Received: 19 July 2024 | Accepted: 15 January 2025

DOI: 10.1002/tpg2.70006

The Plant Genome

ORIGINAL ARTICLE

Special Section: Multi-Omics Prediction in Plant Breeding

Using genome-wide associations and host-by-pathogen predictions to identify allelic interactions that control disease resistance

Owen Hudson[1] | **Jeremy Brawner**[1,2]

[1]Department of Plant Pathology, University of Florida, Gainesville, Florida, USA

[2]Genics Ltd., Saint Lucia, Queensland, Australia

Correspondence
Owen Hudson, Department of Plant Pathology, University of Florida, Gainesville, FL, USA.
Email: owen.hudson@ufl.edu

Assigned to Associate Editor Diego Jarquin.

Funding information
U.S. Department of Agriculture, Grant/Award Number: 1022115; National Institute of Food and Agriculture, Grant/Award Numbers: 13117320, Project No. FLA-PLP-006039

Abstract

Characterizing the molecular mechanisms underlying disease symptom expression has been used to improve human health and disease resistance in crops and animal breeds. Quantitative trait loci and genome-wide association studies (GWAS) are widely used to identify genomic regions that are involved in disease progression. This study extends traditional GWAS significance tests of host and pathogen marker main effects by utilizing dual-genome reaction norm models to evaluate the importance of host-single nucleotide polymorphism (SNP) by pathogen-SNP interactions. Disease symptom severity data from Fusarium ear rot (FER) on maize (*Zea mays* L.) is used to demonstrate the use of both genomes in genomic selection models for breeding and the identification of loci that interact across organisms to impact FER disease development. Dual genome prediction models improved heritability estimates, error variances, and model accuracy while providing predictions for host-by-pathogen interactions that may be used to test the significance of SNP–SNP interactions. Independent GWAS for maize and *Fusarium* populations identified significantly associated loci and predictions that were used to evaluate the importance of interactions using two different association tests. Predictions from dual genome models were used to evaluate the significance of the SNP–SNP interactions that may be associated with population structure or polygenic effects. As well, association tests incorporating host and pathogen markers in models that also included genomic relationship matrices were used to account for population structure. Subsequent evaluation of protein–protein interactions from candidate genes near the interacting SNPs provides a further in silico evaluation method to expedite the identification of interacting genes.

Abbreviations: BLUP, best linear unbiased predictions; FER, Fusarium ear rot; GRM, genomic relationship matrix; GS, genomic selection; GWAS, genome-wide association study; ipTM, interface predicted template modeling; LD, linkage disequilibrium; LRR, leucine-rich repeats; QTL, quantitative trait locus; RLK, receptor-like kinase; SNP, single nucleotide polymorphism.

Plain Language Summary

Breeding for disease resistance primarily uses large host plant populations to estimate breeding values of certain plant genotypes. Combining DNA marker data from both host and pathogen populations allows for the identification of markers and genes that impact the progression of disease and those that interact to affect disease severity. Additionally, by including pathogen genomic data into breeding models, the accuracy of disease resistance and virulence predictions improves. This model can also be applied to predict the success of host plant genotypes in novel pathogen environments.

1 | INTRODUCTION

Characterizing the molecular mechanisms that underlie the expression of disease symptoms has been used to improve our understanding of human health, accelerate the development of new crop cultivars, and improve most domesticated animal breeds. Identification of the genetic regions involved in disease progression has benefited from the identification of quantitative trait loci (QTLs) that are associated with disease (Aguet et al., 2023). This has been particularly useful for the selection of individuals within crosses between well-characterized parental lines (Haley & Visscher, 1998; Villanueva et al., 2004). An alternative experimental approach to identifying markers that are predictive of phenotypes, such as disease symptoms following pathogen challenge, uses genetically diverse populations to provide data for genome-wide associate studies (GWAS) (Wijmenga, 2021). GWAS builds upon the genomic prediction framework to test for the significance of marker effects with phenotypes of interest, after accounting for similarity among the individuals with observations (Visscher & Goddard, 2019; Yang et al., 2010). GWAS uses DNA markers, often large panels of single nucleotide polymorphism (SNP) markers, to identify nucleotide or indel variants that are within or near genes impacting the phenotypic expression of traits (Abdellaoui et al., 2023). Incorporating individual variants and relatedness estimates from genome-wide sets of markers into prediction models has been used to improve predictions for binary and quantitative traits, ranging from resistance to partial resistance or tolerance (Crossa et al., 2017; Eller et al., 2008; Meuwissen et al., 2001; Visscher et al., 2021; Yang et al., 2010).

Similar approaches with diverse pathogen populations have also been used to identify regions of the genome that regulate disease and to predict pathogen virulence (Chen et al., 2024; Demirjian et al., 2023). Most association studies and prediction models target disease resistance in host populations, with very few studies evaluating pathogen virulence. Even fewer simultaneously assess both sides of the pathosystem, although this comprehensive approach can improve predictions for individuals and characterize genes linked to disease development (Bartha et al., 2013; Bartoli & Roux, 2017; Märkle et al., 2021; M. Wang et al., 2018; Zhang et al., 2021). When crop breeding programs artificially inoculate diverse plant populations to screen for disease resistance, a single pathogen isolate/genotype is typically used to produce phenotypic scores of disease severity (Cooper et al., 2019; Maschietto et al., 2017). However, natural pathogen populations change from season to season, and different or new pathogen genotypes may range in virulence on certain host plant genotypes, leading to disease outbreaks in plant genotypes thought to be resistant (Chitwood-Brown et al., 2021; Singh et al., 2008). To develop more robust forms of disease resistance, host-by-pathogen interactions may be included in prediction models in the same way that environmental data are included to predict genotype-by-environment ($G \times E$) interactions (Crossa et al., 2017; Jarquín et al., 2014; Purcell, 2002; Tang et al., 2022).

Disease progression and severity are controlled by interacting proteins produced by both host and pathogen genes; markers near these resistance and virulence genes may be identified with traditional GWAS. GWAS has been used to identify markers that are physically near or in linkage disequilibrium (LD) with genes regulating production of these proteins. An extension of the GWAS single marker regression model to incorporate host and pathogen marker effects, as well as their two-way interaction, provides an approach to estimate the statistical significance of host SNP by pathogen SNP interactions. Identification of genes (or loci) that interact across both organisms in a pathosystem would expedite the comparative or functional genomics methods that are often used for this purpose. In many pathosystems, disease resistance frequently occurs by a plant receptor identifying a matching pathogen virulence factor and mounting a corresponding disease response—this method aims to detect loci that interact in this manner. Models that incorporate genotypic data from both host and pathogen populations were developed to predict changes in disease phenotype when specific combinations of host and pathogen interact (Hudson, Resende et al., 2024).

Here, we extend these prediction models to estimate the significance of host and pathogen marker main effects, as well as the importance of their interactions. These statistical tests provided a short list of genes in LD with associated markers, which facilitated the complex and time-consuming process of quantifying the importance of protein–protein interactions that may mediate disease development (Bryant et al., 2022; Seong & Krasileva, 2021).

Core Ideas

- Combining disease symptom phenotypes with genome-wide DNA markers from host and pathogen improves the accuracy of genomic predictions.
- Including both host and pathogen populations permits predictions of host, pathogen, and host–pathogen interaction effects.
- Two-way genome-wide association study may be used to identify host single nucleotide polymorphism (SNP) by pathogen SNP interactions and may be used to expedite the detection of disease-associated genes.
- Increased accuracy and efficiency of identifying markers near host and pathogen genes can improve disease resistance breeding efforts.

1.1 | Characterizing host-by-pathogen interactions

We provide methods to identify interacting loci using the *Fusarium verticillioides* (previously *F. moniliforme* Sheldon, teleomorph *Gibberella moniliformis* Wineland) and *Zea mays* L. (maize) pathosystem that causes Fusarium ear rot (FER) disease in maize as an example by which other pathosystems can be examined. Maize is grown in 165 countries and is the second most widely grown and third most consumed cereal on earth (Erenstein et al., 2022). FER is endemic to all maize-growing regions worldwide and contributes to significant yield losses through physical damage of ears or the rejection of stored grains when the pathogen's mycotoxin contaminants are detected (Munkvold, 2003). Many *F. verticillioides* isolates produce fumonisins (a group of mycotoxins) as secondary metabolites, which are thought to be a defense response against other competing microbes and herbivores (Picot et al., 2010; Proctor et al., 2006). Consumption of fumonisin-contaminated maize causes deadly diseases in livestock such as horses, pigs, and chickens and is a suspected carcinogen in humans (Bucci & Howard, 1996). Variation in pathogenicity is evident, with *F. verticillioides* surviving as a hemibiotrophic pathogen that lives asymptomatically as an endophyte of maize plants and seeds before certain conditions induce a switch to pathogenic behavior and toxin production (Bacon et al., 2008; Duncan & Howard, 2010).

Because multiple methods of infection exist and not all are well characterized or fully understood, control of FER is difficult with host genetic resistance being the most effective method in reducing FER (Holland et al., 2020; Hung & Holland, 2012; Robertson-Hoyt et al., 2007). No fully resistant maize genotypes have been discovered, and resistance to FER ranges widely across maize genotypes and is considered to be a quantitative rather than a binomial trait (Clements et al., 2004; Hung & Holland, 2012; Robertson et al., 2006). Previous genomic prediction models using different maize populations have shown that QTLs associated with resistance primarily have small effects that differ across maize genotypes (Robertson-Hoyt et al., 2006; Zila et al., 2013). Heritability estimates for FER resistance range widely in the literature, with data from natural infections typically presenting a lower heritability when compared to estimates from controlled inoculation experiments (Butoto et al., 2021; Mesterházy et al., 2012; Robertson-Hoyt et al., 2006). However, realized gains in FER resistance for commercial maize varieties have been lower than expected, and phenotypic variation in symptom development and toxin production is evident across years and among maize varieties (Parsons & Munkvold, 2012). GWAS to determine loci associated with FER resistance have identified loci with small additive effects (1%–3%) on disease severity (Zila et al., 2014). QTLs associated with FER symptom development have not been confirmed molecularly and have been difficult to validate with GWAS in other maize populations (Holland et al., 2020; Robertson-Hoyt et al., 2006).

