

Genome-Wide Association and Metabolic Pathway Analysis of Corn Earworm Resistance in Maize

Marilyn L. Warburton,* Erika D. Womack, Juliet D. Tang, Adam Thrash, J. Spencer Smith, Wenwei Xu, Seth C. Murray, and W. Paul Williams

ABSTRACT

Maize (*Zea mays mays* L.) is a staple crop of economic, industrial, and food security importance. Damage to the growing ears by corn earworm [*Helicoverpa zea* (Boddie)] is a major economic burden and increases secondary fungal infections and mycotoxin levels. To identify biochemical pathways associated with native resistance mechanisms, a genome-wide association analysis was performed, followed by pathway analysis using a gene-set enrichment-based approach. The gene-set enrichment exposed the cumulative effects of genes in pathways to identify those that contributed the most to resistance. Single nucleotide polymorphism–trait associations were linked to genes including transcription factors, protein kinases, hormone-responsive proteins, hydrolases, pectinases, xyloglucanases, and the flavonol synthase gene (in the maysin biosynthesis pathway). The most significantly associated metabolic pathways identified included those that modified cell wall components, especially homogalacturonan, wax esters, and fatty acids; those involved in antibiosis, especially 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA), flavonoids, and phenolics; and those involved in plant growth, including N uptake and energy production. The pathways identified in this study, and especially the cell wall-associated pathways, identified here for the first time, provide clues to resistance mechanisms that could guide the identification of new resistant ideotypes and candidate genes for creation of resistant maize germplasm via selection of natural variants or gene editing.

Core Ideas

- Metabolic pathway analysis increases the utility of genome-wide association study data to identify useful genes and mechanisms.
- Corn earworm (CEW) resistance mechanisms include physical barriers and antibiosis factors.
- Results can be used to create a CEW-resistant ideotype or to improve marker-assisted selection.

MAIZE is the highest-yielding and most economically important crop in the United States, boasting multibillion dollar annual revenue. It is also the main staple food crop in many countries, and total global maize production exceeds 875 million tons (Ranum et al., 2014). The corn earworm (CEW) (*Helicoverpa zea*) is a major pest throughout maize growing regions in the southeastern United States and in Central and South America (Butron and Widstrom, 2001). A related moth, *Helicoverpa armigera*, also known as the African cotton bollworm, is widespread across Europe, Asia, Africa, Australia, and Oceania, and has been recently observed in Brazil (Tay et al., 2013), prompting alarm among maize growers across the Americas. Other lepidopteran insects also threaten maize yields, including the fall armyworm (*Spodoptera frugiperda*), an important economic insect pest of maize

M.L. Warburton, E.D. Womack, J.S. Smith, W.P. Williams, USDA-ARS Corn Host Plant Resistance Research Unit, Mississippi State, MS; J.D. Tang, USDA Forest Service Forest Products Lab., Durability and Wood Protection, Starkville, MS; A. Thrash, Inst. for Genomics, Biocomputing & Biotechnology Starkville, MS; W. Xu, Dep. of Soil and Crop Sciences, Texas A&M Univ., Lubbock, TX; S.C. Murray, Dep. of Soil and Crop Sciences, Texas A&M Univ., College Station, TX. Received 4 Aug. 2017. Accepted 27 Oct. 2017.
*Corresponding author (marilyn.warburton@ars.usda.gov).

Abbreviations: CEW, corn earworm; ES, enrichment score; GBS, genotyping by sequencing; GWAS, genome-wide association study; QTL, quantitative trait locus; SNP, single nucleotide polymorphism.

Plant Genome 11:170069
doi: 10.3835/plantgenome2017.08.0069

© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
This is an open access article distributed under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of the Americas and now an invasive species spreading across West and Central Africa (Goergen et al., 2016). The implications of these new global lepidopteran threats to maize crops are disturbing, and the introduction of other invasive species cannot be ruled out.

Corn earworm larvae initially feed on exposed maize silks, but older larvae prefer to feed on the developing kernels of the ear and may burrow through the silks to consume them (McMillian et al., 1978). Corn earworm larvae are difficult to control with insecticide sprays, as the worm spends so little time outside the ear, and this may lead to substantial feeding damage. This damage promotes fungal ear rot, which can lead to decreased yield and grain quality and an increase in hazardous food safety risks (Dowd, 1998). For example, insect pests may act as a vector of the fungus *Aspergillus flavus*, which produces aflatoxin, a known human and animal carcinogen that severely reduces the market value of the grain and quality for consumption (Bhatnagar-Mathur et al., 2015). Transgenic host plant resistance using protein genes from the bacterium *Bacillus thuringiensis* has proven operative against a range of insects, particularly those in the order Lepidoptera (Chen et al., 2016). Although plants expressing *B. thuringiensis* toxins successfully protect agricultural crops against CEW and other lepidopterans, the practice of a single and widely used method for pest control applies considerable selection pressure on these pests to evolve resistance to the *B. thuringiensis* control mechanisms, which has now been reported in CEW on maize (Dively et al., 2016). Additionally, transgenic crops have been the subject of worldwide controversy and are not accepted in some markets (Wong and Chan, 2016). The development of CEW resistance to *B. thuringiensis* warrants refocused attention on endogenous resistance mechanisms for crop protection.

