

Article

Integrated Transcriptome and GWAS Analysis to Identify Candidate Genes for *Ustilago maydis* Resistance in Maize

Bingyu Yin ^{1,†}, Linjie Xu ^{1,†}, Jianping Li ¹, Yunxiao Zheng ^{1,2}, Weibin Song ² , Peng Hou ³, Liying Zhu ¹, Xiaoyan Jia ¹, Yongfeng Zhao ¹, Wei Song ^{4,*} and Jinjie Guo ^{1,*}

¹ State Key Laboratory of North China Crop Improvement and Regulation, North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Hebei Sub-Center of National Maize Improvement Center of China, College of Agronomy, Hebei Agricultural University, Baoding 071000, China; yinbibgyu139@163.com (B.Y.); 15230867602@163.com (L.X.); lijianping1161997@163.com (J.L.); zyx406939067@163.com (Y.Z.); zhuliyang73@163.com (L.Z.); xiaoyanjia@163.com (X.J.); 13582266897@126.com (Y.Z.)

² State Key Laboratory of Maize Bio-Breeding, National Maize Improvement Center, Department of Plant Genetics and Breeding, China Agricultural University, Beijing 100193, China; songwb@cau.edu.cn

³ Key Laboratory of Crop Physiology and Ecology, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Ministry of Agriculture and Rural Affairs, Beijing 100081, China; houpeng@caas.cn

⁴ Key Laboratory of Crop Genetics and Breeding of Hebei Province, Institute of Cereal and Oil Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang 050051, China

* Correspondence: sw1717@126.com (W.S.); guojinjie512@163.com (J.G.)

† These authors contributed equally to this work.

Abstract: Maize *Ustilago maydis* is a disease that severely affects maize yield and quality. In this paper, we employed transcriptome sequencing and GWAS analysis to identify candidate genes and reveal disease-resistant germplasm resources, thereby laying the foundation for further analysis of the molecular mechanism of maize *Ustilago maydis* resistance and genetic improvement. The results of transcriptome sequencing revealed that a considerable number of receptor kinase genes, signal-transduction-related protein genes, redox-response-related genes, WRKYs, and P450s genes were significantly upregulated. There was a wide range of mutations of *Ustilago maydis* in maize inbred lines. Thirty-two high-resistance maize inbred lines were selected, and 16 SNPs were significantly associated with the disease index. By integrating the results of GWAS and RNA-seq, five genes related to disease resistance were identified, encoding the chitinase 1 protein, fatty acid elongase (FAE), IAA9, GATA TF8, and EREB94, respectively. It provides a certain reference for the cloning of maize anti-tumor smut genes and the breeding of new varieties.

Keywords: maize; *Ustilago maydis*; transcriptome sequencing; genome-wide association analysis; candidate genes



Citation: Yin, B.; Xu, L.; Li, J.; Zheng, Y.; Song, W.; Hou, P.; Zhu, L.; Jia, X.; Zhao, Y.; Song, W.; et al. Integrated Transcriptome and GWAS Analysis to Identify Candidate Genes for *Ustilago maydis* Resistance in Maize.

Agriculture **2024**, *14*, 958. <https://doi.org/10.3390/agriculture14060958>

Academic Editor: Thomas Thomidis

Received: 15 May 2024

Revised: 6 June 2024

Accepted: 12 June 2024

Published: 19 June 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/>).

1. Introduction

Maize (*Zea mays* L.) *Ustilago maydis* occurs in over 100 countries and regions worldwide. It is estimated that approximately 10% of the annual agricultural yield loss is caused by plant fungal diseases [1]. Given a global production of approximately 840 million tons of maize per year, the yield loss due to maize *Ustilago maydis* is estimated to range from 2 to 20%, which is equivalent to the food loss of 26 million to 262 million people [2]. Maize is one of the most significant crops in China, with the seedling-development stage representing a crucial period for the entire growth cycle. During this period, the plant is particularly susceptible to pathogen attack, which not only impedes the normal growth and development of the organ but also may result in a significant reduction in survival rate, which in turn affects the final yield. Infestation by *Ustilago maydis* is capable of causing diseases in different parts of the plant, seriously threatening the yield and quality of maize.

RNA sequencing (RNA-seq) is a technique that involves the collection of all RNAs transcribed from a specific tissue or cell at a particular stage of development or functional state. Some studies have compared the transcriptomes of resistant and susceptible lines in an early response to maize-fusarium verticillioides (*Fusarium verticillioides*) and identified numerous transcripts enriched in plant immune-response-related pathways, including phytohormone signaling, phenylpropanoid biosynthesis, and cytochrome P450 metabolism, which were found to be differentially expressed [3,4]. In a study of resistance, *Fusarium graminearum* was found to increase resistance to *Ustilago maydis* in maize by activating induced systemic resistance (ISR) in maize. RNA-seq analysis demonstrated that phenylpropanoid biosynthesis, amino acid metabolism, and plant-pathogen interaction pathways were enriched in the roots. The application of gene set enrichment analysis in GWAS has been successful in revealing multiple clusters of genes that act together for resistance traits. For instance, in a study of maize kernels, GWAS revealed four key loci and 16 candidate genes that confer resistance to aflatoxin disease in kernels [5]. Furthermore, another GWAS study identified 14 single nucleotide polymorphisms (SNPs) associated with maize big blotch resistance, which are closely related to resistance to the disease caused by *Aspergillus oryzae* [6]. GWAS has also demonstrated its utility in revealing genomic loci and the allelic variants controlling lethal necrosis disease resistance in maize [7].

In this paper, gene-expression profiling of the maize B73 inbred line cultivar after *Ustilago maydis* inoculation was performed using transcriptome sequencing to identify differentially expressed genes (DEGs) that are significant in the early response to pathogen infestation. Additionally, a GWAS of 167 maize inbred line samples, including the maize B73 inbred line, was conducted to identify the genetic markers significantly associated with disease resistance. By comparing the DEGs identified through transcriptomics analysis with the candidate genes analyzed by GWAS, an attempt was made to identify disease-resistance genes that are highly correlated with resistance to *Ustilago maydis* disease with high confidence, as well as to screen for disease-resistant germplasm resources. This provides a certain research basis and reference for the molecular breeding and genetic improvement of maize resistance to *Ustilago maydis* disease.

2. Materials and Methods

2.1. Maize Materials and Test Strains

The maize inbred line B73, provided by Hebei Agricultural University, was utilized for high-throughput transcriptome sequencing analysis. The 167 maize inbred lines employed for screening and GWAS analysis of *Ustilago maydis* resistance materials encompass a broad genetic base, including numerous exemplary backbone inbred lines utilized in breeding. This establishes a foundation for germplasm resources for selecting and breeding excellent disease-resistant materials and for fully exploring new genetic loci (Supplementary Table S1).

Ustilago maydis strain SG200 is a genetically engineered pathogenic virulent strain that is capable of forming infective mycelium without prior haploid coordination of different genotypes [8]. Additionally, the pathogen causes disease in maize seedlings under room-temperature conditions and can form tumors on leaves in less than two weeks after infection [9,10].

2.2. Corn Planting and Inoculation with *Ustilago maydis* SG200

The 167 maize inbred lines, with 128 plants each, in three replicates, were grown simultaneously in the greenhouse and cultured to the three-leaf and one-heart stages. Meanwhile, the fungus solution was cultured, with the fungus solution inoculated on that replicate in batches, and classified according to the degree of fungus infestation on the 8th day to classify the disease resistance, and to collect the phenotypes of the 167 maize inbred lines needed for the calculation of the GWAS.

A suitable quantity of corn seeds should be buried in a 32-hole planting tray containing a mixture of nutrient soil and vermiculite. After planting, a sufficient amount of nutrient

solution should be applied. The seeds were placed in a culture room at 25–28 °C with a photoperiod of 14 h light/10 h dark and 70% relative humidity and watered regularly to ensure normal growth and development.

The seeding infections are carried out by syringe with 300–500 µL of the inoculum cell suspension into the interior of the leaf whorl. The injection site chosen is approx. 1 cm above the soil, which is about 2.5 to 3 cm above the basal plant meristem and is the juvenile stem. The leaf sheaths of the first and second leaf and the leaf blades of the third and fourth leaf, which are immersed into the whorl, are pierced by the syringe halfway onto the center of the stem cylinder, which later shows an infection mark after symptom development. Once the inoculum is seen on the inner whorl of the leaves, the seedling is known to be successfully infected. Plants are kept in controlled growth conditions at 28/22 °C.

The primary culture of pathogenic bacteria, namely the activated *Ustilago maydis*, was inoculated on a PDA (PDA is a common microbiology basal medium made from potato extract and glucose for the growth of fungi) culture plate prepared in advance and placed in a constant-temperature incubator at 28 °C for approximately a week of dark culture. This was conducted until the mycelium grew all over the plate. The plate was then placed in the refrigerator at 4 °C to be saved for future use. The recultivation of pathogenic bacteria involves the selection of one or two marginal clumps of newly cultured *Ustilago maydis* colonies and their inoculation into 4 mL of YEPSL (YEP medium is the standard medium for the maintenance and propagation of *Schizosaccharomyces cerevisiae* and *Saccharomyces cerevisiae*, and the main components are yeast extract powder and peptone. YEPS medium is YEP medium with sorbitol added, which can be abbreviated as YPS medium.) liquid medium. These are then placed into a constant temperature oscillator (180 r/min, 28 °C) for approximately 20 h of cultivation. The preparation of the inoculation solution was a process where, one day prior to inoculation, 50 µL of the cultured bacterial solution from the previous step should be taken and added to 50 mL of YEPSL liquid medium. This mixture should then be incubated for approximately 10 h at 28 °C with shaking and expansion.

The cultured SG200 bacterial solution was centrifuged at 1500 × g for 10 min. The supernatant was poured off, and the pellet was eluted with sterile distilled water. The solution was then diluted with sterilized ultrapure water until the OD600 was approximately 1.0. The inoculation was conducted by injecting the maize seedling, which had been cultivated for eight days, with a 1 mL sterile syringe with the needle removed. The time of inoculation was chosen to be in the morning, and the same batch of the identified materials should be completed on the same day. The phenotypes were observed at 0, 2, 4, 6, and 8 days post-inoculation, respectively, and the onset of disease was recorded on the 8th day. Three biological replicates were set for each time period of treatment.

2.3. RNA Extraction and Real-Time Fluorescence Quantitative qPCR Assay

Maize B73 seedlings at the three-leaf stage, grown in a light culture room for eight days, were sampled at 0 d and 2 d after the treatment with *Ustilago maydis* and wrapped in tin foil labeled with the treatment time. The samples were then rapidly frozen in liquid nitrogen for 10 s and ground, and RNA was extracted using the Rapid General Purpose Plant RNA Extraction Kit of Beijing Huayuoyueyang Company (Beijing, China). The detailed steps were referenced in the instruction manual of the extraction kit. The experiment was conducted in triplicate. The extracted RNA was analyzed by a NanoDrop microspectrophotometer to determine its concentration and purity. The A260/280 ratio was approximately 2.0, indicating that the extracted RNA was of high purity. The extracted RNA was stored in an ultra-low temperature refrigerator at −80 °C for backup.

The extracted RNA was reverse transcribed into cDNA in accordance with the instructions provided in the Takara (Beijing, China) Reverse Transcription Kit, and then, qPCR was performed to detect the gene-expression level. The primer information, reaction setup, and routine PCR conditions for qRT-PCR are shown (Supplementary Table S2). The qZmActin gene was used as an internal reference for fluorescence quantification, and the experiment was carried out in three biological replicates. The relative expression of the genes was

calculated according to the $2^{-\Delta\Delta CT}$ method and analyzed for the significance of differences using Student’s *t*-test (** $p < 0.001$).

2.4. Transcriptome Sequencing

Data were analyzed using Microsoft Excel 2019. Three biological replicates were performed. qPCR detected significantly up-regulated expression of SA signaling-pathway-tagged genes, and then, the same batch of RNA samples were sent to Beijing Group Biotechnology Co., Ltd. (Beijing, China) for transcriptome sequencing analysis. TBtools software (version v1.09876) was used to analyze the expression profiles of the differentially expressed genes.

2.5. Identification and Functional Annotation of Differentially Expressed Genes

In order to identify the genes related to maize disease resistance, genes were screened for differentially expressed genes (DEGs) with a differential gene-expression multiplicity of 1 fold or more ($|\log_2FC| > 1$) and $FDR < 0.05$ [11]. To further characterize the functions of DEGs, the GO and KEGG databases were used to annotate and functionally analyze the DEGs of maize B73.