In this study, we first used a dual-genome prediction model incorporating genomic relatedness matrices (GRMs) from both the host and pathogen populations along with the covariance of host-by-pathogen interactions in a reaction norm model to provide predictions of disease symptoms that are associated with host resistance, pathogen virulence, and each host-by-pathogen interaction. We then ran two independent GWAS for the maize and *Fusarium* populations to generate shortlists of significantly associated loci to test the significance of marker interactions. Predictions for all tested host-by-pathogen combinations were then combined with marker calls for all significantly associated loci so that analysis of variance (ANOVA) provided an additional test for the significance of SNP main effects as well as their interactions across organisms. To protect against biases caused by population structure or polygenic effects, the SNP identified in the one-way GWAS was also incorporated in the dual genome reaction norm model. Using both approaches to evaluate all combinations of significantly associated loci

identified SNP–SNP interactions across our maize and *Fusarium* populations. Genes and corresponding proteins from these SNPs were then identified and evaluated for their likelihood of involvement in pathogenicity and resistance by searching the relevant literature. An additional in silico validation step was performed by estimating the strength of protein–protein interactions using the predicted protein structures from the amino acid sequences of these genes. While results were promising, without molecular validations of these proteins, this research should be understood as a proof of concept for dual-genome interaction predictions that may be used to focus comparative or functional approaches for identifying interacting host and pathogen genes using real-world data.

2 | MATERIALS AND METHODS

2.1 | Overview

The diagram in Figure S1 outlines the steps involved in generating the phenotypic and genotypic data used to fit the dual genome prediction models and identify markers significantly associated with disease symptoms before evaluating the significance of SNP–SNP interactions. Additional information regarding genomic selection (GS) model design can be found in Hudson et al. (2024).

2.2 | Host population

The maize population used in this study contained 158 varieties, comprised of 113 sweet corn and 45 field corn varieties that were selected to maximize genotypic diversity (Gage et al., 2020; Hu et al., 2021; Peixoto et al., 2024; Rogers et al., 2022). Twenty seeds of each genotype were sown in three replicated blocks of the field trial to provide 30 ears of suitable quality for disease inoculation and subsequent phenotyping of disease severity. The 30 ears of each variety were harvested 21–24 days after 50% of the plants reached their silking date. Ears from each variety were harvested and stored at room temperature overnight prior to inoculation.

2.3 | Pathogen population

The *Fusarium* population consisted of 43 *Fusarium* isolates, 41 of which were *F. verticillioides*, and two of the isolates could not be placed to species but formed an out-group clade near to the *F. verticillioides* principal clade (Hudson et al., 2024). The isolates were primarily sourced from the Agricultural Research Service culture collection (NRRL) with additional isolates provided by other researchers and collected from the field in Florida. Isolates from the field were isolated from plant material by sterilizing kernels and plating on potato dextrose agar. Spore suspensions of each fungal isolate were made by performing single spore isolations and growing them on potato dextrose agar for 10 days at room temperature before adding four 1-cm plugs from the plate to 45 mL of one-third strength potato dextrose broth that was shaken continuously for 5 days at 300 rpm. The spore suspensions were then filtered through sterilized cheesecloth to remove the agar and any mycelia. Suspensions were quantified and, depending on the concentration, diluted or concentrated to 5×10^5 spores/mL.

2.4 | Inoculations and incubation

Each set of 30 ears was divided into 10 sets of three, and each set of three ears was inoculated with one of the *Fusarium* isolates. This provided three replicates of the same variety-by-isolate combination, where each variety was inoculated with 10 different isolates in a circular design (Huber et al., 1992).

Inoculation of maize ears with *Fusarium* isolates began by surface sterilizing each ear husk with 70% ethanol and puncturing the ear through the husk with a sterile 18G needle to a 2-cm depth or until impassible resistance was reached. The needle used to puncture the ear was removed, and a second needle, attached to a Prima Tech 5cc adjustable dose vaccinator, was inserted into the hole, and 0.75 mL of the spore solution was dispensed into the wound. The needle was then removed, and the ear was enclosed in a pre-labeled plastic bag that was sealed to prevent contamination. Ears were stored in a growth chamber with a 12:12 h light:dark schedule, and the temperature was set to 30˚C to provide conditions that were near optimal for *F. verticillioides* growth (Samapundo et al., 2005). Ears were incubated for 7 days, then removed from the bag, husked, shanks and silks were removed, and any minor damage to the ear not associated with the inoculation was removed with hand shears. In cases where poor pollination (<50%) resulted in a lack of kernels, severe physical damage was caused by insects, or where additional infections were seen and may have led to an inaccurate phenotype, the ear was discarded.

2.5 | Phenotyping

Using the circular design, the 6794 potential variety-by-isolate pairs were reduced to 1400 pairs for inoculations and phenotyping of disease severity. Three ears provided replications for the host and pathogen combinations that were observed. Disease severity was assessed by counting the number of kernels with visual disease symptoms. Rather than using the subjective scales (1–9) that have been used in other FER studies, the number of infected kernels was

used to provide a quantitative measure with a larger range of values. Images of the entire surface of the unhusked ear were also obtained for reviewing data points using a custom-designed image acquisition machine called The Ear Unwrapper (Hudson et al., 2023).

2.6 | Genomic data

The maize marker data were derived from two separate sequencing projects. The sweet corn varieties were a subset of a larger population that was resequenced (Hu et al., 2021), assembled, and aligned to the reference B73 RefGen v3 from MaizeGDB (Schnable et al., 2009) using Bowtie2 (Langmead & Salzberg, 2012) with variants called using Freebayes (Garrison & Marth, 2012). The field corn varieties were obtained from the cyverse database, using the unimputed HapMap 3.1.1 VCF files, which had used B73 AGP v3 as the reference for variant calling (Qi Sun *Zea mays* haplotype map) (Bukowski et al., 2018). SNPs from the two populations were normalized using bcftools v1.15 (Danecek et al., 2011) "norm" function, and the "isec" function with the $n = 2$ option was used to generate a list of SNPs that were present in both files. The files were merged and then filtered using "view-R" so they contained only the SNPs that were listed in the isec output with 3.21 million SNPs retained. The resulting VCF for all tested maize varieties was used to generate the marker dosage data that were used to calculate a GRM with vcftools using the "–012" option (Danecek et al., 2011).

The genomes of all *Fusarium* isolates were sequenced using pair-end Illumina short-read sequencing and assembled de novo using Spades (Bankevich et al., 2012) before aligning to the *F. verticillioides* 7600 reference genome assembly (ASM14955v1). SNP calls from Freebayes were merging into a single VCF file with vcftools. Variants were initially filtered using Samtools (Li et al., 2009) quality filter of $q > 20$ to produce a total of 2.4 million SNPs. The VCF was then filtered for alleles with minor allele frequency > 0.05 and genotypes with >50% missing calls were removed (Gezan et al., 2021). LD was calculated for both populations using PLINKv 2.0 (Chang et al., 2015; Purcell et al., 2007).

2.7 | Genomic relationship matrices

Marker data were used to derive similarity matrices for hosts and pathogens that were incorporated into mixed models as GRM. Following quality control, the maize marker data from vcftools (0, 1, 2) were read into R and quality checked using ASRgenomics (Gezan et al., 2021) with a minimum minor allele frequency of 0.05 and the threshold for missing genotypes set to 50%. No additional SNPs were dropped, so the total number was 154,000 for maize. The AGHmatrix (Amadeu et al., 2016) package was used to produce the GRM using the Van Raden method (VanRaden, 2008). The maize GRM can be found in Figure S2.

For the *Fusarium* isolates, two different marker sets were used to create a mapping file to identify the SNP used to create the GRM. The mapping file used marker data from 2.4 million SNPs that were filtered using ASRgenomics with a minimum minor allele frequency of 0.05 and the threshold for missing genotypes set to 50%, which reduced the final number of markers to 1.5 million. Those SNPs were further filtered in PLINK v2.0 using the "haploid" option that was designed to produce relatedness estimates for the 0/1 gene content coding of the male X chromosome (Druet & Legarra, 2020). The PLINK-generated GRM was estimated using 386,000 *Fusarium* SNPs. The *Fusarium* GRM can be found in Figure S3.

2.8 | Dual genome prediction models

The maize and *Fusarium* GRMs were combined to build a host–pathogen similarity matrix (HPSM) for incorporation into a reaction norm model described herein as the dual genome prediction model. The dual genome prediction model was fit to provide predictions of resistance for maize varieties and virulence predictions for *Fusarium* isolates as well as restricted maximum likelihood-derived variance component estimates. Variance components were used to calculate genetic parameters that approximate the relative importance of host and pathogen main effects and their interactions. The complete linear model was defined as follows:

$$\mathbf{y}_{ij} = \mu + H_i + P_j + HP_{ij} + E_n,$$

where $\mathbf{y}_{ij}$ is the vector of n phenotypic observations; μ is the overall mean lesion size; H_i is the deviation of random host effect for i maize varieties that are independent and identically distributed following a normal distribution with a mean of zero and a variance associated with host effects, $\sim N(0,\sigma^2{}_H)$; P_j is the random pathogen effect of j *Fusarium* isolate effects, $\sim N(0,\sigma^2{}_P)$; HP_{ij} is the random interaction effect associated with the combination of host i and pathogen j, $\sim N(0,\sigma^2{}_{HP})$; and E_n is the error associated with each individual maize ear, $\sim N(0,\sigma^2{}_E)$. The GRMs for host and pathogen populations were then incorporated in the linear models as variance/covariance matrices along with the HPSM to account for similarities in interaction effects.

The variance of genetic effects in models that incorporate relatedness among experimental treatments is associated with genomic relationship matrices that connect observations using the correlation of host ($\mathbf{G}_\mathrm{H}$) and pathogen ($\mathbf{G}_\mathrm{P}$) effects rather than the design matrix that implies treatments

are independent and unrelated. When GRMs replace design matrices, this equates to observations that are normally distributed with a mean of zero and the GRM specifying the covariance matrix of observations that are distributed with the variance among genetic effects, $\sim N(0, \mathbf{G}_i\ \sigma^2{}_i)$. For the host-by-pathogen interaction effects, a reaction norm model was used so that the covariance matrix describing the similarity of each host and pathogen pair was estimated as the Hadamard product (represented by "•") of relatedness matrices $\mathbf{G}_H$ and $\mathbf{G}_P$. To obtain a covariance matrix that aligned the host-by-pathogen interaction effects, the GRMs were expanded to align host and pathogen pairs with their corresponding interaction, where $\mathbf{G}_{HP} = (\mathbf{Z}_H\mathbf{G}_H\mathbf{Z}_H') \bullet (\mathbf{Z}_P\mathbf{G}_P\ \mathbf{Z}_P')$, and $\mathbf{Z}$ is the design matrix used to align the host and pathogen treatment levels of the relatedness matrix with the associated interaction term (Crossa et al., 2022; Jarquín et al., 2014). Using the reaction norm model, the best linear unbiased predictions (BLUPs) were estimated for each maize variety, *Fusarium* isolate, and each variety-by-isolate combination. The predictions for any combinations of maize variety and *Fusarium* isolate were calculated as the sum of the main effects and their specific interaction.

