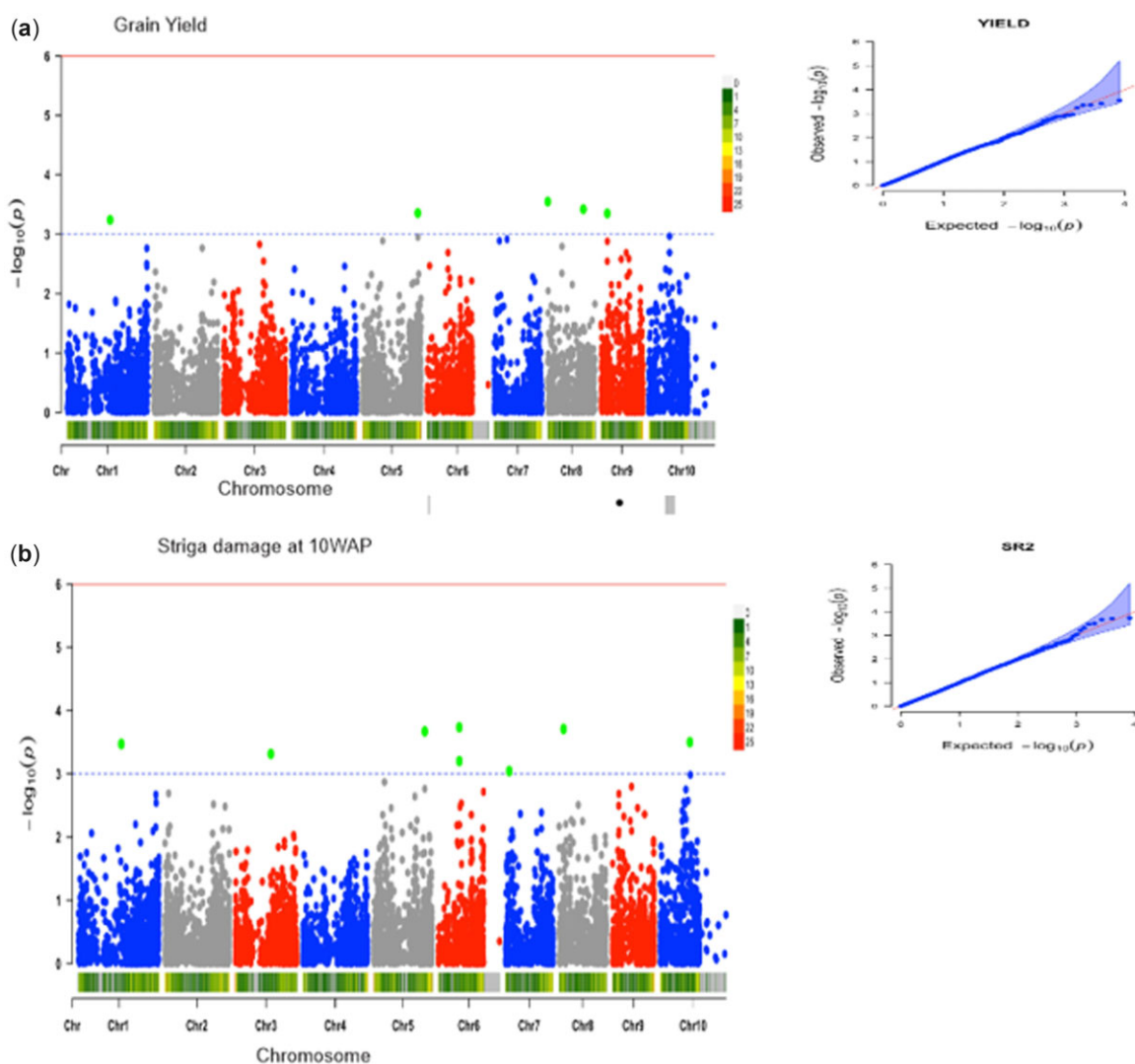
Fig. 3. The Manhattan and Q–Q plots of the SNP-based association mapping for (a) grain yield and (b) *Striga* damage at 10 WAP under artificial *Striga* infestation.
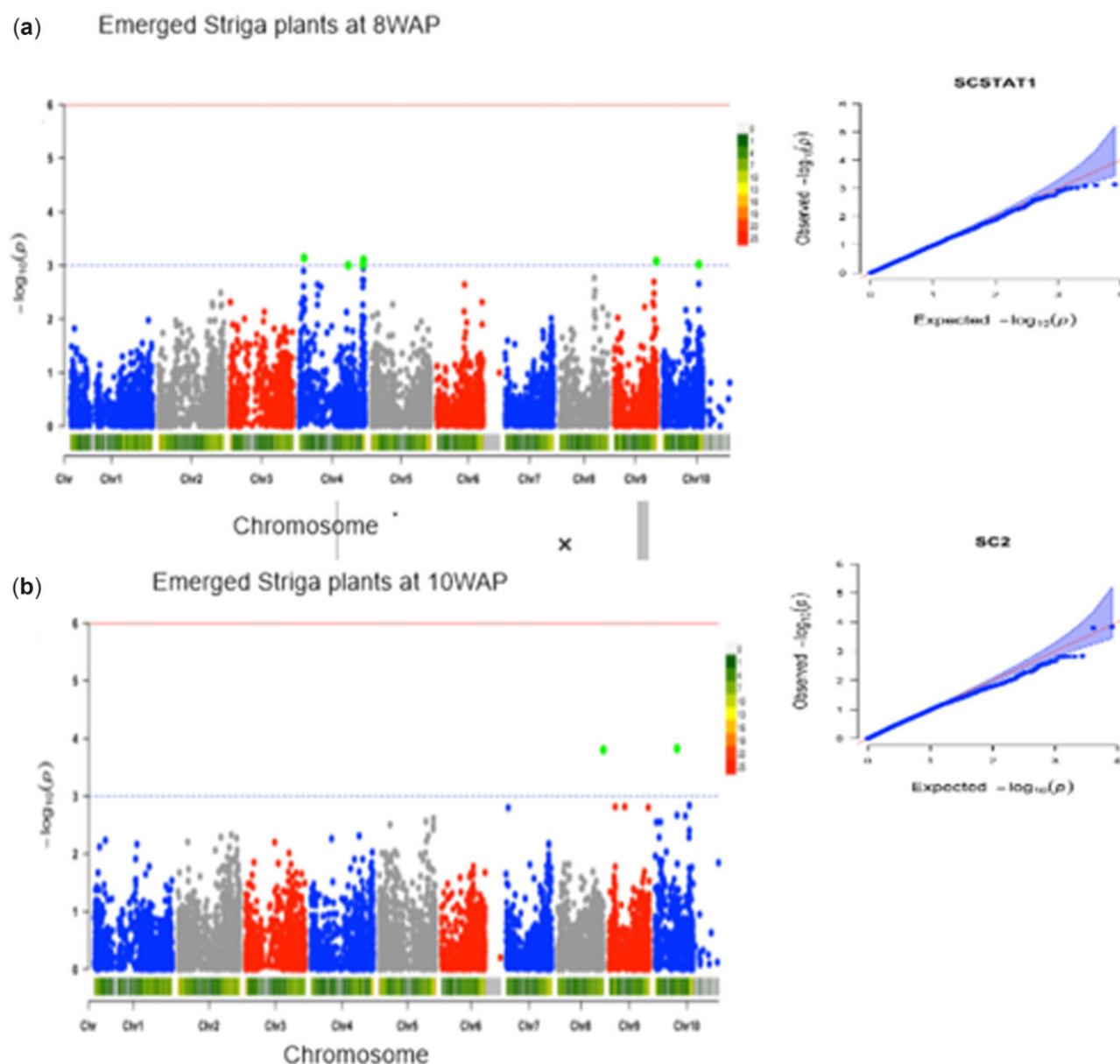
Fig. 4. The Manhattan and Q–Q plots of the SNP-based association mapping for (a) number of emerged *Striga* plants at 8 WAP (b) number of emerged *Striga* plants at 10 WAP under artificial *Striga* infestation.