2.6. Phenotypic Characterization and Phenotypic Data Processing of Maize Self-Inoculation with *Ustilago maydis*

2.6.1. Seedling Disease Grading Criteria

Referring to Jörg Kämper, Krishna Mohan Pathi, and Ruan’s method and improving it [12–14], symptom scoring was conducted on the 8th day after infection. Symptoms were classified into four categories based on symptom severity (Table 1), and the number of maize plants in different susceptibility classes was counted.

Table 1. Classification of symptoms in infected maize seedlings at seedling stage.

Level of Illness Condition Level	Symptomatic Behavior Performance Symptoms
1	The plants show no signs of infection.
2	Mild infestation with greenish discoloration of infected leaves at the inoculation site
3	Moderate infestation, with small rice-like verrucae on leaves, leaf sheaths, etc.
4	Heavy infestation with severe deformation of leaves, stems, or stem bases with more and larger protruding verrucae

2.6.2. Seedling-Resistance Evaluation Criteria

The disease index was used as the raw data for GWAS in this study. The seedling-resistance criteria were determined by Shi Jing et al. in the maize verticillium black powder disease-resistance study (Table 2) [15].

Table 2. Evaluation criteria of seeding stage resistance to corn smut.

Sickness Index Condition Index	Capability of Resisting Condition Index	
0–15.0	Highly resistant	(high anti-HR)
15.1–30.0	Resistant	(Anti-R)
30.1–50.0	Moderately resistant	(CAGMR)
50.1–70.0	Susceptible	(Sense S)
70.1–100.0	Highly susceptible	(High-sensitivity HS)

Disease index = $100 \times \sum (\text{number of diseased plants at each level} \times \text{representative value of each level}) / (\text{total number of plants surveyed} \times \text{representative value of the highest level})$.

The phenotypic data obtained were processed and analyzed using the statistical software Microsoft Office Excel 2019 to evaluate 167 maize inbred lines for resistance to downy mildew at the seedling stage.

2.7. Genotyping and Analysis

2.7.1. Genotype Analysis

The genomic DNA of maize was extracted using the CTAB method [16]; DNA samples with qualified testing quality were subjected to double-ended 100 bp sequencing using the Illumina standard process. The sequencing data were quality-controlled and compared with the maize reference genome (B73RefGen_V4) to obtain the raw genotype data. The quality-control screening of SNP markers was performed by PLINK software (version 1.9b), and the screening criteria were that the minimum allele frequency (MAF) was greater than 0.05, the deletion rate was less than 20%, and 368,219 high-quality SNP markers were obtained after screening. SNP markers dispersed to each chromosome after quality control screening were used in the subsequent after-quality-control screening. The SNP markers were dispersed on each chromosome and used for subsequent population-structure analysis and association analysis.

2.7.2. Population-Structure Analysis

On the basis of the screened SNPs, ADMIXTURE (version: 1.3.0; parameters: default parameters) software analyzed the population structure of the original genotypes, predicted the optimal subpopulation of the population, set the number of populations k from 1 to 10, and extracted the cross-validation error rate (CV error) for different k values from the result files. The value of K was determined based on the minimum cross-validation error rate, and this value of K was the optimal number of subpopulations. Based on the modeled coefficient of the number of subpopulations, K , 167 autologous lines were divided into associated populations.

2.8. Linkage-Disequilibrium Analysis

Linkage disequilibrium (LD) is the determination of marker density required for linkage analysis by LD decay distance. In some cases, genes do not segregate completely independently of each other, especially when two loci of a linkage coexist on the same chromosome. When assessing the degree of linkage disequilibrium between two single nucleotide polymorphisms (SNPs), the r^2 value is commonly used as a quantitative indicator. When the r^2 value is lower than 0.2, it indicates that the two SNPs are co-inherited and may be co-localized on the same quantitative trait locus (QTL) [17]. When the r^2 value reaches 0, it indicates that the two loci are in complete linkage equilibrium; relatively, when the r^2 value reaches 1, it indicates that the two loci are in complete linkage disequilibrium. The r^2 value changes between 0 and 1 reflect the degree of LD. The higher the value is, the stronger the LD between the loci is, and the closer the linkage is accordingly. In this study, the r^2 value was mainly used as an important reference for determining the search range of candidate genes and a quantitative index for evaluating the strength of inter-locus linkage. By accurately calculating the r^2 value, researchers were able to effectively identify and focus on specific gene regions with high linkage disequilibrium [18], thus providing a solid foundation for subsequent gene-function analysis and trait association studies.

2.9. Genome-Wide Association Analysis

The genome of 167 self-crosses was analyzed using SNP markers and six modeling approaches, including BLINK, FarmCPU, GLM, MLM, MLMM, and SUPER, which are available in the GAPIT software (version 3.1). Additionally, mrMLM and FASTmrML were employed. The M, FASTmrEMMA, ISIS EM-BLASSO, and pLARmEB models of the mrMLM software (version 5.01) were employed, while the pKWmEB model was used to analyze genome-wide associations between disease indices and *Ustilago maydis* infection in maize seedlings. The results of the GAPIT software analysis were employed to calculate the

effective number of independent markers (N) using the GEC software tool. The suggested $(1/N)$ p -value threshold was set to control the genome-wide type-1 error rate, in accordance with the recommendations set forth. Significantly associated SNP loci were identified using a significance threshold of 4.89 ($p = 1/N$). All parameters of the analysis process of mrMLM software were set as default values [19]. Significantly associated loci that co-localized by two or more models were further screened.

2.10. Candidate Gene Identification and Functional Annotation

For the detected loci significantly associated with maize downy mildew resistance, the annotation information was found through the maize B73 inbred line RefGen_V4, and the corresponding biological information was found on maizeGDB (<https://www.maizegdb.org/>) and NCBI (<https://www.ncbi.nlm.nih.gov/>). Based on the biological information, we screened the candidate genes related to the resistance to powdery mildew.

3. Results

3.1. Pathogenesis of Maize B73 Inbred Line Inoculated with *Ustilago maydis* SG200 and Detection of Induced ZmPRs

Ustilago maydis is a heterozygous mating fungus. The infestation cycle of *Ustilago maydis* in maize is initiated by spore recognition and the fusion of spores with affinity mating phenotypes, which leads to a morphological transition from yeast-like haploid cells to diploid filamentous bodies. The *Ustilago maydis* strain used in this study is a genetically modified pathogenic strain, SG200, which is capable of forming infective hyphae without prior a haploid mating of different genotypes. It is also capable of pathogenizing maize plantlets under room-temperature conditions, which can lead to the formation of tumors on leaves less than two weeks after infection [13]. The result was that maize B73, inoculated with SG200, exhibited no visible symptoms on day 2, displayed visible greenish symptoms from day 4, and formed a verruca-like structure in the leaf by day 8 (Figure 1A). The absence of visible symptoms in the leaves is considered a critical period for the infestation and growth of *Ustilago maydis* in maize, as well as for the identification of genes related to signaling and hormones during the early stages of maize induction by *Ustilago maydis*. Indeed, the subsequent qPCR results demonstrated that the downstream genes of the defense response, including the disease-course-related protein genes ZmPR3, ZmPR4, and ZmPR5, were significantly upregulated (Figure 1B–D). Therefore, RNA samples from maize B73 on days 0 and 2 after inoculation with SG200 were selected for transcriptome sequencing analysis.

3.2. Expression Analysis of SG200-Induced Early Differential Genes in RNA-seq Screening Maize B73 Inbred Line

In order to systematically screen the regulatory network of SG200-induced early genes in B73 at the genome-wide level, transcriptome sequencing was performed to analyze the samples of B73 2 days after SG200 infection. RNA-seq technology was employed to analyze the transcriptome data and clustered heatmaps (Figure 2A), and volcano plots (Figure 2B) were constructed to visualize the dynamic changes in gene expression. The analysis revealed that a total of 1876 genes exhibited significant upregulation in expression in the 2d_vs_0d comparison, while 1251 genes exhibited significant downregulation (fold-change ≥ 1 , FDR ≤ 0.05 ; Figure 2B). These findings indicate that the SG200 treatment was able to rapidly activate the expression of a large number of genes in the early response of the maize B73 inbred line. The identification of these differentially expressed genes provides valuable information to unravel key gene networks during the interaction of maize with *Ustilago maydis* and lays the foundation for further functional studies.

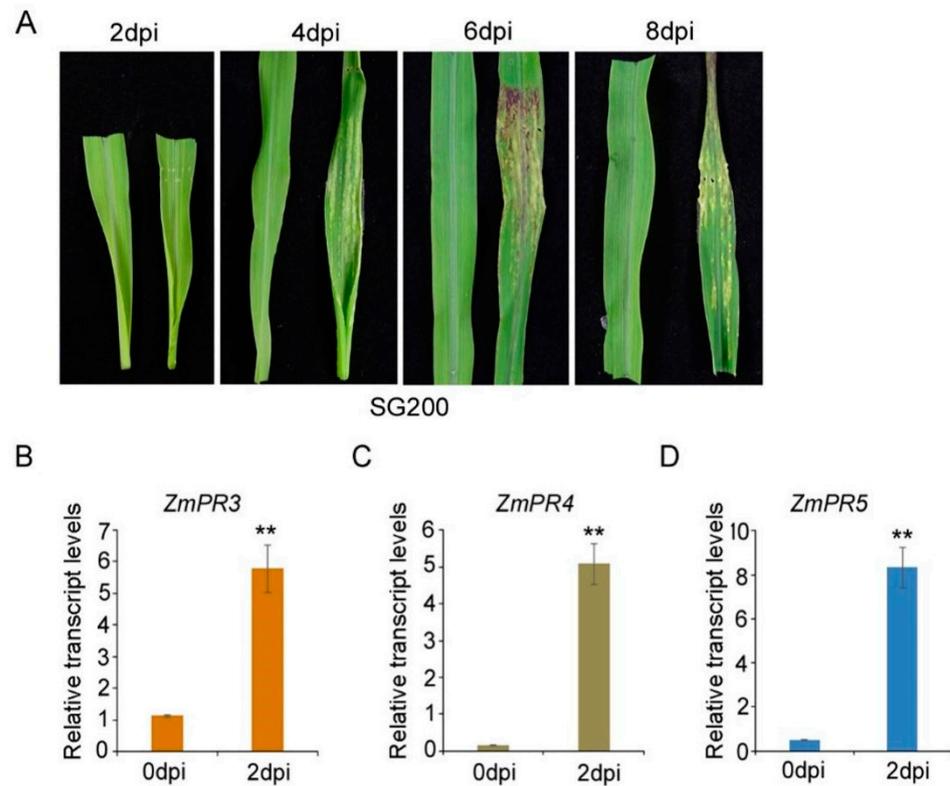


Figure 1. Disease symptoms of B73 after infection with *U. maydis* SG200 and qPCR analysis of *ZmPRs* expression. (A) Disease phenotypes of Maize B73 infected with *U. maydis* SG200. (B–D) All expression levels were normalized to *ZmACTIN*. This experiment was repeated three times with similar results. *p*-values were calculated by Student’s *t*-test (** *p* < 0.001).

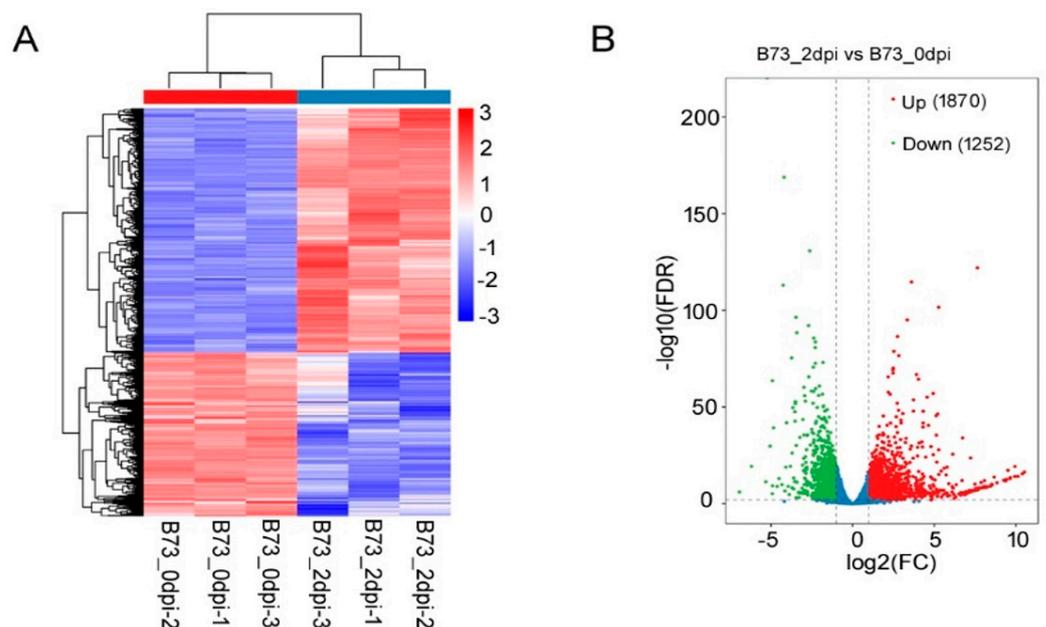


Figure 2. Transcriptome analysis of B73 after infection with *U. maydis* SG200 (A,B). Cluster thermographic and volcano plot analysis of *U. maydis* SG200-induced differentially expressed genes. The dashed line represents significance, and those below the dashed line are also significantly differentially expressed, and those above the dashed line are also significantly differentially expressed.