2.8.1 | Genome-wide associations

A GWAS was performed for both the host (maize) and the pathogen (*Fusarium*) to identify significantly associated loci. Traditional GWAS analyses for host and pathogen were performed using ASreml and ASRgwas (Butler et al., n.d.) in R with the GRM included for maize ($\mathbf{G}_H$) and Fusarium GRM ($\mathbf{G}_P$). The model used for association testing of the maize population's 154,000 markers that remained after quality control included the first three principal components to account for population structure with random effects included to account for *Fusarium* isolates and the maize-by-*Fusarium* interactions. The model used for association testing of the *Fusarium* population's 386,000 markers that remained after quality control included the first four principal components to account for population structure with random effects included to account for maize varieties and the maize variety by Fusarium isolate interactions. Using a threshold of 0.0005 for the *p*-values evaluating the significance of marker effects, 18 maize markers and 28 *Fusarium* markers were identified. Associated markers were then filtered for collinearity prior to incorporation as fixed effects in a linear model for evaluation using backward selection. Backward selection of markers reduced the set of candidate SNP to 14 maize markers and 19 *Fusarium* markers (Table 1). Code and necessary files for a reduced demonstration, including both association tests, are provided in the Supporting Information called "R demo materials."

2.8.2 | Tests of significance for host–pathogen marker interactions

To assess the significance of the host and pathogen main effects and marker interaction effects, we used two methods. Our first method used genomic relationships in each marker prediction, potentially identifying markers fixed in the host and pathogen populations and not separated from family effects (Table S1). The host-by-pathogen BLUPs from the reaction norm model were used as the response for a two-way ANOVA. Each prediction for all observed host-by-pathogen combinations was used as the dependent variables, and allele dosages for each pair of markers identified in the GWAS were used as the independent variables. The second method incorporated the GRMs independently of the marker predictions to account for population structure and eliminate family effects (Table S2). This was done by fixing the marker interaction effects and performing a Wald test on each pair of markers. Each unique pair of host and pathogen markers that were associated with disease severity from the one-way GWASs was tested for significance in this manner. The marker interactions and their corresponding significance values from the results of the first method were visualized using a three-dimensional Manhattan plot of significance (Figure 2A). This plot is 14 × 14, as five *Fusarium* markers and one maize marker were not predicted to significantly interact with any other marker and were dropped. The gray plane represents the significance level of 0.05 and peaks above that point are more significant (lower *p*-value).

To visualize the marker interactions, the marginal means of the interaction BLUPs were calculated to show the change in predicted lesion size for each maize and *Fusarium* marker. Each significant interaction was plotted to show the change in resistance/susceptibility across significantly interacting SNP pairs. Because maize is diploid, there are three options at any given locus: 0, 1, and 2, where 0 identifies the reference allele, 1 is a heterozygote, and 2 identifies the alternate allele. As *Fusarium* is haploid, there are only two options for a locus, 0 and 1, indicating reference or alternative allele, respectively. The three maize marker options are on the *x*-axis, the predicted effect on lesion size is on the *y*-axis, and the *Fusarium* markers appear as two different colored lines. Following a single line (blue or red) across the *x*-axis estimates the effect of one *Fusarium* SNP on three different maize nucleotide variants (0, 1, and 2). If the slope of the line is negative, maize varieties with alternate alleles are predicted to be more resistant to FER. Looking across the two lines at the same *x*-axis point (or maize allele) shows the variant effect at the given locus in *Fusarium*. Whichever line has a greater *y*-axis value is predicted to be more virulent against the corresponding maize variety.

TABLE 1 All significant loci from single GWAS.

Maize marker	Marker #	Chromosome	Position	Effect	*p*-value	Nearest gene
1_199648699	1	1	199,648,699	7.36	0.0000422	*GRMZM2G478417-A0A317YHY4-BZIP transcription factor TRAB1 OR *GRMZM5G881887-O49163-NADPH HC-toxin reductase 1
2_48494867	2	2	48,494,867	8.76	0.0000155	*GRMZM2G115395-B4FG32-F-box domain-containing protein
4_7395427	3	4	7,395,427	14.04	0.0000486	GRMZM2G426415-Hydroquinone Glycosyltransferase—K7TZ17 OR GRMZM2G122706-DUF629 domain-containing protein—A0A1D6PR59
4_8905798	4	4	8,905,798	−6.56	0.0000332	Zm00001eb134550-Aminotran_1_2 domain-containing protein
4_48834749	5	4	48,834,749	−5.54	0.0000474	*GRMZM2G156737-K7U2G7-AP2-EREBP transcription factor OR GRMZM2G051374-B4FPS4-RNase H type-1 domain-containing protein
5_46004713	6	5	46,004,713	−9.56	0.0000511	GRMZM2G129304-C0PHE9-ACT domain-containing protein ACR OR *GRMZM2G095631-K7U2B4-HEAT repeat family protein (has LLR)
6_61884169	7	6	61,884,169	−9.89	0.0000010	*GRMZM2G461716-Aspartyl protease—C4J5P0
6_154973451	8	6	154,973,451	−12.58	0.0000011	GRMZM2G163440-C4JBK6-Pentatricopeptide repeat-containing protein OR *GRMZM2G174585-Putative LRRRLK—A0A1D6M6T3
7_52501993	9	7	52,501,993	−10.02	0.0000023	*Zm00001d019585-has a zinc finger CChC domain—GRMZM2G414935
7_152959225	10	7	152,959,225	−8.63	0.0000099	GRMZM2G048907-P0C1M0-ATP synthase subunit gamma, chloroplastic (Inceptin)
7_172143953	11	7	172,143,953	6.35	0.0000279	*GRMZM2G041595-Q49B42-H(+)/Pi cotransporter (transmembrane) OR *GRMZM2G047995-B4FYV7-Fungal lipase-like domain-containing protein
8_49353655	12	8	49,353,655	15.87	0.0000487	GRMZM2G083836-B6SVU1-asparagine–tRNA ligase
9_65139022	13	9	65,139,022	6.32	0.0000407	*GRMZM2G148400-B4FA15-Glycosyl hydrolase superfamily protein
9_85722220	14	9	85,722,220	−15.32	0.0000337	GRMZM2G079484-Putative cytochrome P450 superfamily protein
***Fusarium* marker**	**Marker #**	**Chromosome**	**Position**	**Effect**	***p*-value**	**Nearest gene**
NC_031675.1_3709173	1	1	3,709,173	−5.93	0.000166	*FVEG_00846-W7LNL0-IDI-2 (Isopentenyl-Diphosphate Delta Isomerase 2) (effector) OR FVEG_00845- W7LNR9-Uncharacterized
NC_031676.1_131350	2	2	131,350	4.23	0.000420	FVEG_15739-W7MBH6-Alpha/beta hydrolase fold-3 domain-containing protein OR *FVEG_05846-W7LZY2-Glucose 1-dehydrogenase
NC_031676.1_4321742	NA	2	4,321,742	−5.35	0.000546	FVEG_03569-W7M1M8-Uncharacterized protein
NC_031677.1_427688	3	3	427,688	5.10	0.000420	FVEG_12662-W7NDR2-Transcription factor domain-containing protein
NC_031677.1_654133	4	3	654,133	4.93	0.000294	*FVEG_12578-W7MT90-Neutral ceramidase
NC_031679.1_1235606	5	5	1,235,606	5.58	0.000338	*FVEG_02990-W7LZ33-WEE/WEE-UNCLASSIFIED protein kinase (phosphotransferase)
NC_031679.1_4094419	NA	5	4,094,419	4.08	0.000573	FVEG_09396-W7MQP7-Lactoylglutathione lyase
NC_031680.1_340806	6	6	340,806	−5.31	0.000453	FVEG_13197-W7NFV1-AAA+ ATPase domain-containing protein
NC_031680.1_3288082	7	6	3,288,082	4.28	0.000485	*FVEG_01749-W7LGL2-Rhodopsin domain-containing protein
NC_031681.1_107262	8	7	107,262	4.25	0.000528	FVEG_06531-W7M4J0-Amine oxidase domain-containing protein
NC_031681.1_583383	9	7	583,383	3.99	0.000473	*FVEG_06718-W7MNA4-Spindle assembly checkpoint component MAD1

(Continues)

TABLE 1 (Continued)

Fusarium marker	Marker #	Chromosome	Position	Effect	*p*-value	Nearest gene
NC_031681.1_2974539	10	7	2,974,539	5.85	0.000044	FVEG_11714-W7MQ65-N1-acetylpolyamine oxidase
NC_031682.1_1602078	NA	8	1,602,078	3.98	0.000545	*FVEG_07235-W7M5Y9-NmrA-like domain-containing protein
NC_031682.1_2412241	11	8	2,412,241	5.90	0.000247	*FVEG_13405-W7MVS1–Laccase
NC_031682.1_2647309	NA	8	2,647,309	−5.23	0.000052	FVEG_13507-W7MW18-dTDP-glucose 4,6-dehydratase
NC_031684.1_779152	12	10	779,152	−5.04	0.000132	*FVEG_16453-W7MNU9-beta-glucosidase
NC_031684.1_1938363	13	10	1,938,363	4.07	0.000535	FVEG_08410-W7MW76-Peptidase C14 caspase domain-containing protein
NW_017387869.1_19023	14	17	19,023	5.48	0.000537	NA—similar to FPOA_00001-has helicase domain
NW_017387865.1_30294	NA	13	30,294	6.71	0.000375	FVEG_14068-W7MXZ5-Uncharacterized protein

Note: Nearest gene includes the coded protein and any predicted domains. Marker # refers to the three-dimensional Manhattan plot in Figure 2A. Asterisks (*) indicate genes likely involved in disease progression.
Abbreviation: GWAS, genome-wide association study; LRR, leucine-rich repeats; RLK, receptor-like kinase.

