

Article

Utilizing Two Populations Derived from Tropical Maize for Genome-Wide Association Analysis of Banded Leaf and Sheath Blight Resistance

Shaoxiong Li ¹, Fuyan Jiang ², Yaqi Bi ², Xingfu Yin ², Linzhuo Li ¹, Xingjie Zhang ¹, Jinfeng Li ¹, Meichen Liu ¹, Ranjan K. Shaw ² and Xingming Fan ^{2,*} 

¹ College of Agriculture, Yunnan University, Kunming 650500, China; 15987701739@163.com (S.L.); lilinzhuo0606@163.com (L.L.); xingjiezhang2022@163.com (X.Z.); jinfengli1020@163.com (J.L.); shirleyliu1028@163.com (M.L.)

² Institute of Food Crops, Yunnan Academy of Agricultural Sciences, Kunming 650205, China; jiangfuyansxx@126.com (F.J.); biyq122627@163.com (Y.B.); xingfuyin626@163.com (X.Y.); ranjanshaw@gmail.com (R.K.S.)

* Correspondence: xingmingfan@163.com

Abstract: Banded leaf and sheath blight (BLSB) in maize is a soil-borne fungal disease caused by *Rhizoctonia solani* Kühn, resulting in significant yield losses. Investigating the genes responsible for regulating resistance to BLSB is crucial for yield enhancement. In this study, a multiparent maize population was developed, comprising two recombinant inbred line (RIL) populations totaling 442 F8RILs. The populations were generated by crossing two tropical inbred lines, CML444 and NK40-1, known for their BLSB resistance, as female parents, with the high-yielding but BLSB-susceptible inbred line Ye107 serving as the common male parent. Subsequently, we utilized 562,212 high-quality single nucleotide polymorphisms (SNPs) generated through genotyping-by-sequencing (GBS) for a comprehensive genome-wide association study (GWAS) aimed at identifying genes responsible for BLSB resistance. The objectives of this study were to (1) identify SNPs associated with BLSB resistance through genome-wide association analyses, (2) explore candidate genes regulating BLSB resistance in maize, and (3) investigate pathways involved in BLSB resistance and discover key candidate genes through Gene Ontology (GO) analysis. The GWAS analysis revealed nineteen SNPs significantly associated with BLSB that were consistently identified across four environments in the GWAS, with phenotypic variation explained (PVE) ranging from 2.48% to 11.71%. Screening a 40 kb region upstream and downstream of the significant SNPs revealed several potential candidate genes. By integrating information from maize GDB and the NCBI, we identified five novel candidate genes, namely, *Zm00001d009723*, *Zm00001d009975*, *Zm00001d009566*, *Zm00001d009567*, located on chromosome 8, and *Zm00001d026376*, on chromosome 10, related to BLSB resistance. These candidate genes exhibit association with various aspects, including maize cell membrane proteins and cell immune proteins, as well as connections to cell metabolism, transport, transcriptional regulation, and structural proteins. These proteins and biochemical processes play crucial roles in maize defense against BLSB. When *Rhizoctonia solani* invades maize plants, it induces the expression of genes encoding specific proteins and regulates corresponding metabolic pathways to thwart the invasion of this fungus. The present study significantly contributes to our understanding of the genetic basis of BLSB resistance in maize, offering valuable insights into novel candidate genes that could be instrumental in future breeding efforts to develop maize varieties with enhanced BLSB resistance.



Citation: Li, S.; Jiang, F.; Bi, Y.; Yin, X.; Li, L.; Zhang, X.; Li, J.; Liu, M.; Shaw, R.K.; Fan, X. Utilizing Two Populations Derived from Tropical Maize for Genome-Wide Association Analysis of Banded Leaf and Sheath Blight Resistance. *Plants* **2024**, *13*, 456. <https://doi.org/10.3390/plants13030456>

Academic Editor: Maryam Alavi-Shoshtari

Received: 11 January 2024

Revised: 1 February 2024

Accepted: 2 February 2024

Published: 4 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: maize; BLSB; *Rhizoctonia solani*; GWAS; candidate gene

1. Introduction

Maize (*Zea mays* L.) stands as a primary cereal crop cultivated extensively worldwide, playing a dual role as a staple food source and a vital contributor to fodder and biofuel

production. However, maize production faces persistent challenges from both biotic and abiotic stresses, resulting in substantial disruptions to yield and product quality [1]. Notably, banded leaf and sheath blight (BLSB), caused by the fungus *Rhizoctonia solani*, has emerged as a severe threat to maize crops. BLSB is a widespread disease that poses a significant threat to maize plants globally. This disease holds the potential to cause yield reductions ranging from 11% to 40% and, under severe conditions, can even lead to complete crop failure [2–4]. According to Chen [5], Songyang County in Zhejiang Province experiences a significant outbreak of BLSB every autumn. In 1983, the average incidence rate of BLSB was 41.2%, leading to a 9.3% loss in maize yield. By 1985, the incidence rate surged to 70.8%, causing a 16.5% yield loss. Tan et al. [6] documented a significant BLSB outbreak in the western mountains of Hubei Province in 1987. A survey conducted in Zigui County revealed that out of 185,000 SI of maize, 80,000 SI were affected, with 15,000 acres experiencing severe infection. The average incidence rate was 98.5%, with some plots recording a 100% incidence rate. The average disease index was 51.4%, peaking at 64.8%, resulting in an average yield loss of 18.8%. The primary causative agent, *Rhizoctonia solani* f. sp. *Sasakii* Exner, is prevalent in maize growing regions worldwide, with particular prominence in the southern parts of China and maize-growing areas in Southeast Asian countries [4]. BLSB, induced by *Rhizoctonia solani*, typically manifests in the leaf sheaths of the first to third nodes of maize plants. The lesions progress gradually upward, extending eventually to the corn ears and husks, where the characteristic sclerotia form at the lesion sites. *Rhizoctonia solani* can disrupt the vascular tissues of maize plants, impeding the normal transport of water and nutrients and resulting in a significant reduction in maize yield. This disease tends to thrive in warm and humid climatic conditions, especially when the temperature is maintained at 28 ± 2 °C and the relative humidity exceeds 88% [7,8].

R. solani, a soil-borne fungus with widespread geographical distribution and a broad host range, poses a significant challenge during maize cultivation. Currently, the primary method for controlling BLSB involves the use of chemical fungicides. However, this approach is becoming increasingly environmentally unsustainable and less acceptable to consumers due to heightened concerns about food safety and environmental protection. Consequently, the cultivation or development of maize varieties resistant to BLSB has become crucial. Despite this urgency, most existing maize varieties remain susceptible to *R. solani*, while only a few maize germplasms exhibit high resistance to BLSB. This limitation significantly hampers both the study of resistance mechanisms and the development of disease-resistant maize varieties [9].

Genome-wide association study (GWAS) is a highly effective mapping approach based on the principle of linkage disequilibrium (LD) for precisely mapping quantitative trait loci (QTL) that control complex traits. Initially proposed in 1996 [10], this approach was originally applied to analyze complex traits associated with human diseases. Through this method, researchers can precisely locate genomic regions linked to specific traits in crops, thereby advancing our understanding of the intricate mechanisms controlling these traits [11,12]. GWAS has emerged as a robust and efficient tool in genetic research, playing a pivotal role in deciphering the genetic basis of complex traits in crops [13,14]. In recent years, scholars have increasingly employed GWAS to investigate loci controlling various traits in maize, including plant height (PH), ear length (EH) [10], yield [15], disease resistance [16], and grain dehydration [17]. Through GWAS, Li Ning [18] successfully identified a gene, *ZmFBL4*, associated with banded leaf and sheath blight. Numerous studies have successfully identified QTLs related to crucial agronomic traits, thereby showcasing the practicality and efficiency of GWAS [19–22]. In crops like maize, GWAS has proven invaluable in exploring key genes and identifying potential disease-resistant genes. The progression of these studies unequivocally underscores GWAS as a powerful tool for effectively establishing associations between markers and phenotypes.

In recent years, significant progress has been made in investigating genetic resources related to resistance against BLSB in maize. Researchers have identified maize varieties, such as Jinyu 506 [23], Zhongke Yu505 [24], and Endan801 [25], that exhibit enhanced

resistance to BLSB. Wu et al. [26] reported that the novel sweet corn variety ‘Huawang Sweet No.7’ displays high resistance to leaf blight and moderate resistance to BLSB. However, currently, maize varieties identified with relative resistance to *R. solani* are predominantly moderately resistant, with no highly resistant or immune maize varieties having been discovered. Therefore, it is crucial to explore additional maize germplasm resources exhibiting higher resistance to BLSB and understand the genetic mechanisms underlying such resistance, and this holds significant importance for breeding maize varieties resistant to BLSB [27]. In this study, we utilized the tropical inbred lines CML444 and NK40-1 as resistant parents, crossing them with the excellent maize inbred line Ye107 from the Reid heterotic group to develop a multi-parent population comprising 442 F8 recombinant inbred lines (RILs). Subsequently, the multi-parent population underwent a GWAS analysis. The study’s objectives were to (1) identify SNPs associated with BLSB resistance through genome-wide association analysis, (2) explore candidate genes regulating BLSB resistance in maize, and (3) investigate pathways involved in BLSB resistance and discover key candidate genes through Gene Ontology (GO) analysis.

2. Results

2.1. Phenotypic Data Analysis

Screening of the two RIL populations, pop1 and pop2, against BLSB was conducted at two locations; Yanshan County, Yunnan Province, in the years 2021 and 2022, and Jinghong City, Yunnan Province, in the year 2022, with corresponding data collected. Descriptive statistics of the phenotypic data for the two RIL populations screened against BLSB are presented in Table 1. The coefficients of variation (CV) for pop2 across three environments were 0.34, 0.39, and 0.31, respectively. For pop1, the CV values across three environments were 0.45, 0.46, and 0.48. The absolute values of skewness and kurtosis for both pop1 and pop2 populations were less than 1, indicating a minor degree of bias. The frequency distribution of phenotypes for pop2 and pop1 closely resembled a normal distribution. Under the three environmental conditions, the broad-sense heritability of the disease severity index for BLSB was high in pop2 and pop1 at 0.93 and 0.92, respectively. Both the genotype \times environment interaction variances were statistically significant. ($p \leq 0.01$) (Table 1) The high heritability of disease severity levels and genotype \times environment interaction variances were significant underscores of the reliability of identifying BLSB-resistant genes through GWAS.

Table 1. Descriptive statistics of the two RIL subpopulations for reactions against BLSB.

Populations (Sample Number.)	Env.	Range	Mean	SD	CV (%)	Skewness	Kurtosis	Variations			H ² (%)
								G	E	G*E	
Pop2 (221)	JH 22 W	1–9	5.59	1.922	0.34	−0.518	0.713	3.42 *	0.25 *	0.3 *	0.93
	YS 21 S	1–9	5.25	2.073	0.39	0.163	−0.061				
	YS 22 S	1–9	5.98	1.865	0.31	−0.408	0.757				
Pop1 (221)	JH 22 W	1–9	4.66	2.18	0.45	−0.035	−0.586	4.14 *	0.49 *	0.118 *	0.92
	YS 21 S	1–9	3.62	2.252	0.46	0.661	−0.133				
	YS 22 S	1–9	4.58	2.226	0.48	−0.04	−0.638				

W represents winter, S represents summer. Env. represents environment. Range represents the range of disease grade of plants within the population. SD represents standard deviation. CV represents coefficient of variation. H² represents heritability. * represents $p < 0.05$.