3.3. GO and KEGG Functional Enrichment Analysis of DEGs

In order to investigate the biological functions of key genes induced by SG200 early in B73, the identified DEGs were subjected to GO function enrichment analysis. GO was mainly divided into three categories: molecular function, biological process, and cellular component. The biological process category encompasses the following processes: defense response to a bacterium, hydrogen peroxide catabolic process, response to stimulus, response to oxidative stress, and the cellular component category (Figure 3). The biological functions of the identified DEGs were investigated using GO function enrichment analysis. GO was divided into three main categories: molecular function, biological process, and cellular component. The biological process category included the following functions: response to stimulus, response to oxidative stress, carbohydrate metabolic process, etc. The cellular component category included the following functions: membrane, extracellular region, microtubule, and Golgi membrane. The molecular function category included the following functions: oxidoreductase activity, transferase activity, transferring acyl groups other than amino-acyl groups, transmembrane transporter activity, and transmembrane metabolic process. The groups include the following: transmembrane transporter activity, oxidoreductase activity, peroxidase activity, etc. The results of these DEGs indicated that, in the early stage of SG200 infestation in maize, SG200 induced a significant increase in the number of transmembrane transporters. In the early stages of SG200 infestation in maize, a significant number of disease-resistance genes were induced. It was postulated that a common set of DEGs might constitute a significant portion of the immune transcriptional reprogramming mechanism.

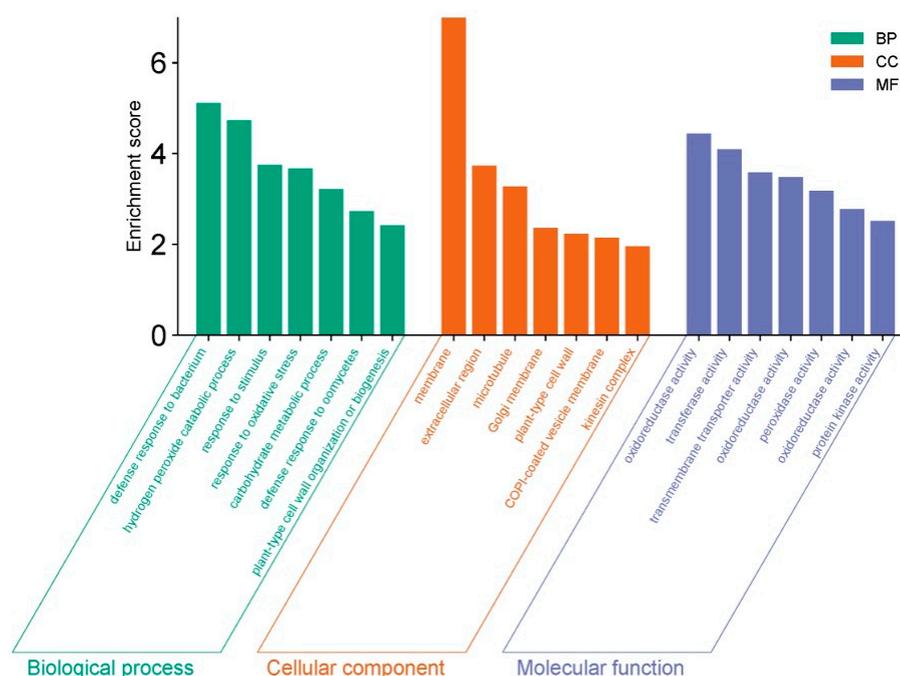


Figure 3. DEGs GO pathway diagram of corn B73 2d_vs_0d treated by SG200.

KEGG analysis further elucidated the biological functions of DEGs. The analysis revealed that 720 differentially expressed genes were significantly enriched in 10 pathways, which were mainly involved in the biosynthesis of secondary metabolites, metabolic pathways, biosynthesis of amino acids, stilbenoid, diarylheptane, and curcumin. The biosynthesis of amino acids; stilbenoid, diarylheptanoid, and gingerol biosynthesis; cysteine and methionine metabolism; keratin and other metabolites; cutin, suberine, and wax biosynthesis; linoleic acid metabolism, biosynthesis of various plant secondary metabolites; and the biosynthesis of various plant metabolites were identified. The biosynthesis of various plant secondary metabolites, phenylpropanoid biosynthesis, terpene alkaloids,

piperidine, and pyridine alkaloid biosynthesis were also observed. The unigenes involved in metabolic pathways are the most numerous, with 325, followed by the biosynthetic pathways of secondary metabolites, with 230 (Figure 4).

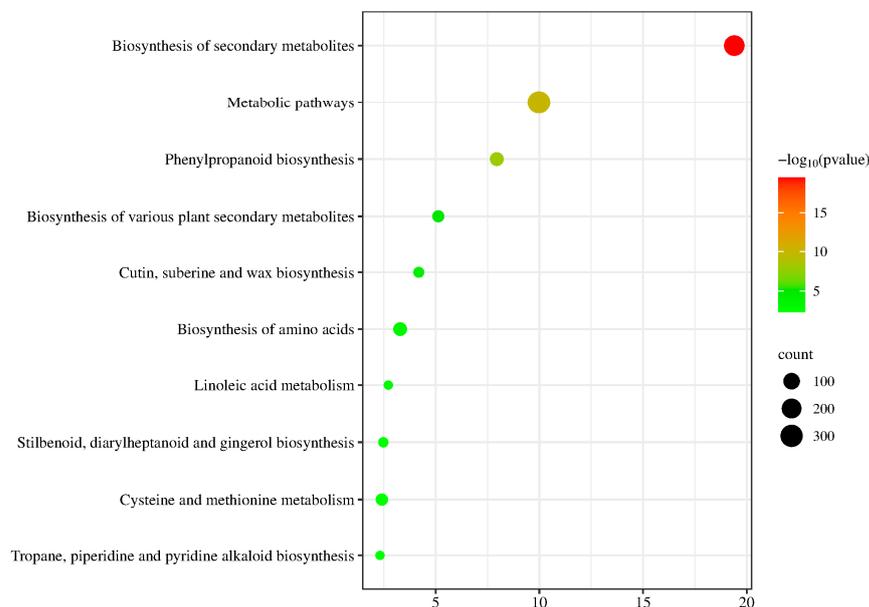


Figure 4. DEGs KEGG enrichment analysis of corn B73 2d_vs_0d treated with SG200.

3.4. A Large Number of Receptor Kinases and Signal-Transduction Proteins Are Markedly Induced

The majority of pattern-recognition receptors that mediate the PTI response (The pattern-triggered immunity (PTI) response is triggered at the plant's cell surface by the recognition of microbe-derived molecules known as microbe- or pathogen-associated molecular patterns or molecules derived from compromised host cells called damage-associated molecular patterns. Membrane-localized receptor proteins, known as pattern recognition receptors, are responsible for this recognition) on the plasma membranes of plant cells are receptor-like kinases. A search of the RNA-seq database for this study revealed the presence of 17 receptor-like kinase genes that exhibited a strong induction of expression. In contrast, no up-regulation of receptor genes for NLRs mediating ETI was observed. This finding is consistent with previous reports that maize resistance to *Ustilago maydis* is more of a quantitative shape of multiple genes rather than gene-to-gene translational resistance (Supplementary Table S3). In addition to the considerable number of receptor-like kinase genes that were found to be induced, 12 genes associated with signaling were also found to be significantly up-regulated, including 6 Ras-like protein genes, 3 MAPK/MAPKKK kinase genes, and 3 CDPK kinase genes associated with calcium signaling. The induction of Ras-like proteins is mediated by guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs) (upstream signaling) to plasma membrane receptors. The MAPK cascade response is involved in multiple signaling pathways and plays an important role in plant disease-resistance responses. Calcium is a crucial second messenger in the plant disease-resistance response, and CDPKs are capable of detecting alterations in intracellular Ca^{2+} concentration, phosphorylating downstream target proteins, and enhancing the plant disease-resistance response [20,21]. The *ZmBAK1* gene, which plays a pivotal role in the disease-resistance response, and two transcription factors situated downstream of the *ZmBAK1*-mediated signaling pathway, *ZmBZR7* and *ZmBZR10*, were also identified as components of the induced receptor-like kinase genes [22].

BAK1 was originally discovered as a key component of brassinosteroid (BR) signaling. *BAK1* is sensed by cell-surface receptors, which, in turn, triggers a signaling cascade, leading to inhibition of the protein kinase *BIN2* and activation of the transcription factor

BES1/BZR1, which directly regulates the expression of thousands of downstream responsive genes. Numerous studies have further shown that BAK1 acts as a co-receptor and forms a complex with FLS2 to regulate the PTI pathway and cell death [22,23]. To further validate the credibility of the RNA-seq results, three differentially expressed genes, *ZmBAK1*, *ZmBZR7*, and *ZmBZR10*, were validated using real-time fluorescence quantitative qRT-PCR. The results showed that the expression of *ZmBAK1*, *ZmBZR7*, and *ZmBZR10* genes was significantly up-regulated by SG200-induced expression (Figure 5A–C), and the above results were consistent with the RNA-Seq results, which indicated that the RNA-Seq results were credible. To analyze the affinity of maize *ZmBZR7* and *ZmBZR10* with other species, a phylogenetic tree analysis was conducted, with reported BZRs in *Arabidopsis thaliana*, rice, and other species by using MEGA6.0 software, and the results showed that *ZmBZR7* and *ZmBZR10* had high homology with OsBZR1 (Figure 5D). A bioinformatics analysis showed that there was a possible transmembrane structural domain of the *ZmBAK1* protein (Figure 5E). The next study on the function and mechanism of action of *ZmBAK1* and *ZmBZR7/ZmBZR10* genes will provide a new entry point to analyze the mechanism of action of *Ustilago maydis* disease resistance in maize.

3.5. Selecting of Maize Seedling Disease-Resistance-Related Genes by GWAS Analysis

One-hundred sixty-seven inbred maize lines grown in the greenhouse were inoculated with *Ustilago maydis* at the three-leaf and one-heart stages. Based on the disease development results at 8 days after inoculation, most of the varieties had similar disease development compared to the control B73, and only a few varieties showed resistant phenotypes (Table 3). The seedling resistance was calculated based on the values of disease-development phenotypes, and the seedling-resistance evaluation was carried out on different maize inbred lines. Thirty-two materials, including Fanrong 2, Dan 340, Zheng 58, B100, and K14, were selected as highly resistant varieties, accounting for 19% of the total materials identified (Figure 6). The 32 materials, including Fanrong 2, Dan 340, Zheng 58, B100, and K14, belonged to highly resistant varieties, accounting for 19% of the total materials identified (Figure 6). There were 29 resistant materials, accounting for 17% of the total materials identified, including Zong 3, DF24, Tuck 488, Huang ye Si 3, etc. There were 50 moderately resistant materials, including B73, 36 susceptible materials, and 20 highly susceptible materials, accounting for 30%, 30%, and 30% of the total materials identified, respectively. They accounted for 30%, 22%, and 12% of the total materials identified. In this study, the phenotypic values caused by *Ustilago maydis* approximately followed a normal distribution, which is suitable for further analysis.

Table 3. Disease-resistance classification of 167 maize inbred lines.