2.8.3 | Marker analysis

The locations of all significant markers were used to determine the nearest putative gene in both genomes. When the marker was near two genes, both genes and their predicted function were included in the list of candidate genes (Table 1). Due to the much larger genome size and fewer SNPs in the maize marker set, only genes labeled as "high confidence" and no transposable elements were listed for the maize markers. If more than two genes were identified within 100 kb of the marker region, then the two most likely to be involved in disease were included in the list following a literature search or evaluation of structural components. If no genes were found within 100 kb of the marker, then the nearest marker was selected. Genes of interest were then investigated for their potential involvement in pathogenicity or disease resistance. First, a literature search was used to determine if that gene had been previously identified as disease-related (a likely pathogenicity or resistance factor for pathogen and host, respectively), giving greater significance to studies containing molecular validations. Next, we performed a structural analysis of each protein product to determine if the protein contained domains that were known to be involved in pathogenicity or resistance. Examples for plant resistance domains include nucleotide binding sites, leucine-rich repeats (LRR), receptor-like kinase (RLK), and coiled coil or toll/interleukin domains. For *Fusarium*, we used SignalP to predict genes with signal peptides and then EffectorP to predict effectors based on conserved pathogenicity-related motifs. Pathogen genes that were not predicted to be effectors were submitted to the pathogen–host interaction database BLAST tool (PHIB-BLAST) to determine if similar sequences were associated with pathogenicity in other pathosystems (Urban et al., 2022). Finally, protein sequences for each gene that was considered a probable pathogenicity or resistance factor were used to predict their structure using AlphaFold. The strength of the interaction between each resistance and pathogenicity-associated protein pair was then estimated with AlphaFold-Multimer via AlphaPulldown v0.30.7 (Yu et al., 2023).

3 | RESULTS

3.1 | Phenotypic data

The data, viewed as the average value of the replications (range = 3–9 replications) for each unique host-by-pathogen combination, were skewed to the right but normally distributed according to a Shapiro–Wilk test of normality, producing a p-value = 0.071 (W = 0.999). Heritability estimates produced from the dual genome selection model were

0.25 for the host, 0.07 for the pathogen, and the proportion of variance accounted for by the host-by-pathogen interaction was 0.09. These estimates were slightly lower than an alternative GS model that used 50 *Fusarium* isolates (Hudson et al., 2024) and were similar in magnitude to what has been found in other dual genome prediction models (Bartoli & Roux, 2017; Märkle et al., 2021; M. Wang et al., 2018; Zhang et al., 2021).

3.2 | Population structure

The maize population consisted of two distinct clusters, as shown in the principal component analysis plot (Figure S4A), one cluster containing the sweet corn varieties and the other cluster containing the field corn varieties. There was a third grouping of three varieties, which were similar but dissimilar to the other groups. This defined the $K = 3$ option that was used to account for population structure in one-way GWAS, which was reinforced by the screeplot (Figure S4B). The *Fusarium* population also formed three clusters (Figure S5A), although the grouping was not as well defined as it was for maize. One cluster did show a potential pattern of recently cultured isolates from the central and southern United States (Florida, North Carolina, Nebraska, and Tennessee). The *Fusarium* screeplot (Figure S5B) did not distinguish $K = 3$ as clearly as the maize screeplot; however, it was the obvious minimum number and increasing K did not result in discovery of novel loci that were not collinear. Due to the selection of varieties to include a high level of genetic diversity in the maize population, the analysis of LD revealed an average $r^2 = 0.34$ and 14 kb between each SNP with the LD decay decreasing slowly to reach $r^2 < 0.2$ at 750 kb.

3.3 | Host GWAS and pathogen GWAS

For both association studies (host and pathogen), the Bonferroni-defined significance threshold (Bonferroni = 0.05/number of markers) did not identify any markers to be associated with disease severity. When the significance threshold was reduced (5×10^{-5} for host and 5×10^{-4} for pathogen), 14 maize SNPs were associated with FER (Figure 1A). The most significant marker, 6_61884169, had a p-value of 1×10^{-6} (Table 1). In the pathogen GWAS, 19 *Fusarium* SNPs were associated with FER (Figure 1B), the most significant of which was NC_031681.1_2974539 marker, which had a p-value of 4.4×10^{-5} (Table 1). Overall, the maize SNPs were more significantly associated with FER than the *Fusarium* SNPs, as was expected with the lower heritability estimate for fusarium.

3.4 | Host–pathogen interaction analysis

Table 1 presents all significant (and not colinear) markers obtained from each single GWAS of host and pathogen. These markers were tested for significant host-by-pathogen interactions, and the results are displayed in Table S1. From the 14 maize SNPs and 19 *Fusarium* SNPs found to be associated with FER, there were a possible 266 unique SNP interactions, and 45 of these interactions were statistically significant (p-value < 0.05) using the two-way ANOVA (Table S1). Of the 14 maize and 19 *Fusarium* SNPs, 12 of the maize SNPs and 15 of the 19 *Fusarium* SNPs were found to have a significant SNP–SNP interaction. The dual-genome Manhattan plot (Figure 2A) displays the 196 most significant SNP interactions, 45 of which were significant and shown as peaks above the gray plane, which indicates significance of 0.05. The two most significant interactions identified a *Fusarium* SNP (NW_017387869.1_19023) with no predicted gene within a large distance (>10Kb) of the SNP.

Using the complete reaction norm model with SNP pairs and their interaction included as fixed effects, 39 of the tested 266 interactions were found to be significant (Table S2). Between the marker combinations that were significantly associated using both models, only four interactions were the same and are highlighted in pink in Tables S1 and S2.

3.5 | SNP interactions

We plotted the change in predicted effect across both maize and *Fusarium* variants for each predicted interaction. Figure 2B shows the interaction plot of markers 6_154973451 (a putative LRR-RLK) in maize and NC_031677.1_654133 (a neutral ceramidase) in *Fusarium*. Figure 2C shows the protein–protein interaction of these two proteins as predicted by AlphaFold Multimer via AlphaPulldown. The interface predicted template modeling (ipTM) score was 0.698 for this interaction (ranging 0–1), and the predicted aligned error plot is provided in Figure S6. The average ipTM score for all other predicted interactions was 0.26.

4 | DISCUSSION

The modern cultivars of maize and wheat that provided the foundation for the Green Revolution's increase in global food production were developed following intense selection for disease resistance with no clear understanding of the interactions between host and pathogen (Borlaug & Dowswell, 1995; Khush, 1999). Progress in developing controlled disease screening systems, genotyping, experimental design,

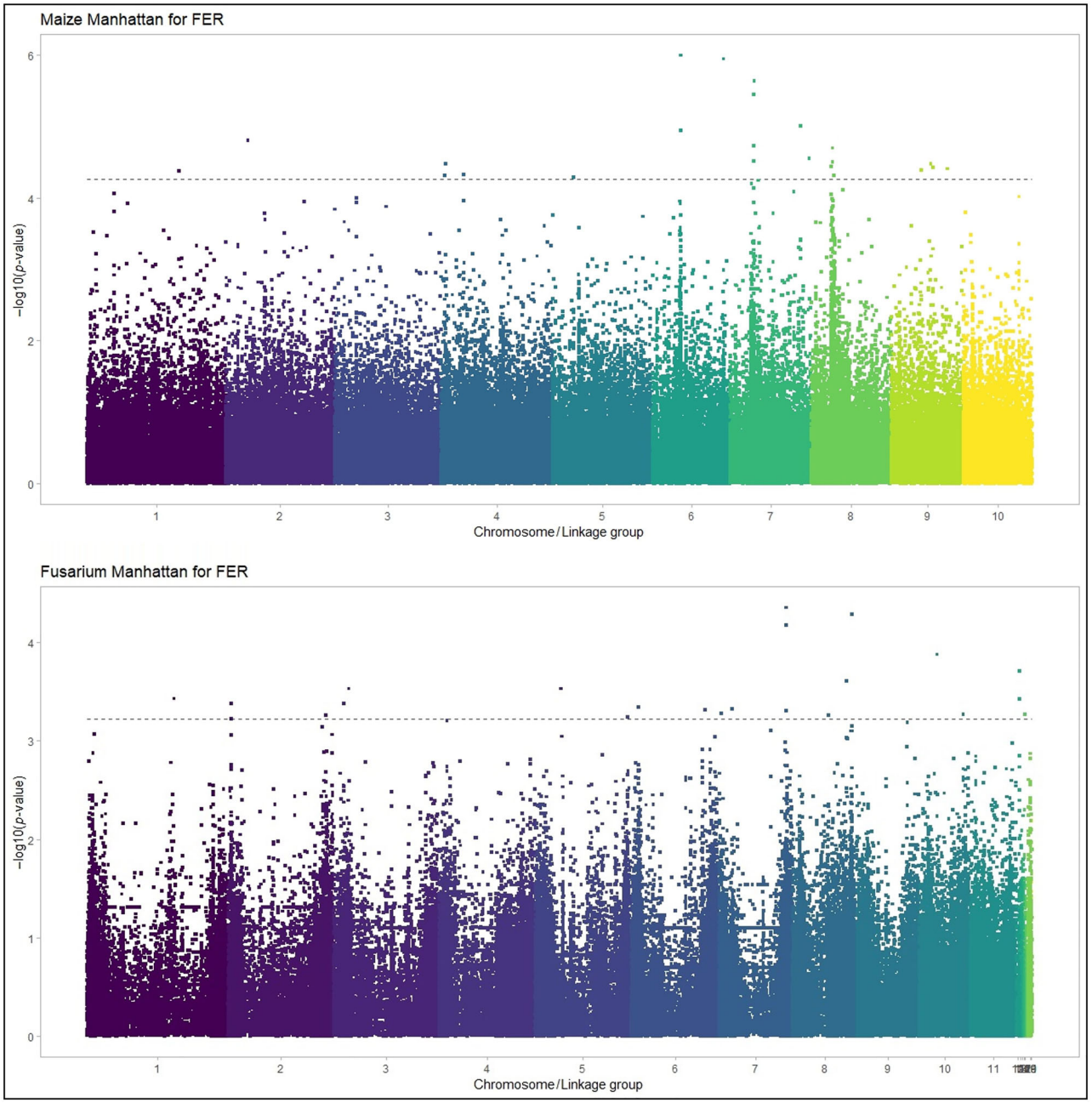

FIGURE 1 Manhattan plots of one-way GWAS. (Top) Maize GWAS Manhattan plot from the "standard" dataset. Dashed line indicating significance = 5×10^{-5}. (Bottom) *Fusarium* GWAS Manhattan plot from the standard dataset. Dashed line indicating significance level = 5×10^{-4}. Bzip, basic leucine zipper; dTDP, thymidine diphosphate; FER, Fusarium ear rot; LLR, leucine rich repeat; MAD, mitotic arrest development; NA, not available.

imaging, and molecular phenotyping provides opportunities to improve the accuracy of selection and accelerate the fixation of beneficial alleles in breeding populations (Merrick et al., 2021; J. Poland & Rutkoski, 2016; S. Ye et al., 2020).