As depicted in Figure 1 and Table 2, the disease scores of the majority of the RILs were more concentrated at the lower end in pop1 group than the pop2 group, indicating lower disease severity levels in pop1 than in pop2. This suggests that pop1 exhibited greater resistance to BLSB than pop2. Regarding environmental variations, significant differences in disease resistance between pop1 were observed at the YS21 environment compared to JH22 and YS22 environments, with pop2 showing no significant differences across the three

environments. Notably, between pop1 and pop2 exhibited significant differences in disease resistance at the three environments JH22, YS21, and YS22.

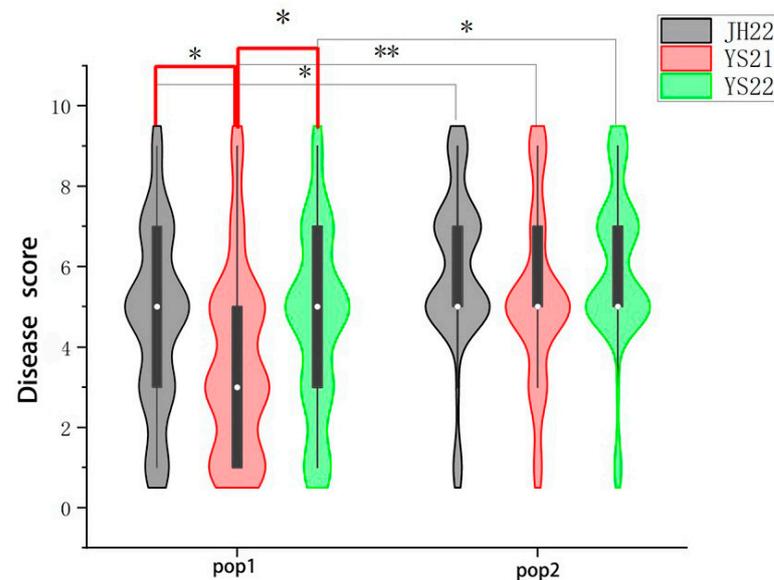


Figure 1. The violin plot of phenotypic distribution of RILs for pop1 and pop2 to BLSB response in three different environments (JH22, YS21, YS22). The asterisk above indicates the result of Student's *t* test, * represents $p < 0.05$, ** represents $p < 0.01$.

Table 2. Analysis of variance for pop1 and pop2 evaluated. * represents $p < 0.05$.

Pop	df	Sum of Squares	F-Value	<i>p</i>	Error
Pop1	2	5.60	14.9	0.0227 *	0.52
Pop2	2	4.12	7.98	0.056	0.38

The correlation analysis for BLSB across different environments within the same population is depicted in Figure 2. The correlation coefficients for disease severity levels in pop1 between the JH22, YS21, and YS22 environments were 0.90, 0.92, and 0.89, respectively. In pop2, the correlation coefficients for disease severity across the three environments (JH22, YS21, and YS22) were 0.89, 0.95, and 0.87, respectively. The consistently high correlation coefficients among the three environments indicate a significant and stable response of the RILs of pop1 and pop2 to BLSB across different environments, ensuring the reliability of the phenotyping for subsequent GWAS analyses.

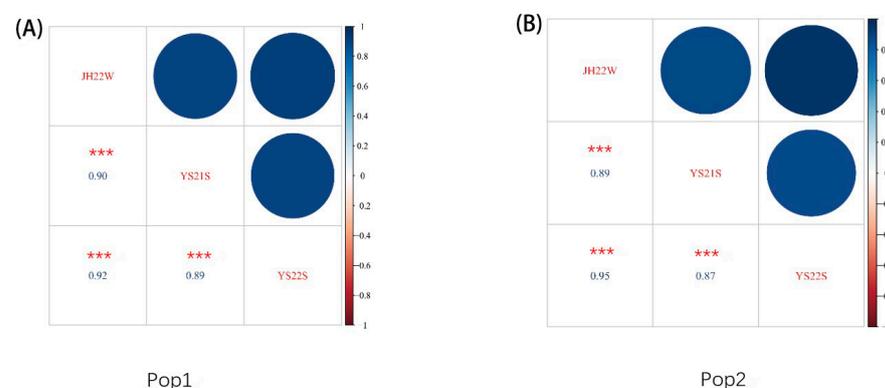


Figure 2. Correlation of pop1 and pop2 to BLSB response across three environments. (A) Correlation of pop1 to BLSB response between the JH22, YS21, and YS22 environments; (B) correlation of pop2 to BLSB response between the JH22, YS21, and YS22 environments. *** indicates $p < 0.001$.

2.2. SNP Characterization, LD Decay Distance, and Population Structure

Through Genotyping-by-Sequencing (GBS), a total of 562,212 high-quality genome-wide SNPs were identified, distributed across all ten chromosomes of maize. The heatmap in Figure 3A illustrates the marker density of SNPs across the maize chromosomes. The number of SNPs identified on chromosome 1 to 10 were as follows: 78,609, 64,672, 63,621, 72,674, 54,053, 46,111, 50,815, 47,367, 42,777, and 41,513. The highest number of SNPs was located on chromosome 1, while chromosome 9 had the fewest. The density of SNPs per mega base (Mb) for chromosomes 1 to 10 were 259.44, 265.05, 270.73, 294.23, 242.39, 265.01, 279.20, 261.70, 269.04, and 276.75, respectively. SNPs were evenly distributed across the chromosomes. In the filtered SNP dataset, the average missing rate was 0.12, and the average minor allele frequency (MAF) was 0.16, indicating the suitability of this dataset for subsequent GWAS (Figure 3B,C). We employed 562,212 SNPs to assess the linkage disequilibrium (LD) decay in the association mapping population. LD in the association mapping panel was estimated using the r^2 of pairwise combinations of SNPs across the chromosomes, with the minimum threshold values set at 0.2. The estimated range of physical distance at which LD decayed was approximately 40 kb (Figure 3D). In comparison with other studies, the LD decay distance of the present study exceeded that reported in the literature. In tropical maize, the average LD decay distance across all ten chromosomes is documented as 8.14 kb [28]. For subtropical maize, at $r^2 = 0.2$, the mean LD decay distance between chromosomes is approximately 5 to 10 kb [29]. In the research conducted by Wu et al. [30], it was observed that the linkage disequilibrium (LD) value in temperate maize was determined to be 391 kb. Rapid LD decay indicates greater genetic diversity, fostering a more favorable genetic relationship among populations. Given that our experiment involves a population comprising both tropical and temperate maize inbreds, the LD decay rate is slow compared to tropical maize.

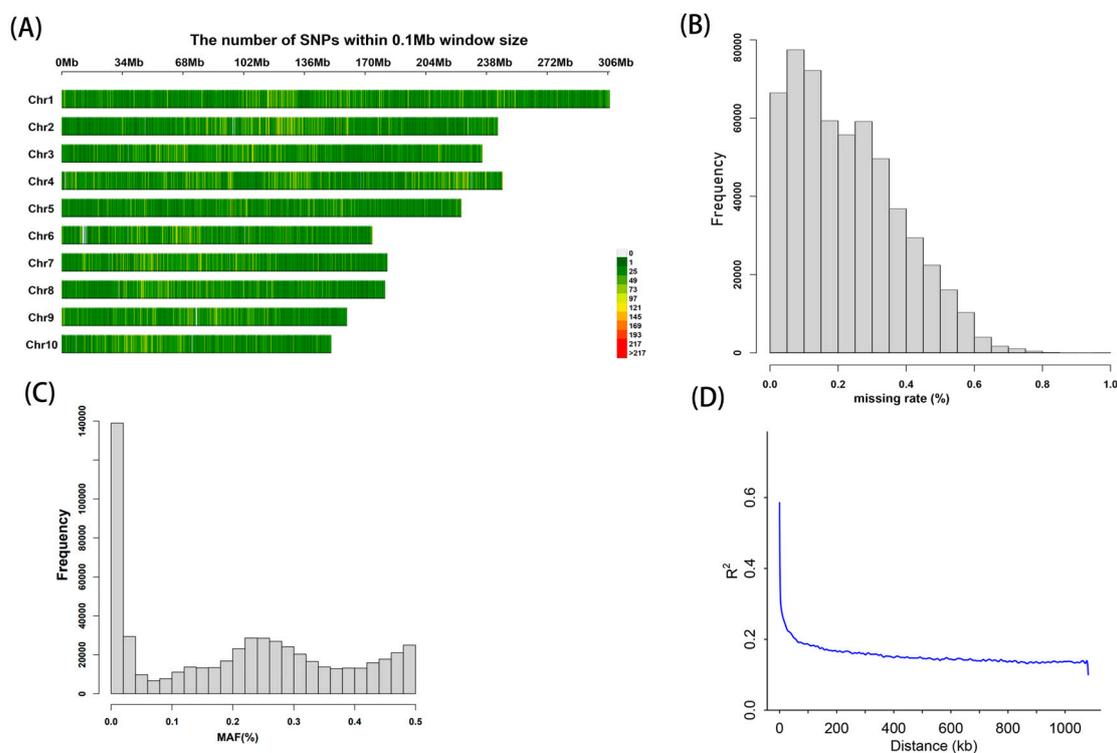


Figure 3. Attenuation map of genotype diversity and LD. (A) Chromosome-specific SNP density within 1-Mb intervals. The range of number of SNPs is indicated by a green to red scale. (B) Frequency distribution of missing data points of SNPs. (C) Distribution of minor allele frequency (MAF) of the RILs. (D) Genome-wide LD decay (r^2) over physical distance (Kb) across all the chromosomes in 442 maize RILs.

The population structure analysis results are depicted in Figure 4. Overall, the results of population structure, principal component analysis (PCA), and genetic distance or relatedness (Figure 4A–C) are consistent. In the PCA plot, scattered points may arise from within-population heterogeneity or outliers. Based on lineage or genetic background, the RILs can be divided into two major clusters. At $K = 2$ (Figure 4C), the population structure of the RILs becomes clear, with 221 RILs in pop1 and 221 in pop2. The phylogenetic tree analysis also revealed two genetic clusters, aligning with the relatedness-based population structure. The pairwise relatedness distribution chart for the 442 RILs (Figure 4D) further separates them into two distinct groups.