factors such as temperature, rainfall, and soil types in the different test environments (Badu-Apraku and Lum 2007; Badu-Apraku *et al.* 2010).

The moderate to high heritability estimates observed for grain yield and other *Striga* adaptive traits implied the efficiency of SNP in the maize panel in the identification of true associations between the markers and putative genes (Adewale *et al.* 2020). The inverse relationship between grain yield and number of emerged *Striga* plants and *Striga* damage syndrome ratings implied that increased number of emerged *Striga* plants led to a lower grain yield (Menkir *et al.* 2012) and that lower *Striga* damage syndrome ratings led to an increased grain yield (Gowda *et al.* 2021). Previous studies recommended the simultaneous selection of reduced *Striga* damage syndrome ratings and the number of emerged *Striga* plants as indicators for *Striga* resistance under field infestation (Menkir and Kling 2007; Kim 1994; Kling *et al.* 2000).

Genome-wide association studies

In the GWAS analysis, the population structure information was used to correct possible false discovery. Q–Q plots were generated by comparing the observed and expected P-values under the null hypothesis of no associations to determine how well the models accounted for the population structure. The results revealed that majority of points in the Q–Q plots were aligned on the diagonal line for all the measured traits indicating that the model successfully accounted for population structure and familiar relationships in the GWAS analysis. The whole-genome scan for phenotypic and allelic variation in Maize *Striga* resistance identified 9 genomic regions on chromosomes 10, 9, 8, 7, 5, 4, 3, and 1 with significant −log10 values. At a threshold of −log (p) = 3, a total of 22 markers were identified to be significantly associated with *Striga* damage, number of emerged *Striga* plants, ear aspect, and grain yield under *Striga* infestation. Information on the SNP
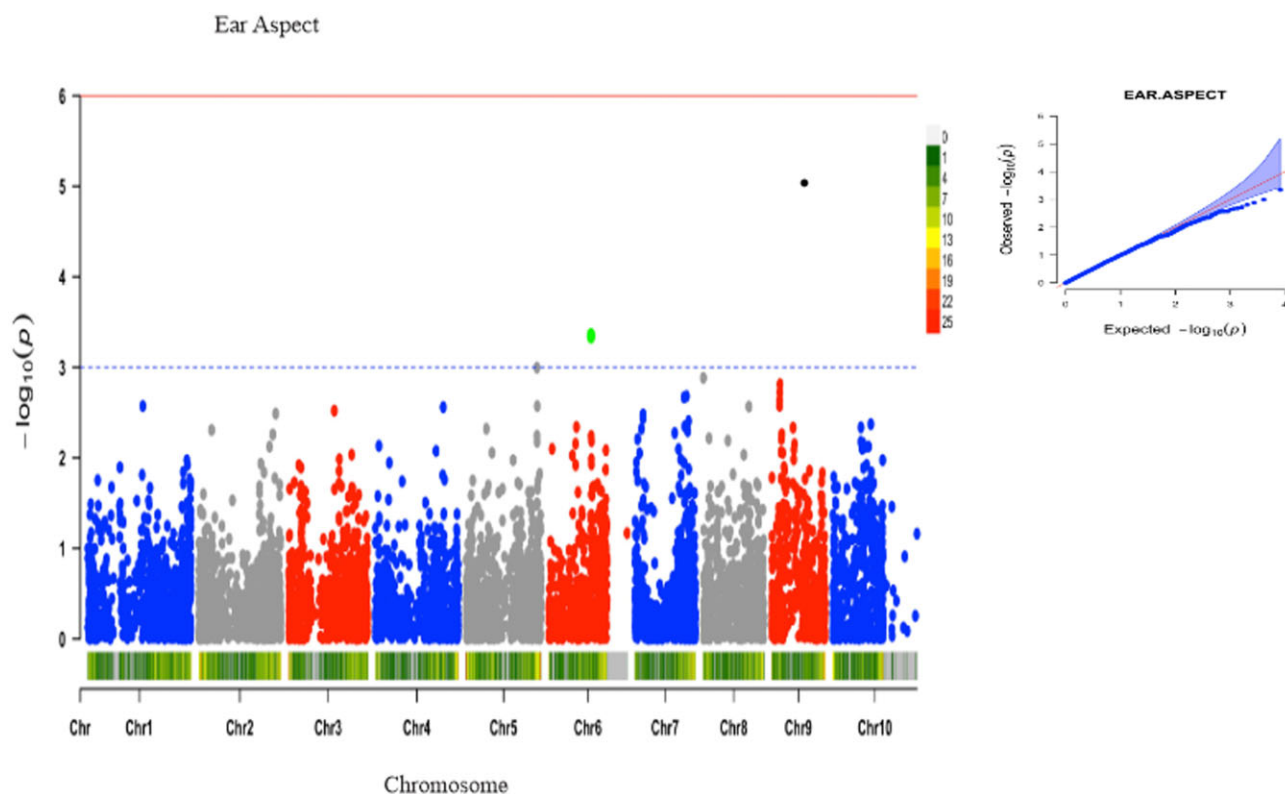
Fig. 5. The Manhattan and Q–Q plots of the SNP-based association mapping of ear aspect under artificial *Striga* infestation.

