Nisa *et al. BMC Genomics* (2024) 25:760
https://doi.org/10.1186/s12864-024-10655-x

BMC Genomics

RESEARCH Open Access

Insights into maydis leaf blight resistance in maize: a comprehensive genome-wide association study in sub-tropics of India

Wajhat- un- Nisa[1], Surinder Sandhu[1*], Sudha Krishnan Nair[2], Harleen Kaur[1], Ashok Kumar[3], Zerka Rashid[2], Gajanan Saykhedkar[2] and Yogesh Vikal[4]

Abstract

Background In the face of contemporary climatic vulnerabilities and escalating global temperatures, the prevalence of maydis leaf blight (MLB) poses a potential threat to maize production. This study endeavours to discern marker-trait associations and elucidate the candidate genes that underlie resistance to MLB in maize by employing a diverse panel comprising 336 lines. The panel was screening for MLB across four environments, employing standard artificial inoculation techniques. Genome-wide association studies (GWAS) and haplotype analysis were conducted utilizing a total of 128,490 SNPs obtained from genotyping-by-sequencing (GBS).

Results GWAS identified 26 highly significant SNPs associated with MLB resistance, among the markers examined. Seven of these SNPs, reported in novel chromosomal bins (9.06, 5.01, 9.01, 7.04, 4.06, 1.04, and 6.05) were associated with genes: bzip23, NAGS1, CDPK7, aspartic proteinase NEP-2, VQ4, and Wun1, which were characterized for their roles in diminishing fungal activity, fortifying defence mechanisms against necrotrophic pathogens, modulating phyto-hormone signalling, and orchestrating oxidative burst responses. Gene mining approach identified 22 potential candidate genes associated with SNPs due to their functional relevance to resistance against necrotrophic pathogens. Notably, bin 8.06, which hosts five SNPs, showed a connection to defense-regulating genes against MLB, indicating the potential formation of a functional gene cluster that triggers a cascade of reactions against MLB. In silico studies revealed gene expression levels exceeding ten fragments per kilobase million (FPKM) for most genes and demonstrated coexpression among all candidate genes in the coexpression network. Haplotype regression analysis revealed the association of 13 common significant haplotypes at Bonferroni ≤ 0.05. The phenotypic variance explained by these significant haplotypes ranged from low to moderate, suggesting a breeding strategy that combines multiple resistance alleles to enhance resistance to MLB. Additionally, one particular haplotype block (Hap_8.3) was found to consist of two SNPs (S8_152715134, S8_152460815) identified in GWAS with 9.45% variation explained (PVE).

Conclusion The identified SNPs/ haplotypes associated with the trait of interest contribute to the enrichment of allelic diversity and hold direct applicability in Genomics Assisted Breeding for enhancing MLB resistance in maize.

Keywords Maize, Maize leaf blight, Genome wide association studies, SNP, Haplotype analysis

*Correspondence:
Surinder Sandhu
surindersandhu@pau.edu

Full list of author information is available at the end of the article

Introduction

Maize stands as a cornerstone among global cereal crops, holding particular significance in regions like Sub-Saharan Africa, Latin America, and key Asian nations, contributing over 20% of food calories [1]. Notably, Asia, boasting eight major maize-producing countries, plays a pivotal role, producing 28% of the world's maize [2]. However, maize cultivation in the Asian tropics faces formidable challenges, particularly during the monsoon season, where diseases such as Maydis Leaf Blight (MLB) pose a substantial threat. MLB, caused by the necrotrophic fungus *Bipolaris maydis* also known as *Drechslera maydis*, has emerged as a major economic concern, inducing yield losses of up to 30% in warm and humid conditions [3]. The historical significance of MLB, notably its devastating outbreak in the USA in 1970, underscores its global impact [4].

The fungus responsible for MLB exhibits three physiological races, with race 'O' being cosmopolitan and highly aggressive. Race 'O' and race 'T' produce phytotoxins, Hm-O and Hm-T toxins, respectively, leading to varying degrees of virulence [5]. MLB, characterized by spindle-shaped lesions on leaves, adversely affects photosynthetic activity, resulting in significant yield reduction [6]. Resistance to MLB is complex, involving quantitative inheritance with additive and recessive gene actions [7]. Despite its economic impact, our understanding of the resistance mechanisms against necrotrophic pathogens, especially MLB, lags behind [8].

This study addresses the critical gap in our knowledge by employing advanced genomics tools, particularly genome wide association studies (GWAS), to unravel the genetic basis of MLB resistance in maize [9]. QTL mapping using Traditional biparental mapping populations has consistently proven to be a powerful approach for identifying loci that co-segregate with the trait of interest within the research population and is known to detect rare variant along with the identification of source of favourable QTL alleles [10] However, It can only test the diversity of segregating alleles between the parental strains, and the mapping resolution depends on the number of recombination events that occur during population development, as discussed by Mitchell-Olds et al. [11]. Additionally, markers are often sparse due to limited recombination events. Indeed, GWAS have become a powerful tool for understanding the genetic basis of various traits and identifying causative loci and genes. GWAS, investigates the associations between genetic markers, and phenotypes of interest across a diverse set of unrelated individuals or lines (unrelated individuals means distantly related and heterogeneous individuals) of a diverse collection [12]. In association mapping panels, historical recombinations that have accumulated over generations, along with long-standing LD established over dozens or even hundreds of generations, persist among the representative accessions. This persistence contributes to improved resolution in association analysis [13] facilitated by the rapid decay of LD.

To date, 140 QTLs have been reported against the disease, with chromosome 3 harbouring the maximum number of QTLs (at 25 loci). Bin 3.04 was reported in many studies [14–19]; as possessing a major QTL and was validated in different genetic backgrounds. In addition, *rhm1*, a recessive gene with a large effect, was reported in bin 6.01 for race O of *C. heterostrophus* (*Bipolaris maydis*) [17]. Further investigation fine-mapped this gene to 8.56 kb region, within which resides a sole potential candidate gene named lysine histidine transporter 1 (*LHT1*), as revealed by Zhao et al. [20]. Additionally, *Zea maize ascorbate peroxidase* gene (*ZMAPx1*) has recently been identified to promote MLB resistance [21]. Another gene *ZmCCoAOMT2*, was reported to play a crucial role in imparting quantitative resistance to MLB [22] in multiple disease resistance QTL *qMdr9.02*.

Certain GWAS have been conducted for MLB, providing valuable insights into the genetic basis of resistance [23 ,24, 25] in American, and Chinese environments. However, uptill now association mapping studies are not conducted in the subtropics of India. The CIMMYT Asia Association Mapping (CAAM) panel, is the collection of diverse set of lines, of diverse origin, from tropics/sub-tropics of Asia having geographical adaptation to our region and represent ample of genetic diversity for various traits including MLB. Well adapted resistant variants will not only add to allelic diversity but can also serve as potential donors to breed for MLB tolerant cultivars, a disease of economic importance in India.

Haplotype is defined as a set of nearby genomic structural variations, such as polymorphic SNPs, with a strong linkage disequilibrium (LD) between them. The use of haplotypes are known to overcome the biallelic limitation of SNPs, enhance the efficiency of identifying QTLs, and offer insights into genetic determinants that individual or independent marker approaches may miss [26]. Therefore, this research aims to contribute valuable insights into the genomic landscape of MLB resistance, providing a foundation for marker-assisted breeding programs with the following objectives: evaluating the diverse array of the association mapping panel (AMP) for their responses to MLB through multilocational artificial screening, identify marker -trait associations and candidate genes controlling quantitatively inherited MLB resistance through genomic wide association studies and perform haplotype analysis Furthermore, the in-depth in silico characterization of identified candidate genes to enrich our comprehension of the defence mechanisms deployed against MLB [27]. This study not only addresses a critical agricultural challenge but also lays the groundwork for

harnessing genomic tools for sustainable maize production by developing resilient maize varieties [28]. in vulnerable regions.

Materials and methods

Planting material

A subset of CAAM panel, consisting of 359 tropical/subtropical inbred lines (Table S1) was acquired from CIMMYT, Hyderabad. The panel underwent rigorous evaluation through artificial inoculation to assess resistance against MLB. This panel comprises maize inbred lines derived from various subtropical and tropical pools within CIMMYT populations from diverse maize programs. Specifically tailored for Asian environments, the panel not only showcases tolerance to abiotic stresses like drought, high temperature, and excess moisture but also exhibits resistance to biotic stresses such as downy mildew, underscoring its diversity and suitability for mapping relevant traits in the region [29].

Phenotypic evaluation of the mapping panel

The CAAM panel underwent evaluation at two locations with high MLB incidence in the Punjab State of India, characterized by a humid subtropical climate. These locations were Punjab Agricultural University (PAU) in Ludhiana and the Regional Research Station, PAU, in Gurdaspur, during the *Kharif* seasons of 2020 (Y1) and 2021 (Y2). These two locations represent distinct agro-climatic zones within Punjab State: Ludhiana falls within the central plain zone (30.9°N;75.85°E; 733 mm/year rainfall), while Gurdaspur is situated in the sub-mountain undulated zone (32.04°N; 75.40°E; 1167.8 mm/year rainfall). The experimental design employed an alpha lattice pattern with two replications in each environment. Each entry was planted in paired rows maintaining a spacing of 60 cm between rows and 20 cm between individual plants.

Preparation of mass culture and inoculation procedure

The most virulent isolate of *Drechslera maydis* (*Dm1*) was selected for mass culture. Mass multiplication of fungal culture was performed on sterile sorghum grains (*Sorghum bicolor* L.) following the methods of Lim [30]. Inoculated flasks containing sorghum grains were incubated at 25±2 °C for 15 days until the grains were uniformly covered with fungal growth. The impregnated sorghum grains were dried by spreading them on a clean paper sheet in the shade at room temperature. After drying, fine powder of these grains were prepared with the help of a mixer grinder. Whorl inoculations were performed by placing 2gm of powdered grains comprising of fungal isolate in the whorls of each plant at 35–40 days after sowing (DAS). Adequate moisture for a longer period to permit spore germination was obtained by spraying 10–12 ml of water in the whorls using a sprayer. To avoid the maximum day temperature (to avoid mortality by direct exposure to sun sight) during the incubation period, inoculation was performed in the late afternoon (4–6 p.m.)

Data recording

The disease reaction data were recorded on 10 plants from plot of two rows. randomly avoiding border plants following a scale of 1–9 of Hooda et al. [31] at two intervals, viz., 45 days after inoculation (DAI) and 55 DAI. Phenotypes with a rating of 1 or 3 had yellow-brown chlorotic lesion that do not intersect with each other, whereas phenotypes with ratings of 6–9 had large elongated necrotic lesions.

Statistical analysis of phenotypic data

All phenotypic data analyses were carried out in META-R (Multienvironment Trial Analysis with R for Windows) version 6.0 developed by CIMMYT [32]. BLUPs were calculated across the individual environments (E1, E2, E3, and E4) and for the data from both locations, i.e., Ep1 (Ludhiana) and Ep2 (Gurdaspur). All three datasets were used for GWAS analysis. The linear models were implemented from the package lme4 of R in META-R to calculate the BLUPs and variance components. The following linear model was used for analyzing the individual environments:

$$Yijk = \mu + Repi + Blockj\ (Repi) + Genk + ijk$$

where *Yijk* is the MLB severity, representing phenotypic performance of the *k*th genotype at the *j*th block in the *i*th replication, μ is the overall mean effect, Rep_i is the effect of the *i*th replicate, $\text{Block}_j(\text{Rep}_i)$ is the effect of the *j*th incomplete block within the *i*th replicate, Gen_k is the effect of the *k*th genotype and ε_{ijk} is the effect of the error associated with the *i*th replication, *j*th incomplete block, and *k*th genotype, which is assumed to be normal with mean zero and variance. For a combined analysis across years, the following linear model was used:

$$\begin{aligned} Yijkl = &\mu + Env_i + Rep_j\,(Env_i) + Block_k\,(Env_i Rep_i) \\ &+ Gen_l + Env_i \times Gen_l + \varepsilon_{ijkl} \end{aligned}$$

$\text{Env}_i$is the effect of the *ith* environment and $\text{Env}_i\times$ $\text{Gen}_l$is the environment × genotype (G ×E) interaction, In both models, all effects, except the overall mean, are declared to be random and normal with a mean of zero and effect-specific variances. The random assumption for the genotype effects allowed us to calculate BLUPs and broad-sense heritability.

$$H^2 = \sigma^2 g / \left(\sigma^2 g + \sigma^2 ge/k + \sigma^2 e/rk\right)$$

where $\sigma^2 g$, $\sigma^2 ge$ and $\sigma^2 e$ are the number of genotypes, genotype-environment interactions and error variance components, respectively; and r and k are the number of replications within each environment and number of environments, respectively.

DNA isolation and genotyping

Genotyping was conducted on 336 lines using the genotyping by sequencing (GBS) platform (GBSv2.7) [33] at the Institute of Genomic Diversity, Cornell University, Ithaca, USA. The genomic DNA underwent digestion with the *ApekI* restriction enzyme, and GBS libraries were constructed in a 96-plex format before being sequenced on the Illumina HiSeq. 2000 [34]. SNP calling was executed using the TASSEL GBS pipeline [35], with B73 serving as the reference genome. The FILLIN method in TASSEL 5.0 was employed to partially impute missing genotypic data. The partially imputed dataset comprised 955,690 SNPs distributed across all chromosomes. For our GWAS study, filtering criteria of a call rate > 0.9, Minor Allele Frequency (MAF) > 0.05, and heterozygosity < 30% were applied, resulting in a refined dataset of 128,490 SNPs. Principal Component Analysis (PCA), kinship matrix, and linkage disequilibrium (LD) were calculated using this curated set of SNPs.

Principal component (PC) analysis, kinship and LD analysis

PCA [36], kinship, and LD analysis was conducted in Genomic Association and Prediction Integrated Tool version 3 (GAPIT) [37]. A three-dimensional plot of principal components was drawn to visualize the possible population stratification among the samples. A scree plot was generated to determine the number of principal components to be included in the GWAS. The kinship matrix was generated with the Van-raden algorithm and was visualized as a heatmap. The LD was estimated by using all the markers and their neighbouring markers as pairwise r^2 values (the squared correlation among alleles at two SNPs). The LD decay was plotted as r^2 values between SNPs against the physical distances (kb) between SNPs at $r^2 = 0.1$.

Genome-Wide Association Mapping: GWAS was performed on the BLUP values obtained for the final disease score across environments (E1, E2, E3, and E4) and on the pooled dataset of two years at location 1 (Ludhiana) and location 2 (Gurdaspur) (Ep1 and Ep2 respectively) on 336 inbred lines (and as supplementary on individual environments) . GBS was used for genotyping to generate 128,490 SNPs and used for GWAS mapping The SNPs were distributed across all ten maize chromosomes. A density plot was constructed for chromosome wise SNPs within 1 Mb window (Fig. 1a, b).