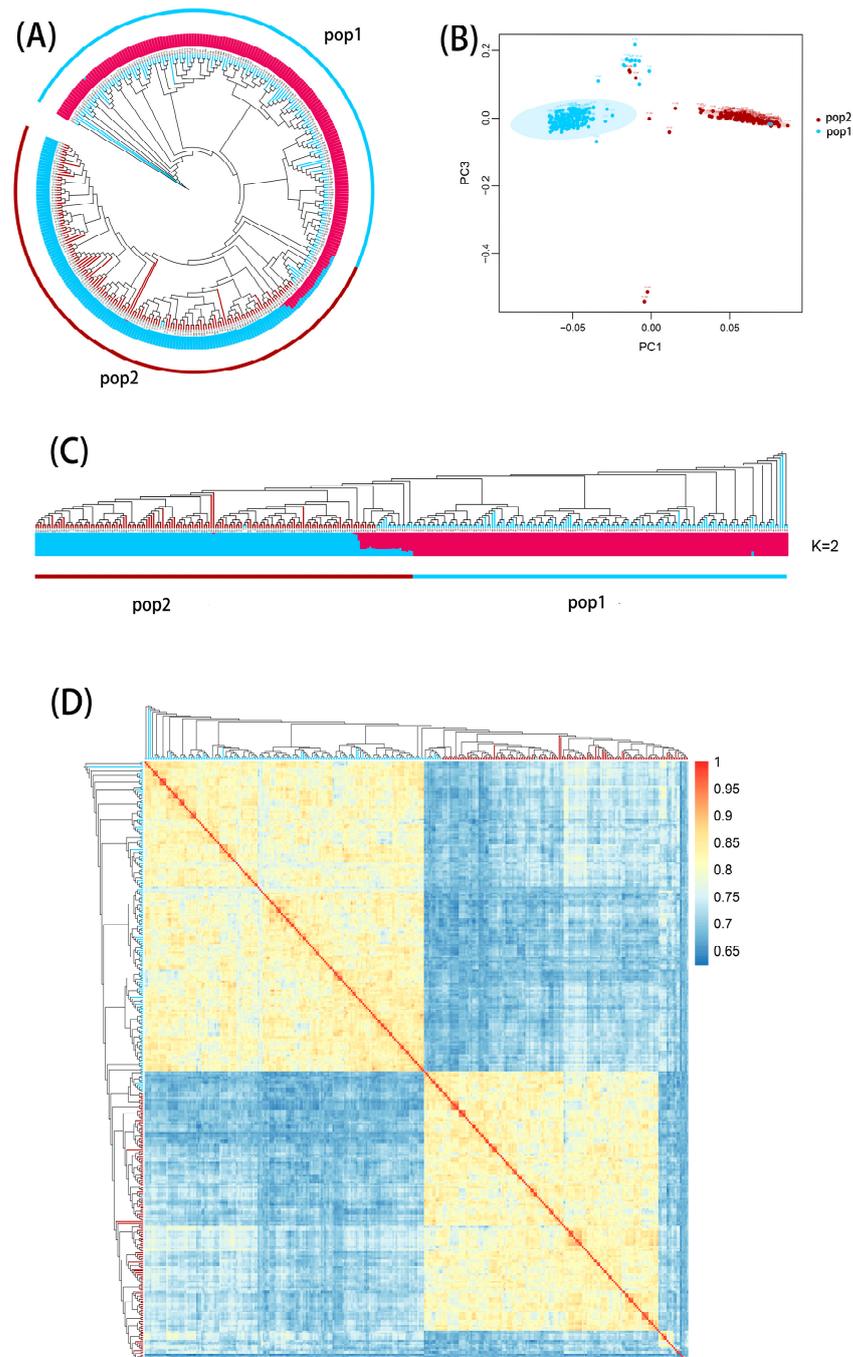


Figure 4. Genetic diversity analysis. (A) Phylogenetic tree of 442 RILs. (B) Principal component analysis. (C) Bayesian clustering plots of 442 maize inbred lines at $K = 2$. (D) Pairwise distribution of 442 maize RILs.

2.3. Genome-Wide Association Analysis for BLSB Resistance

The F-test revealed a significant difference between the RIL subpopulations for BLSB resistance in the two environments, Jinghong and Yanshan, and the data tend to follow a normal distribution (Figure 1). Based on the phenotypic data for BLSB disease severity of 442 RILs in the multi-parent population, and the 562,212 high-quality SNPs, GWAS analysis was conducted to identify SNPs associated with BLSB resistance. The association analysis utilized the Mixed Linear Model (MLM), which considers both the relatedness (K) and population structure as covariates to control false positives arising due to the effect of population stratification. The MLM analysis was conducted with significance levels set at $-\log_{10}(p) > 4.5$. As population structure can influence association analysis, a quantile–quantile (Q–Q) plot was employed to assess how well the adopted model accounted for population structure. The Q–Q plot indicated that this model effectively controls the population structure, with the negative log of association probability values plotted against the observed negative log of association probability values under the null hypothesis of no association. While the MLM helps reduce false-positives, it may lead to the identification of some false-negative markers. This was evident in the present study, as several significant SNPs were below the expected values in the Q–Q plot, indicating the presence of false-negatives. BLUP estimation was performed using the Lme package in RStudio (R-4.3.1) BLUP integrated disease severity levels across environments and utilized 562,212 high-quality SNPs, principal component analysis, and the Kinship matrix for GWAS analysis.

In the JH22W environment, a total of 13 SNPs significantly associated with BLSB resistance (with a significant threshold of $-\log_{10}(p) = 4.5$) were identified. These SNPs were distributed across chromosomes 2, 3, 4, 5, 6, 8, and 10 (Figure 5A). The significant SNPs explained a phenotypic variance (PVE) ranging from 3.18% to 11.71%, with an average of 5.02%. Among all the significant SNPs, chromosome 5 had the highest number (five SNPs), followed by chromosome 8 (four SNPs), while chromosomes 10 and 4 each had three SNPs, chromosome 3 had two SNPs, and the remaining chromosomes had only one.

In the YS21S environment, a total of 18 SNPs significantly associated with resistance to BLSB (with a significant threshold of $-\log_{10}(p) = 4.5$) were identified. These SNPs were distributed across chromosomes 1, 3, 4, 5, 6, 7, 8, and 10 (Figure 5B). The significant SNPs explained a phenotypic variance (PVE) ranging from 2.74% to 11.71%, with an average of 6.69%. Among all significant SNPs, chromosome 5 had the highest number (eight SNPs), while chromosomes 1, 6, and 7 had the fewest, each with one SNP.

In the YS22S environment, a total of 12 SNPs significantly associated with resistance to BLSB (with a significant threshold of $-\log_{10}(p) = 4.5$) were identified. These SNPs were distributed across chromosomes 1, 4, 6, 8, and 10 (Figure 5C). The significant SNPs explained a phenotypic variance (PVE) ranging from 4.26% to 11.71%, with an average of 8.33%. Among all significant SNPs, chromosome 8 had the highest number (five SNPs), while chromosome 4 had the fewest, with one SNP.

In the BLUP analysis, a total of 17 SNPs significantly associated with resistance to BLSB (with a significant threshold of $-\log_{10}(p) = 4.5$) were identified. These SNPs were distributed across chromosomes 1, 2, 3, 4, 5, 6, 8, and 10 (Figure 5D). The significant SNPs explained a phenotypic variance (PVE) ranging from 2.48% to 11.71%, with an average of 6.67%. Among all significant SNPs, chromosome 8 had the highest number (ten SNPs), while chromosomes 1 and 2 had the fewest, each with one SNP.

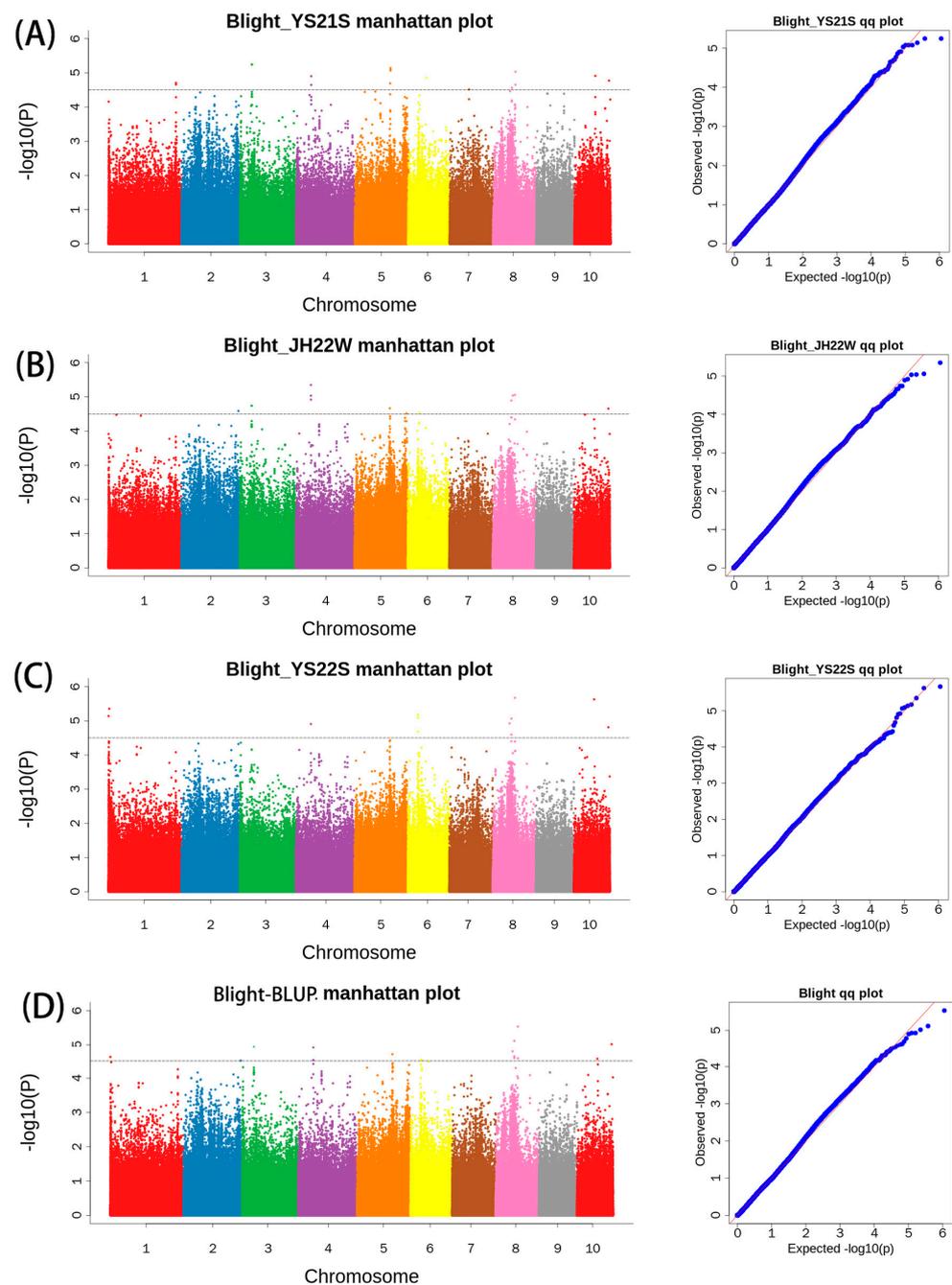


Figure 5. Manhattan map (**left**) and Q–Q plots (**right**) of YS21 (**A**), JH22 (**B**), YS22 (**C**), BLUP (**D**) show SNPs associated with BLSB resistance. Each dot on the left figures represents an SNP, and the black line represents the threshold of $<1 \times 10^{-4}$. Different colors represent different chromosomes. The red lines on the right figures are the trend lines to which the ideal Q–Q plot in each case should correspond.

In summary, a total of 19 SNPs significantly associated with BLSB were identified across the four environments, using a significant threshold of $-\log_{10}(p) = 4.5$ (Table 3). These SNPs serve as valuable candidates for further gene identification.

Table 3. List of 19 significantly associated SNPs identified across four environments.

No.	Chr	Physical Position	Threshold ($-\log_{10}(p) = 4.5$)	Allele	Environments
1	1	442415	4.62	C/T	✦✦
2	1	282946785	4.71	G/A	✦
3	2	239027952	4.51	C/T	✦✦
4	3	50067223	4.92	A/G	✦
5	4	63379553	4.90	C/T	✦
6	4	63306693	4.90	T/A	✦
7	5	147529611	4.69	T/C	✦✦
8	5	148963157	5.13	C/T	✦
9	5	147757336	4.66	T/C	✦
10	5	217509599	4.51	C/T	✦
11	6	42410465	4.52	A/G	✦
12	6	46848365	4.53	C/T	✦✦
13	8	70647932	4.78	C/T	✦✦
14	8	77905708	5.11	C/T	✦✦✦✦
15	8	93109610	5.53	T/G	✦✦✦✦
16	8	83921201	5.04	C/T	✦
17	10	144634116	4.81	G/C	✦✦✦✦
18	10	87410402	4.91	C/T	✦
19	10	84908655	5.62	C/T	✦

No. represents number, chr represents chromosome. Environments represents the point repeatedly located in the four environments, ✦✦✦✦ represents the SNPs identified across the four environments repeatedly located, ✦✦ represents the SNPs identified in both environments repeatedly located, ✦ represents the SNPs identified in one environment.