markers from this study could accelerate the use of genomics-informed selection techniques in breeding *Striga* resistant maize cultivars.

Several QTL associated with various *Striga* resistance indicator traits have been reported in earlier studies. For example, Badu-Apraku *et al.* (2020) identified 116 QTLs associated with 4 *Striga* resistance traits using biparental population derived from 2 early maturing maize lines. A similar study by Badu-Apraku *et al.* (2020) revealed 14 QTLs associated with 3 *Striga* resistance traits. Adewale *et al.* (2020) reported 24 markers significantly associated with *Striga* damage, number of emerged *Striga* plants, number of ears per plant, ear aspect, and grain yield. Also, Gowda *et al.* (2021) identified a total of 57 significant markers for *Striga* resistant traits and yield distributed across the maize genome and controlled by a few major and many minor genes. Stanley *et al.* (2021) identified 30 significant SNPs that were significantly associated with 3 *Striga* resistance trait. Among these studies, S8_17232945 which was detected for *Striga* damage syndrome rating in this study was found to overlap with 1 QTL reported by Badu-Apraku *et al.* (2020). The differential QTL mappings observed in these experiments could be attributed to differences in the genetic materials used for the studies (Kaur *et al.* 2021).

The QTL analysis in this study, and previous studies provided information on the chromosonal regions controlling *Striga* resistance, which can be crucial to breeding *Striga* resistance cultivars through marker-assisted breeding. However, these QTLs have not been utilized in maize *Striga* resistance breeding due to factors, such as limited marker–trait association, a low number of markers used in mapping, small phenotypic variance explained, differences in the genetic backgrounds, and environmental effects (William *et al.* 2007; Tuberosa 2012). A meta-QTL analysis of the results of these findings can be employed to refine the number and position of the QTLs to identify stable QTLs (Sheoran *et al.* 2022). Meta-QTL

analysis has been conducted to successfully locate the regions in the genomes of various traits in different crops. In maize, Chen *et al.* (2017) subjected 999 QTLs to meta-QTL analysis and obtained a total of 76 MQTLs across the maize genome. Three potential candidate genes (GRMZM2G359974, GRMZM2G301884, and GRMZM2 G083894) were associated with kernel size and weight within 3 MQTL using regional association mapping. Guo *et al.* (2018) using a total of 428 individual QTLs for 23 root-related traits identified 53 Meta-QTLs (MQTLs) retrieved over 10 maize chromosomes. Three maize genes (GRMZM5G813206, GRMZM2G167220, and GRMZ M2G467069) that could play important roles on lateral root and crown root development of maize were also identified. Wang *et al.* (2022) carried out MQTL analysis using 282 QTLs from 25 experiments and identified 11 and 34 MQTLs associated with grain dry matter and low grain water content, respectively. Sheoran *et al.* (2022) using a total of 542 QTLs, detected 32 mfeta-QTL possessing 1,907 candidate genes for different abiotic (drought, water logging, heat, and cold) stresses. Furthermore, MQTL 2.1, 5.1, 5.2, 5.6, 7.1, 9.1, and 9.2 were found to control different stress-related traits for combined abiotic stress tolerance. Meta-QTL analyses have also been conducted in crops such as Barley (Zhang *et al.* 2017; Khahani *et al.* 2019), Wheat (Soriano *et al.* 2021; Venske *et al.* 2019), Rice (Sandhu *et al.* 2021; Prakash *et al.* 2022), and Cotton (Said *et al.* 2013; Xu *et al.* 2020).