Sickness Index Condition Index	Assortment Variety
highly resistant	Fanrong 2; LH74; Rock 38; DM07; 18; Dan 340; W668; M3; Zheng 58; Lo1125; B100; N138; L-1; Jingnuo 2; DH65232 (DH9); XOP2; Ay420; Dai 6; 444; ND246; Tai 184; K14; P25; Deer 65; Q1261; 1313; OQ603; ChengZi2142; 8982; Cheng698-3; e220; H21
resistant	A632; He3; 6M502; bt1; b67; SC24-1; 91huang15; 757; Ys06; df24; 7236; Phn37; Ill12e; W9706; 08-64; Tuck488; K12; mbsj; W172; Phn82; Wil900; GY3; Lh123ht; D864; LH128; Huano IV3; Spoke 8527; HHe01; La2-4
moderately resistant	B73; Bai197; lbb15; W117; 9702; Qiong51; K514; Phh93; 85bai64; Wenhuang31413; Z31b; Php55; df27; 806a; Huangc; B12; Cr1ht; 1538; D15; 764; Along 812; Heng31; Lh162; 802; D857; Tuck 52106; Hai9-21; DH40; N42; ChangK; PHP85; 3335; 78002A; H66/6; Su75; D33A; Zheng32; PHJ33; Wobai; eQun4; NS501; Shen135; K22; DM101B; US68113; N68a; Drought21; 9711; SC14; B95.
susceptible	B127; Llegacy67; H114; Mo17; Lh59; 268; 3H2; W499; Phw51; Chang7-2; Gan41; DF20; D1139; M1016; M131-5; 17564; B78; D883; 78371A; VA26; Elect6; 68202; 624; Hb8229; 707; 68122; B394; SC11-1; DH65232; 1614; PHG86; LH156; 478; Long 72; E600; Medium 741

Table 3. Cont.

Sickness Index Condition Index	Assortment Variety
high susceptible	1069; Lh132; BCC03; Q381; CT109; 1145; Tuck8112; t24; 3189; S53; P136; Lx9801; Fang citing; 7903e; D88; Y9961; Brigade 28; e588; Zong548-1521; Wil901

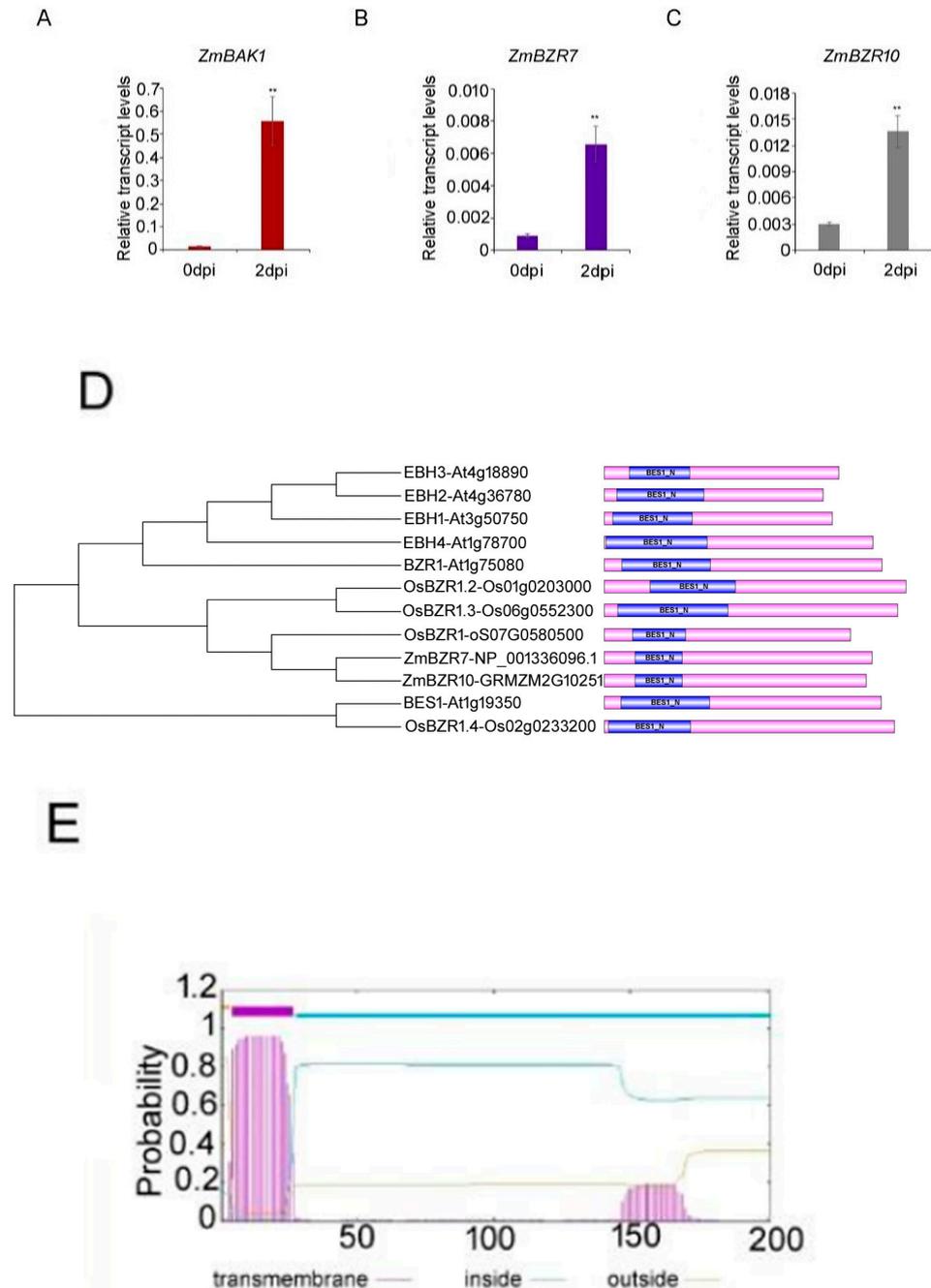


Figure 5. qRT-PCR validation of key proteins screened for oleuropein lactone signaling in the transcriptome and their bioinformatics analysis. (A–C) qRT-PCR experiments were performed using the *ZmACTIN* gene as an internal reference gene for fluorescence quantification, and the experiments were repeated three times to obtain similar results. Student’s *t*-test (** $p < 0.001$). (D) Analysis of *ZmBZR7* and *ZmBZR10* evolutionary trees. (E) Analysis of *ZmBAK1* transmembrane structural domains. The purple area is the conserved area and can be genetically modified.

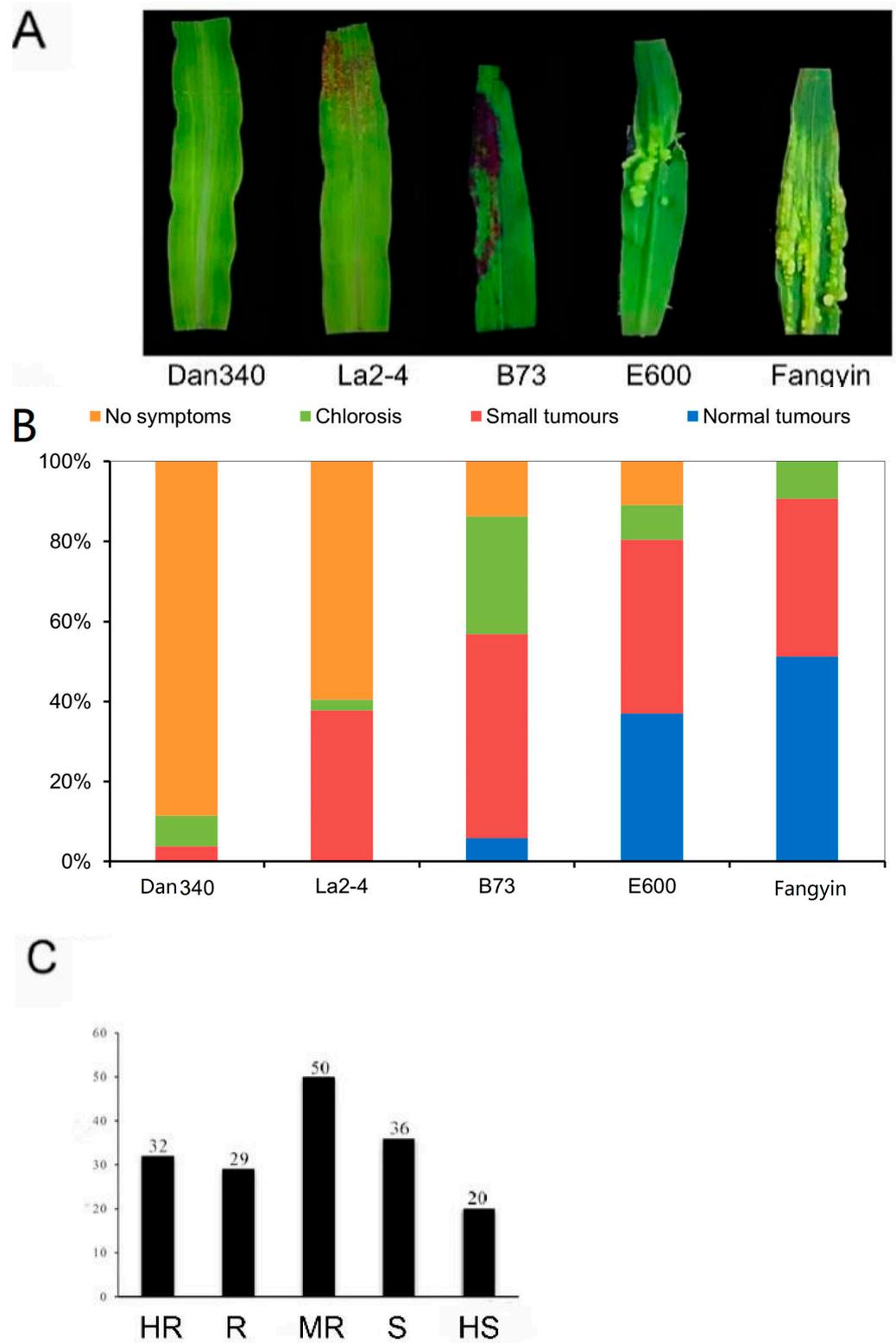
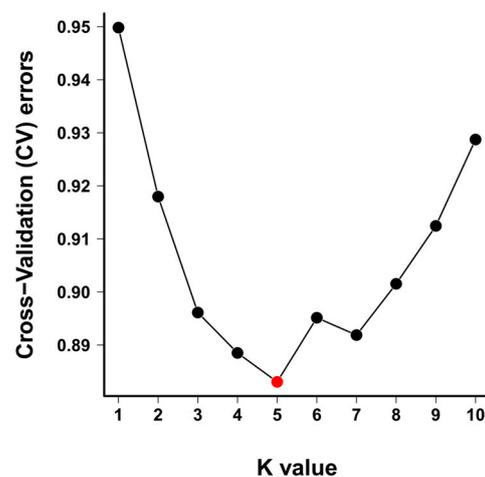


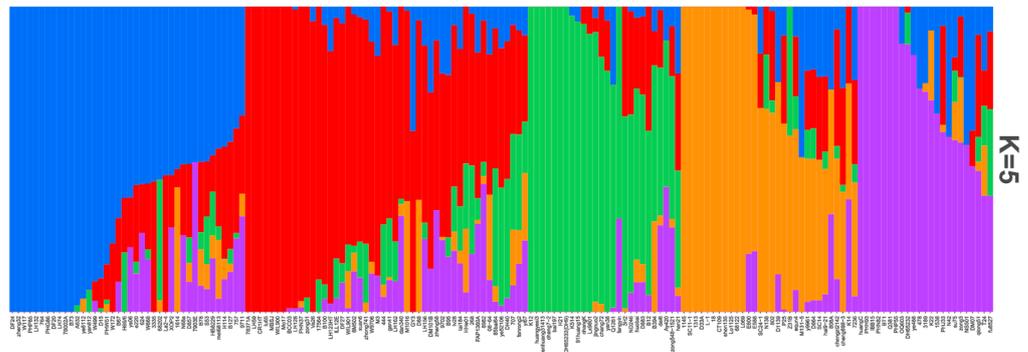
Figure 6. Distribution map of disease levels of maize tumoral melanosis. (A) Onset phenotype of maize inbred line after 8 days of SG200 treatment. (B) Disease grade distribution of maize inbred line after 8 days of SG200 treatment (C) Disease index statistics of 167 inbred lines.