GWAS was performed using the Bayesian-information and linkage disequilibrium iteratively nested keyway (BLINK) package as implemented in GAPIT version 3 in the R version 4.2.1 software environment [38, 39]. BLINK approach employs a multi-locus model for the evaluation of markers distributed across the genome and conducts two fixed-effect models iteratively. One model tests one marker at a time with multiple associated markers as fixed effects to account for population structure, and the other model tests the covariate markers to control spurious associations [40]. The GWAS results were visualized by plotting $-\log_{10}P$ values as Manhattan plots. The values are plotted against the chromosomal position of the SNPs in GAPIT V.3. The quantile–quantile plots (Q–Q plots) in BLINK represented observed versus expected negative $\log_{10}P$ value that deciphered the severity of inflation test statistics. The set Bonferroni-corrected threshold at $P < 0.1$ was very stringent; therefore, the suggestive or exploratory P value threshold to control the genome-wide type 1 error rate was estimated as $< 9.0 \times 10^{-5}$ for identifying the significant SNPs for MLB from the set of 128,490 markers and considered as the significance cut-off for the association [41, 42].The final number of significant SNPs/MTAs was chosen after accounting for SNPs which were consistent or common across Ep1, Ep2 and combined environmentsThe allelic effects were depicted for four associated markers/SNP. It was determined by using disease score data of 336 inbred lines for both alleles (major and minor) of the SNP. The effect was represented as difference between the groups by box plots using Kruskal-Wallis test. The test statistic H (chi^2) is computed as follows:

$$H = \frac{12}{n(n+1)} \left(\sum_g \frac{T_g^2}{n_g} \right) - 3(n+1)$$

where n_g is the number of elements in group g, n is the total number of elements, and T_g is the sum of ranks in group g. The test was done to see whether the effect of alleles differ significantly in resistant and susceptible lines for the disease score. The analysis was performed in Past V.4.13 [43] and DATA Table (2023) [44].

Haplotype regression analysis

Haplotype regression analysis was executed in SNP & Variation Suite (SVS) Version 8.6.0 (SVS, Golden Helix, Inc., Bozeman, MT, www.Goldenhelix.com). SNPs within the bottom 0.1 percentile of the distribution in GWAS for all three datasets were selected for haplotype detection and trait regression. Haplotype frequency estimation was done using the Expectation Maximisation (EM) algorithm with 50 EM iterations [45]. EM is an iterative optimisation method that uses machine learning algorthims

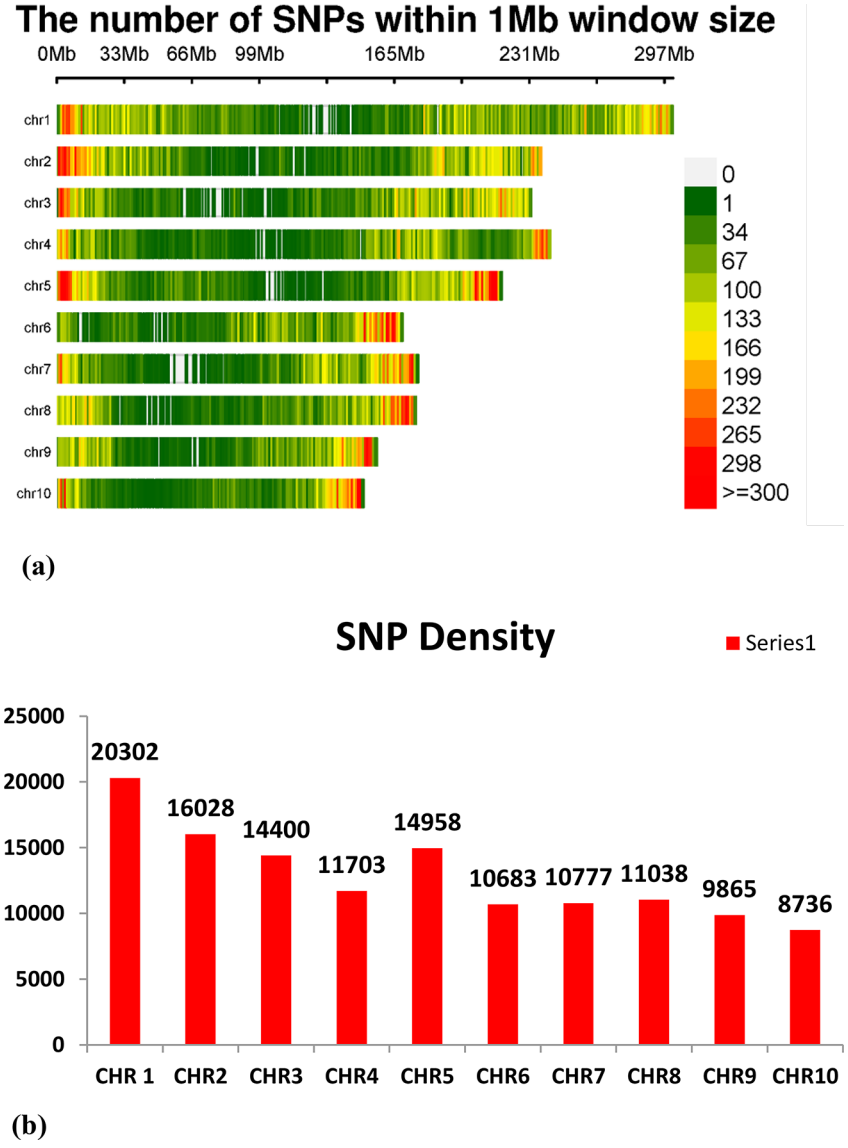

Fig. 1 **(a)** Density plot representing chromosome wise SNPs within 1 Mb window. The horizontal axis of the density plot shows the chromosome length (Mb), and the colours reflect the SNP density distribution, with white indicating the lowest number of SNPs and strong red representing a greater number or density of SNPs. **(b)** Distribution of markers on all 10 chromosomes of maize

to find maximum likelihood and is known to handle missing data. EM convergence tolerance of 0.0001 and a frequency threshold of 0.01 was used. To minimise the historical recombination, haplotype blocks were detected based on the block defining algorithm [46] because the regions with little evidence of recombination among common alleles is considered for assembling the diversity and forms a biological basis of objectively defining haplotype blocks. Regression analysis was carried out with the haplotypes detected, based on step-wise regression of the MLB BLUP estimates of all three datasets with forward elimination at Bonferroni value cut off ≤ 0.05. Haplotype regression analysis uses the expected number of copies of the haplotypes considering genotype as explanatory variable. It has high computational efficiency then Bayesian methods.

Candidate gene mining, in silico expression analysis and interaction predictions

The SNP markers significantly associated with the trait were searched in MaizeGDB (http://www.maizegdb.org) against the reference genome B73_RefGen_V2 to find the physical position of the identified markers and flanking genes. A gene with a marker located within it or the closest high-confidence gene within 0.9 kb flanking of the SNP's physical position was considered as the associated gene to that marker. Information about these genes was gathered from NCBI (https://www.ncbi.nlm.

nih.gov/entrez) and the MaizeGDB (www.maizegdb.org) database. The functions of the predicted candidate genes were reviewed to establish their significance in enhancing resistance. Details on the gene families and Gene Ontology (GO) terms of the candidate genes were obtained from PLAZA (version 5). In silico expression analysis of the genes was carried out using publicly available expression data, presented as fragments per kilobase of transcript per million mapped reads (FPKM), sourced from the q-Teller Maize GDB database (http://www.maizegdb.org). This data, compiled by Walley et al. [47], originated from a comprehensive gene expression atlas constructed through mRNA sequencing (mRNA-seq) involving three biological replicates from 23 distinct tissues. Our analysis specifically considered leaf tissue from various zones (zone 1: symmetrical, leaf zone 2: stomatal, leaf zone 3: growth, and mature leaf), in addition to the vegetative meristem (16–18 days). The gene expression patterns were visualized as a heatmap using TBtools software [48], following the $\log_2$ transformation of FPKM values [49]. To detect interactions and coexpression among the candidate genes, GeneMANIA (https://genemania.org/) was employed. Query list of candidate genes was used as input. By default geneMania prediction server utilizes adaptive network weighted method to dertermine the network. GeneMania extended the list of query genes with the functionaly similar genes. The default force directed COSE network was formed based on weighted sum of individual data sources. The prediction server utilised databases; Gene expression omnibus (GEO) along with Interpro for co-expression data and BioGRID for physical interaction data [50, 51] **Results**.

Phenotypic evaluation of CAAM panel for maydis leaf blight

The panel displayed significant variations (P value < 0.001) in disease severity on a disease scale ranging from 1 to 9. The disease pressure was high for MLB at both locations, as observed by the disease severity score of ≥7 during 2020 and 2021 under artificial epiphytotic conditions. The environment-wise average disease score (DS) ranged from 2.11 to 7.93 (E1), 3.23–7.08 (E2), 2.94–8.9 (E3), and 2.54–8.07 (E4), whereas the DS ranged from 3.15 to 7.93 across the environments (based on a dataset of four environments). The DS ranged from 2.89 to 7.06 at Ep1 (based on two datasets at location 1) and 3.29–7.34 at Ep2 (based on two datasets at location 2). The estimate of broad sense heritability across the environments (0.79) (E1, E2, E3, E4) as well as for pooled within locations; Ep1 (0.61) and Ep2 (0.76) was moderate to high (Table 1). The environment-wise frequency distribution of panel belonging to four classes of resistance is represented by bar plots (Fig. 2). Maximum number of lines were falling in the class of moderate resistance (DS scale 4–5) and moderate susceptibility (DS scale 6–7) in each of four environments.

Population structure, kinship and linkage disequilibrium (LD) analysis

Principal component analysis and kinship analysis of the association panel were conducted using a filtered set of SNPs (128,490). The first three principal components (PCs) encapsulated most of the genetic variation, as illustrated in Fig. 3a. The pairwise relative kinship matrix of the 336 genotypes revealed a low levels of genetic relatedness within the panel, (Fig. 3b). Genome wide LD plot displayed the LD decay of 0.9 at $r^2 = 0.1$ (Fig. 3c).

GWAS for MLB resistance

GWAS was performed with a subset of 128,490 SNPs following BLINK model after a rigorous quality check. This model corrects both kinship (K) and population structure (Q), as depicted by the least genomic inflation deciphered from the Q–Q plot (Fig. 4a, b, c). The distribution of SNPs across chromosomes showed greater density at the ends and lower density at the centromeric regions, with chromosome 1 having the highest number of SNPs (20302) and chromosome 10 having the lowest number of SNPs (8736). The P value threshold was ($< 9.0 \times 10^{-5}$) for identifying the significant SNPs for MLB In each dataset, Manhattan plots (Fig. 4d, e, f) were generated to display the $-\log_{10}P$ value of each SNP from the association study. At P value $< 9.0 \times 10^{-5}$, 13 SNPs were detected across environments (E1, E2, E3, and

Table 1 Variance components and descriptive statistics of CAAM panel for MLB disease score

	Heritability (h^2)	Genotype variance	Residual variance	Mean	Range	LSD	CV
Ludhiana (Ep1)	0.61	0.96***	0.47	5.35	2.89–7.06	2.17	12.87
Gurdaspur (Ep2)	0.76	0.80***	0.63	4.94	3.29–7.34	1.43	16.04
ACROSS	0.79	0.85***	0.56	5.16	3.15–7.93	1.33	14.46
LDH 20 (E1)	0.93	1.29***	0.20	5.04	2.11–7.93	0.87	8.91
GRD20 (E2	0.72	0.75***	0.59	4.90	3.23–7.08	1.55	15.72
LDH21(E3)	0.87	2.54***	0.75	5.65	2.94–8.5	1.71	15.36
GRD21(E4)	0.78	1.23***	0.69	4.98	2.54–8.07	1.65	16.62

***$p < 0.001$, Ep1- pooled environment Ludhiana (Ldh); Ep2- pooled environment Gurdaspur (Grd), across-Combined over environments (E1: Ldh 2020, E2: Grd 2020, E3:Ldh 2021, E4: Grd 2021), h^2-broad sense heritability, LSD- Least square distance, *cv*- coefficient of variation

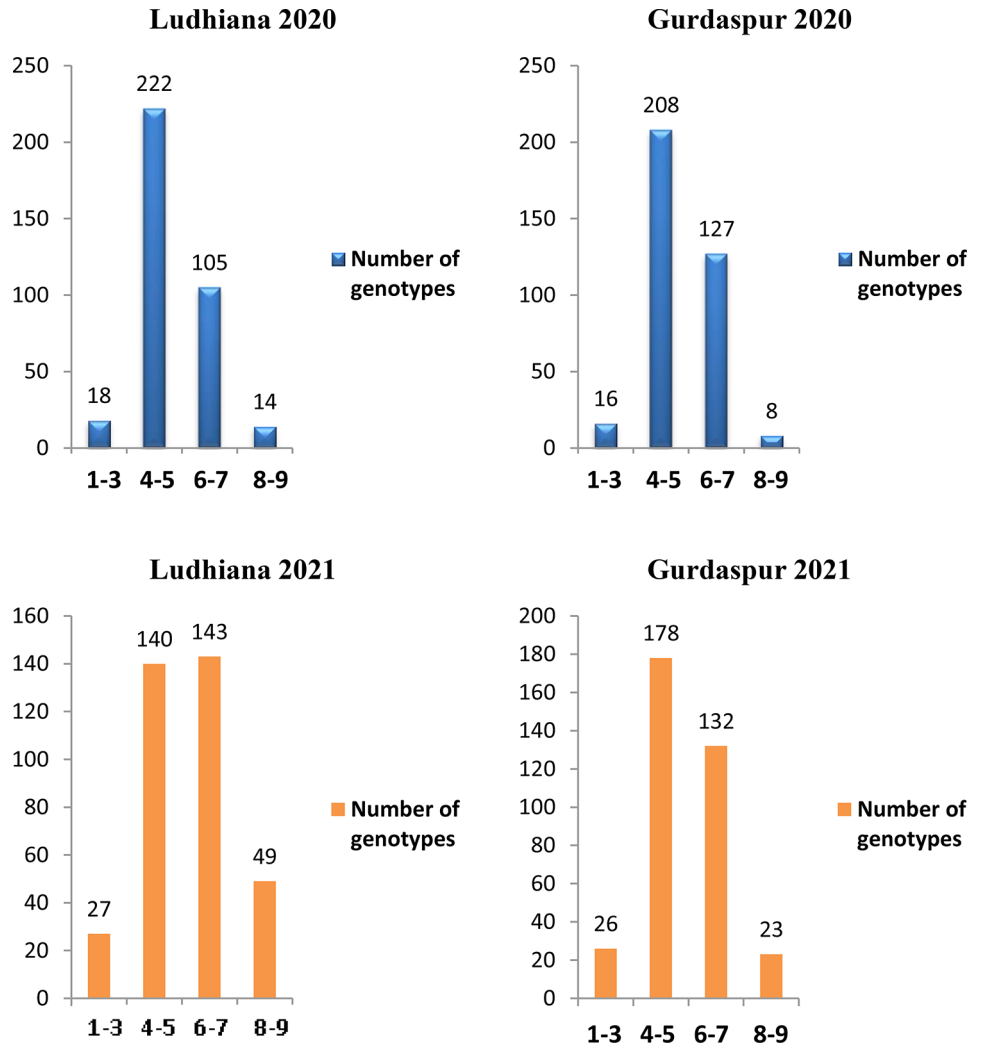

Fig. 2 Frequency distribution of genotypes in different classes of resistance for MLB (1-3-resistant; 4-5-moderate resistant; 6-7-moderate susceptible; 8-9-susceptible)