2.4. Candidate Genes Revealed by GWAS

Using the B73 RefGen_v4 reference genome, 23 candidate genes were identified within a 40 Kb region upstream and downstream of the 19 SNPs significantly associated with BLSB (Table 4). Table 3 provides functional annotations for these candidate genes sourced from maizeGDB. These genes are associated with maize cell membrane proteins, cell immune proteins, and cell metabolism, transport, transcriptional regulation, and structural proteins. Based on functional annotations and genes that were consistently identified across different environments (Table 3), five candidate genes with the potential to have a role in regulating BLSB resistance were identified in this study. Three of these candidate genes, namely *Zm00001d009723* (chr8:77905708) (Figure 6B, Table 4), *Zm00001d009975* (chr8:93109610) (Figure 6C, Table 4), and *Zm00001d026376* (chr10:144634116) (Figure 6D, Table 4), may play a direct or indirect role in influencing BLSB resistance in maize, with phenotypic variance (PVE) of 7.54%, 11.71%, and 4.14%, respectively. These three genes were consistently identified across four distinct environments (Table 3). The first two candidate genes are located on chromosome 8, while the subsequent one is located on chromosome 10. The functions of these three genes are primarily related to protein binding, defense against various stress stimuli, resistance to fungi, and transmembrane protein transport. Additionally, two candidate genes (*Zm00001d009566* and *Zm00001d009567*) placed in close proximity to the significant SNP positioned at chr8:70647932 were consistently identified in two environments (Figure 6A, Table 4). This SNP explained a phenotypic variance of 5.31% and both genes, located on chromosome 8, are associated with cell immune defense.

The 23 genes listed in Table 4 were subjected to Gene Ontology (GO) enrichment analysis (Figure 7A). These genes were associated with a total of 30 GO terms, with protein binding and catalytic activity being the most annotated GO terms in the Molecular Function (MF) category. In the Cellular Component (CC) category, the GO terms included cell membrane, intracellular membrane, cell, and intracellular, with intracellular being the most annotated term. In the Biological Process (BP) category, GO terms such as biosynthetic process, cellular process, macromolecule, organic substance, and protein metabolic process were encompassed, with cellular process emerging as the most frequently annotated term. Notably, within this category, the term GO:0009834~plant-type secondary cell wall biogene-

sis exhibited enrichment for two genes (Figure 7B), associated with the process of secondary cell wall formation in plants. These two genes, *Zm00001d009566* and *Zm00001d0095670*, align with previously identified candidates found in only two environments.

Table 4. Candidate genes identified through screening of the significantly associated SNPs and functional annotation.

Chr	Position	Gene ID	PVE	Protein	Function
1	442415	Zm00001d027254	10.41%	Uncharacterized	unknown
1	282946785	Zm00001d034064	8.90%	PAT complex subunit Asterix	protein insertion into ER membrane
2	239027952	Zm00001d007758	4.45%	Overlapping homologous superfamilies	intracellular protein transport
3	50067223	Zm00001d040563	5.54%	Uncharacterized	
4	63379553	Zm00001d050063	2.48%	Zinc finger, RING-type	overlapping homologous superfamilies
4	63306693	Zm00001d050062	6.57%	Uncharacterized	unknown
5	147529611	Zm00001d016156	7.59%	1,3-beta-glucan synthase component FKS1-like, domain-1 InterPro entry	putative callose synthase 8
5	148963157	Zm00001d016183	5.15%	Coatomer, epsilon subunit	retrograde vesicle-mediated transport, Golgi to endoplasmic reticulum
5	147757336	Zm00001d016161	5.08%	CRIB domain	overlapping homologous superfamilies
5	217509599	Zm00001d018257	3.45%	unknown	unknown
		Zm00001d018258		SANT/Myb domain	overlapping homologous superfamilies
		Zm00001d018259		Ubiquitin-like protein Atg12	autophagosome assembly
		Zm00001d018260		Glutamine-Leucine-Glutamine, QLQ	regulation of DNA-templated transcription
6	42410465	Zm00001d035715	7.54%	Transposase, Tnp1/En/Spm-like	unknown
6	46848365	Zm00001d035769	3.40%	Chorismate mutase, AroQ class, eukaryotic type InterPro entry	aromatic amino acid family biosynthetic process
8	70647932	Zm00001d009566	5.31%	FAS1 domain	bacterial immunogenic protein MPT70 (1 FAS1 domain)
		Zm00001d009567		FAS1 domain	bacterial immunogenic protein MPT70 (1 FAS1 domain)
8	77905708	Zm00001d009723	7.54%	SET domain	protein binding
8	83921201	Zm00001d009823	5.01%	Lateral organ boundaries, LOB	unknown
8	93109610	Zm00001d009975	11.71%	PADRE domain	this domain is associated with plant defense upon diverse stress stimulus and has a role in disease resistance to fungus
10	144634116	Zm00001d026376	4.14%	WD40 repeat	protein binding
10	87410402	Zm00001d024778	6.10%	Proteolipid membrane potential modulator	transmembrane transport
10	84908655	Zm00001d024717	10.81%	unknown	unknown

Chr represents chromosome, PVE represents phenotypic variation explained.

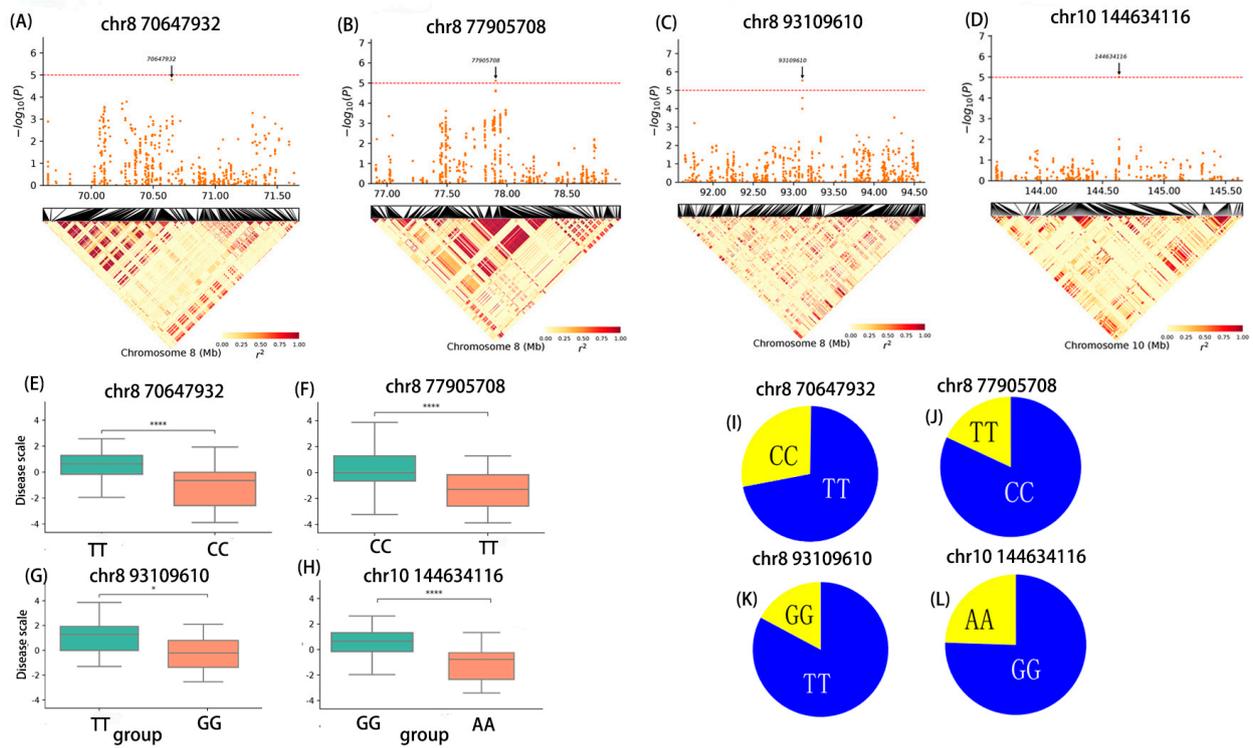


Figure 6. Identification of four SNPs associated with BLSB resistance, including chr8:70647932, chr8:77905708, chr8:93109610, and chr10:144634116. Manhattan plot and LD heatmaps of chr8:70647932 (A), chr8:77905708 (B), chr8:93109610 (C), chr10:144634116 (D). (E–H), respectively, illustrate the difference in the corresponding phenotype between the two groups of RILs. The TT haplotype of the SNP chr8:70647932 is substituted to CC, the CC haplotype of the SNP-chr8: 77905708 is substituted to TT, the TT haplotype of the SNP-chr8:93109610 is substituted to GG, the GG haplotype of the SNP-chr10:144634116 is substituted to AA. Haplotype distribution of SNP-chr8:70647932 (I), SNP-chr8:77905708 (J), SNP-chr8:93109610 (K), SNP-chr10: 144634116 (L) in the multiparent population. * represents significant $p < 0.05$, *** represents highly significant $p < 0.0001$.

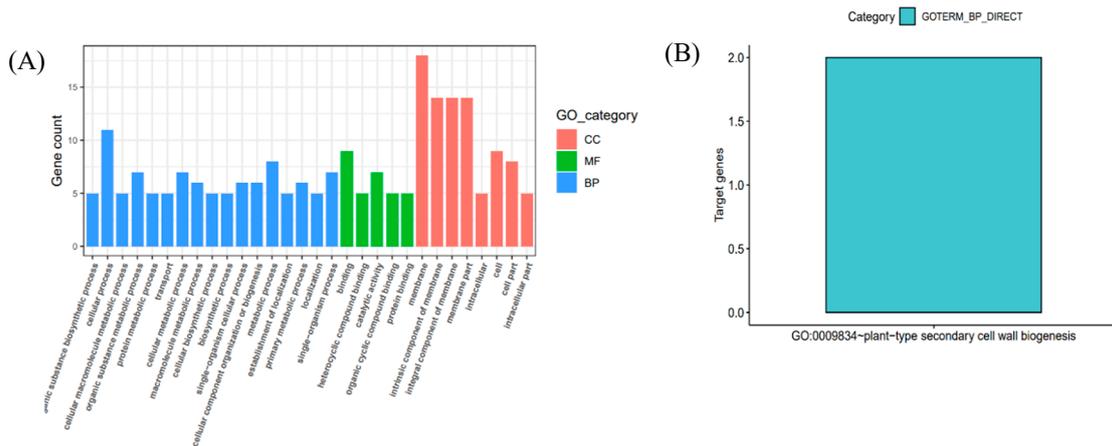


Figure 7. (A) GO annotation of candidate genes. (B) GO enrichment analysis of gene enrichment map with $p < 0.05$.