Conclusion

The 141 QPM inbred lines used for this study displayed high genetic variability in response to the *Striga*-related traits. The GWASs conducted in this research revealed 22 SNPs that were significantly associated with *Striga* resistance adaptive traits. The identified SNP markers after validation would be invaluable for molecular breeding for maize *Striga* resistance in SSA. A meta-

QTL analysis should be employed for the identification of stable QTLs from existing genomic regions that have been associated with yield and other *Striga* resistance traits. This will facilitate the detection of putative genes underlying *Striga* resistance in maize.

Data availability

The datasets used in the present study are available at the IITA CKAN repository. Genotypic data of 141 extra-early maturing quality protein maize inbred lines: https://doi.org/10.25502/k6f5-t130/d. Phenotypic data of 169 extra-early maturing quality protein maize evaluated under *Striga* conditions across 3 environments, 2019–2020: https://doi.org/10.25502/14yp-5w46/d. Supplemental material is available at G3 online.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: GO and BB; Methodology: GO, BB, PA, and QO; Formal analysis: PA and GO; Investigation: GO; Original draft writing: GO; Review and editing: OA, PA, MA, and QO; Supervision: BB, OA, and MA. All authors have read and agreed to the published version of the manuscript.

Literature cited

Adewale SA, Badu-Apraku B, Akinwale RO, Agre AP, Gedil M, Garcia-Oliveira AL. Genome-wide association study of Striga resistance in early maturing white tropical maize inbred lines. BMC Plant Biol. 2020;20:203. https://doi.org/10.1186/s12870-020-02360-0

Akaogu IC, Badu-Apraku B, Tongoona P, Ceballos H, Gracen V, Offei SK, Dzidzienyo D. Inheritance of Striga hermonthica adaptive traits in an early-maturing white maize inbred line containing resistance genes from Zea diploperennis. Plant Breed. 2019;138(5): 546–552.

Alvarado G, Rodríguez FM, Pacheco A, Burgueño J, Crossa J, Vargas M, Pérez-Rodríguez P, Lopez-Cruz MA. META-R: A software to analyze data from multi-environment plant breeding trials. Crop J. 2020. https://doi.org/10.1016/j.cj.2020.03.010.

Azmach G, Gedil M, Menkir A, Spillane C. Marker-trait association analysis of functional gene markers for provitamin A levels across diverse tropical yellow maize inbred lines. BMC Plant Biol. 2013;13:227.

Babiker AGT. Striga: the spreading scourge in Africa. Regul Plant Growth Dev. 2007;42:74–87.

Badu-Apraku B, Oyekunle M, Akinwale RO, Lum AF. Combining Ability of Early-maturing White Maize Inbreds under Stress and Nonstress Environments. Agronomy Journal. 2011;103(2): 544–557.

Badu-Apraku B, Adewale S, Paterne AA, Gedil M, Toyinbo J, Asiedu R. Identification of QTLs for grain yield and other traits in tropical maize under Striga infestation. PLoS One. 2020;15(9):e0239205.

Badu-Apraku B, Akinwale RO, Fakorede MAB. Selection of early maturing maize inbred lines for hybrid production using multiple traits under Striga-infested and Striga-free environments. Maydica. 2010;55(3):261.

Badu-Apraku B, Fakorede MAB, Talabi AO, Oyekunle M, Akaogu IC, Akinwale RO, Annor B, Melaku G, Fasanmade Y, Aderounmu M. Gene action and heterotic groups of early white quality protein maize inbreds under multiple stress environments. Crop Sci. 2016;56(1):183–199.

Badu-Apraku B, Fakorede MAB. Advances in Genetic Enhancement of Early and Extra-Early Maize for Sub-Saharan Africa. Switzerland: Springer; 2017.

Badu-Apraku B, Lum AF. Agronomic performance of Striga resistant early-maturing maize varieties and inbred lines in the savannas of West and Central Africa. Crop Sci. 2007;47(2):737–748.