E4), with *P* value ranging from 1.6×10^{-6} to 8.3×10^{-5} and an effect size ranging from −0.46 to 0.43. Six SNPs were detected in the pooled environment Ep1, for which the *P* value ranged from 9.1×10^{-6} to 7.1×10^{-5} and the effect size ranged from −0.34 to 0.33. Fifteen SNPs were detected in the pooled environment Ep2, for which the *P* value ranged from 3.8×10^{-6} to 8.4×10^{-5}, and the effect size ranged from −0.33 to 0.44. Six SNPs overlapped between across environment analysis and location-wise analysis, whereas only one SNP (S8_155841067) was found to overlap between Ep1 and Ep2. Five SNPs (S8_152460815, S5_140936401, S3_156792785, S8_155841067, and S8_162518701) overlapped in both Ep2 and across environments analysis (E, E2, E3, and E4), and three SNPs (S6_130006038, S8_155841067, and S1_232344813) overlapped among Ep1 and across environments analysis (E, E2, E3, and E4). (Table 2). Of the 34 identified SNPs in the present study, 26 SNPs were considered significant after accounting for common SNPs according to the cumulative analysis of the datasets. SNP S8_155841067 showed the strongest association with the lowest *P* value and was reportedly most stable SNP across (combined), pooled, and individual environments (E2 and E3) as well (Table S2 a, b, c). Chromosome 8 harboured the highest number of significant SNPs (six SNPs) in two different chromosomal bins: 8.06 (5 SNPs) and 8.01 (1 SNP). Seven SNPs that exhibited a significant association with MLB were reported from novel chromosomal bins, viz., 9.01 (S9_8243435), 7.04 (S7_161657633), 9.06 (S9_141454813), 6.05 (S6_141510514), 5.01 (S5_3412526), 1.04 (S1_52252512), and 4.06 (S4_166482019). The allelic effects of the four significant (Fig. 5) and common SNPs reported from combined environments and pooled environments (Ep1, Ep2) were examined using the Kruskal-Wallis test. These SNPs were harboured in the chromosomal bins 8.06 (3)

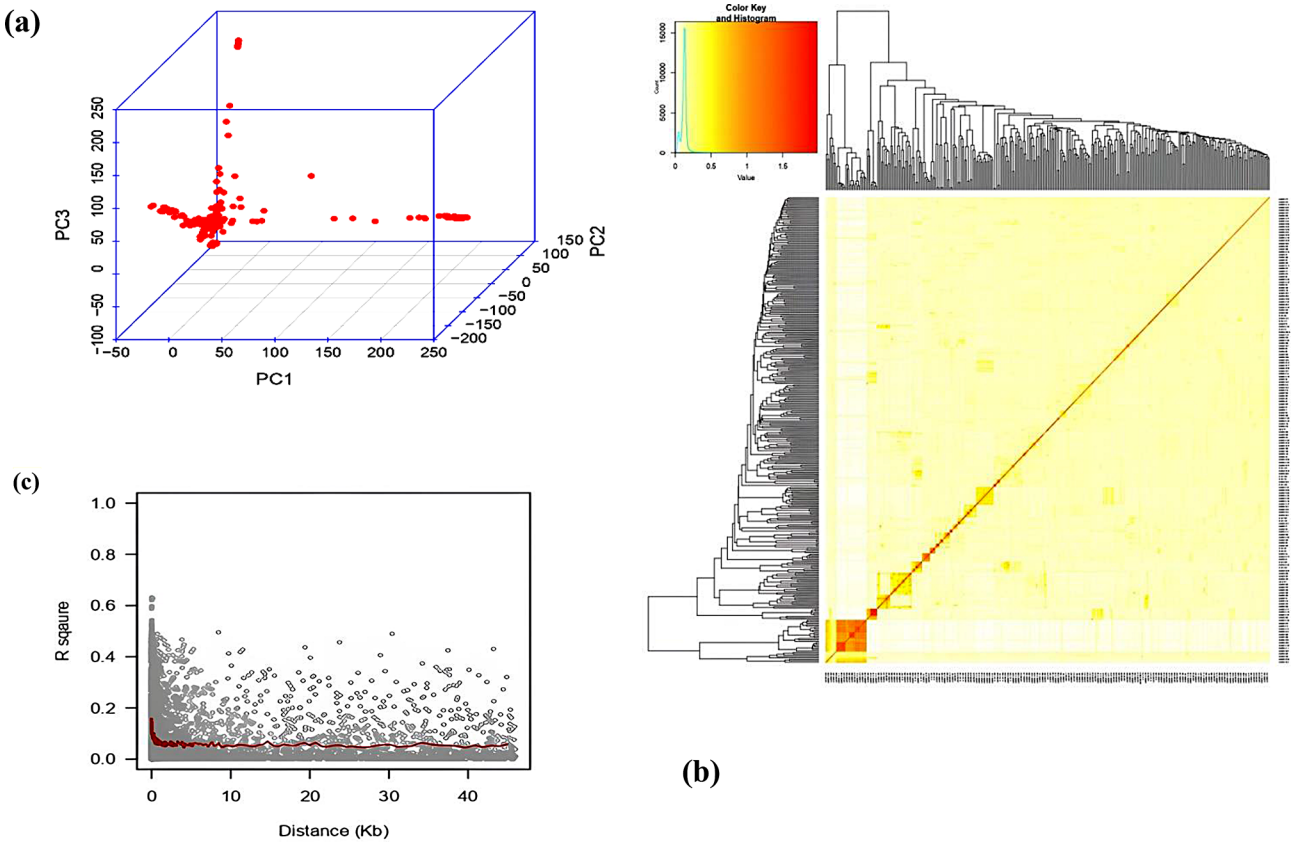

Fig. 3 Principal component analysis (PCA) and Kinship matrix results (**a**) Three-dimensional plot of the first three principal components, (**b**) Kinship heat map depicting genetic distance between 336 inbreds. The colour histogram indicates the distribution of coefficients of co-ancestry, with white to yellow hue for our inbred lines representing lower genetic relatedness as per the colour key, whereas as stronger red represents higher relatedness, (**c**) Linkage disequilibrium (LD) decay plot in CAAM panel based on 128,490 SNPs

and 6.01 (1), which were reported to be enriched with QTLs for resistance against MLB and other diseases. The chi-square values and probability (*P*) values indicated presence of significant phenotypic differences in plant MLB score for allelic effects of all four SNPs (Table 3). Box and whisker plots were employed to illustrate the significant allelic effects of the SNPs for MLB resistance (Fig. 5). For all four SNPs, homozygous combination of favourable alleles were reported in resistant lines with lowest disease score (<3.0), e.g. allele 'CC' (homozygous) for SNP S8_155841067. In contrast, lines with the 'TT ' allele (homozygous) exhibited mean disease scores ranging from 7.9 to 8.0, indicating their susceptibility to MLB.

Haplotype detection and regression analysis for the trait

A set of 188 SNPs in the bottom 0.1 percentile distribution in GWAS study of across, Ep1 and Ep2 were used for haplotype detection. The analysis identified 75 haplotype blocks across the 10 chromosomes. Haplotype Regression Analysis (HTR) was carried out with 75 haplotypes on MLB BLUP estimates of three individual datasets separately. For across environments 31 haplotype blocks were identified at Bonferroni value≤0.05 that explained 4.19–16.05% phenotypic variance. Twenty one significant haplotype blocks were identified for Ep1 with explained phenotypic variance of 3.45–10.56% and 36 haplotype blocks detected were associated with MLB resistance in Ep2 explaining phenotypic variance of 3.78–16.44% (Table S1). Thirteen common significant (Bonferroni value≤0.05) haplotype blocks were identified in HTR analysis of Ep1 and Ep2 environments.

These haplotype blocks which include 2–5 SNPs were spread on seven chromosomes (1, 2, 3, 4, 8, 9 and 10 and the proportion of variance explained by these common blocks ranged from3.45–10.57% (Table 4). These commonly identified significant haplotype blocks were subsequently compared with the SNPs identified in GWAS and candidate genes reported. Among them, one particular haplotype block (Hap_8.3) was found to consist two SNPs (S8_152715134, S8_152460815) identified in GWAS.

Functional annotation of the candidate genes

Twenty-five unique candidate genes were identified for MLB resistance. These genes, were found to possess functional domains associated with biotic stress tolerance. Three candidate gene models (3) were associated

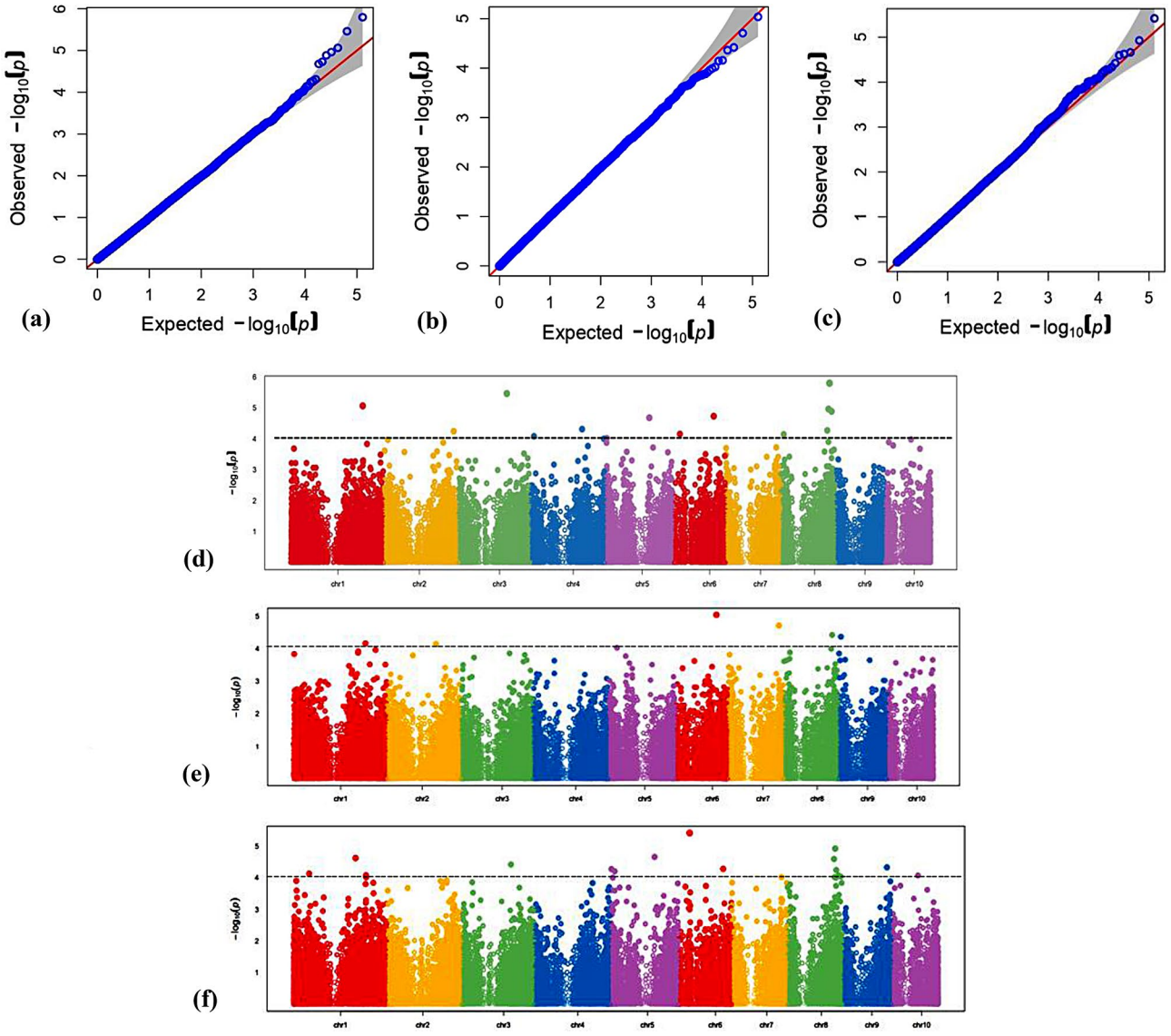

Fig. 4 Quantile–quantile plot (Q–Q plot) of MLB resistance from (**a**) the across environments, (**b**) pooled analysis from two years of evaluation at Env 1 (Ep1), and (**c**) pooled analysis from two years of evaluation at Env 2 (Ep2), representing observed versus expected negative $\log_{10}P$ values deciphering the severity of inflation test statistics. Manhattan plot of GWAS analysis of MLB resistance from (**d**) the across-environment analysis, (**e**) pooled analysis from two years of evaluation at Env 1 (Ep1), and (**f**) pooled analysis from two years of evaluation at Env 2 (Ep2) with the individual SNPs on all chromosomes on the X-axis and $-\log_{10}P$ values of each SNP on the Y-axis. The different colours represent the 10 chromosomes of *Zea mays* L. The black line marks the threshold for an exploratory *P* value

with SNP S8_155841067 (lowest *P* value) on chromosome 8 (bin 8.06), i.e., B3 domain-containing protein (gene harbouring the SNP), SPA1 protein (suppressor of Phya-105 1), and RNA helicase (ATP-dependent helicase rhp16) (genes flanking the physical distance of 0.9-1 kb from the position of SNP) Additional gene models reported from bin 8.06 included the ABC transporter protein family member, AP-4 complex (subunit epsilon), abi20-ABI3-VP1-transcription factor 20, and AP2/ERF domain-containing protein. Moreover, one gene, PR5-like receptor kinase, was identified in bin 8.01. These genes may be functionally relevant for defending against necrotrophic fungi. Genes with different DNA-binding domains mediated by jasmonic acid (JA) signalling, leading to activation of defence, were also identified. For example, the MYB DNA-binding domain superfamily and Indole-3-pyruvate monooxygenase YUCCA1 were found on chromosome 6 (6.01). The basic helix-loop-helix (bHLH) DNA-binding superfamily protein-producing gene was found on chromosome 5 (5.04), the lipolytic acyl hydrolase (LAH)/patatin protein producing gene was reported on chromosome 1 (1.06), and the genes Wun1 and vq4-VQ motif transcription factor 4 were reported in bins 6.05 and 1.04. Apart from the above genes, we