The haplotypes of the aforementioned SNPs are illustrated in Figure 6. For instance, the SNP at chr8:77905708 is in the intron of the candidate gene *Zm00001d009723*, with a C–T substitution (Figure 6F,J). The SNP at chr8:93109610 is positioned at 18,578 bp downstream of the candidate gene *Zm00001d009975*, with a T–G substitution (Figure 6G,K). The SNP at chr10:144634116 is located in the intron of the candidate gene *Zm00001d026376*, with G–A

substitution (Figure 6H,L). The SNP at chr8:70647932 is situated 13,849 bp downstream of the candidate genes *Zm00001d009566* and *Zm00001d009567*, with a T-G substitution (Figure 6E,I). The base changes, including (C/T), (T/G), (G/A), and (T/C), at the SNP sites lead to significant differences in base compositions among RILs. These indicated that the TT, GG, AA, and CC haplotypes of the SNPs positioned at chr8:77905708, chr8:93109610, chr10:144634116, and at chr8:70647932, respectively, are favorable for reducing BLSB.

3. Discussion

3.1. The Feasibility of GWAS in this Experiment

Genetic studies of complex traits in plants have historically been challenging due to the influence of multiple minor-effect genes, often modulated by diverse genetic backgrounds or environmental conditions. Against this backdrop, genome-wide association study (GWAS) has emerged as a potent tool for dissecting the genetic factors underlying complex quantitative traits. The identification of genes resistant to bacterial leaf and sheath blight of maize (BLSB) proves to be a pertinent focus for GWAS, given its economic impact as a complex maize disease [31]. In this study, we employed 562,212 SNPs to identify 19 potential loci controlling resistance to BLSB in the multiparent population. Over the past years, GWAS has delivered promising results by uncovering novel disease-resistant loci [32–34]. In contrast to traditional bi-parental mapping, this approach is not influenced by the quantity of recombination events and is not constrained by the reduction in the number of allelic variants present in the parents. This method can effectively capture allelic variants underrepresented in the population [35] and accelerate the selection of allelic combinations suitable for the target environment in the donor pool [36]. However, factors such as phenotypic variation, population size, and population substructure can also influence the precision of GWAS [37]. Our results indicate that resistance to BLSB in this population is controlled by polygenes with small to moderate phenotypic effects. Therefore, the authors assert that the GWAS method is highly applicable for identifying molecular loci related to resistance against BLSB in maize and exploring candidate genes.

3.2. Comparison of the Results of the Present Study with the Previous Results

Previous studies have identified several candidate genes associated with resistance to BLSB. Li et al. [18] conducted GWAS and discovered a significant association between *ZmFBL41*, (on chromosome 4), encoding an F-box protein, and resistance to BLSB in maize. Using a natural population of 318 different inbred lines, the authors identified 28 SNPs located on chromosomes 1, 4, 7, and 8. The mechanism involves the interaction of *ZmFBL41* with *ZmCAD*, leading to the degradation of *ZmCAD* and disruption of lignin synthesis. This degradation weakens the plant cell wall, making it more susceptible to fungal colonization, highlighting the crucial role of the cell wall and lignin in plant immunity. The candidate genes identified in this study, *Zm00001d009566* and *Zm00001d009567*, are associated with secondary wall formation in plants and are linked to cell immune defense, indicating the credibility of these two genes.

Cao et al. [1] employed RNA sequencing to analyze the gene expression profiles of maize infected with different virulent strains of BLSB. They identified numerous differentially expressed genes, including those related to defense. Gene ontology (GO) terms related to defense were enriched in both low virulence (LVS) and high virulence (HVS) regulatory genes, such as “response to stimulus”, “defense response”, “cell wall organization”, “response to stress”, “response to biotic stimulus”, and “response to hormones”. The gene *Zm00001d009975* identified in this study exhibits similar responses, participating in defense against various stress stimuli and playing a role in resistance to fungal diseases. Additionally, a set of core genes was discovered, among which the overexpressed genes *zmac41* and *ZmBAK1* enhance maize resistance to BLSB. Chen et al. [38], through meta-analysis, identified 15 “consistency” QTLs controlling resistance to BLSB in maize, distributed on chromosomes 2, 4, 6, 8, 9, and 10. The authors also identified five hotspots (Bin2.02, 2.06, 6.02, 9.0, 9.04) for BLSB resistance. Yang [39] constructed a genetic linkage

map using 125 SSR markers in a backcross (BC1) population (CML270 × 478) × CML270 developed by crossing the resistant maize inbred line CML270 with the susceptible inbred line 478. Through composite interval mapping, they detected three major QTLs for BLSB resistance, two on chromosome 1 and one on chromosome 7, explaining 18% and 20% of the phenotypic variation, respectively. Lin et al. [40] identified 11 BLSB resistance-related QTLs based on disease index, distributed on chromosomes 2, 4, 5, 8, and 9, with additive effects that explained phenotypic variance ranging from 0.0165 to 0.0545 and 20.81% to 7.29%, respectively. Chen et al. [41] analyzed QTLs and genetic effects using composite interval mapping, identifying four BLSB-resistant QTLs on chromosomes 6, 7, and 10. Two QTLs on chromosome 6 explained 12.63% and 0.27% of the phenotypic variance, while one QTL each on chromosomes 7 and 10 explained 15.21% and 5.42%, respectively. Adhikari et al. [42] conducted a population analysis using 76 polymorphic microsatellite markers. Single-marker analysis (SMA) identified markers linked to the BLSB-resistant QTL using linear regression and maximum likelihood analysis. Analysis of variance (ANOVA) revealed one major QTL on chromosome 5 and four minor QTLs on chromosomes 1, 3, 4, and 8. Liu [43], using disease index as the resistance indicator, employed composite interval mapping (LOD > 2) to locate 11 QTLs related to BLSB resistance, distributed on chromosomes 2, 4, 5, 8, and 9. The additive effects ranged from 0.0165 to 0.0545, and the phenotypic variance explained varied between 2.81% and 7.29%.

In this study, the identified genes are located on chromosomes 8 and 10, explaining a phenotypic variance ranging from 5.31% to 11.71%. While these chromosomal locations align with previous research, there are disparities in the interval. Possible explanations include: (1) the use of diverse germplasms or varieties in different studies, leading to significant variations in response to BLSB infection. This study used tropical germplasms, potentially enhancing genetic diversity; (2) substantial environmental impact on BLSB, as observed in the locations of Yanshan County and Jinghong City in Yunnan, classified as high-temperature- and high-humidity-prone areas; (3) variations in sample size and statistical methods across studies may affect GWAS accuracy. This study involved a total of 442 RILs, exceeding the sample sizes used in previous studies [44–48]. These considerations underscore the importance of accounting for genetic diversity, environmental factors, and methodological variations when interpreting GWAS results, highlighting the robustness of this study's approach with its diverse tropical germplasm and larger sample size.

3.3. Functional Annotation of Genes Identified through GWAS

In this study, we compared the positions of SNPs significantly associated with BLSB resistance, consistently identified across four environments, with previous studies. However, no overlapping regions were found. This indicates that the BLSB resistance genes, namely *Zm00001d009723*, *Zm00001d009975*, *Zm00001d026377*, *Zm00001d009566*, and *Zm00001d009567*, identified in this study could be novel candidate genes related to BLSB resistance. These candidate genes were functionally annotated to decipher their potential role in disease resistance.

Zm00001d009723 encodes the SET domain, characterized by a multi-helix structure with irregularly arranged long α -helices and short α -helices. This domain is widely present in various organisms, including plants, yeast, mammals, bacteria, and viruses [49]. In addition to its crucial role in chromatin structure and gene activity regulation, the SET domain may also be associated with plant disease defense as it participates in regulating gene expression and chromatin modification [50].

Zm00001d009975 encodes the PADRE domain, which is associated with plant disease defense, particularly resistant to fungi. The PADRE contains a conserved sequence motif at the N-terminus and an intrinsically disordered region with multiple phosphorylation sites at the C-terminus. Typically found in small single-domain proteins with a conserved sequence motif and an intrinsically disordered region with multiple phosphorylation sites, this domain may play a role in defense against various stress stimuli, maintaining plant health and disease resistance [51].

Zm00001d026376 encodes the WD40 repeat, a short motif of 40 amino acids or fewer, usually terminating with a Trp-Asp (W-D) dipeptide. This motif typically comprises 7–8 blade-like β -propeller folds but may contain 4 to 16 repeat units [52]. WD-repeat proteins are widely distributed in various organisms and participate in diverse functions such as signal transduction, transcription regulation, cell cycle control, and apoptosis. They may act as platforms for protein complex assembly or mediators for protein–protein interactions, potentially involved in plant disease defense, such as the assembly of complexes like G proteins, TAFII transcription factors, and E3 ubiquitin ligases [53,54]. Some proteins containing WD40 repeat sequences may participate in plant processes related to diseases, as observed in *Arabidopsis*, where proteins with WD40 repeat sequences play a crucial regulatory role, including resistance and defense responses against pathogens. These proteins may play essential roles in signaling pathways, initiating the plant’s immune response against diseases.

Zm00001d009566 and *Zm00001d009567* both encode the FAS1 domain, an extracellular module typically containing around 140 amino acid residues, found in plants, animals, and bacteria. This domain features a novel fold composed of seven-chain β -wedges and at least five α -helices, along with conserved elements related to asparagine [55–57]. Proteins known to contain the FAS1 domain include the human TGF- β -induced Ig-H3 protein, volvox’s major cell adhesion protein, *Arabidopsis*’ fasciclin-like arabinogalactan protein, mammalian stabilin protein, human extracellular matrix protein periostin, and bacterial immunogenic protein MPT70, among others. Although there is no detailed description of the disease defense-related aspects of the FAS1 domain, it can be speculated that this domain may play a role in cell adhesion and interactions, possibly related to the plant’s defense mechanisms. Cell adhesion and interactions in plants play a crucial role in disease defense as they can impact pathogen invasion and the plant’s immune response [58–62].

The presence of these genes and domains suggests their potential roles in plant disease defense and resistance mechanisms, involving critical biological processes such as gene regulation, protein interactions, immune responses, and cell adhesion. Further investigation into the functions of these genes and domains will contribute to a better understanding of how plants respond to various pathogens and stresses, ultimately enhancing their disease resistance capabilities.

4. Materials and Methods

4.1. Plant Materials and Population Construction

This study, conducted by our group in 2012, employed the tropical maize inbred lines CML444 and NK40-1, known for their resistance to banded leaf and sheath Blight (BLSB), as female parents. These were crossed with the temperate maize inbred line Ye107, which is susceptible to BLSB. The pedigrees, heterotic groups, and ecotypes of the three parents are detailed in Table 5. The disease scale was derived from field surveys, and the survey methodology followed the guidelines outlined in Table 6. Through a single-seed descent method over eight generations, two recombinant inbred line (RIL) subpopulations were generated: pop1 (CML444 \times Ye107) and pop2 (NK40-1 \times Ye107), each consisting of 221 RILs. All 442 RILs from both subpopulations were planted in Yanshan County, Yunnan Province (23°19′–23°59′ N, 103°35′–104°45′ E) in 2021 and in Yanshan County and Jinghong City (21°27′–22°36′ N, 100°25′–101°31′ E) in 2022. The experimental design followed a Randomized Complete Block Design (RCBD) with three replications per group. Each row consisted of 14 plants, with a row length of 4 m, row spacing of 0.70 m, and inter-plant spacing of 25 cm. Standard agronomical practices were followed to manage the maize experimental sites [63].

Table 5. Parental lines used to develop the multiparent population.

Parent	Pedigree	Heterotic Group	Ecotype	Disease Scale
Ye107	Derived from US hybrid DeKalb XL80	Reid	Temperate	9
CML444	P43C9-1-1-1-1-BBBB-1-1-2-5-1(DH)	nonReid	Tropical	3
NK40-1	Derived from US hybrid	Reid	Tropical	3

DH, doubled haploid.