Relationships between plants and microbes are controlled by interacting proteins, whether that be symbiotic mutualists or phytopathogenic attacks. Classically, plant resistance (R) gene products that recognize the pathogen a/virulence (avr) protein, the plant will mount a defense response, but if the pathogen protein goes undetected, the disease can progress (Biezen & Jones, 1998). However, protein–protein interactions govern less specific interactions, such as general plant cell wall receptors that mount a generalized defense response when in contact with microbe-associated molecular patterns. Quantitative disease resistance (QDR), such as in FER, is thought to be controlled by a combination of specific and general genetic interactions, but to what extent each is responsible is still unknown. Using methods

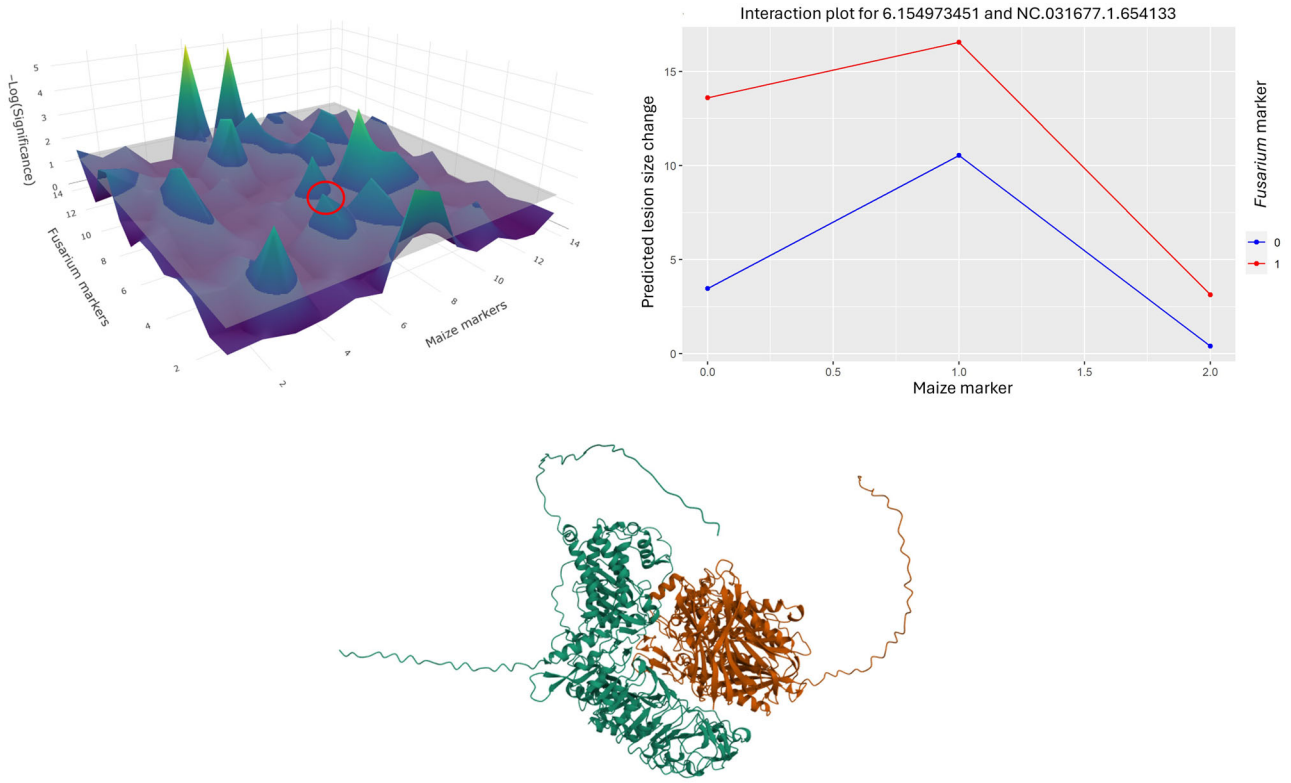

FIGURE 2 Interaction analysis. (Top left) three-dimensional Manhattan plot of the top 14 maize and *Fusarium* single nucleotide polymorphisms (SNPs) identified by associating SNP with breeding value predictions. Significance is displayed as the −log (*p*-value); thus, the gray plane at $Z = 1.3$ indicates a significance level of 0.05. Circled is the SNP–SNP (and protein–protein) interaction investigated in the top right and bottom figures, between markers 6_154973451 and NC_031677.1_654133. (Top right) Interaction plot of the indicated markers. The maize marker/allele is displayed as 0, 1, or 2, where 0 is identical to the reference, 1 is heterozygous to the reference, and 2 is homozygous alternate to the reference. The haploid *Fusarium* marker is differentiated in colors such that 0 is blue and is identical to the reference and 1 is red and is alternate to the reference. The *Y* is the predicted lesion size change following the interaction of a given combination of markers/alleles. (Bottom) The protein–protein interaction structure of the gene products from the indicated markers. Green is the maize protein (GRMZM2G174585—a putative leucine-rich repeats [LRR]-receptor-like kinase [RLK]) and orange is the *Fusarium* protein (FVEG_12578—a neutral ceramidase).

such as the dual-genome model described here, any statistically significant combination of loci can be identified, whether it be general or specific (Corwin & Kliebenstein, 2017; Gou et al., 2023; J. A. Poland et al., 2009). Identification of interacting SNPs provides a first step in prioritizing genes for functional studies and providing and understanding their function for improving disease resistance in crops. We present this study as a methodology that can be applied to a wide range of pathosystems, those with specific R–avr gene interactions, and general defense-eliciting protein–receptor interactions that contribute to QDR.

4.1 | Moving from one-way to two-way GWAS

Traditional single-genome GWAS facilitates the identification of QTLs; however, they do not associate any allele, SNP, or QTL with any other gene target, nor do they validate the result. Molecular assays can validate the interaction between individual protein–protein pairs but are time- and resource-intensive and require the targets to be identified previously. GWAS using diverse host and pathogen populations may be used to identify candidate genes and estimate the significance of their interactions to prioritize functional validation studies (Y. Wu et al., 2023). Our goal was to use a sparse testing design to produce sufficient phenotypic data to associate markers from the pathogen (*Fusarium*) and markers from the host (maize) that indicate an interaction between organisms occurring at two specific loci. We used FER because it is a quantitative pathosystem and many studies have found different alleles associated with disease resistance and susceptibility (da Silva et al., 2022; Ju et al., 2017; Lanubile et al., 2017; Liao, 2023; Liao et al., 2023; Robertson-Hoyt et al., 2006; L. Wu et al., 2024). While we identified markers and genes predicted to interact in the FER pathosystem, the purpose of this manuscript is to demonstrate the potential of the dual-genome prediction and GWAS

as a process for identifying genes involved in disease resistance and susceptibility. More phenotypic data from larger populations will also be required before markers could be used in applied breeding programs, and any protein–protein interaction would be validated molecularly using methods such as co-immunoprecipitation, pull-down assays, or yeast two-hybrid systems.

Rather than evaluating all 154,000 maize markers by 386,000 *Fusarium* markers in a complete dual-GWAS, we selected subsets of SNP from each to evaluate interactions in SNP that had significant main effects. By only using the significantly associated markers of each traditional GWAS, the final interaction analysis required the evaluation of a 14×19 marker matrix describing 266 interactions rather than all 231 billion possible interactions. While modeling all possible marker interactions is the optimal method to detect significance, the computational power required is currently above most, perhaps all, high-performance computing systems.

The estimated heritability of maize for the dataset was 0.25, which was slightly lower than our previous estimate (0.29), which used a *Fusarium* population that included other species of *Fusarium* (Hudson et al., 2024). The other *Fusarium* species typically did not produce disease symptoms, which inflated the estimate of genetic variance and would likely have biased association tests if the out-species were included in the GWAS. This led to lower estimates of pathogen heritability (0.07) and the proportion of variance accounted for by host–pathogen interactions (0.09), which were 0.11 and 0.15, respectively, when all isolates were included. The lower significance values of markers in the pathogen GWAS compared to the host GWAS reflect these genetic parameter estimates. Also, the proportion of the significantly associated genes from the *Fusarium* GWAS that were considered "likely related to pathogenicity" was lower (43%) than the proportion of genes from the maize GWAS that were considered "likely related to disease resistance" (55%). Our previous GS models used a 10-fold cross-validation method to measure model accuracy, resulting in accuracy rates proportional to the aforementioned heritability estimates: host = 0.33, pathogen = 0.20, and host–pathogen interaction = 0.31 (Hudson et al., 2024). Because of this, it is likely that the proportion of the markers associated with pathogenicity or resistance is inflated, with fewer genes being truly causal of the trait. However, the error variance declined from 67% and 85% in the individual host and pathogen GS models, respectively, to 46% in the dual selection model that included both host and pathogen populations. Improvements in model fit were achieved by including host, pathogen, and host–pathogen interactions, which provided a means to identify candidate genes near associated markers within host and pathogen genomes and determine the significance of each SNP–SNP interaction. This novel approach may increase the efficiency of disease resistance breeding programs that are exposed to diverse pathogen populations.