Table 2 Significant SNPs for MLB resistance identified in CAAM panel in different environments

	SNP ©	Chr	Pos (B73 V_2)	*P*.value	MAF	Effect	Allele
Across Environments (E1, E2, E3, E4)	**S8_155841067**	chr8	155,841,067	1.6×10^{-6}	0.08	0.430147	C/T
	S3_156792785	chr3	156,792,785	3.4×10^{-6}	0.10119	-0.46926	T/G
	S1_232344813	chr1	232,344,813	8.6×10^{-6}	0.214286	-0.33994	G/A
	S8_152715134	chr8	152,715,134	1.0×10^{-5}	0.16369	-0.35597	T/A
	S8_162518701	chr8	162,518,701	1.3×10^{-5}	0.331845	-0.26475	G/C
	S6_130006038	chr6	130,006,038	1.8×10^{-5}	0.300595	-0.2915	G/A
	S5_140936401	chr5	140,936,401	2.1×10^{-5}	0.25744	0.266323	C/G
	S4_166482019	chr4	166,482,019	4.8×10^{-5}	0.125	0.371895	C/T
	S8_148676841	chr8	148,676,841	5.3×10^{-5}	0.116071	-0.33136	C/A
	S2_223252193	chr2	223,252,193	5.7×10^{-5}	0.13244	-0.3614	G/C
	S6_21316804	chr6	21,316,804	7.0×10^{-5}	0.171131	-0.3094	T/G
	S8_8887701	chr8	8,887,701	7.2×10^{-5}	0.275298	0.250936	G/A
	S4_11836688	chr4	11,836,688	8.3×10^{-5}	0.107143	0.405554	C/T
Ep1	**S6_130006038**	chr6	130,006,038	9.1×10^{-5}	0.300595	-0.2899	G/A
	S7_161657633	chr7	161,657,633	1.9×10^{-5}	0.19494	0.332909	A/T
	S8_155841067	chr8	155,841,067	3.8×10^{-5}	0.080357	0.35661	C/T
	S9_8243435	chr9	8,243,435	4.3×10^{-5}	0.181548	-0.34259	G/T
	S1_232344813	chr1	232,344,813	6.9×10^{-5}	0.214286	-0.2929	G/A
	S2_157608147	chr2	157,608,147	7.1×10^{-5}	0.389881	-0.19765	A/G
Ep2	**S6_34825812**	chr6	34,825,812	3.8×10^{-5}	0.083333	0.439454	C/T
	S8_152460815	chr8	1.52E+08	1.1×10^{-5}	0.074405	0.449882	T/A
	S5_140936401	chr5	1.41E+08	2.1×10^{-5}	0.25744	0.259789	C/G
	S1_200269986	chr1	2E+08	2.3×10^{-5}	0.08631	0.385707	A/G
	S8_148676841	chr8	1.49E+08	2.5×10^{-5}	0.116071	-0.33722	C/A
	S3_156792785	chr3	1.57E+08	3.8×10^{-5}	0.10119	-0.40914	T/G
	S9_141454813	chr9	1.41E+08	4.6×10^{-5}	0.25744	-0.29864	T/A
	S6_141510514	chr6	1.42E+08	5.2×10^{-5}	0.071429	-0.4189	G/A
	S5_3412526	chr5	3,412,526	5.3×10^{-5}	0.263393	0.254293	A/G
	S8_155841067	chr8	1.56E+08	5.6×10^{-5}	0.080357	0.355416	C/T
	S5_13798307	chr5	13,798,307	6.2×10^{-5}	0.395833	-0.23036	G/T
	S1_52252512	chr1	52,252,512	7.4×10^{-5}	0.090774	0.366379	A/C
	S1_233546091	chr1	2.34E+08	8.3×10^{-5}	0.165179	-0.33468	C/T
	S8_162518701	chr8	1.63E+08	8.3×10^{-5}	0.331845	-0.23456	G/C
	S10_83669175	chr10	83,669,175	8.4×10^{-5}	0.061012	0.442897	G/T

© MAF - minor allele frequency, effect- SNP effect + represents major allele as favorable allele and – represents minor allele as favorable allele, Ep1 represents pooled dataset of two years at E1(Ludhiana), Ep2 represents pooled dataset of two years at E2 (Gurdaspur)

identified one NBS-LRR defence protein-encoding gene, PIK6-NP, on chromosome 10 (10.3). Genes Brassinosteroid-insensitive 1-associated receptor kinase and ubiquitin protein (ligase-binding) were reported on chromosome 1 (1.08). Gene Sterol 3-beta-glucosyltransferase UGT80A2 was reported on chromosome 2 (2.08). Other genes associated with the significant SNPs that are identified in different chromosomal bins were CLAVATA3 embryo surrounding region-related-16 (9.06), calcium-dependent protein kinase 7 (5.01), bZIP transcription factor 23 (9.01), putative amino acid acetyl-transferase NAGS1 (7.04), aspartic proteinase nepenthesin-2 (4.06), vq4-VQ motif-transcription factor 4 (1.04), the wound-induced protein Wun1, and the auxin responsive Aux/IAA family member (6.05). These genes have been functionally annotated in the literature for the induction of pathogenesis-related (PR) gene expression and therefore can be considered to prompt defence against MLB (Table 5).

In silico expression and gene interaction analysis of putative candidate genes

Gene expression data were specifically curated for 22 candidate genes, as detailed in Fig. 6a, b. The remaining genes (3) did not have available expression data across the four leaf stages or for the vegetative meristem in the database. Genes GRMZM2G033413, GRMZM5G813007, GRMZM2G031352, AC210013.4_FG014, GRMZM2G061602, GRMZM2G164787, GRMZM2G031584, and GRMZM2G313737 exhibited increased expression in all four stages of the leaf and vegetative meristem, whereas GRMZM2G033413/bZIP

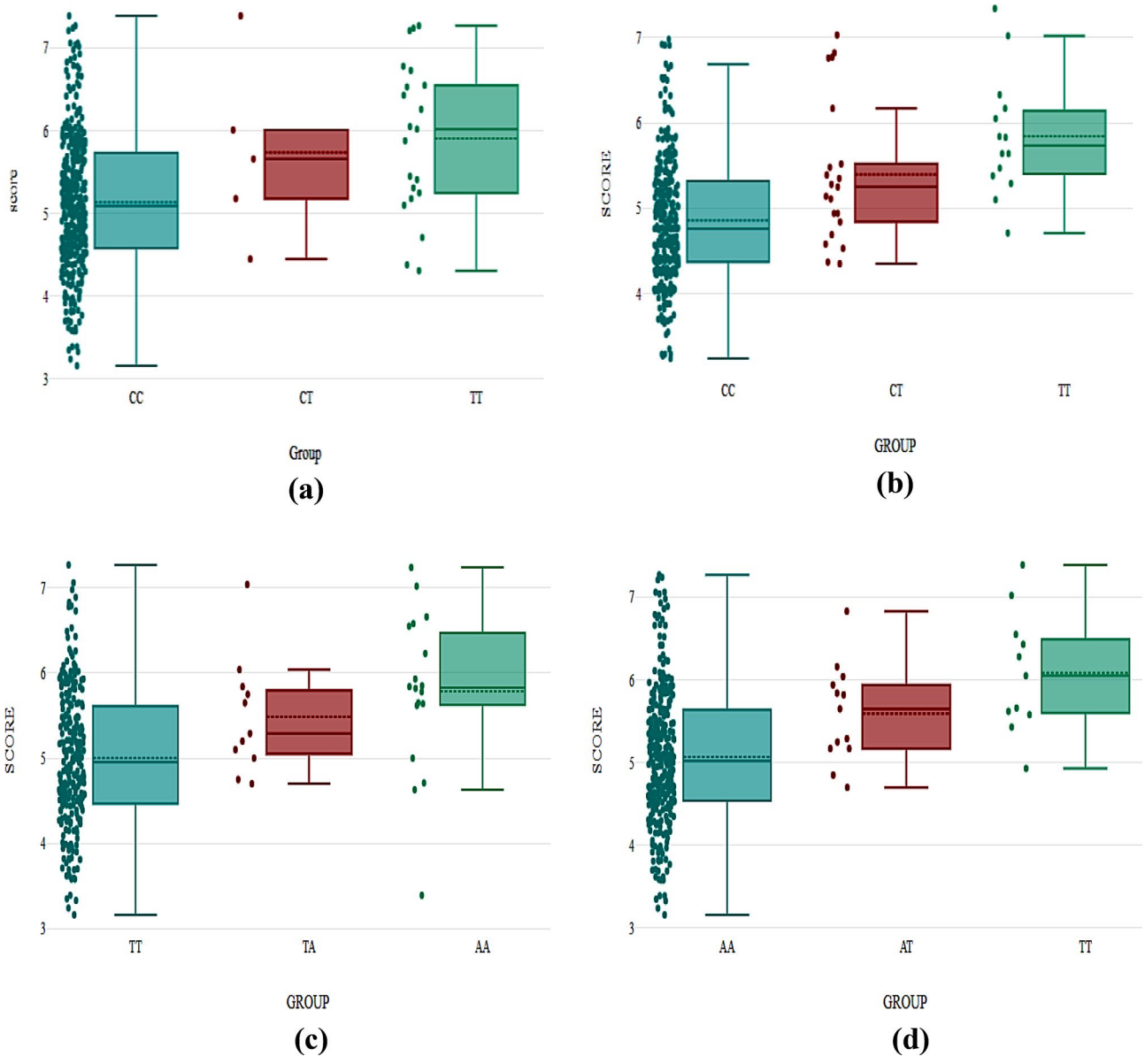

Fig. 5 Box-plots depicting allelic effect differences of SNPs/ MTAs based on Kruskal-wallis non-parametric test; (**a**) S8_ 155841067, (**b**) S6_34825812, (**c**) S8_152715134, and (**d**) S8_152460815)

Table 3 List of SNPs/MTAs with significant differences between allelic effects on the basis of Kruskal–Wallis test

SNP	Allele	H (chi2)	*P* value
S8_ 155,841,067	C/T	14.09	0.00087***
S8_152715134	T/A	10.66	0.004849**
S6_34825812	C/T	24.61	4.52E-06***
S8_152460815	T/A	19.36	6.24E-05***

*<0.05, **<0.01, ***<0.001

transcription factor 23 had the highest expression among all the genes, followed by GRMZM5G813007/amino-acid acetyltransferase *NAGS1*. The other genes exhibited expression only in some leaf zones. The identified candidate genes (22) were queried into the GeneMANIA web server for functional prediction. The physical interaction and coexpression between the genes in the network were 8.98% and 88.70%, respectively (Fig. 7a, b). The genes SPA1/ GRMZM2G061602 and ABI3/GRMZM2G313737 were identified on the same chromosome, Chr8 (8.06), in close proximity to the single SNP S8_155841067. Notably, these genes have been reported to physically interact and coexpress within gene networks. The gene network extended its association with other genes, e.g., between the genes AT3G56880.1 (candidate gene) and WRKY75 (from the network). Both of these genes act as positive jasmonate-mediated regulators of plant basal defence against necrotrophic fungal pathogens and were reported

Table 4 Common haplotypes identified at Ludhiana (Ep1) and Gurdaspur (Ep2) using haplotype trend regression analysis for MLB resistance in CAAM panel

Haplotype block	Chromosome	Markers Used	*P*-Value	*R* Squared (%)	Bonferroni *P*- Value	FDR	Selected Regressors	Environments©
Hap_1.1	1	S1_11475895, S1_11850802, S1_11991587	9.7×10^{-5}	7.266	7.2×10^{-3}	2.9×10^{-4}	GTT, ACT	Ep-2
			5.5×10^{-4}	4.745	4.1×10^{-2}	2.2×10^{-3}	GTT	Ep-1
Hap_1.2	1	S1_289686730, S1_289688084	1.4×10^{-4}	5.138	1.0×10^{-2}	3.7×10^{-4}	AT	Ep-2
			3.2×10^{-4}	4.610	2.4×10^{-2}	1.6×10^{-3}	TC	Ep-1
Hap_2.1	2	S2_1353042, S2_1553601	9.1×10^{-8}	9.150	6.8×10^{-6}	2.2×10^{-6}	GC	Ep-2
			1.3×10^{-4}	4.767	1.0×10^{-2}	1.4×10^{-3}	GC	Ep-1
Hap_2.2	2	S2_12027414, S2_12241580	3.6×10^{-5}	6.560	2.7×10^{-3}	1.3×10^{-4}	GC	Ep-2
			1.5×10^{-5}	7.167	1.1×10^{-3}	3.8×10^{-4}	GC	Ep-1
Hap_2.3	2	S2_205904685, S2_205905808	6.5×10^{-6}	6.161	4.9×10^{-4}	4.9×10^{-5}	GA	Ep-2
			2.8×10^{-4}	4.038	2.1×10^{-3}	1.6×10^{-3}	GA	Ep-1
Hap_3	3	S3_56134621, S3_56812535	1.6×10^{-5}	7.322	1.2×10^{-3}	7.9×10^{-5}	TA	Ep-2
			1.9×10^{-4}	5.553	1.4×10^{-2}	1.7×10^{-3}	TA	Ep-1
Hap_4	4	S4_149745183, S4_149899657	7.2×10^{-5}	5.052	5.4×10^{-3}	2.6×10^{-4}	CC	Ep-2
			3.2×10^{-4}	4.175	2.4×10^{-2}	1.7×10^{-3}	CC	Ep-1
Hap_8.1	8	S8_3292504, S8_3427103, S8_4258284	5.2×10^{-6}	8.895	3.9×10^{-4}	4.9×10^{-5}	TGG, GCA	Ep-2
			2.1×10^{-4}	5.114	1.5×10^{-2}	1.5×10^{-3}	GCA	Ep-1
Hap_8.2	8	S8_8731001, S8_8731102	2.7×10^{-8}	8.938	2.0×10^{-6}	1.0×10^{-6}	CG	Ep-2
			6.5×10^{-4}	3.459	4.9×10^{-2}	2.3×10^{-3}	CG	Ep-1
Hap_8.3	8	S8_151346456, S8_152460815, S8_152715134	7.4×10^{-6}	9.353	5.5×10^{-4}	5.0×10^{-5}	GAT	Ep-2
			5.9×10^{-4}	5.597	4.4×10^{-2}	2.3×10^{-3}	GAT	Ep-1
Hap_9	9	S9_73520507, S9_74919639, S9_89451669, S9_89598809, S9_99083878	3.7×10^{-4}	6.279	2.7×10^{-2}	8.9×10^{-4}	ACGGC	Ep-2
			$2.7\times10-4$	6.561	2.0×10^{-2}	1.8×10^{-3}	ACGGC	Ep-1
Hap_10.1	10	S10_12104511, S10_14001955, S10_14522387	5.8×10^{-4}	3.783	4.3×10^{-2}	1.2×10^{-3}	CAT	Ep-2
			4.8×10^{-9}	10.570	3.6×10^{-7}	3.6×10^{-7}	CAT	Ep-1
Hap_10.2	10	S10_140737865, S10_141006825	2.8×10^{-5}	5.655	2.0×10^{-3}	1.2×10^{-4}	AG	Ep-2
			2.7×10^{-4}	4.292	2.0×10^{-2}	1.7×10^{-3}	AG	Ep-1

©Ep1 represents pooled dataset of two years at E1(Ludhiana), Ep2 represents pooled dataset of two years at E2 (Gurdaspur)

to be coexpressed in the network. While the remaining identified candidate genes did not display direct relationships with each other, they did exhibit associations with genes belonging to the same family. For instance, UBQ10/GRMZM2G164787 with UBQ13 and BZIP23/GRMZM2G033413 with BZIP53 were related to the shared protein domain family.