Table 6. Disease scoring for BLSB.

Scale	Reaction Category	Disease Index	Symptoms
0	Immune (IM)	0	Symptom-free throughout the entire plant
1	Highly Resistant (HR)	0.1~20.0	Disease manifestation on the fourth leaf sheath below the ear and subsequent lower leaf sheaths
3	Resistant (R)	20.1~40.0	Disease manifestation on the third leaf sheath below the ear and subsequent lower leaf sheaths
5	Moderately Resistant (MR)	40.1~60.0	Disease manifestation on the second leaf sheath below the ear and subsequent lower leaf sheaths
7	Susceptible (S)	60.1~80.0	Disease manifestation on the first leaf sheath below the ear and subsequent lower leaf sheaths
9	Highly Susceptible (HS)	80.1~100.0	Disease symptoms manifest in the leaf sheaths above the ear.

4.2. Disease Scoring and Calculation of Disease Index

BLSB screening of RILs of the multiparent populations was conducted in field conditions at two sites, Yanshan County (YS) and Jinghong city (JH), Yunnan Province, China, during the summer of 2021 and 2022 and the winter of 2022. The resistance to BLSB was assessed starting in the 4th week after maize dispersal. The BLSB resistance score was determined for each RIL population based on the percentage of total leaf area infected by BLSB. The screening was performed to assess BLSB resistance levels when the temperature ranged between 20 and 25 °C, and high relative humidity (over 81%) prevailed in the summer in the YS region and in the winter in the JH region of Yunnan province. These conditions created a favorable environment for the growth and spread of *Rhizoctonia solani* [7,8].

The disease severity classification for BLSB followed symptom identification criteria outlined in the ‘Manual of Corn Diseases and Insect Pests’ by Wang Xiaoming and Dai Fachao [64]. Severity assessment involved determining the largest leaf sheath lesion, with a disease score assigned to each plant on a scale of 0 to 9, as detailed in Table 6.

4.3. Phenotyping and Statistical Analysis

The phenotypic data were subjected to descriptive statistical analyses using SPSS (SPSS Statistics 26) and ORGIN (Origin 2022) software. Measures such as mean, minimum, maximum, standard deviation (SD), coefficient of variation (CV), skewness, and kurtosis were calculated. Phenotypic data frequency distribution was conducted using SPSS software. Kurtosis and skewness were employed to assess the normality of the frequency distribution. Pearson’s correlation analysis was performed using the Corrplot function in R. Broad-sense heritability was calculated following the approach outlined by Knapp et al. [65,66].

$$H2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ge}^2/e + \sigma_e^2/re} \times 100\%$$

where σ_g^2 refers to the genetic variance, σ_{ge}^2 refers to the variance of genotype-environment interaction, σ_e^2 refers to residuals, e refers to number of environments (the sum across years and locations), and r refers to the number of replications [65].

Additionally, we employed the lme4 version 1.1–30 [67] R package for calculating BLUP. The calculation formula used was

Multi-locations:

$$Y_{ijkl} = \mu + Line_i + Loc_j + (Line \times Loc)_{ij} + Rep(Loc)_j + \varepsilon_{ijkl}$$

One-location:

$$Y_{ilk} = \mu + Line_i + \varepsilon_{ilk}$$

where Y_{ijkl} , μ , $Line_i$, and Loc_l represent the BLSB phenotype values of each plant, intercept, i th line effects, and l th location effects, respectively. Rep_j represents the j th replication effect, and ε_{ijkl} represents the random effects. $(Line \times Loc)_{ij}$ is used to display the interaction of the i th line at the j th location, and $Rep(Loc)_j$ shows the nested effect of the j th replication within the l th location. The Y_{ilk} is the BLSB phenotype value of each ear at one location, and ε_{ilk} represents the random effects.

The calculated BLUP values, along with the average phenotypic data from YS and JH, were used for the subsequent GWAS.

4.4. DNA Extraction and Genotyping-by-Sequencing (GBS)

The genomic DNA was extracted from seedling leaves of each F8RIL using the cetyl trimethyl ammonium bromide (CTAB) method [68]. Subsequently, the isolated genomic DNA from each F8RIL was digested with the restriction endonucleases PstI and MspI (New England BioLabs, Ipswich, MA, USA), followed by ligation with barcode adapters using the T4 ligase (New England BioLabs). GBS DNA libraries were constructed and sequenced following the GBS protocol [69].

All ligated samples were pooled and purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA). Polymerase chain reaction (PCR) amplification was carried out using primers complementary to both adaptors. Finally, the PCR products were purified and quantified using the Qubit dsDNA HS Assay Kit (Life Technologies, Grand Island, NY, USA). After selecting PCR products of 200–300 bp size using an Egel system (Life Technologies), the library concentration was estimated with a Qubit 2.0 fluorometer and the Qubit dsDNA HS Assay Kit (Life Technologies). Sequencing was performed on an Ion Proton sequencer (Life Technologies, software version 5.10.1) with P1v3 chips. The final reads were generated using TASSEL v5.0 (https://github.com/Euphrasiologist/GBS_V2_Tassel5, (accessed on 6 March 2022)) [70]. Before TASSEL analysis, 80 poly (A) bases were appended to the 3' ends of all sequencing reads. SNPs were generated using the Genome Analysis Toolkit software (GATK-3.8) [71] by aligning with the maize B73 reference genome (B73_V4, ftp://ftp.ensemblgenomes.org/pub/plants/release-37/fasta/zea_mays/DNA, (accessed on 6 March 2022)) [72]. Following the SNP calling in the RILs, the quality of each SNP was assessed based on criteria such as minor allele frequency (MAF), the percentage of missing data points, and linkage disequilibrium. Plink v 1.9 [73] was utilized to filter the SNPs, with the parameters set to $-\text{geno } 0.2$ and $-\text{maf } 0.05$, to exclude loci with deletion rates above 10% and loci with minimum allele frequencies below 5%. In total, 562,212 high-quality SNPs were generated and annotated using the ANNOVAR software tool (v2013-05-20) [74].

4.5. Linkage Disequilibrium (LD), Population Structure, and LD Block Analysis

During population structure analysis, we employed a model-based clustering algorithm in ADMIXTURE Version 1.3 [75]. Preliminary analysis involved multiple runs with consecutive K values ranging from 1 to 12, with a five-fold cross-validation procedure for each K value. The most likely K value was determined using the cross-validation value of ADMIXTURE. Inbred lines with a membership probability greater than 0.5 were assigned to corresponding clusters and visualized using TBtools software v1.098727 [76]. Principal component analyses (PCA) and cluster analyses were performed in R. Linkage Disequilibrium (LD) was evaluated using Pop LD decay 3.40 software (<https://github.com/BGI-shenzhen/PopLDdecay> (accessed on 23 June 2023)) and Perl scripts [77]. LD decay was investigated by plotting pairwise r^2 values against physical

distance in base pairs, with the critical r^2 for LD decay determined by keeping the minimum threshold values at 0.2, indicating the maximum significant physical distance at which LD occurs.

4.6. Genome-Wide Association Study

GWAS was conducted using the efficient mixed-model association (EMME) analysis method in the GEMMA (genome-wide efficient mixed-model association) package [78]. The analysis used the following mixed-model approach:

$$y = Xa + Sb + Km + e;$$

where y represents the phenotype, and a and b are fixed effects denoting marker effects and non-marker effects, respectively. m represents unknown random effects. The incidence matrices for a , b , and m are denoted by X , S , and K , respectively, while e is a vector of random residual effects. To account for population structure, the top three principal components (PCs) were used to construct the S matrix, and the kinship (K) matrix was built using the matrix of simple matching coefficients. The genetic relationship between individuals was modeled as a random effect using the K matrix. A significant p -value threshold of $p < 1 \times 10^{-6}$ was set to control the type 1 error during association analysis.

We used PLINK [73] to calculate independent markers with the parameter -indep-pairwise 50 5 0.2. A significance threshold of $-\log_{10}(p) > 4.5$, calculated using the formula $-\log_{10}(1/\text{SNP numbers})$, was employed to identify significant SNPs associated with maize BLSB resistance. SNP loci meeting or exceeding the threshold were extracted using bedtools v1.7 [79] Based on the B73 v4 reference genome and annotation information, candidate genes associated with BLSB in maize were identified within a 40 kb upstream and downstream region of the significantly associated SNPs. The decision to use a 40 kb region for screening candidate genes was based on a plateau observed in the LD decay plot, where the r^2 values plateaued at 40 kb.

4.7. Gene Predictive Analysis

Based on the estimation of LD decay rate, genes situated within 40 kb upstream and downstream of the significant SNPs underwent functional annotation. Subsequently, relevant gene functional annotation information was retrieved from the maizeGDB database (<http://www.maizegdb.org> (accessed on 8 August 2023)). The determination of physical positions for both genes and SNPs relied on the maize B73 RefGen_V4 reference genome. GO analysis for the identified candidate genes from the GWAS was executed utilizing the DAVID tool (<https://www.DAVID.com/tools> (accessed on 10 August 2023)) [80] and the agrigo tool (<https://www.agrigo.com/tools> (accessed on 10 August 2023)) [81]. Data visualization was conducted using R Studio. This annotation process aided in further validation and identification of candidate genes associated with the target traits acquired through GWAS analysis. The p -values for GO analyses were computed using the following formula [82].

$$P = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i} \times \binom{N-M}{n-i}}{\binom{N}{n}}$$

N : Total number of samples;

M : The number of “specific categories” in the total sample;

n : A random number drawn from the total sample;

k : Take the number of samples that belong to a “specific category”.

4.8. Haplotype Analysis

Haploview v4.2 software was used to analyze the genes consistently detected in multiple environments or those with functions related to BLSB.

5. Conclusions

In this study, two F8 RIL subpopulations were developed by crossing the inbred parent Ye107 with inbred lines CML444 and NK40-1. A GWAS, utilizing 562,212 high-quality SNPs, identified nineteen SNPs significantly associated with BLSB resistance. Subsequently, 23 candidate genes related to BLSB resistance were identified within a 40 kb upstream and downstream region of these SNPs. By employing genome databases such as maizeGBD and NCBI, five novel candidate genes, including *Zm00001d009723*, *Zm00001d009975*, *Zm00001d026376*, *Zm00001d009566*, and *Zm00001d009567*, were validated to confirm their potential roles in BLSB resistance. These genes are implicated in diverse activities encompassing cellular metabolism, transport, transcriptional regulation, and structural proteins. They are involved in various functions, such as protein binding, defense against multiple stress stimuli, resistance to fungal infections, protein transmembrane transport, and enhancement of immune defense. In conclusion, the insights gained from this study could significantly contribute to a deeper comprehension of the regulatory mechanisms underlying BLSB resistance in maize. The identified SNPs and candidate genes hold promise for aiding breeders in selecting maize varieties with enhanced BLSB resistance.