Badu-Apraku B, Menkir A, Lum AF. Assessment of genetic diversity in extra-early Striga resistant tropical inbred lines using multivariate analyses of agronomic data. J Plant Breed Genet. 2005;59: 67–80.

Badu-Apraku B, Yallou CG, Alidu H, Talabi AO, Akaogu IC, Annor B, Adeoti A. Genetic improvement of extra-early maize cultivars for grain yield and Striga resistance during three breeding eras. Crop Sci. 2016;56(5):2564–2578.

Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67(1):133–199.

Bebawi FF, Mutwali EM. Witchweed management by sorghum-sudan grass seed size and stage of harvest. Agron J. 1991;83(5): 781–786.

Chen L, An Y, Li Y-X, Li C, Shi Y, Song Y, Zhang D, Wang T, Li Y. Candidate loci for yield-related traits in maize revealed by a combination of metaQTL analysis and regional association mapping. Front Plant Sci. 2017;8:2190.

De Groote H, Wangare L, Kanampiu F, Odendo M, Friesen D. Potential Markets for Herbicide Resistant Maize Seed for Striga Control in Africa (No. 724–201649107). 2005. doi: 10.22004/ag.econ.24690.

Earl DA, von Holdt BM. Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. Conservation Genet Resour. 2012;4(2):359–361.

Ejeta G. The Striga scourge in Africa: a growing pandemic. In: Ejeta G, Gressel J, editors. Integrating New Technologies for Striga Control towards Ending the Witch Hunt. London: World Scientific Publishing; 2007. p. 3–16.

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–2620.

Gao L, Turner MK, Chao S, Kolmer J, Anderson JA. Genome wide association study of seedling and adult plant leaf rust resistance in elite spring wheat breeding lines. PLoS One. 2016;11(2):e0148671.

Gowda M, Makumbi D, Das B, Nyaga C, Kosgei T, Crossa J, Beyene Y, Montesinos-López OA, Olsen MS, Prasanna BM. Genetic dissection of Striga hermonthica (Del.) Benth. resistance via genome-wide association and genomic prediction in tropical maize germplasm. Theor Appl Genet. 2021;134(3):941–958.

Gressel J, Hanafi A, Head G, Marasas W, Obilana AB, Ochanda J, Souissi T, Tzotzos G. Major heretofore intractable biotic constraints to African food security that may be amenable to novel biotechnological solutions. Crop Prot. 2004;23(8):661–689.

Hallauer AR, Carena MJ, de Miranda Filho JB. *Quantitative genetics in maize breeding.* Vol. 6. Springer Science & Business Media, 2010.

Hearne SJ. Control—the *Striga* conundrum. Pest Manag Sci. 2009;65(5):603–614.

Jaccoud D, Peng K, Feinstein D, Kilian A. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 2001;29(4):E25. doi:0.1093/nar/29.4.e25.

Kaur S, Rakshit S, Choudhary M, Das AK, Kumar RR. Meta-analysis of QTLs associated with popping traits in maize (Zea mays L.). PLoS One. 2021;16(8):e0256389. doi: 10.1371/journal.pone.0256389.

Khahani B, Tavakol E, Shariati JV. Genome-wide meta-analysis on yield and yield-related QTLs in barley (*Hordeum vulgare* L.). Mol Breeding. 2019;39(4):1–16.

Khan ZR, Midega CAO, Pittchar JO, Murage AW, Birkett MA, Bruce TJA, Pickett JA. Achieving food security for one million sub-Saharan African poor through push-pull innovation by 2020. Philos Trans R Soc Lond B Biol Sci. 2014;369(1639):20120284. doi: 10.1098/rstb.2012.0284.

Kibe M, Nair SK, Das B, Bright JM, Makumbi D, Kinyua J, Suresh LM, Beyene Y, Olsen MS, Prasanna BM, *et al.* Genetic dissection of resistance to gray leaf spot by combining genome-wide association, linkage mapping, and genomic prediction in tropical maize germplasm. Front Plant Sci. 2020;11:1602. doi: 10.3389/fpls.2020.572027.

Kim SK, Winslow MD. Progress in breeding maize for *Striga*-tolerance/resistance at IITA 1991. In: Ransom JK, editor. Proceeding of 5th Intl. Symp. Parasitic Weeds. Nairobi, Ibadan: IITA; 1991. p. 494–499.