Discussion

Maize leaf blight, caused by the necrotrophic pathogen *Bipolaris maydis*, presents a significant global threat to maize cultivation. This polycyclic disease becomes epidemic under favourable conditions, and phenotyping under artificial epiphytotic conditions with high disease pressure proves to be cost-intensive. Therefore, a profound understanding of host plant resistance (HPR) is imperative to identify molecular markers for MLB resistance and enhance the efficiency of developing resistant tropical and subtropical maize germplasm [52]. The present study leveraged high heritability (0.79) for the MLB disease score, based on pooled data from the CAAM panel, in four environments, suggesting the possibility of accurate phenotypic selection to breed for MLB resistance in maize [53, 54]. The CAAM panel exhibited lower genetic relatedness, rapid linkage disequilibrium (LD) decay, and a moderate population structure. Moderate population structure in CIMMYT Asia tropical and sub-tropical lines was reported in previous studies [55]. George et al. [56] corroborated this observation and reported substantial diversity in tropical and subtropical lines in the Asian region, rendering it challenging to establish clear-cut distinctions into well-defined clusters. Warburton et al. [57] suggested that this could be due to the fact that the populations from where Asian lines were derived had a heterogeneous nature with larger diversity within, than between source populations. It is known that LD decays more rapidly in tropical maize germplasm (1 kb) than in temperate germplasm (10 kb), but faster LD decay rates have been reported in some tropical diversity panels [58].

Identification of 26 SNPs significantly associated with MLB with low to moderate effect sizes across all 10

Table 5 Gene annotation of MLB resistance associated SNPs with their nearest region (colocalized gene)

SNP	Chr. Bin location	Putative candidate gene	©Environments	Gene/protein name	Function	Crops	Relative expression in leaf and vegetative meristem (FPKM)	References
]S8_155841067	8.06	GRMZM2G313737	Across, Ep1, Ep2	B3 domain containing protein	Activation of basal defence response by Mitogen activated Protein kinase (MAPK) mediated phosphorylation of RAV family protein (B3 domain family) against *Botrytis cinerea* and *R. solani* .	*Arabidopsis thaliana* (*AtRAV1* gene), Tomato (*SlRAV1* gene), wheat	Leaf zone1(10.81) Leaf zone 2 (7.26) Leaf zone 3 (15.58) Mature leaf (1.27) Veg. meristem (11.44)	[76]
	8.06	GRMZM2G061602		Protein SUPPRESSOR OF PHYA-105 1/SPA 1	Serine /threonine protein kinase activity, Involved in regulating circadian rhythms. Circadian clock predicts the natural infection time of pathogens and regulates MAMP-triggered basal immunity	*Arabidopsis thaliana* (RPP4-mediated cell death against fungus *Hyaloperonospoa arabidopsidis*),	Leaf zone1(3.53) Leaf zone 2 (2.33) Leaf zone 3 (1.44) Mature leaf (8.5) Veg. meristem (5.21)	[74-89]
	8.06	GRMZM2G313833		ATP-dependent helicase rhp16q	DEAD-box RNA helicase functions in the regulation of defence responses against *Magnaporthe grisea*, responsible for programmed cell death and RNA splicing	Rice (*OsBIRH1*), tomato (*SlDEAD35*),*Arabidopsis thaliana*	Leaf zone1(3.53) Leaf zone 2 (2.33) Leaf zone 3 (1.44) Mature leaf (8.5) Veg. meristem (5.21)	[90, 91]
S3_156792785	3.05	GRMZM2G031584	Ep1, Ep2	DNA J heat shock N-terminal domain containing protein	HSP40 or J-proteins, which recognizes an unfolded substrate and delivers it to DnaK, accelerates ATP hydrolysis of HSP70s, and induces conformational changes in stable chaperone proteins	Rice, Arabidopsis	Leaf zone1(4.29) Leaf zone 2 (3.81) Leaf zone 3 (1.44) Mature leaf (8.5) Veg. meristem (5.21)	[92, 93]
	3.05	GRMZM2G031528		Heavy metal detox superfamily protein	Cross talk between heavy metals and defence responses against Fungi, prompt resistance through ROS signaling	Arabidopsis (*HIPP27*), Rice, Pepper	Leaf zone1(15.44) Leaf zone 2 (9.28) Leaf zone 3 (2.98) Mature leaf (0) Veg. meristem (30.02)	[94]
S6_34825812	6.01	GRMZM2G013581	Ep2	MYB DNA-Binding domain superfamily	MYB TFs can get activated in response to signaling molecule jasmonic acid leading to overexpression of resistance genes and defence response is regulated through H^2o^2 accumulation	Maize (*RS2* gene) Wheat (*TaMYB29* gene) Arabidopsis (AtMYB108)	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (1.19) Mature leaf (0) Veg. meristem (6.92)	[95, 96]

Table 5 (continued)

SNP	Chr. Bin location	Putative candidate gene	©Environments	Gene/protein name	Function	Crops	Relative expression in leaf and vegetative meristem (FPKM)	References
S1_232344813	1.08	GRMZM2G164787	Ep1, Ep2	Ubiquitin protein ligase binding	Regulation of defence proteins NBS-LRR by ubiquitination (E3 ubiquitin ligases) leading to effector triggered immunity (ETI)	Arabidopsis, Rice	Leaf zone1(5.19) Leaf zone 2 (3.19) Leaf zone 3 (1.86) Mature leaf (1.94) Veg. meristem (9.08)	[97]
S8_152715134	8.06	GRMZM2G014089	Across	ABC transporter B family member 4	Transport secondary metabolites: defence molecules	Arabidopsis (*AtPDR12* gene)	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (0) Mature leaf (3.32) Veg. meristem (0)	[98]
S8_162518701	8.06	GRMZM2G031352	Across, Ep2	AP-4 complex subunit epsilon	Plant immunity associated with hypersensitive cell death	Arabidopsis	Leaf zone1(15.89) Leaf zone 2 (18.94) Leaf zone 3 (22.49) Mature leaf (7.77) Veg. meristem (13.77)	[99, 73]
S5_140936401	5.04	GRMZM2G064638	Across, Ep2	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein/ BHLH 139	These proteins are regulated by Jasmonic acid to prompt defence response	Arabidopsis, tomato	Leaf zone1(4.66) Leaf zone 2 (3.81) Leaf zone 3 (0) Mature leaf (0) Veg. meristem (1.47)	[100]
S4_166482019	4.06 (novel)	GRMZM2G444623	Across	Aspartic proteinase nepenthesin-2	Role as fungal-resistance genes, reduces the activity of fungal phytases (enzymes that dephosphorylate the phytic acid substrate)	barley *(HvNEP-1)*	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (3.12) Mature leaf (1.88) Veg. meristem (0)	[86]
S8_148676841	8.06	GRMZM2G130586	Across, Ep2	Uncharacterized	-			
S2_223252193	2.06	GRMZM2G160515	Across	Uncharacterized	-			
S6_21316804	6.01	GRMZM2G108501	Across	Indole – 3-pyruvate monooxygenase YUCCA1	Induces auxin biosynthesis, Inhibits virus replication. Induces SAR	Arabidopsis, Rice	-	[101]
S8_8887701	8.01	GRMZM2G079219	Across	PR5-like receptor kinase/ rust resistance kinase Lr10	The predicted ligand-binding domain of PR5like receptor kinase is related to a family of plant defense proteins, suggesting a possible role for the receptor in the perception of microbial signals.	*Arabidopsis thaliana*,tomato (*Cf9* gene), Rice	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (0) Mature leaf (3.24) Veg. meristem (0)	[102, 103]
S4_11836688	4.03	GRMZM2G495768	Across	Uncharacterized	-			

Table 5 (continued)

SNP	Chr. Bin location	Putative candidate gene	©Environments	Gene/protein name	Function	Crops	Relative expression in leaf and vegetative meristem (FPKM)	References
S7_161657633	7.04(novel)	GRMZM5G813007	Ep1	Putative amino-acid acetyltransferase NAGS1	Synthesis of L Arginine. Arginine metaboloism plays role in pathogen defence	Arabidopsis, Tomato	Leaf zone1(17.79) Leaf zone 2 (15.31) Leaf zone 3 (13.46) Mature leaf (11.58) Veg. meristem (14.55)	[104]
S9_8243435	9.01(novel)	GRMZM2G033413	Ep1	bZIP transcription factor 23	Induction of pathogenesis related (PR) gene expression, HR and cell death	Arabidopsis (*AtbZIP10*), cotton, Rice, soyabean	Leaf zone1(25.98) Leaf zone 2 (16.4) Leaf zone 3 (11.8) Mature leaf (10.95) Veg. meristem (34.95)	[105]
S2 157,608,147	2.06	GRMZM2G009626	Ep1	Sterol 3-beta-glucosyltransferase UGT80A2	Plays a key role in plant innate immunity by regulating nutrient efflux into the apoplast	Arabidopsis, cotton, Oat	Leaf zone1(4.78) Leaf zone 2 (3.73) Leaf zone 3 (2.40) Mature leaf (7.76) Veg. meristem (2.29)	[106]
S6_130006038	6.01	GRMZM5G864847	Across, Ep1	Auxin responsive Aux/IAA family member	Auxin repressors repress auxin signalling in plants which act as a defence strategy during infection. Crosstalk between JA and Auxin prompts defence against nectrotrops	Arabidopsis	Leaf zone1(1.96) Leaf zone 2 (1.86) Leaf zone 3 (0) Mature leaf (0) Veg. meristem (3.91)	[107]
S8_152460815	8.06	GRMZM2G065538	Ep2	AP2/ERF and B3 domain-containing protein	Regulates the expression of defense-related genes and the activities of antioxidant enzymes	Wheat, Rice, Tomato, soyabean	Leaf zone1(0) Leaf zone 2 (1.33) Leaf zone 3 (0) Mature leaf () Veg. meristem (2.78)	108
S1_200269986	1.07	GRMZM5G814389	Ep2	Lipolytic acyl hydrolase (LAH) / patatin protein class	Act as defence against plant parasites mediated by JA signalling	Arabidopsis	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (4.75) Mature leaf (0) Veg. meristem (0)	[109]
S9_141454813	9.06 (novel)	GRMZM2G043892	Ep2	CLAVATA3 embryo surrounding region-related-16	Diverse roles in plant growth and development; involved with cell signalling during pathogen attack	Potato, Arabidopsis (*AtCLV3*), soyabean, rice	-	[110]
S6_141510514	6.05 (novel)	GRMZM2G357834	Ep2	Wound-induced protein, Wun1	Induced by Wounding, methyl jasmonate, and pathogen infection to reinforce cell wall composition after wounding	Potato, Tobacco	Leaf zone1(0) Leaf zone 2 (1.43) Leaf zone 3 (2.92) Mature leaf (12.61) Veg. meristem (0)	[82]

Table 5 (continued)

SNP	Chr. Bin location	Putative candidate gene	©Environments	Gene/protein name	Function	Crops	Relative expression in leaf and vegetative meristem (FPKM)	References
S5_3412526	5.01(novel)	AC210013.4_FG014	Ep2	Calcium-dependent protein kinase 7	Plant innate immunity, including regulation of oxidative burst	Tobbaco, tomato, Rice	Leaf zone1(4.14) Leaf zone 2 (6.87) Leaf zone 3 (7.65) Mature leaf (37.28) Veg. meristem (4.71)	[84]
S5_13798307	5.02	GRMZM2G033489	Ep2	Uncharacterized	-			
S1_52252512	1.04(novel)	GRMZM2G128644	Ep2	vq4 - VQ motif-transcription factor4	SA and/or JA-mediated defense responses against pathogens.	Rice, Arabidopsis, soyabean	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (0) Mature leaf (1.35) Veg. meristem (0)	[111]
S1_233546091	1.08	GRMZM2G001845	Ep2	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1	ABA signalling protein serine/threonine kinase activity	Arabidopsis,	-	112
S10_83669175	10.3	GRMZM2G060054	Ep2	Disease resistance protein PIK6-NP	NBS-LRR (defence related protein), Pik allele characterized for resistance against Rice Blast. Triggering hypersensitivity response	Rice	Leaf zone1(0) Leaf zone 2 (0) Leaf zone 3 (2.46) Mature leaf (0) Veg. meristem (0)	[113]

©Ep1 represents pooled dataset of two years at E1(Ludhiana), Ep2 represents pooled dataset of two years at E2 (Gurdaspur)