4.2 | Alternative tests of association

Using the first association test method of associating the total genetic value with SNP effects, all 266 marker interactions were tested to determine their significance, noting these markers may be associated with structure or polygenic effects. This resulted in 15 of the 19 *Fusarium* markers and 10 of the 14 maize markers providing significant (<0.05) interaction estimates that are listed in Table S1 (all markers tested are in Table 1). Quantile–quantile plots are shown in Figures S7 and S8 for *Fusarium* and maize, respectively. As association tests using genetic effects from prediction models have been shown to inflate the number of false positive associations, an additional test that conditions on population structure and polygenic effects by incorporating GRMs into the model was used and results are presented in Table S2 (Yang et al., 2010, and Abdellaoui et al., 2023). Examining both tests provides a similar approach to what was recommended for family-based association tests that utilized a two-step procedure to identify associations with both among family and within-family effects. Here, we discuss results from associations with total genetic effects, which may be associated with population structure or polygenic effects. Several of these interactions were found to be significant for both association tests and are of particular interest for further investigation, and methods to further prioritize these genes are discussed below. All 266 marker interactions were tested to determine their significance using two different methods that differed in their treatment of population structure and family effects. Results from using the first method (Table S1) are often not repeatable as association tests using genetic effects from prediction models as the response have been shown to inflate the number of false positive associations. To reflect this potential, we used a second method, which conditions on population structure and polygenic effects by incorporating GRMs into the model, and results are presented in Table S2 (Yang et al., 2010 and a 15 years of GWAS paper). Examining both tests provides a similar approach to what was recommended for family-based association tests that utilized a two-step procedure to identify associations with both among family and within-family effects (Beben et al., 2009; Laird & Lange, 2006). These methods produced different results regarding which marker pairs were significant, only four marker pairs being significant in both models.

We did a cursory analysis of the predicted marker interactions to determine if any were of particular interest by identifying the alleles closest to each significant SNP and searching the literature for pathological relevance. If the subsequent allelic interaction was between two alleles that had relevance to disease resistance or susceptibility, we further explore the relationship using AlphaFold Multimer software to predict the likelihood of two predicted proteins interacting directly. This was done to provide an additional in

silico method by which identification of trait-associated markers/alleles can be made more efficient and is not intended to validate our results. This was done for the results produced using both association tests described above. It was interesting to note that the most significant interacting *Fusarium* marker from both association tests was located on a non-chromosomal location, which was not associated with a particular allele. It was also found in one of the four interactions that were the same across the two association test results.

Nearly 60 protein interactions were modeled using AlphaFold Multimer, where the ipTM score range is 0–1 and interactions below 0.6 are likely failed predictions, scores 0.8 and higher represent high-quality predictions, and values between 0.6 and 0.8 are a gray zone. The average ipTM score of tested interactions was 0.28 but the highest ipTM score of all tested predictions, 0.69, came from the interaction between a putative LRR-RLK in maize and a neutral ceramidase in *Fusarium* and is displayed in Figure 2C. These two proteins are unlikely to physically interact, as ceramidases primarily act intracellularly while RLKs act as receptors on the outside of the cell. However, it is interesting to note that while ceramides are associated with stress-related responses and apoptosis, fumonisin acts to inhibit ceramide synthase enzymes, preventing ceramidase enzymes from performing their normal function and leading to an accumulation of sphinganine and sphingosine, cytotoxicity, and potentially apoptosis (Berkey et al., 2012; Parveen et al., 2019; Stockmann-Juvala & Savolainen, 2008; Zhu et al., 2023).

Because FER resistance is highly quantitative, the markers and genes identified in Table 1 and Tables S1 and S2 likely constitute only a small portion of many genetic factors that control FER disease severity. The above-described marker interactions were used to identify nearby genes that may produce interacting proteins that are involved in different phases of disease development. Compared to other GWAS and GS models on FER, our populations were small and although genetic diversity was high, the disease screening provided low heritability estimates as would be expected when additional variation is accounted for by interactions in linear models. However, the improvement in heritability estimates, error variances, and model accuracy gained from adding the pathogen population and the host–pathogen interaction matrix into the model suggests that accounting for variation among pathogen isolates could improve predictions from studies spanning environments or years where different pathogen populations impact the host populations. As well, incorporating similarities among varieties and isolates in prediction models provides a means to connect different controlled inoculation studies that use different varieties and isolates.

5 | CONCLUSIONS

This research was undertaken to provide a proof of concept method with a real-world dataset that demonstrates the integration of disparate sources of data into a disease screening program designed to develop crops with more durable forms of disease resistance. Using genomic connectivity to combine studies into larger populations may be used to generate much larger phenotypic data points and may be used to optimize the use of dual-genome models for disease resistance breeding. For FER specifically, molecular validation of interacting genes near QTLs and additional studies measuring fumonisin production would provide additional practical utility. Because high-throughput plant phenotyping and machine learning tools continue to improve our ability to rapidly acquire massive amounts of phenotypic data, disease screening platforms that improve in their ability to incorporate these multi-omic data sources provide an approach to deliver different forms of disease resistance to breeding programs.

AUTHOR CONTRIBUTIONS

Owen Hudson: Data curation; formal analysis; investigation; methodology; software; validation; visualization; writing—original draft; writing—review and editing. **Jeremy Brawner**: Conceptualization; formal analysis; methodology; resources; software; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

Microbial strains used in this work were provided by the USDA-ARS Culture Collection (NRRL), Dr. Heather Hallen-Adams, Dr. Heather Kelly, Dr. Autumn McLaughlin, Dr. Jim Holland, and Eric Butoto; USDA-ARS GRIN for maize seed; Drs. Anuj Sharma, Joseph Gage, and Jose Huguet-Tapia for bioinformatic assistance; Dr. Kristen Leach for help in the field; Dr. Marcio Resende and Dr. Marco Peixoto for assistance with maize genotype data; Dr. Nick Dufault and Kristin Beckham for providing room for inoculations.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Field corn genomic data are available from Cyverse or Panzea as Maize 282 AGPv3 from Bukowski et al. (2018). Sweet corn genomic data can be requested from the Resende Sweet corn lab at the University of Florida and the reference genome can be found in Hu et al. (2021). All Fusarium genomes have been uploaded to NCBI under Bioproject ID number: PRJNA1047553. Phenotypic data can be found in the Supporting Information of Hudson et al. (2024). An abbreviated example is provided in the Supporting Information

(R demo materials) to this manuscript with the R code and the necessary data files.

ORCID

Owen Hudson https://orcid.org/0000-0002-7409-2118
Jeremy Brawner https://orcid.org/0000-0003-3877-6253

REFERENCES

Abdellaoui, A., Yengo, L., Verweij, K. J. H., & Visscher, P. M. (2023). 15 years of GWAS discovery: Realizing the promise. *American Journal of Human Genetics*, *110*(2), 179–194. https://doi.org/10.1016/j.ajhg.2022.12.011

Aguet, F., Alasoo, K., Li, Y. I., Battle, A., Im, H. K., Montgomery, S. B., & Lappalainen, T. (2023). Molecular quantitative trait loci. *Nature Reviews Methods Primers*, *3*(1), 1–22. https://doi.org/10.1038/s43586-022-00188-6

Amadeu, R. R., Cellon, C., Olmstead, J. W., Garcia, A. A., Resende, M. F., Jr., & Muñoz, P. R. (2016). AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: A blueberry example. *The Plant Genome*, *9*(3), plantgenome2016-01. https://doi.org/10.3835/plantgenome2016.01.0009

Bacon, C., Glenn, A., & Yates, I. (2008). *Fusarium verticillioides*: Managing the endophytic association with maize for reduced fumonisins accumulation. *Toxin Reviews*, *27*(3–4), 411–446. https://doi.org/10.1080/15569540802497889

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., & Prjibelski, A. D. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bartha, I., Carlson, J. M., Brumme, C. J., McLaren, P. J., Brumme, Z. L., John, M., Haas, D. W., Martinez-Picado, J., Dalmau, J., & López-Galíndez, C. (2013). A genome-to-genome analysis of associations between human genetic variation, HIV-1 sequence diversity, and viral control. *eLife*, *2*, e01123. https://doi.org/10.7554/eLife.01123

Bartoli, C., & Roux, F. (2017). Genome-wide association studies in plant pathosystems: Toward an ecological genomics approach. *Frontiers in Plant Science*, *8*, Article 763. https://doi.org/10.3389/fpls.2017.00763

Beben, B., Visscher, P., & McRae, A. (2009). Family-based genome-wide association studies. *Pharmacogenomics*, *10*(2), 181–190. https://doi.org/10.2217/14622416.10.2.181

Berkey, R., Bendigeri, D., & Xiao, S. (2012). Sphingolipids and plant defense/disease: The "death" connection and beyond. *Frontiers in Plant Science*, *3*, Article 68. https://doi.org/10.3389/fpls.2012.00068

Biezen, E. A. V. D., & Jones, J. D. G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, *23*(12), 454–456. https://doi.org/10.1016/S0968-0004(98)01311-5

Borlaug, N. E., & Dowswell, C. R. (1995). Mobilising science and technology to get agriculture moving in Africa. *Development Policy Review*, *13*(2), 115–129. https://doi.org/10.1111/j.1467-7679.1995.tb00084.x

Bryant, P., Pozzati, G., & Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. *Nature Communications*, *13*(1), Article 1265. https://doi.org/10.1038/s41467-022-28865-w

Bucci, T. J., & Howard, P. C. (1996). Effect of fumonisin mycotoxins in animals. *Journal of Toxicology: Toxin Reviews*, *15*(3), 293–302. https://doi.org/10.3109/15569549609016449

Bukowski, R., Guo, X., Lu, Y., Zou, C., He, B., Rong, Z., Wang, B., Xu, D., Yang, B., Xie, C., Fan, L., Gao, S., Xu, X., Zhang, G., Li, Y., Jiao, Y., Doebley, J. F., Ross-Ibarra, J., Lorant, A., … Xu, Y. (2018). Construction of the third-generation *Zea mays* haplotype map. *GigaScience*, *7*(4), gix134. https://doi.org/10.1093/gigascience/gix134

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2009). *ASReml-R reference manual*. The State of Queensland, Department of Primary Industries and Fisheries.