3.6. Population Structure and LD Analysis of 167 Natural Population Maize Materials

In this study, 368,219 high-quality SNP markers were selected for population structure and LD analysis. Population-structure analysis of the original genotypes was performed using ADMIXTURE (version: 1.3.0; parameter: default parameter) software to predict the optimal grouping of the populations, and the cross-validation error rate (CV error) in the results was visualized to determine the optimal grouping of the populations. The number of populations, K , was set from 1 to 10, and the cross-validation error rate (CV error) for different values of k was extracted from the result file. The lowest value of the cross-validation error rate (CV error) was found at $K = 5$ (Figure 7), indicating that the 167 self-crosses could be roughly classified into five subpopulations. These subpopulations were designated as follows: blue for the first subpopulation; red for the second subpopulation; green for the third subpopulation; orange for the fourth subpopulation; and purple for the fifth subpopulation.



(A)



(B)

Figure 7. Grouping of 167 maize inbred lines. (A) Determine the lowest point of the K value according to the CV value, as the optimal number of groups (B) using ADMIXTURE to calculate the population results.

When the cutoff for LD decay was set to 0.20, the average decay distance of the 167 maize inbred lines was about 260 Kb (Figure 8). With the help of the B73RefGen_V4 reference genome database on the maizeGDB website, candidate genes were searched within 260 Kb of each extension upstream and downstream of the SNP locus.

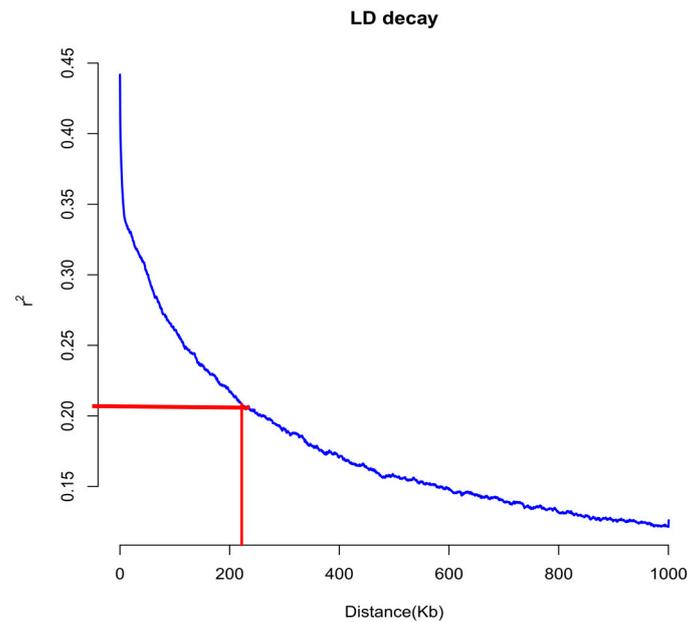


Figure 8. LD decay plots of 167 maize inbred lines.

3.7. Genome-Wide Association Analysis of Maize Resistance to *Ustilago maydis* SG200

The GWAS of the disease index in maize inbred lines resistant to *Ustilago maydis* SG200 was performed using the BLINK, FarmCPU, GLM, MLM, MLMM, and SUPER models of GAPIT software (Figure 9). Through the combined application of these models, we identified a total of 21 genetic marker loci significantly associated with disease index. Among the models used, the multi-locus model MLMM showed the highest detection efficacy, successfully localizing seven significantly associated loci. This was followed by the multi-locus model FarmCPU and the single-locus SUPER model, both of which identified four significant loci. In addition, the multi-locus model BLINK model detected three significant loci, while the unit point models GLM and MLM identified two and one significant SNPs, respectively. Of particular note, the significant SNPs detected by the unit point models GLM and MLM, as well as the chr4_128393468 and chr4_128393500 loci localized by the SUPER model loci, were also confirmed in duplicate in the multi-locus model. These loci were concentrated on chromosomes 4 and 7 (Supplementary Table S4). Among them, chr4_128393468 on chromosome 4 was present in all six computational models, explaining 17.30% of the phenotypic variation. The locus chr4_128393500 was localized by four models, and its phenotypic contribution was 18.07%.

Four SNP loci were co-localized by two or more models, namely chr4_128393468, chr4_128393500, chr7_88700440, and chr7_88695930 (Figure 10). In particular, two loci, chr4_128393468 and chr4_128393500, on chromosome 4 showed a high degree of co-localization, recognized by six and four different genetic models, respectively. This finding suggests that these loci may play key roles in regulating certain biological processes. In addition, the chr7_88700440 and chr7_88695930 loci located on chromosome 7 were also co-localized by three and two models, respectively, further emphasizing their potential importance in genetic studies.

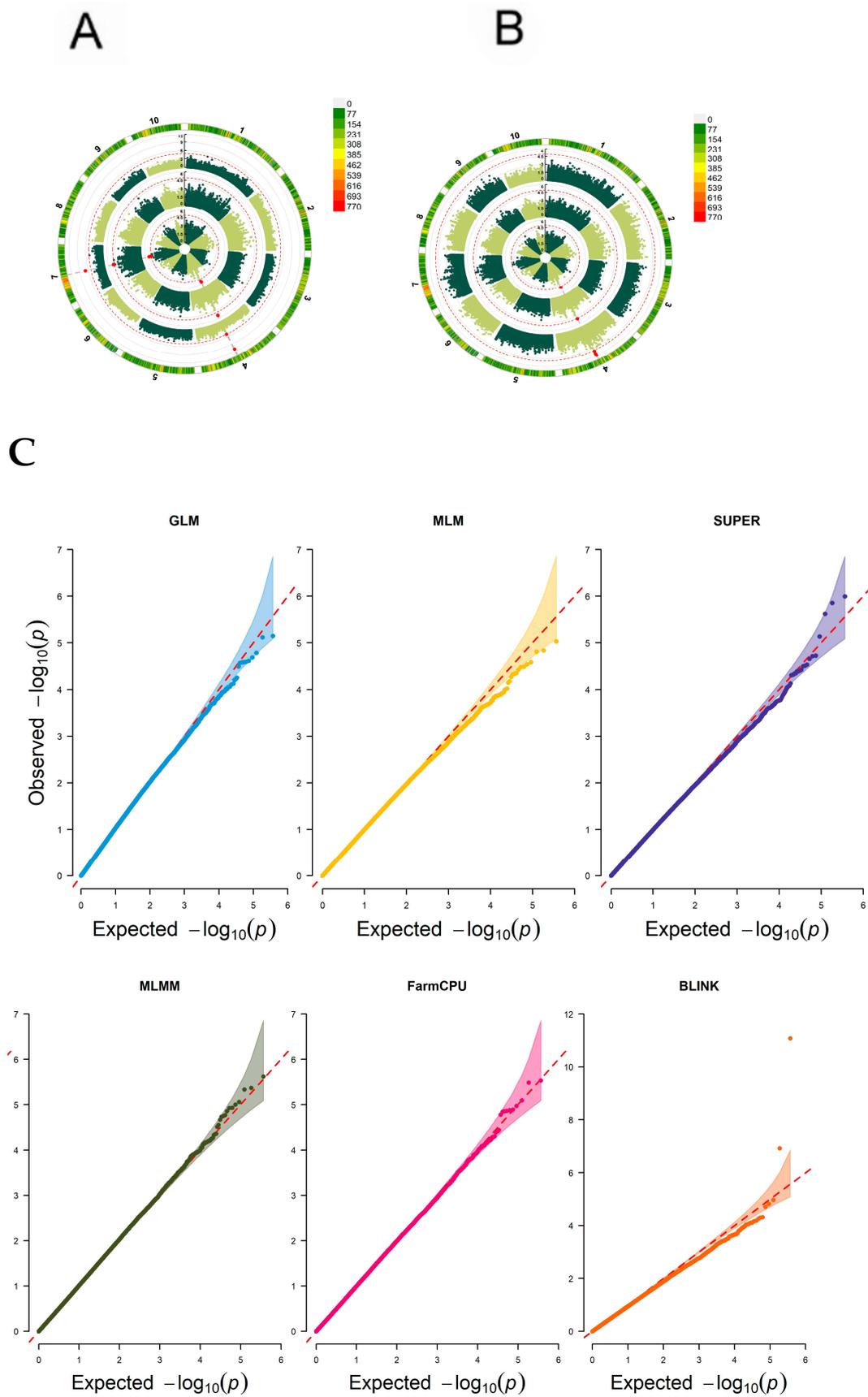


Figure 9. Six models were used for correlation analysis of disease index. (A) GLM, MLM, SUPER, (B) MLMM, FarmCPU, BLINK, and (C) QQ diagram for correlation analysis using six models.

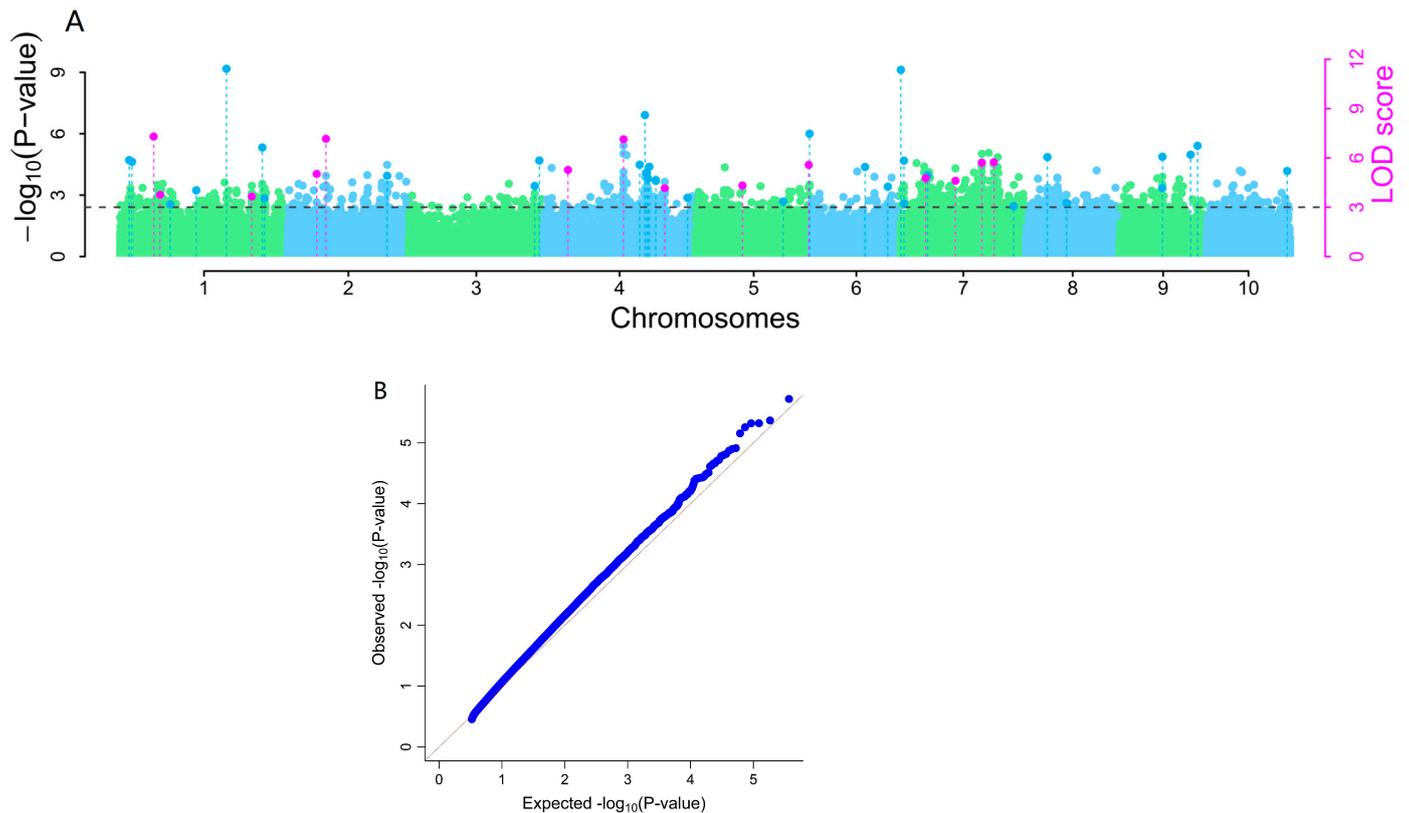


Figure 10. MrMLM software was used to analyze the association between disease severity index (A) Manhattan map. In this figure, the light blue loci are insignificant SNPS, the dark blue loci are SNPS detected by one model, and the pink loci are SNPS detected by two or more methods. (B) QQ chart.