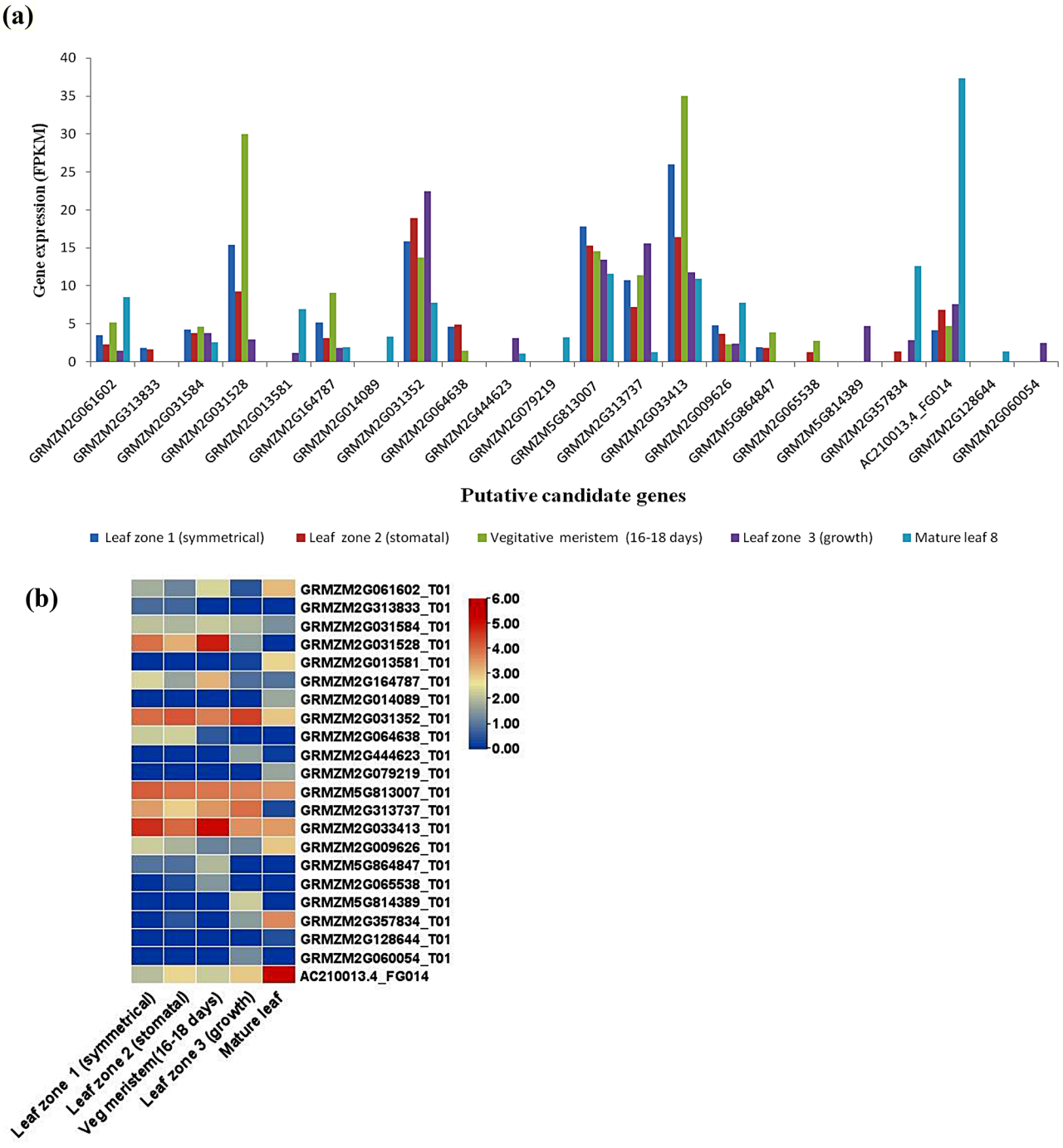

Fig. 6 (**a**) Graphical representation of expression data as Fragments per kilobase of transcript per million map reads of 22 candidate genes, (**b**) Heatmap exhibiting the expression patterns (Log_2 transformed FPKM values) of candidate genes in leaf tissue at four different stages and in the vegetative meristem; blue represents lower values, and red represents higher values

chromosomes suggested that resistance to MLB is governed by multiple quantitative trait nucleotides (QTNs) with small effects. Our observation of the quantitative nature of MLB has been reported in earlier studies also [14–15]]. Chromosome-specific analysis reveals crucial genomic regions that are important for disease resistance in general, and resistance to MLB in particular. Chromosomal bin 8.06, found in our study, comprised five SNPs. The physical coordinates of these identified SNPs colocalized with the QTL *qMSR8* (151.45 to 166.98 Mb), which was identified from the same AM panel and validated for charcoal rot (caused by a necrotrophic pathogen) [54]. This bin also harbours QTLs for other important diseases; GLS, NCLB, common rust,

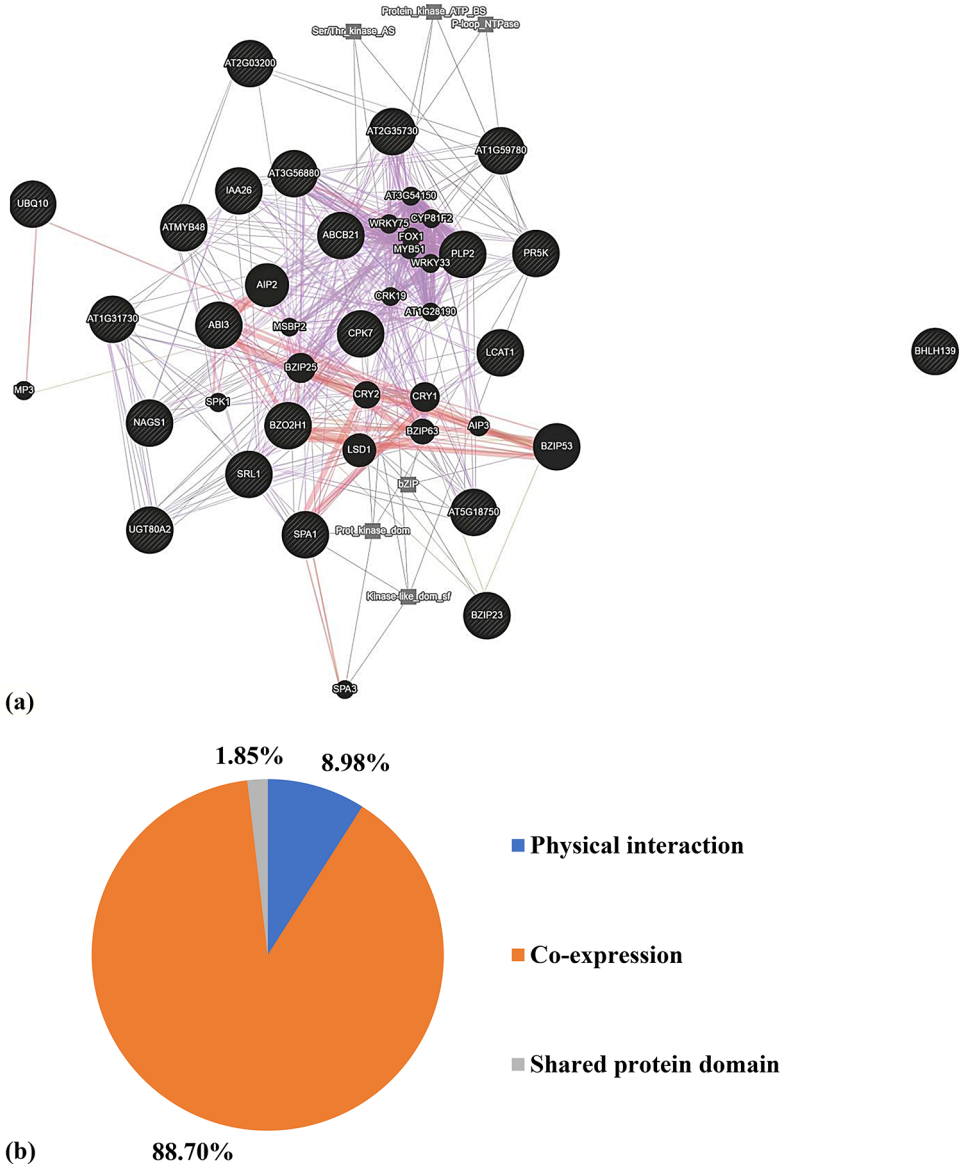

Fig. 7 (**a**) Illustration of the physical interaction and co-expression network of candidate genes and allied genes depicted by GeneMANIA. Pink lines signify physical interactions among the candidate genes and other genes from the same family, while light purple lines represent the co-expression of these genes. (**b**) Percentage of each category of interactions among the genes

and smut [59, 60]. Furthermore, comprehensive meta-QTL analysis revealed the presence of a cluster of QTLs on chromosome 8, accompanied by significant consensus QTLs associated with MLB, NCLB, and GLS, all located within a narrow confidence interval [61]. Based on these well-aligned reports, we suggest further studies on the significant associations we have detected in this chromosomal bin for validation and deployment efforts to combat MLB effectively. Furthermore, Bin 8.01 (comprising SNP S8_8887701) corresponds to a previous study reporting a QTL (*qAUDPC8.1*) in the Indian germplasm for MLB, emphasizing relevance of bin 8.01 in MLB resistance [62]. Chromosome 3 has been reported to harbour the maximum number of QTLs (at 25 loci) for resistance to MLB specifically in bin 3.04 [15–17, 19],] possessing major QTLs which are validated in different genetic backgrounds. A meta-QTL study highlighted the significance of genomic regions within bins 3.04–3.08 for MLB resistance 61]]. We identified one SNP (S3_156792785) in bin 3.05 in our GWAS study. This bin is also recognized for harbouring stable genomic regions linked to other diseases caused by necrotrophs, e.g., fusarium ear rot [63].

We identified two SNPs within chromosomal bin 6.01 (S6_21316804 and S6_34825812). Bin 6.01 is recognized as a hotspot for resistance against various viral diseases,

and MLB [64]. A minor QTL associated with MLB resistance was reported in this bin [41]. Moreover, the *rhm*1 gene, known to confer complete MLB resistance against race 'O', is situated within or in close proximity to bin 6.01 [17]. Our findings suggest the possibility of a broader spectrum of allelic variation against MLB within bin 6.01. Three SNPs were reported on chromosome 5, among which one SNP (S5_140936401) was reported in bin 5.04. Previous studies [22, 65],] have identified significant SNPs associated with MLB resistance within bin 5.04. Additionally, bin 5.04 has been reported to host resistance against NCLB, GLS, and MLB, as documented by Martins et al. [66]. A single SNP was identified on chromosome 10 (10.3), which remarkably correspond with earlier research findings where a disease QTL (dQTL) for MLB was reported. This finding emanated from two B73-resistant NILs (NC292 and NC330) against MLB, which were developed by repeated backcrossing with elite source of MLB resistance (NC250P), further reinforcing the significance of this genomic region in conferring resistance [66]. Two SNPs identified to be significantly associated with MLB resistance in this study (S1_233546091 and S1_232344813) on Chromosome 1 (bin 1.08), colocalized with earlier reported SNPs against NCLB [36] and Fusarium stalk rot [67] in the same panel. Moreover, in previous studies, no SNPs or QTLs colocalized with the physical coordinates of SNPs reported in the seven novel chromosomal bins associated with MLB (Table 6).

The majority of the haplotype blocks identified from our haplotype detection analysis were formed with two or three SNPs. The size of these haplotypes is intricately linked to the level of linkage disequilibrium (LD) within the population under study, as elucidated by Slatkin (2008) [68]. A rapid decay of LD results in the formation of smaller haplotype [69, 70]. A single haplotype block (Hap_8.3) exhibiting a significant effect was identified within chromosomal bin 8.06, which corroborated with our GWAS findings. This haplotype was found to account for approximately 9.3% (Ep2) and 5.3% (Ep1) of the variation observed for the trait respectively in our study. Two haplotype blocks, reported in two novel bins, 2.00 (Hap_2.1) & 1.11 (Hap_1.2) explained PVE of 9.1%, 5.7% for Ep2, and 5.1%, 4.6% for Ep1, respectively. The use of

Table 6 Summary of chromosomal location of Maydis leaf blight (MLB) resistance quantitative trait loci (QTL)/SNPs in maize from previous studies

S. No.	Bins[a] /Chromosomes Reporting MLB QTL/SNPs	Markers used	Populations	References
1	1.02, 1.07, 1.09, 2.04, 3.03, 3.04, 4.09, 5.01, 7.02, 9.05 and 10.04	SSR	RILs (Mo 17 × B73)	[13]
2	1.06, 1.08, 1.09, 2.09, 3.04, 3.06, 6.00, 7.02 and 8.03	SSR	RILs (B73 × Mo17)	[18]
3	1.08–1.09, 2.06–2.07, 3.04, 3.07, 3.09, 6.06 and 9.03–9.04	SSR	RILs (NC300 × B104)	[14]
4	1.03, 1.05, 1.06, 1.10, 2.04, 3.04, 3.06, 4.02–4.03, 6.02, 7.03 and 8.02–8.03	SSR	Advanced intercross RILs (B73 × Mo17)	[15]
5	2.07, 3.04, 6.01 and 8.05	SSR	RILs (H99 × B73) and (B73 × B52)	[16]
6	1.09, 2.05–2.06, 3.03, 5.05–5.06, 6.01, 9.02 and 10.03	SSR and SNPs	NILs (NC292 × B73) and (NC330 × B73)	[66]
7	1.05–1.06, 1.08–1.09, 2.04, 2.09, 3.04–3.05, 8.05 and 10.05	SSR and SNPs	RILs (KI14 × B73)	[113]
8	1.10, 2.03, 3.03, 3.04, 8.06, 8.05, 9.03, 9.04, 9.05 and 10.04	SSR	RILs (B73 × CML254), (CML254 × B97) and (B97 × Ki14)	[17]
9	3, 4, 6, 8, 9 and 10	SSR	$F_{2:3}$ (T14 × T4)	[114]
10	6 (6.01) Gene *rhm*	RFLP (UMC85 and p144)	F3(RH95*rhm* × B73)	[115]
11	1.03/1.04, 1.07, 1.09, 1.06, 1.05, 2.02/2.03, 2.04, 2.05, 3.03, 3.04, 3.05, 3.06, 3.09, 4.00/4.01, 4.05, 4.09, 5.03, 5.04, 5.06, 5.07, 6.01, 6.06, 7.00, 7.01, 7.03, 8.03, 8.06/8.07, 9.02, 9.03/9.04, 9.04, 9.07, 10.03 and 10.07	SNPs	NAM	[22]
12	1.09, 2.05–2.06, 3.03, 6.01 and 9.02	SSR	NILs (NC292 × B73) and (NC330 × B73)	[116]
13	2.04, 3.04, 3.05 and 8.05	SSR	Teosinte NILs from 10 populations	[117]
14	1, 2, 3, 4, 5, 6, 7, 8, 9 and 10	SNP	NAM	[65]
15	3.08, 9.03, 8.01, 8.03, and 8.01	SSR	RIL (LM5 × CM140)	[62]
16	1,2,3,4,5,6,7,8,9,10	SNP	ROAM, 8 RIL populations, and 513 diversity maize inbred lines	[25]
17	1,2,3,4,5,6,7,8,9,10	SNP	NAM	[24]
18	**9.06**, **5.01**, **9.01**, **7.04**, **4.06**, **1.04**, **and 6.05**	**SNP**	**CAAM**	

[a]Chromosome bin location of QTL peak on 1 of the 10 chromosomes of the maize genome. Bins divide the genetic map into 100 approximately equal segments of approximately 20 centiMorgans between two fixed Core Marker. The segments are designated with the chromosome number followed by a two-digit decimal (e.g., 1.00, 1.01, 1.02, etc.)