Kim SK. Breeding maize for *Striga* tolerance and the development of a field infestation technique. In: Kim SK, editor. Combating Striga in Africa. Proceedings of the Workshop by IITA, ICRISAT and IDRC, Aug. 22–24, 1988. Ibadan: IITA; 1991. p. 96–108.

Kim SK. Genetics of maize tolerance of *Striga hermonthica*. Crop Sci. 1994;34(4):900–907.

Kim SK. Horizontal resistance: core to a research breakthrough to combat *Striga* in Africa. Integr Pest Manag Rev. 1996;1(4):229–249.

Kling JG, Fajemisin JM, Badu-Apraku B, Diallo A, Menkir A, Melake-Berhan A. *Striga* resistance breeding in maize. In: Haussmann BIG, editor. Breeding for Striga Resistance in Cereals. Proc. of a Workshop Held at IITA, Ibadan, Nigeria. Weikersheim, Germany: Margraf Verlag; 2000. p. 103–118.

Li X, Zhou Z, Ding J, Wu Y, Zhou B, Wang R, Ma J, Wang S, Zhang X, Xia Z, *et al.* Combined Linkage and Association Mapping Reveals QTL and Candidate Genes for Plant and Ear Height in Maize. Front Plant Sci. 2016;7:833.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. GAPIT: genome association and prediction integrated tools. Bioinformatics. 2012;28(18):2397–2399. doi:10.1093/bioinformatics/bts444.

Liu HJ, Yan J. Crop genome-wide association study: a harvest of biological relevance. Plant J. 2019;97(1):8. doi:10.1111/tpj.14139.

Mandumbu R, Mutengwa C, Mabasa S, Mwenje E. Challenges to the exploitation of host plant resistance for Striga management in cereals and legumes by farmers in sub-Saharan Africa: a review. Acta Agriculturae Scandinavica, Section B — Soil & Plant Science. 2019;69(1):82–88.

Menkir A, Meseka S. Genetic Improvement in Resistance to Striga in Tropical Maize Hybrids. Crop Science. 2019;59(6):2484–2497.

Menkir A, Kling JG. Response to recurrent selection for resistance to *Striga hermonthica* (Del.) Benth in a tropical maize population. Crop Sci. 2007;47(2):674–682.

Menkir A, Makumbi D, Franco J. Assessment of Reaction Patterns of Hybrids to Striga hermonthica (Del.) Benth. under Artificial Infestation in Kenya and Nigeria. Crop Science. 2012;52(6):2528–2537.

Mogga M, Sibiya J, Shimelis H, Lamo J, Yao N. Diversity analysis and genome-wide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. PLoS One. 2018;13(6):e0198012.

Mohamed K, Bolin J, Musselman L, Peterson A.Genetic diversity of *Striga* and implications for control and modeling future distributions. In: Ejeta G, Gressel J, editors. Integrating New Technologies for Stiga Control: Towards Ending the Witch-Hunt. Singapore: World Sei. Publ.; 2007. p. 71–84.

Nickrent DL, Musselman LJ. Introduction to parasitic flowering plants. Plant Health Instructor. 2004;13:300–315.

Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35(3):526–528.

Parker C, Riches CR. *Striga*, the witchweeds, on cereal crops. In: Parasitic Weeds of the World: Biology and Control. Wallingford (UK): CAB International; 1993. p. 1–74.

Parker C. Parasitic Weeds: A World Challenge. Weed Sci. 2012;60(2):269–276.

Parker C. The parasitic weeds of the orobanchaceae. In: Joel DM, Gressel J, Musselman LJ, . Parasitic Orobanchaceae. Parasitic Mechanisms and Control Strategies. New York (NY): Springer; 2013. 313–343.

Pennisi E. Armed and dangerous. Science (New York, NY). 2010; 327(5967):804–805.

Prakash NR, Lokeshkumar BM, Rathor S, Warraich AS, Yadav S, Vinaykumar NM, Dushynthkumar BM, Krishnamurthy SL, Sharma PC. Meta-analysis and validation of genomic loci governing seedling and reproductive stage salinity tolerance in rice. Physiol Plant. 2022;174(1):13629.Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–959.