To gain a comprehensive understanding of the genetic basis, we employed six additional association models in the mrMLM software, including mrMLM, FASTmrMLM, FASTmrEMMA, ISIS EM-BLASSO, pLARmEB, and pKWmEB. The combined application of these models enabled the identification of significant-association loci on multiple chromosomes of the maize genome. The results of the analysis (Supplementary Table S4), indicate that the models identified a total of 72 significant-association loci on chromosomes 1, 2, 4, 5, 6, 7, 8, 9, and 10. Among these, the ISIS EM-BLASSO model exhibited the highest detection efficacy, with 20 significant-association loci identified. Subsequently, the FASTmrMLM model identified 16 significant-association sites. The pKWmEB and pLARmEB models, respectively, detected 15 and 14 significant-association sites. In contrast, the pKWmEB and pLARmEB models identified seven significant-association sites in the FASTmrEMMA model. Notably, the mrMLM model did not detect any significant-association sites in this analysis. Fourteen loci were identified as co-localized by two or more models in the analysis results of the mrMLM package. This result demonstrates a high degree of consistency and stability in the association signals of these loci across the models, thus enhancing the credibility of these loci as potential disease-resistance loci.

The locus chr7_88700440, located on chromosome 7, was consistently detected in all five different genetic models, showing significant co-localization. The LOD score of this locus was 6.27, a value significantly higher than the conventional threshold, indicating a statistically significant association with the *Ustilago maydis* SG200 disease resistance phenotype in maize inbred lines. Moreover, the phenotypic variation explained by the chr7_88700440 locus ranged from 2.26% to 11.74%. This broad variation in phenotypic variation further highlights that this locus regulates a critical step in the genetics of maize resistance to *Ustilago maydis* SG200 disease. This finding provides crucial insights into the

genetic mechanisms underlying disease resistance and may facilitate the development of valuable molecular markers for future molecular marker-assisted selection (MAS).

In this experiment, we employed the GAPIT software to perform the association analysis. The unit point model, generalized linear model (GLM), mixed linear model (MLM), and super association analysis (SUPER) successfully identified two, one, and four genetic markers with statistically significant associations, respectively. In the analysis of multi-locus models, BLINK, FarmCPU, and multimarker–multimodal mixed model (MLMM) identified three, seven, and four significantly associated loci, respectively. Genome-wide multi-locus association analysis was also performed using mrMLM software. Among the numerous analytical methods employed, ISIS EM-BLASSO demonstrated superior performance, successfully identifying the greatest number of 20 SNPs. FASTmrMLM, in turn, identified 16 SNPs, while the pKWmeB model identified 15 SNPs. The pLARmeB method also achieved a noteworthy success, identifying 14 SNPs. In contrast, the FASTmrEMMA method identified the lowest number of significantly associated loci at seven. However, the mrMLM model did not detect any significant-association loci, which may be due to the occurrence of overfitting, a phenomenon caused by the inclusion of too many variables in the multi-locus model. Among the identified loci, chr4_128393468 was co-localized by six models of GAPIT software and three models of mrMLM software, while chr7_88700440 was co-localized by five models of GAPIT software and four models of mrMLM software. The two loci were consistently localized in multiple models of GAPIT and mrMLM software, suggesting that they play an important role in disease-resistance traits and may be involved in key biological pathways and mechanisms.

3.8. Candidate Gene Mining and Functional Annotation of Candidate Genes for Tumor-Resistant Black Powder Disease

In this study, based on the LD decay distance results, an interval based on $r^2 = 0.20$ was added or subtracted upstream and downstream of the physical position of each SNP locus to define the physical position query region of candidate genes. To further screen for candidate genes associated with plant resistance, we searched for genes encoding proteins within the query region defined above using the online gene browser provided by the maizeGDB website (www.maizegdb.org/gbrowse, 12 December 2023) for the maize B73 genome sequence version V4 (Maize B73RefGen_V4). Meanwhile, to verify the biological relevance of these genes, we analyzed these screened genes against the database on the NCBI website.

A total of 19 candidate genes were screened in the Maize B73RefGen_V4 database based on the significant-association sites in the GAPIT analysis results (Supplementary Table S5). Chr 4_128393468 was co-localized by six models. There were six candidate genes near this locus, respectively, and the identification of these genes provides an important basis for further functional studies and potential biological mechanism exploration. Similarly, another SNP locus chr4_128393500 located on chromosome 4 was co-localized by the four genetic models, while the chr4_128393451 locus ($-\log_{10}P = 5.65$) localized by the SUPER model also showed the same six candidate genes as the chr4_128393468 locus. This result further supports the biological relevance and possible functional importance of these genes. On chromosome 7, near the SNP locus chr7_88700440 co-localized by the three models, we mined eight candidate genes, which were the same as the locus chr7_88695930 co-localized by the two models. The locus chr7_88700232 was localized by the MLMM model ($-\log_{10}P = 5.00$), and chr7_88699020 ($-\log_{10}P = 4.92$) and chr7_88699535 ($-\log_{10}P = 4.91$) localized exactly the same candidate genes. This result suggests that these genes show significant co-localization in multiple models, which increases their biological relevance and may play a key role in regulating specific phenotypes.

In the genetic association analysis performed by applying the mrMLM software, there were 14 significant SNP loci detected by two or more models, and a total of 117 candidate genes were screened (Supplementary Table S6).

In particular, the SNP locus chr7_88700440 located on chromosome 7 was localized in a total of five genetic models, and eight candidate genes were identified in its neighboring region. The co-localization of these genes suggests that they may play important roles in regulating related biological traits. Another SNP locus, chr7_112915305, located on chromosome 7, was localized by four genetic models, and six candidate genes were identified in the region adjacent to it. The discovery of these genes provides new perspectives for understanding the biological traits associated with the locus. In addition, we identified multiple SNP loci co-localized in three genetic models, such as chr1_244281660, chr1_78993263, chr4_128393468, and chr5_220156746, and 17, 7, 6, and 39 candidate genes were identified near these loci, respectively. These results further enriched our understanding of the biological functions of these loci and provided a list of important candidate genes for future studies.

Finally, for those SNP loci that were only detected in two genetic models, such as chr1_59674633, chr2_70835476, chr4_209222445, chr4_47520061, chr5_93407215, chr7_48508221, and chr7_69407281, we also identified five, three, five, eight, five, six, and one candidate genes within their neighboring regions, respectively. Although these loci were only detected in two models, their discovery still provides valuable information for further genetic studies.

A total of 19 genes from the GAPIT analysis results were found to have corresponding homologous sequences in the NCBI database, and detailed functional annotations have been obtained for 13 of these genes. Further analysis revealed that these candidate genes mined by genome-wide association studies (GWAS) had extensive overlap with the genes known to control the corresponding traits previously reported in the literature. This finding provides strong support for validating the reliability of the GWAS results and provides additional evidence for the biological functions of these candidate genes (Supplementary Table S7).

4. Discussion

4.1. Seedling Identification and Resource Evaluation of *Ustilago maydis* Maize

Identifying germplasm resources for *Ustilago maydis* resistance is a crucial preliminary step in the process of mining disease-resistance alleles and realizing disease-resistance breeding. In this study, 167 maize inbred lines were identified for seedling resistance to *Ustilago maydis* by manual injection. The results demonstrated that there were 32 highly resistant materials. Further comparative analysis of lineage origin revealed that the high-resistance phenotype was distributed across several maize germplasm populations, including the Brigade Red Bone group (Dan 340), the Ruide group (Zheng 58), the Tang Si Pingtong group (Huangnuo Si 3), and the P group (18). These findings provide valuable genetic resources for breeding maize varieties with enhanced resistance to *Ustilago maydis*. Consequently, the mining and identification of genes conferring resistance to *Ustilago maydis* in these maize germplasms exhibiting high resistance, and the analysis of their molecular mechanisms of resistance, are of great theoretical and practical importance for the breeding of *Ustilago maydis* maize.

4.2. Comparison between the Disease-Resistance Genes Identified by Transcriptome Analysis and the Results of Previous Studies

Maize plants have developed a series of molecular mechanisms to resist pathogen stress. In the transcriptome analysis, the GO pathway was significantly enriched for defense responses against bacteria (GO:0042742), as well as oxidoreductase activity (GO:0016709), while the secondary metabolite biosynthesis pathway showed extremely high enrichment in the KEGG pathway analysis. These findings underscore the pivotal role of oxidoreductases in the regulation of hydrogen peroxide metabolism.

In this study, a significant up-regulation of the expression of ZmBAK1, a key member of the Brassica napus lactone (brassinosteroids) signaling pathway, and the downstream transcription factors ZmBZR7 and ZmBZR10 was observed in maize. BAK1, as a key

co-receptor in pattern-recognition-receptors (PRRs)-mediated immune signaling, plays a crucial role in the plant's innate immune response. On the surfaces of plant cells, BAK1 binds to a variety of microbe-associated molecular patterns (MAMPs), thereby triggering a series of immune responses. A study demonstrated that the transcriptome profiles of maize plants infected with a virulent strain of *R. solani* for 3 and 5 days were analyzed by RNA-seq. The results demonstrated that *ZmNAC41* and *ZmBAK1* were involved in resistance to *R. solani*. In previous experiments, researchers observed that the two mutants, *bak1-4* and *bkk1-1*, exhibited comparable susceptibility to wild-type plants against infection with Pto DC3000. This was evidenced by the fact that they exhibited a disease-resistant phenotype [24].

4.3. Comparison of GWAS Analysis Localization Results with Previous Work

In this study, the GWAS of maize disease indices after infection with *Ustilago maydis* was carried out, and a series of significant-association loci were successfully identified using the GAPIT software package and were located within the interval of disease-resistance genes that had been reported in previous studies. The two significantly associated loci, chr4_128393468 and chr4_128393500, located on chromosome 4, were both located in the region of 4.05 bin, with phenotypic contributions of 17.30% and 18.07%, respectively, indicating that they play an important role in the regulation of maize resistance to *Ustilago maydis*. Genome-wide linkage analysis was conducted on Mo17 (resistant) and B73 (susceptible) crosses in an advanced hybrid recombinant inbred line population evaluated for resistance to gray spot (GLS) in three environments, and five significant QTL were detected in 1.05, 2.04, 4.05, 9.03, and 9.05 of the population [25]. In addition, we identified two significant-association loci located on chromosome 7, chr7_88700440 and chr7_88695930, both within the bin 7.02 region, with genetic contributions to the phenotype of 11.59% and 11.50%, respectively, which were consistent with the resistance QTLs detected within the same bin region in the previous study. Some studies detected and validated a stable primary QTL in the bin7.02 region of chromosome 7 and found that mutants carrying the disease-resistance gene at locus 7.02 were significantly more resistant to Northern Corn Leaf Blight (NCLB). The levels of cob rot (FER) and fumonisin B1 (FB1) in F3 progeny, through F3 progeny, in different years and at different sowing dates with significant differences were found, with a QTL in bin7.02 conferring maize resistance to Fusarium rot and affecting fumonisin B1 levels [26]. Four molecular markers were linked to maize roughshod disease-resistance genes, *umc1656* (bin6.02), by the SSR-BSA method using a population of 90110 × 478, *umc1401* (bin7.02), *bnlg1823* (bin8.07), and *umc1268* (bin8.07), and it was hypothesized that at least three maize rough shank disease-resistance loci existed in 90110 of GWAS analysis locus results with previous work [27].

In this study, genome-wide association analysis of maize disease indices after infection with *Ustilago maydis* disease was performed. A series of significant-association loci were successfully identified using the GAPIT software package, all of which were located within the interval of disease-resistance genes that have been reported in previous studies. The two significantly associated loci, chr4_128393468 and chr4_128393500, located on chromosome 4, were both located in the region of 4.05 bin, with phenotypic contributions of 17.30% and 18.07%, respectively. This suggests that they play an important role in the regulation of maize resistance to *Ustilago maydis*. In a genome-wide linkage analysis of Mo17 (resistant) and B73 (susceptible) crosses in an advanced hybrid recombinant inbred line population evaluated for resistance to gray spot (GLS) in three environments, five significant quantitative trait loci (QTL) were detected in 1.05, 2.04, 4.05, 9.03, and 9.05 of the population [25]. In addition, two significant-association loci were identified on chromosome 7, chr7_88700440 and chr7_88695930, both within the bin 7.02 region. These loci exhibited genetic contributions to the phenotype of 11.59% and 11.50%, respectively. These findings were consistent with the resistance quantitative trait loci (QTLs) detected within the same bin region in the previous study. A stable primary QTL in the bin 7.02 region of chromosome 7 demonstrated that mutants carrying the disease-resistance gene at

locus 7.02 exhibited significantly greater resistance to Northern Corn Leaf Blight (NCLB). One study observed significant differences in the levels of cob rot (FER) and fumonisin B1 (FB1) in F3 progeny across different years and sowing dates. This led to the identification of a QTL in bin7.02, conferring maize resistance to Fusarium rot and affecting fumonisin B1 levels. Some studies identified four molecular markers linked to maize roughshod disease-resistance genes, *umc1656* (bin6.02), through the SSR-BSA method using a population of 90110 × 478; *umc1401* (bin7.02), *bnlg1823* (bin8.07), and *umc1268* (bin8.07) were identified as potential maize rough shrivel disease-resistance loci in 90110.