RILs: recombinant inbred lines; NILs: near isogenic lines; NAM: nested association mapping, ROAM: Random-Open-parent Association Mapping, CAAM: CIMMYT Asia Association Mapping

haplotypes increases the phenotypic variance explained, and can be beneficial when identifying marker phenotype associations for the genetic dissection of loci underlying the complex trait [71]. Additionally, SNPs/ haplotypes reported in previously unreported genomic regions/bins (9.06, 5.01, 9.01, 7.04, 4.06, 6.05, 2.00, and 1.11) could be unique to the CAAM panel and the environment studied, and could be candidates for enriched allelic diversity associated with MLB resistance.

Twenty-six SNPs associated with MLB resistance in this study were associated with annotated genes with functional domains that were previously reported to influence disease resistance in various crops (Table 5). Genes in Chromosome 8 play pivotal roles in various defence pathways, viz., and activation of basal defence by mitogen-activated protein kinases, serine/threonine protein kinase activity, circadian rhythm-generated basal immunity, hypersensitive cell death response, and transport of secondary metabolites required against necrotrophs, e.g., phytoalexins, especially camalexin (3-thiazol-2-yl-indole), a secondary metabolite toxic to *B. maydis* [73-76]. Based on the predicted co-expression results, co-expression of these genes was detected in the network, and it could be possible that these genes (bin 8.06) may form a cluster, initiating a cascade of reactions against MLB, which warrants further investigation [76]. Furthermore, the expression of these genes exceeded 10-FPKM in all leaf zones. Physical interaction of genes SPA1/GRMZM2G061602 and ABI3/GRMZM2G313737 underscore their role in basal defence response via MAMP responsive MAPK mechanisms [77, 78]. The SPA1 gene further advances the notion of circadian rhythm-generated basal immunity against MLB, which reveals the potential for further studies on such genes in the future [75]. The candidate gene GRMZM2G013581/ MYB DNA binding domain (bin 6.01) was identified in our study. Chen et al. [25] functionally validated gene MYBR92 (encoding a MYB-like transcription factor) against MLB.

Specific genes associated with novel SNPs identified are functionally recognized for their expression in response to cross-talk between jasmonic acid and ethylene, which enhances sensitivity to necrotrophic pathogens [79]. For example, the transcription factor (TF) BZIP 23/GRMZM2G033413 (S9_8243435) is known to modulate the response to various stresses, including abiotic factors and hormone transduction [80]. Another gene, GRMZM2G357834/WUN1, is involved in plant-defence responses regulated by JA and its methyl ester, methyl jasmonate (MeJA) against necrotrophs [81]. GRMZM5G813007/NAGS1 (S7_161657633) is involved in the L-arginine biosynthesis pathway [82]. Arginine serves as a precursor for the synthesis of nitric oxide (NO) and polyamines (PAs), both of which are known to promote defence mechanisms. AC210013.4_FG014/ CDPK7 (S5_3412526) gene has been found to respond to various stimuli, including abscisic acid (ABA), cold, drought, salinity, heat, elicitors, and pathogens [840]. Our study highlighted the possible role of ubiquitination required in facilitating the function of NBS-LRR proteins (promoting effector-triggered immunity), specifically by the GRMZM2G164787/Ubiquitin protein (ligase binding) gene associated with SNP S1_232344813) [84]. Genes E3 Ubiquitin protein (ligase) and CDPK7 have been reported in a previous study [25] as associated with resistance to MLB. Another important gene GRMZM2G444623/aspartic proteinase nepenthesin-2 (S4_166482019) was identified which was reported to reduce the activity of fungal phytases. In Barley, a related gene nepenthesin-1 (HVNEP-1) was discovered that reduced the production of mycotoxin 15-acetyldeoxynivalenol (15-ADON) from *Fusarium graminearum* [85]. It would be worthwhile to investigate the role of NEP2 in MLB resistance in maize. The reported SNPs in genes associated with the JA/ET signalling pathway and other defence mechanisms add depth to our understanding of MLB resistance, to carry forward with independent validation of the candidate genes.

Moreover, contrasting genotypes identified in this study could be used to develop mapping populations for further genetic dissection of the trait, The construction of breeder-friendly Kompetitive allele-specific PCR (KASP) markers for the significant and stable MTAs/ single SNPs identified may facilitate the deployment of these genomic regions through marker-assisted selection in the maize breeding process. In addition, the significant MTAs identified during the current study can be integrated into genomic prediction models to evaluate their potential for selection for MLB. Desirable haplotypes can be used for haplotype-based breeding in maize for MLB resistance through resequencing approach as, the molecular markers that define these favorable haplotypes can be developed and used to select the most desirable combination of haplotypes governing the specific phenotype. Moreover, inbred lines with novel recombination in chromosomal blocks of interest can be selected by haplotype-related markers [86]. The identified important genes may also be validated using functional genomics techniques. However, the potential challenge one can face is impact on the marker/SNP effect which can differ with populations and environments. This challenge arises due to differencs in LD between SNP and QTL in different populations, effect of G x E interaction, and spurious associations [87] Overall, this comprehensive genomic analysis provides valuable insights for targeted breeding strategies to enhance MLB resistance in maize.

Conclusion

In summary, GWAS and haplotype trait regression studies on resistance to MLB in Asia-adapted CAAM panel identified 26 SNPs and 13.haplotypes associated with the trait. The study confirmed the quantitative nature of the resistance with identified variants exhibiting low to moderate effect sizes. But gene annotation and network analysis of the identified variants points to some important genes that are implicated in diverse defence pathways in particular, and stress tolerance in general. Several of the identified variants were located in previously reported chromosomal bins, and some new genomic regions were also identified in this study. This not only enhances our appreciation of allelic diversity but also deepens our understanding of the intricate mechanisms behind resistance to MLB in maize. Additionally, the identification of a number of SNPs and haplotype within chromosomal bin 8.06, which is known to harbour dQTLs/dQTNs for resistance to multiple diseases, underscores its potential to be further investigated for validation and possible deployment of trait markers for resistance to MLB.

Abbreviations

MLB	Maydis leaf blight
CAAM	CIMMYT Asia association mapping panel
GWAS	Genome wide association studies
QTL	Quantitative trait loc

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10655-x.

Supplementary Material 1

Acknowledgements

The authors would like to thank CIMMYT for providing the germplasm and access to GBS data and Director Regional Research Station Gurdaspur, PAU, Ludhiana for providing access to the experimental station. The authors gratefully acknowledge the support received from the One CGIAR Accelerated Breeding Initiative (ABI).

Author contributions

"S.S and S.K.N conceived and supervised the conductance of the experiment; W.U.N conducted the field evaluation and phenotyping; W.U.N carried out the GWAS and in silico analysis, interpreted the results and wrote the manuscript; S.K.N provided the germplasm and GBS data; H.K provided the culture for inoculation; A.K maintained the trial at the Gurdaspur location; Z.K analyzed the phenotypic data and conducted haplotype analysis; and G.S helped in GWAS analysis. W.U.N, S.S, S.K.N, and Y.V finalized the manuscript. All the authors reviewed the manuscript".

Funding

The authors gratefully acknowledge the funding received from the CGIAR Research Program (CRP) on MAIZE for generating the genotypic data used in this study.

Data availability

The experimental data that support the findings of this study have been deposited in the BioStudies database with the primary accession code S-BSST1435.

Declarations

Ethics approval and consent to participate

In compliance with the IUCN Policy Statement on Research, in the present study, the material used was maize (*Zea mays* L.), a cultivated species maintained through conventional breeding. Furthermore, none of the materials utilized in the present research are at risk of extinction. All the experiments were carried out in accordance with relevant guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Dept. of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, India
[2]International Maize and Wheat Improvement Centre (CIMMYT), Hyderabad, India
[3]Regional Research Station, Punjab Agricultural University, Gurdaspur, Ludhiana, India
[4]School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India

Received: 17 May 2024 / Accepted: 23 July 2024
Published online: 05 August 2024

References

1. Shiferaw B, Prasanna B, Hellin J, Banziger M. Crops that feed the world. Past successes and future challenges to the role played by maize in global food security. Food Secur. 2011;3:307–27.
2. Coclanis PA. The golden fuel. 2023 https://aeon.co/essays/what-explains-the-unstoppable-rise-of-maize-in-asia
3. Malik VK, Gogoi R, Hooda KS, Singh M. Identification of multiple disease resistant maize accessions. Indian Phytopathol. 2017;70(1):80–5.
4. Ullstrup AJ. The impacts of the southern corn leaf blight epidemics of 1970–1971. Annu Rev Phytopathol. 1972;10(1):37–50.
5. Singh V, Lakshman DK, Roberts DP, et al. Fungal species causing maize leaf blight in different agro-ecologies in India. Pathogens. 2021;10(12):1621. https://doi.org/10.3390/pathogens10121621.
6. Byrnes KJ, Pataky JK, White DG. Relationships between yield of three maize hybrids and severity of southern leaf blight caused by race O of Bipolaris Maydis. Plant Dis. 1989;73(10):834–40.
7. Hooker AL. Inheritance of chlorotic-lesion resistance to Helminthosporium turcicum in seedling corn. Phytopathology. 1963;53:660–2.
8. Hammond-Kosack KE, Rudd JJ. Plant resistance signalling hijacked by a necrotrophic fungal pathogen. Plant Signal Behav. 2008;3(11):993–5.
9. Atwell S, Huang YS, Vilhjálmsson BJ, et al. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature. 2010;465(7298):627–31.
10. Bernardo R. Breeding for quantitative traits in plants. 3rd ed. Woodbury, Minnesota: Stemma; 2020. Pp. 114–5.
11. Mitchell-Olds T. Complex-trait analysis in plants. Genome Biol. 2010;11:113.
12. Huang X, Han B. Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol. 2014;65:531–51.
13. Rafalski JA. Association genetics in crop improvement. Curr Opin Plant Biol. 2010;13(2):174–80.
14. Carson ML, Stuber CW, Senior ML. Identification and mapping of quantitative trait loci conditioning resistance to southern leaf blight of maize caused by Cochliobolus Heterostrophus race O. Phytopathology. 2004;94(8):862–7.
15. Balint-Kurti PJ, Carson ML. Analysis of quantitative trait loci for resistance to southern leaf blight in juvenile maize. Phytopathology. 2006;96(3):221–5.
16. Balint-Kurti PJ, Zwonitzer JC, Wisser RJ, Carson ML, Oropeza, Rosas MA, Holland JB, Szalma SJ. Precise mapping of quantitative trait loci for resistance to southern leaf blight, caused by *Cochliobolus heterostrophus* race O, and flowering time using advanced intercross maize lines. Genetics. 2007;176:645 – 57.
17. Balint-Kurti PJ, Zwonitzer JC, Pè ME, Pea G, Lee M, Cardinal AJ. Identification of quantitative trait loci for resistance to southern leaf blight and days to