Author Contributions: Conceptualization, X.F.; data curation, J.L., M.L., X.Z. and L.L.; writing—original draft preparation, S.L.; writing—review and editing, F.J., Y.B., X.Y. and R.K.S.; funding acquisition, X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Xingdian Talent Support Program of Yunnan Province, National Natural Science Foundation of China (U2202204), and Seed Industry Joint Laboratory Project of Yunnan Province (202205AR070001).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the reviewers for their comments and the English editor for improving the manuscript's language. We are also grateful to Kuangtian Hui for their help with the figure artwork.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cao, H.; Yang, Z.; Song, S.; Xue, M.; Liang, G.; Li, N. Transcriptome analysis reveals genes potentially related to maize resistance to *Rhizoctonia solani*. *Plant Physiol. Biochem.* **2022**, *193*, 78–89. [[CrossRef](#)]
2. Wu, Q.; Zhang, L.; Xia, H.; Yu, C.; Dou, K.; Li, Y.; Chen, J. Omics for understanding synergistic action of validamycin A and *Trichoderma asperellum* GDFS1009 against maize sheath blight pathogen. *Sci. Rep.* **2017**, *7*, 40140. [[CrossRef](#)]
3. Chen, B.; Han, H.; Hou, J.; Bao, F.; Tan, H.; Lou, X.; Wang, G.; Zhao, F. Control of Maize Sheath Blight and Elicit Induced Systemic Resistance Using *Paenibacillus polymyxa* Strain SF05. *Microorganisms* **2022**, *10*, 1318. [[CrossRef](#)] [[PubMed](#)]
4. Zheng, A.; Lin, R.; Zhang, D.; Qin, P.; Xu, L.; Ai, P.; Ding, L.; Wang, Y.; Chen, Y.; Liu, Y.; et al. The evolution and pathogenic mechanisms of the rice sheath blight pathogen. *Nat. Commun.* **2013**, *4*, 1424. [[CrossRef](#)] [[PubMed](#)]
5. Chen, F. Occurrence and Control of Autumn Corn Sheath Blight. *Plant Prot. Sci.* **1986**, *5*, 27–28.
6. Tan, S.F.; Jiang, C.Y. Investigation on loss rate of corn sheath blight in western Hubei mountainous area. *Plant Prot.* **1988**, *14*, 54.
7. Wekesa, T.B.; Wafula, E.N.; Kavesu, N.; Sangura, R.M. Taxonomical, functional, and cytopathological characterization of *Bacillus* spp. from Lake Magadi, Kenya, against *Rhizoctonia solani* Kühn in *Phaseolus vulgaris* L. *J. Basic Microbiol.* **2023**, *63*, 1293–1304. [[CrossRef](#)] [[PubMed](#)]
8. Qin, X.; Chang, Y.; Wang, Y.; Yang, J.; Nie, S.; Guo, N.; Xu, J.; Cai, F.; Zhao, C.; Fu, Y. *Aspergillus* sp. R3, a new producer for cyclopyrazon acid, inhibits rice sheath blight fungus *Rhizoctonia solani* Kühn. *Physiol. Mol. Plant Pathol.* **2023**, *125*, 102007. [[CrossRef](#)]
9. Zhao, M.J.; Gao, S.B.; Zhang, Z.M.; Rong, T.Z.; Pan, G.T. Initial identification of quantitative trait loci controlling resistance to banded leaf and sheath blight at elongating and heading date in maize. *Fen Zi Xi Bao Sheng Wu Xue Bao* **2006**, *39*, 139–144. [[PubMed](#)]
10. Riedelsheimer, C.; Lisec, J.; Czedik-Eysenberg, A.; Sulpice, R.; Flis, A.; Grieder, C.; Altmann, T.; Stitt, M.; Willmitzer, L.; Melchinger, A.E. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8872–8877. [[CrossRef](#)]

11. Mullins, N.; Kang, J.; Campos, A.I.; Coleman, J.; Edwards, A.C.; Galfalvy, H.; Levey, D.F.; Lori, A.; Shabalin, A.; Starnawska, A.; et al. Dissecting the Shared Genetic Architecture of Suicide Attempt, Psychiatric Disorders, and Known Risk Factors. *Biol. Psychiatry* **2022**, *91*, 313–327. [[CrossRef](#)] [[PubMed](#)]
12. Jiang, S.; Zhang, H.; Ni, P.; Yu, S.; Dong, H.; Zhang, A.; Cao, H.; Zhang, L.; Ruan, Y.; Cui, Z. Genome-Wide Association Study Dissects the Genetic Architecture of Maize Husk Tightness. *Front. Plant Sci.* **2020**, *11*, 861. [[CrossRef](#)] [[PubMed](#)]
13. Carlson, C.H.; Gouker, F.E.; Crowell, C.R.; Evans, L.; Difazio, S.P.; Smart, C.D.; Smart, L.B. Joint linkage and association mapping of complex traits in shrub willow (*Salix purpurea* L.). *Ann. Bot.* **2019**, *124*, 701–716. [[CrossRef](#)] [[PubMed](#)]
14. Mammadov, J.; Sun, X.; Gao, Y.; Ochsenfeld, C.; Bakker, E.; Ren, R.; Flora, J.; Wang, X.; Kumpatla, S.; Meyer, D.; et al. Combining powers of linkage and association mapping for precise dissection of QTL controlling resistance to gray leaf spot disease in maize (*Zea mays* L.). *BMC Genom.* **2015**, *16*, 916. [[CrossRef](#)] [[PubMed](#)]
15. Wang, B.; Lin, Z.; Li, X.; Zhao, Y.; Zhao, B.; Wu, G.; Ma, X.; Wang, H.; Xie, Y.; Li, Q.; et al. Genome-wide selection and genetic improvement during modern maize breeding. *Nat. Genet.* **2020**, *52*, 565–571. [[CrossRef](#)] [[PubMed](#)]
16. Zhao, M.; Liu, S.; Pei, Y.; Jiang, X.; Jaqueth, J.S.; Li, B.; Han, J.; Jeffers, D.; Wang, J.; Song, X. Identification of genetic loci associated with rough dwarf disease resistance in maize by integrating GWAS and linkage mapping. *Plant Sci.* **2022**, *315*, 111100. [[CrossRef](#)] [[PubMed](#)]
17. Li, S.; Zhang, C.; Yang, D.; Lu, M.; Qian, Y.; Jin, F.; Liu, X.; Wang, Y.; Liu, W.; Li, X. Detection of QTNs for kernel moisture concentration and kernel dehydration rate before physiological maturity in maize using multi-locus GWAS. *Sci. Rep.* **2021**, *11*, 1764. [[CrossRef](#)] [[PubMed](#)]
18. Li, N.; Lin, B.; Wang, H.; Li, X.; Yang, F.; Ding, X.; Yan, J.; Chu, Z. Natural variation in ZmFBL41 confers banded leaf and sheath blight resistance in maize. *Nat. Genet.* **2019**, *51*, 1540–1548. [[CrossRef](#)]
19. Yi, G.; Shen, M.; Yuan, J.; Sun, C.; Duan, Z.; Qu, L.; Dou, T.; Ma, M.; Lu, J.; Guo, J.; et al. Genome-wide association study dissects genetic architecture underlying longitudinal egg weights in chickens. *BMC Genom.* **2015**, *16*, 746. [[CrossRef](#)]
20. Shi, X.; Zhou, Z.; Li, W.; Qin, M.; Yang, P.; Hou, J.; Huang, F.; Lei, Z.; Wu, Z.; Wang, J. Genome-wide association study reveals the genetic architecture for calcium accumulation in grains of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biol.* **2022**, *22*, 229. [[CrossRef](#)]
21. Li, H.; Peng, Z.; Yang, X.; Wang, W.; Fu, J.; Wang, J.; Han, Y.; Chai, Y.; Guo, T.; Yang, N.; et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* **2013**, *45*, 43–50. [[CrossRef](#)]
22. Liu, M.; Zhang, M.; Yu, S.; Li, X.; Zhang, A.; Cui, Z.; Dong, X.; Fan, J.; Zhang, L.; Li, C.; et al. A Genome-Wide Association Study Dissects the Genetic Architecture of the Metaxylem Vessel Number in Maize Brace Roots. *Front. Plant Sci.* **2022**, *13*, 847234. [[CrossRef](#)]
23. Zhu, Y.F.; Chen, Z.H.; Ren, H.; Wang, A.Y.; Guo, X.Y. Breeding and Application of New Maize Variety Jinyu 506. *Bull. Agric. Sci. Technol.* **2014**, *9*, 176–178.
24. Fei, J.F.; Sun, Z.; Wang, X.C. Breeding of new maize variety Zhongkeyu 505 and discussion on related commercial breeding problems. *Chin. Agric. Sci. Bull.* **2021**, *37*, 6–13.
25. Du, S.K.; Wang, L.M.; Liu, C.B. Breeding and cultivation techniques of new maize variety Endan 801. *Mod. Agric. Sci. Technol.* **2015**, *1*, 72.
26. Wu, A.; Li, J. A new sweet corn cultivar 'Huawangtian 7'. *Acta Hortic. Sin.* **2018**, *45*, 809–810. [[CrossRef](#)]
27. Pavani, P.; Singh, L.N.; Sinha, B.; Bathula, P. In-vitro Sensitivity Test of Native Trichoderma spp. against Growth of Rhizoctonia solani f.sp. sasakii Causing Banded Leaf and Sheath Blight of Maize in Manipur. *Int. J. Environ. Clim. Chang.* **2022**, *12*, 998–1002. [[CrossRef](#)]
28. Ren, J.; Li, Z.; Wu, P.; Zhang, A.; Liu, Y.; Hu, G.; Cao, S.; Qu, J.; Dhliwayo, T.; Zheng, H.; et al. Genetic Dissection of Quantitative Resistance to Common Rust (*Puccinia sorghi*) in Tropical Maize (*Zea mays* L.) by Combined Genome-Wide Association Study, Linkage Mapping, and Genomic Prediction. *Front. Plant Sci.* **2021**, *12*, 692205. [[CrossRef](#)] [[PubMed](#)]
29. Thirunavukkarasu, N.; Hossain, F.; Shiriga, K.; Mittal, S.; Arora, K.; Rathore, A.; Mohan, S.; Shah, T.; Sharma, R.; Namratha, P.M.; et al. Unraveling the genetic architecture of subtropical maize (*Zea mays* L.) lines to assess their utility in breeding programs. *BMC Genom.* **2013**, *14*, 877. [[CrossRef](#)] [[PubMed](#)]
30. Wu, X.; Li, Y.; Shi, Y.; Song, Y.; Wang, T.; Huang, Y.; Li, Y. Fine genetic characterization of elite maize germplasm using high-throughput SNP genotyping. *Theor. Appl. Genet.* **2014**, *127*, 621–631. [[CrossRef](#)] [[PubMed](#)]
31. Oreiro, E.G.; Grimares, E.K.; Atienza-Grande, G.; Quibod, I.L.; Roman-Reyna, V.; Oliva, R. Genome-Wide Associations and Transcriptional Profiling Reveal ROS Regulation as One Underlying Mechanism of Sheath Blight Resistance in Rice. *Mol. Plant-Microbe Interact.* **2020**, *33*, 212–222. [[CrossRef](#)]
32. Gurung, S.; Mamidi, S.; Bonman, J.M.; Xiong, M.; Brown-Guedira, G.; Adhikari, T.B. Genome-wide association study reveals novel quantitative trait Loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS ONE* **2014**, *9*, e108179. [[CrossRef](#)] [[PubMed](#)]
33. Kump, K.L.; Bradbury, P.J.; Wissler, R.J.; Buckler, E.S.; Belcher, A.R.; Oropeza-Rosas, M.A.; Zwonitzer, J.C.; Kresovich, S.; McMullen, M.D.; Ware, D.; et al. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* **2011**, *43*, 163–168. [[CrossRef](#)] [[PubMed](#)]