Several key SNP loci for resistance to downy mildew were also identified using mrMLM software. The *chr1_59674633* and *chr1_78993263* on chromosome 1 were located within the QTL interval for disease resistance studied by previous research, both located at bin1.04. The QTL used the CIM mapping method and localized one QTL with a contribution of 10.6%, which was located in the marker RFLP *asg3* near the region of chromosome bin1.04 [28]. In this study, significant disease-resistance SNP loci were detected on chromosome 2 (bin2.04 and bin2.05) and chromosome 5 (bin5.04 and bin5.09), respectively. QTL was associated with black sigatoka resistance in the F2:3 population of the T32 (highly resistant genotype) × HC (highly susceptible genotype) cross using the T32 (highly resistant genotype) × HC (highly susceptible genotype) cross to identify QTLs located on the second bin2.04 interval on chromosome 2, a stable and novel QTL for HS resistance [29]. A GWAS on 150 different maize self-crosses under striga-infested and non-infested conditions was conducted and found that most of the identified genes were located in the 2.05 interval and encoded transcription factors, disease-resistance proteins, zinc-finger structural domain proteins, leucine-rich repeat protein kinases, and some pathogenesis-related proteins [30]. Some studies identified a GSR resistance QTL on bin5.04 with a candidate gene encoding DOF5.7 involved in biotic stress processes in plants [31]. In bin5.05, a QTL for fusarium resistance explained 13% of the total phenotypic variation [32]. A study conducted a three-season field trial on a maize RIL population from a cross between CML444 and SC Malawi under GLS stress and found various QTL hotspots for GLS resistance in chromosomes bin7.02 and 7.03 [33]. Notably, this study also detected resistance SNPs in the interval of chromosome 1 bin1.08 and chromosome 4 bin4.09, respectively, and these SNP positions are different from the results of previous studies, which can be regarded as new maize *Ustilago maydis* resistance loci.

4.4. Integration Analysis of GWAS Candidate Genes with RNA-seq DEGs Results

In integrating the gene-expression data obtained from genome-wide association studies and RNA-seq technology, we successfully identified a series of candidate genes that are closely related to specific biological traits. The discovery of these genes not only provides new perspectives for understanding the molecular mechanisms of these traits but also offers potential loci for future crop improvement and disease-resistance breeding (Table 4).

Located on chromosome 1 at the *chr1_244281660* locus, which was pinpointed by three models in the mrMLM package, we identified two notable candidate genes, *Zm00001d032946* and *Zm00001d032948*. These two genes were similarly expressed in the B73 maize cultivar on day 2 after inoculation with *Ustilago maydis* SG200 as compared to the inoculation. The comparison of gene-expression profiles before (day 0) showed significant differential expressions (DEGs), a result that implies that they may play a crucial role in the plant's defense response to the pathogen. The *Zm00001d032948* gene was identified as encoding a key biosynthetic enzyme—the fatty acid elongase. This enzyme plays a central role in the biosynthesis of plant epidermal waxes by regulating the length of fatty acid chains, which in turn affects the structure and function of the waxes, enabling the plant to defend itself against pests and pathogens. The *Zm00001d032946* gene has been shown to have an important role in the defense mechanism of the plant, as has the related protein, chitinase 1 [34,35]. Chitinase 1 may play a role in plant defense responses against pathogens by recognizing and degrading pathogen cell walls, thereby enhancing plant disease resistance.

Table 4. Candidate genes identified by both methods.

SNP	Genome	Functional Notes	Previous Study	Journal
1_244281660	Zm00001d032946	Chitinase 1	Chitinase activities themselves remained unchanged under biotic stress, but their expression was widely upregulated under stress	BMC Plant Biology
1_244281660	Zm00001d032948	fatty acid elongase	Fatty acid elongase activity directly affects the number and chain length of VLCFAs	The Plant Journal
5_220156746	Zm00001d018414	IAA9—auxin-responsive Aux/IAA family member	Overexpression of miR393 inhibits growth hormone signaling and enhances Arabidopsis resistance to the bacterial pathogen <i>Pseudomonas butyrlica</i>	Science
5_220156746	Zm00001d018421	GATA transcription factor 8	AtGATA8 (BME3) positively regulates seed germination in <i>Arabidopsis thaliana</i>	The Plant Journal
7_88700440	Zm00001d020043	AP2-EREBP transcription factor	Rice OsERF922, on the other hand, negatively regulates resistance to rice blast disease	Journal of Experimental Botany

At the chr5_220156746 locus on chromosome 5, which was pinpointed by three models in the mrMLM package, we identified two other candidate genes, Zm00001d018414 and Zm00001d018421. These genes also exhibited significant DEGs in their gene-expression changes 2 days after inoculation. The Zm00001d018414 gene was identified as encoding the IAA9 protein, a member of the growth-hormone-responsive Aux/IAA family. As an important hormone for plant growth and development, growth hormone not only plays a central role in the growth and development process of plants but also plays an important role in the defense mechanism of plants against a variety of pathogens [36,37]. For example, it was found that the growth-hormone signaling pathway was significantly down-regulated after rice black streak dwarf virus (RBSDV) infestation [38]. In addition, inhibition of growth-hormone signaling by overexpression of miR393 enhanced Arabidopsis resistance to the bacterial pathogen *Pseudomonas syringae*. Meanwhile, the Zm00001d018421 gene encodes a member of the GATA transcription factor family, GATA transcription factor 8. The GATA transcription factor family in plants is a major class of transcriptional regulators, among which AtGATA8 (BME3) plays a positive regulatory role in *Arabidopsis* seed germination [39]. Notably, the expression level of PbGATA8 was significantly up-regulated 54.2 fold within 1 h after SA (SA is an important plant defense agent hormone that plays a significant role in plant defense responses, especially in SAR. In addition, SA is also involved in regulating the growth and hair of plants and various stress responses) treatment [40], a change that strongly suggests its potential role in plant disease-resistance response.

In addition, the chr7_88700440 locus was analyzed by a combination of five models from the mrMLM package, as well as three models from the GAPIT package, and we identified the candidate gene Zm00001d020043. This gene was identified as DEGs in gene-expression analysis after inoculation with the bacterium *Aspergillus tumefaciens* SG200. The gene expression of PbGATA8 was also found to be significantly higher than that of PbGATA8 in maize (*Zea mays*), the gene Zm00001d020043 encodes an ethylene response element binding factor, AP2-EREBP transcription factor, which plays a crucial role in regulating plant growth and development and coping with diverse biotic and abiotic stresses. Members of the AP2-EREBP transcription factor family play multidimensional regulatory roles in plant growth and development and adversity response mechanisms [41–43].

These results not only provide important information for understanding the molecular mechanism of plant resistance to *Ustilago maydis* but also provide potential candidate genes for molecular marker-assisted selection and genetic engineering improvement of crop

disease resistance in the future. Future studies should focus on the functional validation of these genes and their specific mechanisms of action in plant growth and development and in response to environmental stresses. Through further functional validation and phenotypic association studies, these genes are expected to become key targets for improving crop disease resistance and developing effective disease management strategies.

5. Conclusions

The gene regulatory network in the maize B73 inbred line at 2 days of induction by *Ustilago maydis* SG200 was systematically screened at the genome-wide level using transcriptome sequencing, and key regulatory genes for maize resistance to *Ustilago maydis* at early stages of *Ustilago maydis*, including the up-regulation of a large number of receptor kinases, signaling-related proteins, redox-response-related genes, WRKYs, and P450s, etc., were screened. These include up-regulated expression of the disease resistance key co-receptor gene ZmBAK1 and its downstream transcription factors ZmBZR7/ZmBZR10. The disease index of 167 maize inbred lines at the seedling stage was evaluated for powdery mildew resistance, and 32 high-resistant germplasm resources, such as Prosperity 2, Dan 340, Zheng 58, B100, and K14 were screened. Association analysis of the disease index using GAPIT and mrMLM software detected 21 and 72 significant-association loci, respectively. A total of nine models were localized to the chr4_128393468 locus and eight models were localized to the chr7_88700440 locus by the two software, which were all located within the previously reported QTL regions associated with maize disease resistance. Integration of GWAS and RNA-seq results yielded five disease-resistance-related genes, Zm00001d032946, Zm00001d032948, Zm00001d018414, Zm00001d018421, and Zm00001d020043. These genes encode chitinase 1 protein, FAE (fatty acid elongase), IAA9, GATA TF8, and EREB94, respectively, and their biological functions and molecular mechanisms in *Ustilago maydis* resistance will be of great importance for improving maize breeding for *Ustilago maydis* resistance. It has important theoretical significance and application value for improving the breeding of maize for *Ustilago maydis* resistance.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/agriculture14060958/s1>, Table S1: 167 catalogues of inbred line materials; Table S2: Fluorescent quantitative qPCR primers and PCR routine; Table S3: Receptor kinases and proteins in signaling transduction screened by transcriptome; Table S4: Significant association sites screened by GAPIT software; Table S5: Significant associated sites screened by mrMLM software; Table S6: GAPIT candidate gene mining and functional annotation; Table S7: Significant association loci and their candidate genes screened by MrMLM software; Table S8: Candidate genes and functional annotations mined by MrMLM software.

Author Contributions: J.G., W.S. (Wei Song), W.S. (Weibin Song) and P.H. designed this study; J.L., Y.Z. (Yunxiao Zheng) and Y.Z. (Yongfeng Zhao) developed the populations; B.Y. and X.J. recorded the data; B.Y., Y.Z. (Yunxiao Zheng) and L.Z. analyzed the data; B.Y. and L.X. drafted the manuscript; B.Y., L.X., P.H., Y.Z. (Yunxiao Zheng), W.S. (Wei Song), W.S. (Weibin Song) and J.G. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Science and Technology Innovation Team of Maize Modern Seed Industry in Hebei, Grant/Award Number: 21326319D; National Key Research and Development Program of China, Grant/Award Number: 2018YFD0300501; State Key Laboratory of North China Crop Improvement and Regulation, Grant/Award Number: NCCIR2021ZZ-104.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: All the data for this study are included in this manuscript.