Butoto, E. N., Marino, T. P., & Holland, J. B. (2021). Effects of artificial inoculation on trait correlations with resistance to Fusarium ear rot and fumonisin contamination in maize. *Crop Science*, *61*(4), 2522–2533. https://doi.org/10.1002/csc2.20551

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), s13742-015-0047-8. https://doi.org/10.1186/s13742-015-0047-8

Chen, C., Keunecke, H., Bemm, F., Gyetvai, G., Neu, E., Kopisch-Obuch, F. J., McDonald, B. A., & Stapley, J. (2024). GWAS reveals a rapidly evolving candidate avirulence effector in the Cercospora leaf spot pathogen. *Molecular Plant Pathology*, *25*(1), e13407. https://doi.org/10.1111/mpp.13407

Chitwood-Brown, J., Vallad, G. E., Lee, T. G., & Hutton, S. F. (2021). Breeding for resistance to Fusarium wilt of tomato: A review. *Genes*, *12*(11), Article 1673. https://doi.org/10.3390/genes12111673

Clements, M. J., Maragos, C. M., Pataky, J. K., & White, D. G. (2004). Sources of resistance to fumonisin accumulation in grain and Fusarium ear and kernel rot of corn. *Phytopathology*, *94*(3), 251–260. https://doi.org/10.1094/PHYTO.2004.94.3.251

Cooper, J. S., Rice, B. R., Shenstone, E. M., Lipka, A. E., & Jamann, T. M. (2019). Genome-wide analysis and prediction of resistance to Goss's wilt in maize. *The Plant Genome*, *12*(2), 180045. https://doi.org/10.3835/plantgenome2018.06.0045

Corwin, J. A., & Kliebenstein, D. J. (2017). Quantitative resistance: More than just perception of a pathogen. *The Plant Cell*, *29*(4), 655–665. https://doi.org/10.1105/tpc.16.00915

Crossa, J., Montesinos-López, O. A., Pérez-Rodríguez, P., Costa-Neto, G., Fritsche-Neto, R., Ortiz, R., Martini, J. W. R., Lillemo, M., Montesinos-López, A., Jarquin, D., Breseghello, F., Cuevas, J., & Rincent, R. (2022). Genome and environment based prediction models and methods of complex traits incorporating genotype × environment interaction. In N. Ahmadi & J. Bartholomé (Eds.), *Genomic prediction of complex traits: Methods and protocols* (pp. 245–283). Springer. https://doi.org/10.1007/978-1-0716-2205-6_9

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., Campos, G. D. L., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

da Silva, K. J., Guimarães, C. T., Tinoco, S. M. D. S., Bernardino, K. D. C., Trindade, R. D. S., Queiroz, V. A. V., da Conceição, R. R. P., Guilhen, J. H. S., de Oliveira, N. T., Damasceno, C. M. B., Noda, R. W., Dias, L. A. D. S., Guimarães, L. J. M., Melo, J. D. O., & Pastina, M. M. (2022). A genome-wide association study investigating fumonisin contamination in a panel of tropical maize elite lines. *Euphytica*, *218*(9), Article 130. https://doi.org/10.1007/s10681-022-03082-0

Demirjian, C., Vailleau, F., Berthomé, R., & Roux, F. (2023). Genome-wide association studies in plant pathosystems: Success or failure? *Trends in Plant Science*, *28*(4), 471–485. https://doi.org/10.1016/j.tplants.2022.11.006

Druet, T., & Legarra, A. (2020). Theoretical and empirical comparisons of expected and realized relationships for the X-chromosome. *Genetics Selection Evolution*, *52*(1), Article 50. https://doi.org/10.1186/s12711-020-00570-6

Duncan, K. E., & Howard, R. J. (2010). Biology of maize kernel infection by *Fusarium verticillioides*. *Molecular Plant-Microbe Interactions*, *23*(1), 6–16. https://doi.org/10.1094/MPMI-23-1-0006

Eller, M. S., Holland, J. B., & Payne, G. A. (2008). Breeding for improved resistance to fumonisin contamination in maize. *Toxin Reviews*, *27*(3–4), 371–389. https://doi.org/10.1080/15569540802450326

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Security*, *14*(5), 1295–1319. https://doi.org/10.1007/s12571-022-01288-7

Gage, J. L., Monier, B., Giri, A., & Buckler, E. S. (2020). Ten years of the maize nested association mapping population: Impact, limitations, and future directions. *The Plant Cell*, *32*(7), 2083–2093. https://doi.org/10.1105/tpc.19.00951

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*. https://doi.org/10.48550/arXiv.1207.3907

Gezan, S., de Oliveira, A., & Murray, D. (2021). *ASRgenomics: An R package with complementary genomic functions*. VSN International.

Gou, M., Balint-Kurti, P., Xu, M., & Yang, Q. (2023). Quantitative disease resistance: Multifaceted players in plant defense. *Journal of Integrative Plant Biology*, *65*(2), 594–610. https://doi.org/10.1111/jipb.13419

Haley, C. S., & Visscher, P. M. (1998). Strategies to utilize marker-quantitative trait loci associations. *Journal of Dairy Science*, *81*, 85–97. https://doi.org/10.3168/jds.S0022-0302(98)70157-2

Hammond-Kosack, K. E., & Jones, J. D. G. (1997). Plant disease resistance genes. *Annual Review of Plant Biology*, *48*(1), 575–607. https://doi.org/10.1146/annurev.arplant.48.1.575

Holland, J. B., Marino, T. P., Manching, H. C., & Wisser, R. J. (2020). Genomic prediction for resistance to Fusarium ear rot and fumonisin contamination in maize. *Crop Science*, *60*(4), 1863–1875. https://doi.org/10.1002/csc2.20163

Hu, Y., Colantonio, V., Müller, B. S., Leach, K. A., Nanni, A., Finegan, C., Wang, B., Baseggio, M., Newton, C. J., & Juhl, E. M. (2021). Genome assembly and population genomic analysis provide insights into the evolution of modern sweet corn. *Nature Communications*, *12*(1), Article 1227. https://doi.org/10.1038/s41467-021-21380-4

Huber, D. A., White, T. L., & Hodge, G. R. (1992). The efficiency of half-sib, half-diallel and circular mating designs in the estimation of genetic parameters in forestry: A simulation. *Forest Science*, *38*(4), 757–776. https://doi.org/10.1093/forestscience/38.4.757

Hudson, O., Hudson, D., Brahmstedt, C., & Brawner, J. (2023). The Ear Unwrapper: A maize ear image acquisition pipeline for disease severity phenotyping. *AgriEngineering*, *5*(3), 1216–1225. https://doi.org/10.3390/agriengineering5030077

Hudson, O., Meinecke, C. D., & Brawner, J. (2024). Comparative genomics of *Fusarium* species causing Fusarium ear rot of maize. *PLoS ONE*, *19*(10), e0306144. https://doi.org/10.1371/journal.pone.0306144

Hudson, O., Resende, M. F. R., Jr., Messina, C., Holland, J. B., & Brawner, J. (2024). Prediction of resistance, virulence, and host-by-pathogen interactions using dual genome prediction models. *Theoretical and Applied Genetics*, *137*(8), Article 196. https://doi.org/10.1007/s00122-024-04698-7

Hung, H.-Y., & Holland, J. B. (2012). Diallel analysis of resistance to Fusarium ear rot and fumonisin contamination in maize. *Crop Science*, *52*(5), 2173–2181. https://doi.org/10.2135/cropsci2012.03.0154

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., & de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics*, *127*(3), 595–607. https://doi.org/10.1007/s00122-013-2243-1

Ju, M., Zhou, Z., Mu, C., Zhang, X., Gao, J., Liang, Y., Chen, J., Wu, Y., Li, X., Wang, S., Wen, J., Yang, L., & Wu, J. (2017). Dissecting the genetic architecture of *Fusarium verticillioides* seed rot resistance in maize by combining QTL mapping and genome-wide association analysis. *Scientific Reports*, *7*(1), Article 46446. https://doi.org/10.1038/srep46446

Khush, G. S. (1999). Green revolution: Preparing for the 21st century. *Genome*, *42*(4), 646–655. https://doi.org/10.1139/g99-044

Laird, N., & Lange, C. (2006). Family-based designed in the age of large-scale gene-association studies. *Nature Reviews Genetics*, *7*(5), 385–394. https://doi.org/10.1038/nrg1839

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lanubile, A., Maschietto, V., Borrelli, V. M., Stagnati, L., Logrieco, A. F., & Marocco, A. (2017). Molecular basis of resistance to Fusarium ear rot in maize. *Frontiers in Plant Science*, *8*, Article 1774. https://doi.org/10.3389/fpls.2017.01774

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, X. (2023). Physical resistance: A different perspective on maize ear rot resistance. *Plant Growth Regulation*, *100*(3), 573–576. https://doi.org/10.1007/s10725-023-00960-y

Liao, X., Sun, J., Li, Q., Ding, W., Zhao, B., Wang, B., Zhou, S., & Wang, H. (2023). ZmSIZ1a and ZmSIZ1b play an indispensable role in resistance against Fusarium ear rot in maize. *Molecular Plant Pathology*, *24*(7), 711–724. https://doi.org/10.1111/mpp.13297

Lin, N.-C., & Martin, G. B. (2007). Pto- and Prf-mediated recognition of AvrPto and AvrPtoB restricts the ability of diverse *Pseudomonas syringae* pathovars to infect tomato. *Molecular Plant-Microbe Interactions*, *20*(7), 806–815. https://doi.org/10.1094/MPMI-20-7-0806

Liu, X.-H., Liang, S., Wei, Y.-Y., Zhu, X.-M., Li, L., Liu, P.-P., Zheng, Q.-X., Zhou, H.-N., Zhang, Y., Mao, L.-J., Fernandes, C. M., Del

Poeta, M., Naqvi, N. I., & Lin, F.-C. (2019). Metabolomics analysis identifies sphingolipids as key signaling moieties in appressorium morphogenesis and function in *Magnaporthe oryzae*. *MBio*, *10*(4), e01467–19. https://doi.org/10.1128/mBio.01467-19

Märkle, H., John, S., Cornille, A., Fields, P. D., & Tellier, A. (2021). Novel genomic approaches to study antagonistic coevolution between hosts and parasites. *Molecular Ecology*, *30*(15), 3660–3676. https://doi.org/10.1111/mec.16001

Maschietto, V., Colombi, C., Pirona, R., Pea, G., Strozzi, F., Marocco, A., Rossini, L., & Lanubile, A. (2017). QTL mapping and candidate genes for resistance to Fusarium ear rot and fumonisin contamination in maize. *BMC Plant Biology*, *17*(1), Article 20. https://doi.org/10.1186/s12870-017-0970-1

Merrick, L. F., Burke, A. B., Chen, X., & Carter, A. H. (2021). Breeding with major and minor genes: Genomic selection for quantitative disease resistance. *Frontiers in Plant Science*, *12*, Article 713667. https://doi.org/10.3389/fpls.2021.713667

Mesterházy, Á., Lemmens, M., & Reid, L. M. (2012). Breeding for resistance to ear rots caused by *Fusarium* spp. in maize—A review. *Plant Breeding*, *131*, 1–19. https://doi.org/10.1111/j.1439-0523.2011.01936.x

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Munkvold, G. P. (2003). Epidemiology of *Fusarium* diseases and their mycotoxins in maize ears. *European Journal of Plant Pathology*, *109*(7), 705–713. https://doi.org/10.1023/A:1026078324268