34. Wen, Z.; Tan, R.; Yuan, J.; Bales, C.; Du, W.; Zhang, S.; Chilvers, M.I.; Schmidt, C.; Song, Q.; Cregan, P.B.; et al. Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. *BMC Genom.* **2014**, *15*, 809. [[CrossRef](#)] [[PubMed](#)]
35. Weigel, D. Natural variation in Arabidopsis: From molecular genetics to ecological genomics. *Plant Physiol.* **2012**, *158*, 2–22. [[CrossRef](#)] [[PubMed](#)]
36. Zhao, K.; Tung, C.W.; Eizenga, G.C.; Wright, M.H.; Ali, M.L.; Price, A.H.; Norton, G.J.; Islam, M.R.; Reynolds, A.; Mezey, J.; et al. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2011**, *2*, 467. [[CrossRef](#)] [[PubMed](#)]
37. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 29. [[CrossRef](#)] [[PubMed](#)]
38. Chen, W.S. Screening of Maize Sheath Blight Resistance Resources and Meta-Analysis of Resistance QTL. Master's Thesis, Sichuan Agricultural University, Chengdu, China, 2014.
39. Yang, H. Studies on Identification and QTL Mapping and Breeding Utilization of Germplasm Resistant to Banded Leaf and Sheath Blight in Maize. Ph.D. Thesis, Sichuan Agricultural University, Chengdu, China, 2005.
40. Lin, H.J.; Liu, C.L.; Shen, Y.O. QTL analysis of maize sheath blight resistance based on RIL population. *J. Nucl. Agric. Sci.* **2013**, *27*, 895–903.
41. Chen, G.P.; Tan, H.; Zheng, D.B.; Yang, L.T. QTL mapping for sheath blight resistance in maize. *Southwest China J. Agric. Sci.* **2009**, *22*, 950–955. [[CrossRef](#)]
42. Sneha Adhikari, A.J.A.K. Identification of QTL for Banded Leaf and Sheath Blight in Teosinte-Derived Maize Population. *Springerlink* **2022**, *11*, 155–163.
43. Liu, C.L. QTL Mapping for the Resistance to Banded Leaf and Sheath Blight and Related Traits on The RIL Population. Master's Thesis, Sichuan Agricultural University, Chengdu, China, 2011.
44. van der Biezen, E.A.; Jones, J.D. The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* **1998**, *8*, R226–R227. [[CrossRef](#)] [[PubMed](#)]
45. Jamann, T.M.; Balint-Kurti, P.J.; Holland, J.B. QTL mapping using high-throughput sequencing. *Methods Mol. Biol.* **2015**, *1284*, 257–285. [[CrossRef](#)] [[PubMed](#)]
46. Wang, J.; Ma, M.C.; Mennie, A.K.; Pettus, J.M.; Xu, Y.; Lin, L.; Traxler, M.G.; Jakoubek, J.; Atanur, S.S.; Aitman, T.J.; et al. Systems biology with high-throughput sequencing reveals genetic mechanisms underlying the metabolic syndrome in the Lyon hypertensive rat. *Circ. Cardiovasc. Genet.* **2015**, *8*, 316–326. [[CrossRef](#)] [[PubMed](#)]
47. Zhang, K.; Fan, G.; Zhang, X.; Zhao, F.; Wei, W.; Du, G.; Feng, X.; Wang, X.; Wang, F.; Song, G.; et al. Identification of QTLs for 14 Agronomically Important Traits in *Setaria italica* Based on SNPs Generated from High-Throughput Sequencing. *G3-Genes Genomes Genet.* **2017**, *7*, 1587–1594. [[CrossRef](#)] [[PubMed](#)]
48. Kaminski, K.P.; Korup, K.; Andersen, M.N.; Sonderkaer, M.; Andersen, M.S.; Kirk, H.G.; Nielsen, K.L. Cytosolic glutamine synthetase is important for photosynthetic efficiency and water use efficiency in potato as revealed by high-throughput sequencing QTL analysis. *Theor. Appl. Genet.* **2015**, *128*, 2143–2153. [[CrossRef](#)] [[PubMed](#)]
49. Jenuwein, T.; Laible, G.; Dorn, R.; Reuter, G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell. Mol. Life Sci.* **1998**, *54*, 80–93. [[CrossRef](#)]
50. Rea, S.; Eisenhaber, F.; O'Carroll, D.; Strahl, B.D.; Sun, Z.W.; Schmid, M.; Opravil, S.; Mechtler, K.; Ponting, C.P.; Allis, C.D.; et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **2000**, *406*, 593–599. [[CrossRef](#)]
51. Didelon, M.; Khafif, M.; Godiard, L.; Barbacci, A.; Raffaele, S. Patterns of Sequence and Expression Diversification Associate Members of the PADRE Gene Family with Response to Fungal Pathogens. *Front. Genet.* **2020**, *11*, 491. [[CrossRef](#)]
52. Jain, B.P.; Pandey, S. WD40 Repeat Proteins: Signalling Scaffold with Diverse Functions. *Protein J.* **2018**, *37*, 391–406. [[CrossRef](#)]
53. Li, D.; Roberts, R. WD-repeat proteins: Structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci.* **2001**, *58*, 2085–2097. [[CrossRef](#)] [[PubMed](#)]
54. Smith, T.F.; Gaitatzes, C.; Saxena, K.; Neer, E.J. The WD repeat: A common architecture for diverse functions. *Trends Biochem.Sci.* **1999**, *24*, 181–185. [[CrossRef](#)]
55. Clout, N.J.; Tisi, D.; Hohenester, E. Novel fold revealed by the structure of a FAS1 domain pair from the insect cell adhesion molecule fasciclin I. *Structure* **2003**, *11*, 197–203. [[CrossRef](#)]
56. Huber, O.; Sumper, M. Algal-CAMs: Isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila* fasciclin I. *Embo J.* **1994**, *13*, 4212–4222. [[CrossRef](#)]
57. Ulstrup, J.C.; Jeansson, S.; Wiker, H.G.; Harboe, M. Relationship of secretion pattern and MPB70 homology with osteoblast-specific factor 2 to osteitis following *Mycobacterium bovis* BCG vaccination. *Infect. Immun.* **1995**, *63*, 672–675. [[CrossRef](#)]
58. Coutu, D.L.; Wu, J.H.; Monette, A.; Rivard, G.E.; Blostein, M.D.; Galipeau, J. Periostin, a member of a novel family of vitamin K-dependent proteins, is expressed by mesenchymal stromal cells. *J. Biol. Chem.* **2008**, *283*, 17991–18001. [[CrossRef](#)]
59. Faik, A.; Abouzouhair, J.; Sarhan, F. Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (*Triticum aestivum*) and rice (*Oryza sativa*): Identification and bioinformatic analyses. *Mol. Genet. Genomics* **2006**, *276*, 478–494. [[CrossRef](#)]
60. Kim, J.E.; Kim, S.J.; Lee, B.H.; Park, R.W.; Kim, K.S.; Kim, I.S. Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-beta-induced gene, betaig-h3. *J. Biol. Chem.* **2000**, *275*, 30907–30915. [[CrossRef](#)] [[PubMed](#)]

61. Kzhyskowska, J.; Gratchev, A.; Martens, J.H.; Pervushina, O.; Mamidi, S.; Johansson, S.; Schledzewski, K.; Hansen, B.; He, X.; Tang, J.; et al. Stabilin-1 localizes to endosomes and the trans-Golgi network in human macrophages and interacts with GGA adaptors. *J. Leukoc. Biol.* **2004**, *76*, 1151–1161. [[CrossRef](#)]
62. Matsumoto, S.; Matsuo, T.; Ohara, N.; Hotokezaka, H.; Naito, M.; Minami, J.; Yamada, T. Cloning and sequencing of a unique antigen MPT70 from Mycobacterium tuberculosis H37Rv and expression in BCG using *E. coli*-mycobacteria shuttle vector. *Scand. J. Immunol.* **1995**, *41*, 281–287. [[CrossRef](#)] [[PubMed](#)]
63. Wang, Y.; Bi, Y.; Jiang, F.; Shaw, R.K.; Sun, J.; Hu, C.; Guo, R.; Fan, X. Mapping and Functional Analysis of QTL for Kernel Number per Row in Tropical and Temperate-Tropical Introgression Lines of Maize (*Zea mays* L.). *Curr. Issues Mol. Biol.* **2023**, *45*, 4416–4430. [[CrossRef](#)] [[PubMed](#)]
64. Dai, F.C.; Wang, X.M. *Field Manual of Corn Pests and Diseases: Identification of Pests and Diseases and Identification of Resistance*; China Agricultural Science and Technology Press: Beijing, China, 2002.
65. Knapp, S.J.; Stroup, W.W.; Ross, W.M. Exact Confidence Intervals for Heritability on a Progeny Mean Basis. *Crop Sci.* **1985**, *25*, 192–194. [[CrossRef](#)]
66. Fisher, R.A. XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth Environ. Sci. Trans. R. Soc. Edinb.* **1919**, *52*, 399–433. [[CrossRef](#)]
67. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]
68. Stewart, C.J.; Via, L.E. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *Biotechniques* **1993**, *14*, 748–750. [[PubMed](#)]
69. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* **2012**, *7*, e32253. [[CrossRef](#)] [[PubMed](#)]
70. Li, C.; Guan, H.; Jing, X.; Li, Y.; Wang, B.; Li, Y.; Liu, X.; Zhang, D.; Liu, C.; Xie, X.; et al. Genomic insights into historical improvement of heterotic groups during modern hybrid maize breeding. *Nat. Plants* **2022**, *8*, 750–763. [[CrossRef](#)] [[PubMed](#)]
71. Mckenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
72. Jiao, Y.; Peluso, P.; Shi, J.; Liang, T.; Stitzer, M.C.; Wang, B.; Campbell, M.S.; Stein, J.C.; Wei, X.; Chin, C.S.; et al. Improved maize reference genome with single-molecule technologies. *Nature* **2017**, *546*, 524–527. [[CrossRef](#)]
73. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
74. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)]
75. Mussmann, S.M.; Douglas, M.R.; Chafin, T.K.; Douglas, M.E. ADMIXPIPE: Population analyses in ADMIXTURE for non-model organisms. *BMC Bioinform.* **2020**, *21*, 337. [[CrossRef](#)] [[PubMed](#)]
76. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant.* **2020**, *13*, 1194–1202. [[CrossRef](#)] [[PubMed](#)]
77. Zhang, C.; Dong, S.; Xu, J.; He, W.; Yang, T. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **2019**, *35*, 1786–1788. [[CrossRef](#)] [[PubMed](#)]
78. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [[CrossRef](#)] [[PubMed](#)]
79. Strable, J.; Wallace, J.G.; Unger-Wallace, E.; Briggs, S.; Bradbury, P.J.; Buckler, E.S.; Vollbrecht, E. Maize YABBY Genes drooping leaf1 and drooping leaf2 Regulate Plant Architecture. *Plant Cell* **2017**, *29*, 1622–1641. [[CrossRef](#)] [[PubMed](#)]
80. Sherman, B.T.; Hao, M.; Qiu, J.; Jiao, X.; Baseler, M.W.; Lane, H.C.; Imamichi, T.; Chang, W. DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **2022**, *50*, W216–W221. [[CrossRef](#)]
81. Tian, T.; Liu, Y.; Yan, H.; You, Q.; Yi, X.; Du, Z.; Xu, W.; Su, Z. agriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **2017**, *45*, W122–W129. [[CrossRef](#)]
82. Chen, L.; Zhang, Y.H.; Wang, S.; Zhang, Y.; Huang, T.; Cai, Y.D. Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS ONE* **2017**, *12*, e184129. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.