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

References

- Lanver, D.; Müller, A.N.; Happel, P.; Schweizer, G.; Haas, F.B.; Franitza, M.; Pellegrin, C.; Reissmann, S.; Altmüller, J.; Rensing, S.A.; et al. The Biotrophic Development of *Ustilago maydis* Studied by RNA-Seq Analysis. *Plant Cell* **2018**, *30*, 300–323. [[CrossRef](#)] [[PubMed](#)]
- Pérez-Rodríguez, F.; González-Prieto, J.M.; Vera-Núñez, J.A.; Ruiz-Medrano, R.; Peña-Cabriales, J.J.; Ruiz-Herrera, J. Wide distribution of the *Ustilago maydis*-bacterium endosymbiosis in naturally infected maize plants. *Plant Signal. Behav.* **2021**, *16*, 1855016. [[CrossRef](#)] [[PubMed](#)]
- Lister, R.; O'Malley, R.C.; Tonti-Filippini, J.; Gregory, B.D.; Berry, C.C.; Millar, A.H.; Ecker, J.R. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **2008**, *133*, 523–536. [[CrossRef](#)]
- Yao, L.; Li, Y.; Ma, C.; Tong, L.; Du, F.; Xu, M. Combined genome-wide association study and transcriptome analysis reveal candidate genes for resistance to *Fusarium* ear rot in maize. *J. Integr. Plant Biol.* **2020**, *62*, 1535–1551. [[CrossRef](#)]
- Han, G.; Li, C.; Xiang, F.; Zhao, Q.; Zhao, Y.; Cai, R.; Cheng, B.; Wang, X.; Tao, F. Genome-wide association study leads to novel genetic insights into resistance to *Aspergillus flavus* in maize kernels. *BMC Plant Biol.* **2020**, *20*, 206. [[CrossRef](#)]
- Rizzardì, D.A.; Peterlini, E.; Scapim, C.A.; Pinto, R.J.B.; Faria, M.V.; Contreras-Soto, R.I. Genome wide association study identifies SNPs associated with northern corn leaf blight caused by *Exserohilum turcicum* in tropical maize germplasm (*Zea mays* L.). *Euphytica* **2022**, *218*, 40–49. [[CrossRef](#)]
- Gowda, M.; Das, B.; Makumbi, D.; Babu, R.; Semagn, K.; Mahuku, G.; Olsen, M.S.; Bright, J.M.; Beyene, Y.; Prasanna, B.M. Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theor. Appl. Genet.* **2015**, *128*, 1957–1968. [[CrossRef](#)] [[PubMed](#)]
- Brefort, T.; Doehlemann, G.; Mendoza-Mendoza, A.; Reissmann, S.; Djamei, A.; Kahmann, R. *Ustilago maydis* as a Pathogen. *Annu. Rev. Phytopathol.* **2009**, *47*, 423–445. [[CrossRef](#)] [[PubMed](#)]
- Steinberg, G.; Perez-Martin, J. *Ustilago maydis*, a new fungal model system for cell biology. *Trends Cell Biol.* **2008**, *18*, 61–67.
- Lanver, D.; Tollot, M.; Schweizer, G.; Lo Presti, L.; Reissmann, S.; Ma, L.S.; Schuster, M.; Tanaka, S.; Liang, L.; Ludwig, N.; et al. *Ustilago maydis* effectors and their impact on virulence. *Nat. Rev. Microbiol.* **2017**, *15*, 409–421. [[CrossRef](#)]
- Kamber, T.; Buchmann, J.P.; Pothier, J.F.; Smits, T.H.M.; Wicker, T.; Duffy, B. Fire blight disease reactome: RNA-seq transcriptional profile of apple host plant defense responses to *Erwinia amylovora* pathogen infection. *Sci. Rep.* **2016**, *6*, 21600–21617. [[CrossRef](#)] [[PubMed](#)]
- Pathi, K.M.; Rink, P.; Budhagatapalli, N.; Betz, R.; Saado, I.; Hiekel, S.; Becker, M.; Djamei, A.; Kumlehn, J. Engineering Smut Resistance in Maize by Site-Directed Mutagenesis of LIPOXYGENASE 3. *Front. Plant Sci.* **2020**, *11*, 543895. [[CrossRef](#)] [[PubMed](#)]
- Kämper, J.; Kahmann, R.; Bölker, M.; Ma, L.-J.; Brefort, T.; Saville, B.J.; Banuett, F.; Kronstad, J.W.; Gold, S.E.; Müller, O.; et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **2006**, *444*, 97–101. [[CrossRef](#)]
- Ruan, X.; Ma, L.; Zhang, Y.; Wang, Q.; Gao, X. Dissection of the Complex Transcription and Metabolism Regulation Networks Associated with Maize Resistance to *Ustilago maydis*. *Genes* **2021**, *12*, 1789. [[CrossRef](#)]
- Shi, J. Biology of Powdery Mildew and Identification of Resistance in Maize. Master's Thesis, Gansu Agricultural University, Lanzhou, China, 2010.
- Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4325. [[CrossRef](#)]
- Sahito, J.H.; Zhang, H.; Gishkori ZG, N.; Ma, C.; Wang, Z.; Ding, D.; Zhang, X.; Tang, J. Advancements and Prospects of Genome-Wide Association Studies (GWAS) in Maize. *Int. J. Mol. Sci.* **2024**, *25*, 1918. [[CrossRef](#)]
- Shah, S.M.S. *Mining Candidate Genes for Low Nitrogen Stress Tolerance in Maize by GWAS and QTL Localization and Transcriptome Analysis*; Chinese Academy of Agricultural Sciences: Beijing, China, 2022.
- Cui, Y.; Zhang, F. The Application of Multi-Locus GWAS for the Detection of Salt-Tolerance Loci in Rice. *Front. Plant Sci.* **2018**, *9*, 1464. [[CrossRef](#)]
- Boutrot, F.; Zipfel, C. Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annu. Rev. Phytopathol.* **2017**, *55*, 257–286. [[CrossRef](#)] [[PubMed](#)]
- Lolle, S.; Stevens, D.; Coaker, G. Plant NLR-triggered immunity: From receptor activation to downstream signaling. *Curr. Opin. Immunol.* **2020**, *62*, 99–105. [[CrossRef](#)]
- Yasuda, S.; Okada, K.; Saijo, Y. A look at plant immunity through the window of the multitasking coreceptor BAK1. *Curr. Opin. Plant Biol.* **2017**, *38*, 10–18. [[CrossRef](#)]
- Wang, Y.; Zhang, Y.; Zhou, R.; Dossa, K.; Yu, J.; Li, D.; Liu, A.; Mmadi, M.A.; Zhang, X.; You, J. Identification and characterization of the bZIP transcription factor family and its expression in response to abiotic stresses in sesame. *PLoS ONE* **2018**, *13*, e0200850. [[CrossRef](#)] [[PubMed](#)]
- Roux, M.; Schwessinger, B.; Albrecht, C.; Chinchilla, D.; Jones, A.; Holton, N.; Malinovskiy, F.G.; Tör, M.; de Vries, S.; Zipfel, C. The Arabidopsis Leucine-Rich Repeat Receptor-Like Kinases BAK1/SERK3 and BKK1/SERK4 Are Required for Innate Immunity to Hemibiotrophic and Biotrophic Pathogens. *Plant Cell* **2011**, *23*, 2440–2455. [[CrossRef](#)] [[PubMed](#)]
- Balint-Kurti, P.J.; Wissler, R.; Zwonitzer, J.C. Use of an Advanced Intercross Line Population for Precise Mapping of Quantitative Trait Loci for Gray Leaf Spot Resistance in Maize. *Crop Sci.* **2008**, *48*, 1696–1704. [[CrossRef](#)]
- Maschietto, V.; Colombi, C.; Pirona, R.; Pea, G.; Strozzi, F.; Marocco, A.; Rossini, L.; Lanubile, A. QTL mapping and candidate genes for resistance to *Fusarium* ear rot and fumonisin contamination in maize. *BMC Plant Biol.* **2017**, *17*, 20. [[CrossRef](#)] [[PubMed](#)]

27. Luan, J.; Wang, F.; Li, Y.; Zhang, B.; Zhang, J. Mapping quantitative trait loci conferring resistance to rice black-streaked virus in maize (*Zea mays* L.). *TAG Theor. Appl. Genet. Theor. Angew. Genet.* **2012**, *125*, 781–791. [[CrossRef](#)] [[PubMed](#)]
28. Lv, X.-L.; Li, X.-H.; Xie, C.-X.; Hao, Z.; Ji, H.; Shi, L.; Zhang, S. Comparative QTL mapping of resistance to sugarcane mosaic virus in maize based on bioinformatics. *Hereditas* **2008**, *30*, 101–108. [[CrossRef](#)] [[PubMed](#)]
29. Qiu, H.; Tan, K.; Li, C.; Shen, T.; Zhang, Z. Identification of QTL for resistance to head smut in maize (*Zea mays* L.). *Euphytica* **2021**, *217*, 185–192. [[CrossRef](#)]
30. Stanley, A.E.; Menkir, A.; Ifie, B.; Paterne, A.A.; Unachukwu, N.N.; Meseke, S.; Mengesha, W.A.; Bossey, B.; Kwadwo, O.; Tongoona, P.B.; et al. Association analysis for resistance to *Striga hermonthica* in diverse tropical maize inbred lines. *Sci. Rep.* **2021**, *11*, 24193–24207. [[CrossRef](#)]
31. Pè, M.E.; Gianfranceschi, L.; Taramino, G.; Tarchini, R.; Angelini, P.; Dani, M.; Binelli, G. Mapping quantitative trait loci (QTLs) for resistance to *Gibberella zeae* infection in maize. *Mol. Gen. Genet. MGG* **1993**, *241*, 11–16.
32. Robertson-Hoyt, L.A.; Jines, M.P.; Balint-Kurti, P.J.; Kleinschmidt, C.E.; White, D.G.; Payne, G.A.; Maragos, C.M.; Molnár, T.L.; Holland, J.B. QTL Mapping for Fusarium Ear Rot and Fumonisin Contamination Resistance in Two Maize Population. *Crop Sci.* **2006**, *46*, 1734–1743. [[CrossRef](#)]
33. Berger, D.K.; Carstens, M.; Korsman, J.N.; Middleton, F.; Kloppers, F.J.; Tongoona, P.; A Myburg, A. Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genet.* **2014**, *15*, 60–72. [[CrossRef](#)] [[PubMed](#)]
34. Bartholomew, E.S.; Black, K.; Feng, Z.; Liu, W.; Shan, N.; Zhang, X.; Wu, L.; Bailey, L.; Zhu, N.; Qi, C.; et al. Comprehensive Analysis of the Chitinase Gene Family in Cucumber (*Cucumis sativus* L.): From Gene Identification and Evolution to Expression in Response to *Fusarium oxysporum*. *Int. J. Mol. Sci.* **2019**, *20*, 5309. [[CrossRef](#)] [[PubMed](#)]
35. Cheng, H.; Shao, Z.; Lu, C.; Duan, D. Genome-wide identification of chitinase genes in *Thalassiosira pseudonana* and analysis of their expression under abiotic stresses. *BMC Plant Biol.* **2021**, *21*, 87–98. [[CrossRef](#)] [[PubMed](#)]
36. Wang, S.; Fu, J. Insights into Auxin Signaling in Plant–Pathogen Interactions. *Front. Plant Sci.* **2011**, *2*, 74. [[CrossRef](#)] [[PubMed](#)]
37. Navarro, L.; Dunoyer, P.; Jay, F.; Arnold, B.; Dharmasiri, N.; Estelle, M.; Voinnet, O.; Jones, J.D.G. A Plant miRNA Contributes to Antibacterial Resistance by Repressing Auxin Signaling. *Science* **2006**, *312*, 436–439. [[CrossRef](#)]
38. Zhang, H.; Tan, X.; Li, L.; He, Y.; Hong, G.; Li, J.; Lin, L.; Cheng, Y.; Yan, F.; Chen, J.; et al. Suppression of auxin signalling promotes rice susceptibility to Rice black streaked dwarf virus infection. *Mol. Plant Pathol.* **2019**, *20*, 1093–1104. [[CrossRef](#)] [[PubMed](#)]
39. Liu, P.-P.; Koizuka, N.; Martin, R.C.; Nonogaki, H. The BME3 (Blue Micropylar End 3) GATA zinc finger transcription factor is a positive regulator of Arabidopsis seed germination. *Plant J.* **2005**, *44*, 960–971. [[CrossRef](#)] [[PubMed](#)]
40. Manzoor, M.A.; Sabir, I.A.; Shah, I.H.; Wang, H.; Yu, Z.; Rasool, F.; Mazhar, M.Z.; Younas, S.; Abdullah, M.; Cai, Y. Comprehensive Comparative Analysis of the GATA Transcription Factors in Four Rosaceae Species and Phytohormonal Response in Chinese Pear (*Pyrus bretschneideri*) Fruit. *Int. J. Mol. Sci.* **2021**, *22*, 12492. [[CrossRef](#)] [[PubMed](#)]
41. Chen, X.; Guo, Z. Tobacco OPBP1 enhances salt tolerance and disease resistance of transgenic rice. *Int. J. Mol. Sci.* **2008**, *9*, 2601–2613. [[CrossRef](#)]
42. Zhao, Y.; Wei, T.; Yin, K.Q.; Chen, Z.; Gu, H.; Qu, L.J.; Qin, G. *Arabidopsis* RAP2.2 plays an important role in plant resistance to *Botrytis cinerea* and ethylene responses. *New Phytol.* **2012**, *195*, 450–460. [[CrossRef](#)]
43. Liu, D.; Chen, X.; Liu, J.; Ye, J.; Guo, Z. The rice ERF transcription factor OsERF922 negatively regulates resistance to Magnaporthe oryzae and salt tolerance. *J. Exp. Bot.* **2012**, *63*, 3899–3911. [[CrossRef](#)] [[PubMed](#)]

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.