Said JI, Lin Z, Zhang X, Song M, Zhang J. A comprehensive meta-QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics. 2013;14(1):776.

Sandhu N, Pruthi G, Prakash Raigar O, Singh MP, Phagna K, Kumar A, Sethi M, Singh J, Ade PA, Saini DK. Meta-QTL analysis in rice and cross-genome talk of the genomic regions controlling nitrogen use efficiency in cereal crops revealing phylogenetic relationship. Front Genet. 2021;12:807210. doi: 10.3389/fgene.2021.807210.

Sheoran S, Gupta M, Kumari S, Kumar S, Rakshit S. Meta-QTL analysis and candidate genes identification for various abiotic stresses in maize (*Zea mays* L.) and their implications in breeding programs. Mol Breeding. 2022;42(5):1–26.

Shikha K, Shahi JP, Vinayan MT, Zaidi PH, Singh AK, Sinha B. Genome-wide association mapping in maize: status and prospects. Biotech. 2021;11(5):1–10.

Sibhatu B. Review on *Striga* weed management. Int J Life Sci Scienti Res. 2016;2(2):110–120.

Soriano J, Colasuonno P, Marcotuli I, Gadaleta A. Meta-QTL analysis and identification of candidate genes for quality, abiotic and biotic stress in durum wheat. Scientific Reports. 2021;11(1):1–5.

Stanley AE, Menkir A, Ifie B, Paterne AA, Unachukwu NN, Meseka S, Mengesha WA, Bossey B, Kwadwo O, Tongoona PB, *et al.* Association analysis for resistance to Striga hermonthica in diverse tropical maize inbred lines. Sci Rep. 2021;11(1):24193.

Sukumaran S, Reynolds MP, Sansaloni C. Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. Front Plant Sci. 2018;9:81.

Tang Y, Liu X, Wang JLM, Wang Q, Tian F. GAPIT Version 2: an enhanced integrated tool for genomic association and prediction. Plant Genome. 2016;9:1–9. doi:10.3835/plantgenome2015.11.0120.

Tuberosa R. Phenotyping for drought tolerance of crops in the genomics era. Front Physiol. 2012;3:347.

Venske E, Dos Santos RS, Farias DdR, Rother V, Da Maia LC, Pegoraro C, Costa De Oliveira A. Meta-Analysis of the QTLome of Fusarium Head Blight Resistance in Bread Wheat: Refining the Current Puzzle. Front Plant Sci. 2019;10:727.

Wang M, Yan J, Zhao J, Song W, Zhang X, Xiao Y, Zheng Y. Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Sci. 2012;196:125–131.

Wang W, Ren Z, Li L, Du Y, Zhou Y, Zhang M, Li Z, Yi F, Duan L. Meta-QTL analysis explores the key genes, especially hormone related genes, involved in the regulation of grain water content and grain dehydration rate in maize. BMC Plant Biol. 2022;22(1):346.

William HM, Trethowan R, Crosby-Galvan EM. Wheat breeding assisted by markers: CIMMYT's experience. Euphytica. 2007;157(3):307–319.

Xu S, Pan Z, Yin F, Yang Q, Lin Z, Wen T, Zhu L, Zhang D, Nie X. Identification of candidate genes controlling fiber quality traits in upland cotton through integration of meta-QTL, significant SNP and transcriptomic data. J Cotton Res. 2020;3(1):1–12.

Xue Y, Warburton ML, Sawkins M, Zhang X, Setter T, Xu Y, Grudloyma P, Gethi J, Ribaut J-M, Li W, *et al.* Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. Theor Appl Genet. 2013;126(10):2587–2596.

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2006;38(2):203–208.

Zhang X, Shabala S, Koutoulis A, Shabala L, Zhou M. Meta-analysis of major QTL for abiotic stress tolerance in barley and implications for barley breeding. Planta. 2017;245(2):283–295.

Zhu XM, Shao XY, Pei YH, Guo XM, Li J, Song XY, Zhao MA. Genetic diversity and genome-wide association study of major ear quantitative traits using high-density SNPs in maize. Front Plant Sci. 2018;9:966.

Communicating editor: A. Lipka