Parsons, M., & Munkvold, G. (2012). Effects of planting date and environmental factors on Fusarium ear rot symptoms and fumonisin B1 accumulation in maize grown in six North American locations. *Plant Pathology*, *61*(6), 1130–1142. https://doi.org/10.1111/j.1365-3059.2011.02590.x

Parveen, F., Bender, D., Law, S., Mishra, V., Chen, C., & Ke, L. (2019). Role of ceramidases in sphingolipid metabolism and human diseases. *Cells*, *8*(12), 1573. https://doi.org/10.3390/cells8121573

Peixoto, M. A., Leach, K. A., Jarquin, D., Flannery, P., Zystro, J., Tracy, W. F., Bhering, L., & Resende, M. F. R. (2024). Utilizing genomic prediction to boost hybrid performance in a sweet corn breeding program. *Frontiers in Plant Science*, *15*, 1293307. https://doi.org/10.3389/fpls.2024.1293307

Picot, A., Barreau, C., Pinson-Gadais, L., Caron, D., Lannou, C., & Richard-Forget, F. (2010). Factors of the *Fusarium verticillioides*-maize environment modulating fumonisin production. *Critical Reviews in Microbiology*, *36*(3), 221–231. https://doi.org/10.3109/10408411003720209

Poland, J., & Rutkoski, J. (2016). Advances and challenges in genomic selection for disease resistance. *Annual Review of Phytopathology*, *54*, 79–98. https://doi.org/10.1146/annurev-phyto-080615-100056

Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C., & Nelson, R. J. (2009). Shades of gray: The world of quantitative disease resistance. *Trends in Plant Science*, *14*(1), 21–29. https://doi.org/10.1016/j.tplants.2008.10.006

Proctor, R. H., Plattner, R. D., Desjardins, A. E., Busman, M., & Butchko, R. A. (2006). Fumonisin production in the maize pathogen *Fusarium verticillioides*: Genetic basis of naturally occurring chemical variation. *Journal of Agricultural and Food Chemistry*, *54*(6), 2424–2430. https://doi.org/10.1021/jf0527706

Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics*, *5*(6), 554–571. https://doi.org/10.1375/twin.5.6.554

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., & Daly, M. J. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Robertson, L. A., Kleinschmidt, C. E., White, D. G., Payne, G. A., Maragos, C. M., & Holland, J. B. (2006). Heritabilities and correlations of Fusarium ear rot resistance and fumonisin contamination resistance in two maize populations. *Crop Science*, *46*(1), 353–361. https://doi.org/10.2135/cropsci2005.0139

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. (2006). QTL mapping for Fusarium ear rot and fumonisin contamination resistance in two maize populations. *Crop Science*, *46*(4), 1734–1743. https://doi.org/10.2135/cropsci2005.12-0450

Robertson-Hoyt, L. A., Kleinschmidt, C. E., White, D. G., Payne, G. A., Maragos, C. M., & Holland, J. B. (2007). Relationships of resistance to Fusarium ear rot and fumonisin contamination with agronomic performance of maize. *Crop Science*, *47*(5), 1770–1778. https://doi.org/10.2135/cropsci2006.10.0676

Rogers, A. R., Bian, Y., Krakowsky, M., Peters, D., Turnbull, C., Nelson, P., & Holland, J. B. (2022). Genomic prediction for the germplasm enhancement of maize project. *The Plant Genome*, *15*(4), e20267. https://doi.org/10.1002/tpg2.20267

Samapundo, S., Devliehgere, F., De Meulenaer, B., & Debevere, J. (2005). Effect of water activity and temperature on growth and the relationship between fumonisin production and the radial growth of *Fusarium verticillioides* and *Fusarium proliferatum* on corn. *Journal of Food Protection*, *68*(5), 1054–1059. https://doi.org/10.4315/0362-028X-68.5.1054

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., … Wilson, R. K. (2009). The B73 maize genome: Complexity, diversity, and dynamics. *Science*, *326*(5956), 1112–1115. https://doi.org/10.1126/science.1178534

Seong, K., & Krasileva, K. V. (2021). Computational structural genomics unravels common folds and novel families in the secretome of fungal phytopathogen *Magnaporthe oryzae*. *Molecular Plant-Microbe Interactions*, *34*(11), 1267–1280. https://doi.org/10.1094/MPMI-03-21-0071-R

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Njau, P., Wanyera, R., Herrera-Foessel, S. A., & Ward, R. W. (2008). Will stem rust destroy the world's wheat crop? In D. L. Sparks (Ed.), *Advances in agronomy* (Vol. 98, pp. 271–309). Academic Press. https://doi.org/10.1016/S0065-2113(08)00205-8

Stockmann-Juvala, S., & Savolainen, K. (2008). A review of the toxic effects and mechanisms of action of fumonisin B1. *Human & Experimental Toxicology*, *27*(11), 799–809. https://doi.org/10.1177/0960327108099525

Tang, Y., You, D., Yi, H., Yang, S., & Zhao, Y. (2022). IPRS: Leveraging gene-environment interaction to reconstruct polygenic risk score. *Frontiers in Genetics*, *13*, Article 801397. https://doi.org/10.3389/fgene.2022.801397

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S. Y., Sahu, J., Iyer, S. V., Khamari, L., De Silva, N., Martinez, M. C., Pedro, H., Yates, A. D., & Hammond-Kosack, K. E. (2022). PHI-base in 2022: A multi-species phenotype database for pathogen–host interactions. *Nucleic Acids Research*, *50*(D1), D837–D847. https://doi.org/10.1093/nar/gkab1037

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Vasquez-Montaño, E., Hoppe, G., Vega, A., Olivares-Yañez, C., & Canessa, P. (2020). Defects in the ferroxidase that participates in the reductive iron assimilation system results in hypervirulence in *Botrytis cinerea*. *MBio*, *11*(4). https://doi.org/10.1128/mbio.01379-20

Villanueva, B., Dekkers, J. C. M., Woolliams, J. A., & Settar, P. (2004). Maximizing genetic gain over multiple generations with quantitative trait locus selection and control of inbreeding. *Journal of Animal Science*, *82*(5), 1305–1314. https://doi.org/10.2527/2004.8251305x

Visscher, P. M., & Goddard, M. E. (2019). From R.A. Fisher's 1918 paper to GWAS a century later. *Genetics*, *211*(4), 1125–1130. https://doi.org/10.1534/genetics.118.301594

Visscher, P. M., Yengo, L., Cox, N. J., & Wray, N. R. (2021). Discovery and implications of polygenicity of common diseases. *Science*, *373*(6562), 1468–1473. https://doi.org/10.1126/science.abi8206

Wang, J., Chen, Y.-L., Li, Y.-K., Chen, D.-K., He, J.-F., & Yao, N. (2021). Functions of sphingolipids in pathogenesis during host–pathogen interactions. *Frontiers in Microbiology*, *12*, Article 701041. https://doi.org/10.3389/fmicb.2021.701041

Wang, M., Roux, F., Bartoli, C., Huard-Chauveau, C., Meyer, C., Lee, H., Roby, D., McPeek, M. S., & Bergelson, J. (2018). Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. *Proceedings of the National Academy of Sciences*, *115*(24), E5440–E5449. https://doi.org/10.1073/pnas.1710980115

Wijmenga, C. (2021). From LD-based mapping to GWAS. *Nature Reviews Genetics*, *22*(8), 480–481. https://doi.org/10.1038/s41576-021-00366-4

Wu, L., Bian, W., Abubakar, Y. S., Lin, J., Yan, H., Zhang, H., Wang, Z., Wu, C., Shim, W., & Lu, G. (2024). FvKex2 is required for development, virulence, and mycotoxin production in *Fusarium verticillioides*. *Applied Microbiology and Biotechnology*, *108*(1), Article 228. https://doi.org/10.1007/s00253-024-13022-8

Wu, Y., Qi, T., Wray, N. R., Visscher, P. M., Zeng, J., & Yang, J. (2023). Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes. *Cell Genomics*, *3*(8), 100344. https://doi.org/10.1016/j.xgen.2023.100344

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608

Ye, S., Li, J., & Zhang, Z. (2020). Multi-omics-data-assisted genomic feature markers preselection improves the accuracy of genomic prediction. *Journal of Animal Science and Biotechnology*, *11*(1), Article 109. https://doi.org/10.1186/s40104-020-00515-5

Yu, D., Chojnowski, G., Rosenthal, M., & Kosinski, J. (2023). AlphaPulldown—A python package for protein–protein interaction screens using AlphaFold-Multimer. *Bioinformatics*, *39*(1), btac749. https://doi.org/10.1093/bioinformatics/btac749

Zhang, F., Hu, Z., Wu, Z., Lu, J., Shi, Y., Xu, J., Wang, X., Wang, J., Zhang, F., Wang, M., Shi, X., Cui, Y., Vera Cruz, C., Zhuo, D., Hu, D., Li, M., Wang, W., Zhao, X., Zheng, T., … Li, Z. (2021). Reciprocal adaptation of rice and *Xanthomonas oryzae* pv. oryzae: Cross-species 2D GWAS reveals the underlying genetics. *The Plant Cell*, *33*(8), 2538–2561. https://doi.org/10.1093/plcell/koab146

Zhu, X.-M., Li, L., Bao, J.-D., Wang, J.-Y., Daskalov, A., Liu, X.-H., Del Poeta, M., & Lin, F.-C. (2023). The biological functions of sphingolipids in plant pathogenic fungi. *PLoS Pathogens*, *19*(11), e1011733. https://doi.org/10.1371/journal.ppat.1011733

Zila, C. T., Ogut, F., Romay, M. C., Gardner, C. A., Buckler, E. S., & Holland, J. B. (2014). Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. *BMC Plant Biology*, *14*(1), Article 372. https://doi.org/10.1186/s12870-014-0372-6

Zila, C. T., Samayoa, L. F., Santiago, R., Butrón, A., & Holland, J. B. (2013). A genome-wide association study reveals genes associated with Fusarium ear rot resistance in a maize core diversity panel. *G3 Genes|Genomes|Genetics*, *3*(11), 2095–2104. https://doi.org/10.1534/g3.113.007328

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hudson, O., & Brawner, J. (2025). Using genome-wide associations and host-by-pathogen predictions to identify allelic interactions that control disease resistance. *The Plant Genome*, *18*, e70006. https://doi.org/10.1002/tpg2.70006