anthesis in two maize recombinant inbred line populations. Phytopathology. 2008;98(3):315–20.
18. Negeri AT, Coles ND, Holland JB, Balint-Kurti PJ. Mapping QTL controlling southern leaf blight resistance by joint analysis of three related recombinant inbred line populations. Crop Sci. 2011;51:1571–9.
19. Balint-Kurti PJ, Carson ML. Analysis of quantitative trait loci for resistance to southern leaf blight in juvenile maize. Phytopathology. 2006;96:221–5.
20. Zhao Y, Lu X, Liu C, Guan H, Zhang M, Li Z, Cai H, Lai J. Identification and fine mapping of rhm1 locus for resistance to Southern corn leaf blight in maize. J Integr Plant Biol. 2012;54:321–29.
21. Zhang J, Jia X, Wang GF, Ma S, Wang S, Yang Q, Wu. L.Ascorbate peroxidase 1 confers resistance to southern corn leaf blight in maize. J Integr Plant Biol. 2022;64:1196–11.
22. Yang Q, He Y, Kabahuma M, Chaya T, Kelly A, Borrego E, Bian Y, El Kasmi F, Yang L, Teixeira P, Kolkman J, Nelson R, Kolomiets M, Dang JL, Wisser R, Caplan J, Li X, Lauter N, Balint-Kurti P. A gene encoding maize caffeoyl-CoA O-methyltransferase confers quantitative resistance to multiple pathogens. Nat Genet. 2017;49(9):1364–72.
23. Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Holland JB. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat Genet. 2011;43:163–68.
24. Bian Y, Yang Q, Balint-Kurti PJ, Wisser RJ, Holland JB. Limits on the reproducibility of marker associations with southern leaf blight resistance in the maize nested association mapping population. BMC Genomics 2014 15:1068.
25. Chen G, Xiao Y, Dai S, Dai Z, Wang X, Li B, Jaqueth JS, Li W, Lai Z, Ding J, Yan J. Genetic basis of resistance to southern corn leaf blight in the maize multi-parent population and diversity panel. Plant Biotechnol J 2023 21:506–20.
26. Contreras-Soto RI, Mora F, de Oliveira MAR, Higashi W, Scapim CA, Schuster I. A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-based haplotype analysis. PLoS ONE. 2017;12(2):e0171105.
27. Kumar S, Singh VP, Saini DK, et al. Meta-QTLs, ortho-MQTLs, and candidate genes for thermotolerance in wheat (Triticum aestivum L). Mol Breed. 2021;41(11):69. https://doi.org/10.1007/s11032-021-01264-7.
28. Kumar B, Choudhary M, Kumar K, et al. Maydis leaf blight of maize: update on status, sustainable management and genetic architecture of its resistance. Physiol Mol Plant Pathol. 2022;121:101889.
29. Zaidi PH, Seetharam K, Krishna G, et al. Genomic regions associated with root traits under drought stress in tropical maize (*Zea mays* L). PLoS ONE. 2016;11(10):e0164340.
30. Lim SM. Heterotic effects of resistance in maize to *Helminthosporium maydis* race O. Phytopathology. 1975;65(10):1117–20.
31. Hooda KS, Bagaria PK, Khokhar M, Kaur H, Rakshit S. Mass Screening techniques for resistance to Maize diseases. ICAR- Indian Institute of Maize Research. Campus: PAU; 2018.
32. Alvarado G, López M, Vargas M et al. 2015. META-R (Multi Environment Trail Analysis With R for Windows). version 4.1. http://hdl.handle.net/11529/10201
33. CIMMYT, Laboratory Protocols. CIMMYT Applied Molecular Genetics Laboratory Protocols. Mexico: CIMMYT; 2001.
34. Elshire RJ, Glaubitz JC, Sun Q, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE. 2011;6(5):e19379.
35. Glaubitz JC, Casstevens TM, Lu F, et al. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE. 2014;9(2):e90346.
36. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38(8):904–9.
37. Wang J, Zhang Z. *GAPIT*. Version 3: boosting power and accuracy for genomic association and prediction. Genomics Proteom Bioinf. 2021;19:787–96.
38. R Core Team. R: A Language and Environment for Statistical Computing; 2021. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/
39. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. GigaScience. 2019;8.
40. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome wide association studies. PLos Genet. 2016;12(2):1005767.
41. Mao H, Wang H, Liu S, et al. A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. Nat Commun. 2015;6:8326.
42. Cui Z, Luo J, Qi C, et al. Genome-wide association study (GWAS) reveals the genetic architecture of four husk traits in maize. BMC Genomics. 2016;17:946.
43. Hammer Ø, Harper D, Ryan P.PAST. Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica. 2001;4:1–9.
44. DATAtab Team. (2023). DATAtab: Online Statistics Calculator. DATAtab e.U. Graz, Austria. URL https://datatab.net.
45. Excoffier LSM, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol. 1995;12(5):921–7.
46. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. Science. 2002;296(5576):2225–9.
47. Walley JW, Sartor RC, Shen Z, et al. Integration of omic networks in a developmental atlas of maize. Science. 2016;353(6301):814–8.
48. Chen C, Chen H, Zhang Y, et al. TB tools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
49. Xia Y. Statistical normalization methods in microbiome data with application to microbiome cancer research. Gut Microbes. 2023;15(2):2244139.
50. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38:214–20. https://doi.org/10.1093/nar/gkq537.
51. Franz M, Rodriguez H, Lopes C, Zuberi K, Montojo J, Bader GD, Morris Q. GeneMANIA update 2018. Nucleic Acids Res. 2018;46:60–4.
52. Yadav OP, Hossain F, Karjagi CG. Genetic improvement of maize in India: retrospect and prospects. Agric Res. 2015;4:325–38.
53. Gowda M, Das B, Makumbi D, et al. Genome –wide association studies and genomic predictions of resistance to maize lethal necrosis disease in tropical maize germplasm. Theor Appl Genet. 2015;128(10):1957–68.
54. Rashid Z, Kaur H, Babu V, Singh PK, Harlapur SI, Nair SK. Identification and validation of genomic regions associated with charcoal rot resistance in tropical maize by genome-wide association and linkage mapping. Front Plant Sci. 2021;12:726767.
55. Rashid Z, Sofi M, Harlapur SI, et al. Genome-wide association studies in tropical maize germplasm reveal novel and known genomic regions for resistance to Northern corn leaf blight. Sci Rep. 2020;10(1):21949. https://doi.org/10.1038/s41598-020-78928-5.
56. George ML, Regalado E, Warburton M, Vasal S, Hoisington D. Genetic diversity of maize inbred lines in relation to downy mildew. Euphytica. 2004;135(2):145–55.
57. Warburton ML, Ribaut JM, Franco J, Crossa J, Dubreuil P, Betrán FJ. Genetic characterization of 218 elite CIMMYT maize inbred lines using RFLP markers. Euphytica. 2005;142(1–2):97–106. https://doi.org/10.1007/s10681-005-0817-y.
58. Romay MC, Millard MJ, Glaubitz JC, et al. Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 2013;14(6):R55.
59. Martins LB, Rucker E, Thomason W, Wisser RJ, Holland JB, Balint-Kurti P. Validation and characterization of maize multiple disease resistance QTL. (Bethesda). 2019;G3(9):2905–12.
60. Bubeck DM, Goodman MM, Beavis WD, Grant D. Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci. 1993;33(4):838–47.
61. Ali F, Pan Q, Chen G, Zahid KR, Yan J. Evidence of multiple disease resistance (MDR) and implication of meta-analysis in marker assisted selection. PLoS ONE. 2013;8(7):e68150.
62. Kaur M, Vikal Y, Kaur H, Pal L, Kaur K, Chawla JS. Mapping quantitative trait loci associated with southern leaf blight resistance in maize (*Zea mays* L). J Phytopathol. 2019;167(10):591–600.
63. Liu Y, Hu G, Zhang A, et al. Genome-wide association study and genomic prediction of Fusarium ear rot resistance in tropical maize germplasm. Crop J. 2021;9(2):325–41.
64. McMullen MD, Simcox KD. Genomic organization of disease and insect resistance genes in maize. Mol Plant Microbe Interact. 1995;8(6):811–5.
65. Li YX, Chen L, Li C, et al. Increased experimental conditions and marker densities identified more genetic loci associated with southern and northern leaf blight resistance in maize. Sci Rep. 2018;8(1):6848.
66. Zwonitzer JC, Bubeck DM, Bhattramakki D, Goodman MM, Arellano C, Balint-Kurti PJ. Use of selection with recurrent backcrossing and QTL mapping to identify loci contributing to southern leaf blight resistance in a highly resistant maize line. Theor Appl Genet. 2009;118(5):911–25.
67. Rashid Z, Babu V, Sharma SS, Singh PK, Nair SK. Identification and validation of a key genomic region on chromosome 6 for resistance to Fusarium stalk rot in tropical maize. Theor Appl Genet. 2022;135(12):4549–63.

68. Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 2008;9(6):477–85. https://doi.org/10.1038/nrg2361.
69. Amaral AJ, Megens HJ, Crooijmans RP, Heuven HC, Groenen MA. Linkage disequilibrium decay and haplotype block structure in the pig. Genetics. 2008;179(1):569–79.
70. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. Plant Genome-US. 2008;1(1):5–20.
71. Lorenz AJ, Hamblin MT, Jannink JL. Performance of single nucleotide polymorphisms versus haplotypes for genome-wide association analysis in barley. PLoS ONE. 2010;5(11):e14079. https://doi.org/10.1371/journal.pone.0014079.
72. Fuji K, Shirakawa M, Shimono Y, et al. The adaptor complex AP-4 regulates vacuolar protein sorting at trans-golgi network by interacting with vacuolar SORTING RECEPTOR 1. Plant Physiol. 2016;170(1):211–9.
73. Wang W, Barnaby JY, Tada Y, et al. Timing of plant immune responses by a central circadian regulator. Nature. 2011;470(7332):110–4.
74. Yazaki K. ABC transporters involved in the transport of plant secondry metabolities. FEBS Lett. 2006;580(4):1183–91.
75. Chandan RK, Kumar R, Swain DM, et al. RAV1 family members function as transcriptional regulators and play a positive role in plant disease resistance. Plant J. 2023;114(1):39–54.
76. Wisser RJ, Balint-Kurti PJ, Nelson RJ. The genetic architecture of disease resistance in maize: a synthesis of published studies. Phytopathology. 2006;96(2):120–9.
77. Meng X, Zhang S. MAPK cascades in plant disease resistance signaling. Annu Rev Phytopathol. 2013;51:245–66.
78. Meshram S, Gogoi R. Look into circadian rhythm in maize during Bipolaris maydis infection. Plant Dis Res. 2021;36(2):101–5.
79. Creelman RA, Mullet JE. Biosynthesis and action of jasmonates in plants. Annu Rev Plant Physiol Plant Mol Biol. 1997;48:355–81.
80. Ma H, Liu C, Li Z, et al. ZmbZIP4 contributes to stress resistance in maize by regulating ABA synthesis and root development. Plant Physiol. 2018;178(2):753–70.
81. Yen SK, Chung MC, Chen PC, Yen HE. Environmental and developmental regulation of the wound-induced cell wall protein WI12 in the halophyte ice plant. Plant Physiol. 2001;127(2):517–28.
82. Anwar R, Mattoo AK, Handa AK. Polyamine interactions with plant hormones: crosstalk at several levels. Int J Hydrog Energy. 2015;38:1039–51.
83. Romeis T, Piedras P, Jones JDG. Resistance gene-dependent activation of a calcium-dependent protein kinase in the plant defense response. Plant Cell. 2000;12(5):803–16.
84. Huang TK, Han CL, Lin SI, et al. Identification of downstream components of ubiquitin-conjugating enzyme phosphate 2 by quantitative membrane proteomics in *Arabidopsis* roots. Plant Cell. 2013;25(10):4044–60.
85. Bekalu ZE, Krogh Madsen C, Dionisio G, et al. Overexpression of nepenthesin HvNEP-1 in barley endosperm reduces Fusarium head blight and mycotoxin accumulation. Agronomy. 2020;10(2):203.
86. Bevan MW, et al. Genomic innovation for crop improvement. Nature. 2017;543:346–54.
87. Bastiaansen JWM, Bovenhuis H, Lopes MS et al. SNP effects depend on genetic and environmental Context. Proceedings, 10th World Congress of Genetics Applied to Livestock Production.2015.
88. Atamian HS, Creux NM, Brown RI, Garner AG, Blackman BK, Harmer SL. Circadian regulation of sunflower heliotropism, floral orientation, and pollinator visits. Science. 2016;353(6299):587–90.
89. Li D, Liu H, Zhang H, Wang X, Song F. OsBIRH1, a DEAD-box RNA helicase with functions in modulating defence responses against pathogen infection and oxidative stress. J Exp Bot. 2008;59(8):2133–46.
90. Sahu PP, Rai NK, Chakraborty S, et al. Tomato cultivar tolerant to Tomato leaf curl New Delhi virus infection induces virus-specific short interfering RNA accumulation and defence-associated host gene expression. Mol Plant Pathol. 2010;11(4):531–44.
91. Kampinga HH, Craig EA. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat Rev Mol Cell Biol. 2010;11(8):579–92.
92. Matsuoka E, Kato N, Hara M. Induction of the heat shock response in Arabidopsis by heat shock protein 70 inhibitor VER-155008. Funct Plant Biol. 2019;46(10):925–32.
93. Li J, Zhang M, Sun J, et al. Heavy Metal Stress-Associated Proteins in Rice and Arabidopsis: genome-wide identification, Phylogenetics, Duplication, and expression profiles analysis. Front Genet. 2020;11:477.
94. Zhu X, Li X, He Q, et al. TaMYB29: a novel R2R3-MYB transcription factor involved in wheat defence against stripe rust. Front Plant Sci. 2021;12:783388.
95. Mengiste T, Chen X, Salmeron J, Dietrich R. The BOTRYTIS SUSCEPTIBLE1 gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. Plant Cell. 2003;15(11):2551–65.
96. Trujillo M, Shirasu K. Ubiquitination in plant immunity. Curr Opin Plant Biol. 2010;13(4):402–8.
97. Hwang JU, Song WY, Hong D, et al. Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. Mol Plant. 2016;9(3):338–55.
98. Hatsugai N, Nakatsuji A, Unten O, et al. Involvement of adapter protein complex 4 in hypersensitive cell death induced by avirulent bacteria. Plant Physiol. 2018;176(2):1824–34.
99. Hao Y, Zong X, Ren P, Qian Y, Fu A. Basic helix-loop-helix (bHLH) transcription factors regulate a wide range of functions in Arabidopsis. Int J Mol Sci. 2021;22(13):7152.
100. Cao X, Yang H, Shang C, Ma S, Liu L, Cheng J. The roles of auxin biosynthesis yucca gene family in plants. Int J Mol Sci. 2019;20(24):6343.
101. Thomma BP, Eggermont K, Penninckx IA, et al. Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. Proc Natl Acad Sci USA. 1998;95(25):15107–11.
102. Hammond-Kosack KE, Tang S, Harrison K, Jones JD. The tomato Cf-9 disease resistance gene functions in tobacco and potato to confer responsiveness to the fungal avirulence gene product avr 9. Plant Cell. 1998;10(8):1251–66.
103. Kalamaki MS, Alexandrou D, Lazari D, et al. Over-expression of a tomato N-acetyl-L-glutamate synthase gene (SlNAGS1) in Arabidopsis thaliana results in high ornithine levels and increased tolerance in salt and drought stresses. J Exp Bot. 2009;60(6):1859–71.
104. Kaminaka H, Näke C, Epple P, et al. bZIP10-LSD1 antagonism modulates basal defense and cell death in Arabidopsis following infection. EMBO J. 2006;25(18):4400–11.
105. Kopischke M, Westphal L, Schneeberger K, et al. Impaired sterol ester synthesis alters the response of Arabidopsis thaliana to Phytophthora infestans. Plant J. 2013;73(3):456–68.
106. Wang D, Pajerowska-Mukhtar K, Culler AH, Dong X. Salicylic acid inhibits pathogen growth in plants through repression of the auxin signaling pathway. Curr Biol. 2007;17(20):1784–90.
107. Xu ZS, Xia LQ, Chen M, et al. Isolation and molecular characterization of the Triticum aestivum L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. Plant Mol Biol. 2007;65(6):719–32.
108. Yang W, Devaiah SP, Pan X, Isaac G, Welti R, Wang X. AtPLAI is an acyl hydrolase involved in basal jasmonic acid production and Arabidopsis resistance to Botrytis Cinerea. J Biol Chem. 2007;282(25):18116–28.
109. Li Z, Liu D, Xia Y, et al. Identification and functional analysis of the CLAVATA3/EMBRYO SURROUNDING REGION (CLE) gene family in wheat. Int J Mol Sci. 2019;20(17):4319.
110. Pecher P, Eschen-Lippold L, Herklotz S, et al. The Arabidopsis thaliana mitogen-activated protein kinases MPK3 and MPK6 target a subclass of VQ-motif-containing proteins to regulate immune responses. New Phytol. 2014;203(2):592–606.
111. Shang Y, Yang D, Ha Y, Hur Y-S, Lee MM, Nam KH. Brassinosteroid-insensitive 1-Associated receptor kinase 1 modulates abscisic acid signaling by inducing PYR1 monomerization and Association with ABI1 in Arabidopsis. Front Plant Sci. 2022;13:849467.
112. Zhai C, Lin F, Dong Z, et al. The isolation and characterization of Pik, a rice blast resistance gene which emerged after rice domestication. New Phytol. 2011;189(1):321–34.
113. Zwonitzer JC, Coles ND, Krakowsky MD, Arellano C, Holland JB, McMullen MD, Balint-Kurti PJ. Mapping resistance quantitative trait loci for three foliar diseases in maize recombinant inbred line population–evidence for multiple disease resistance. Phytopathology. 2010;100:72–9.
114. Pengfei L, Jiang F, Zhang J, Wang H, Wang X. QTL mapping for resistance to southern leaf blight in sweet corn. Afr J Agric Res. 2011;6:197–203.
115. Zaitlin D, DeMars S, Ma Y. Linkage of rhm a recessive gene for resistance to Southern corn leaf blight, RFLP marker loci in maize (Zea mays) seedlings. Genome. 1993;36:555–64.
116. Belcher AR. The physiology and host genetics of quantitative resistance in maize to the fungal pathogen Cochliobolus heterostrophus. Dissertation, North Carolina State University, Raleigh. 2009.

117. Lennon JR, Krakowsky M, Goodman M, Flint-Garcia S, Balint-Kurti PJ. Identification of teosinte alleles for resistance to southern leaf blight in near isogenic maize lines. Crop Sci. 2017;57:1973–83.

Publisher's Note