Natural host plant resistance is an economical, environmentally friendly, and efficient tool for controlling CEW infestation in maize. Maize lines with resistance to this lepidopteran pest have been identified and incorporated into maize breeding programs (Byrne et al., 1997; Widstrom et al., 1988; Widstrom and Snook, 2001; Wiseman and Widstrom, 1992). Host plant resistance to CEW can result from morphological features such as tight covering of the ear (Archer et al., 1994). It can also be caused by antibiosis, which is an interaction between two organisms that is damaging to one of them. Features of antibiosis include chemicals in silks that repel, retard, or kill larvae. Maysin, a C-glycosyl flavone, and related compounds (chlorogenic acid, apimaysin, and 3'-methoxymaysin) confer resistance to silk feeding by CEW in maize by inhibiting larval growth (Cortés-Cruz et al., 2003; Guo et al., 2004; Snook et al., 1994; Wiseman et al., 1992). Sources of resistant germplasm that exhibit these characteristics have been identified and have been associated with the silks, maize kernels, and husks (Widstrom and Snook, 2001; Wilson et al., 1991).

A better understanding of the genetic basis of resistance to the CEW in maize will help breeders develop superior

hybrids. Quantitative trait loci (QTLs) for CEW resistance in maize have been identified from multiple CEW-resistant maize lines. Byrne et al. (1996) identified 13 QTLs in the GT114 × GT119 F₂ mapping population, two of which accounted for almost 60% of the phenotypic variance in specific environments. One QTL was identified as being caused by the pericarp color (*pl*) locus in the flavonoid (maysin) pathway. Using the F₂ GT114 × NC7A mapping population, Lee et al. (1998) found other genes in the maysin synthesis pathway that fell under a major QTL for CEW resistance and concentrations of maysin and apimaysin. These genes include *recessive enhancer of maysin 1* and *red aleurone*. Widstrom et al. (2003) also found a major QTL for concentrations of silk maysin, 3'-methoxymaysin + apimaysin, and chlorogenic acid over the *pl* locus.

A complementary approach to QTL mapping is association mapping, which relies on historical recombination events occurring within many different lineages for the discovery of markers linked to the trait of interest. In a genome-wide association study (GWAS) of maize aflatoxin resistance, an association mapping panel of 287 inbred lines was genotyped by sequencing (GBS) and phenotyped for aflatoxin content in testcrossed replicated field experiments (Warburton et al., 2013). Many single nucleotide polymorphisms (SNPs) associated with aflatoxin accumulation resistance were identified, and these were then used in a metabolic pathway analysis focusing on the combined effects of many genes grouped according to their shared biological function (Tang et al., 2015). This is a promising approach to complement significant SNP-trait associations and study the genetic basis of a trait (Tang et al., 2015; Wang et al., 2007). Originally developed to study differences in gene expression data for important human diseases (Subramanian et al., 2005), pathway analysis has been used with association mapping to find any biological insights missed when focusing on only one or a few genes that have the highest associations with the trait of interest (Tang et al., 2015). In addition, biologically relevant pathways can be used to determine candidate genes for association analysis or to interpret large datasets produced by other high-throughput approaches, including RNA sequencing, proteomics, and metabolomics. This is being done in maize in targeted pathways such as the well-characterized carotenoid pathway (Owens et al., 2014), in functionally related networks (Walley et al., 2016; Baute et al., 2016), and in protein interaction networks (Musungu et al., 2015), among others.

The pathway-based approach with the results of the aflatoxin GWAS study by Tang et al. (2015) successfully combined GWAS data to jointly consider all genetic sequences positively associated with aflatoxin accumulation resistance. This approach revealed 17 high-ranking pathways that represent mechanisms to prevent fungal growth and to prevent fungi from producing deleterious aflatoxin (Tang et al., 2015). These pathways were consistent with known resistance mechanisms, primarily centered on the production of the bioactive plant hormone methyl-jasmonate, known to orchestrate plant responses

to several biotic stresses. The same association mapping panel was also phenotyped for CEW damage in a replicated multilocation field trial. Therefore, to better understand the genetic basis and physiological mechanisms of resistance to CEW in maize, this study was performed to elucidate the biological pathways and associated genes that contribute to CEW resistance via GWAS followed by pathway analysis based on a gene-set enrichment approach. The joint analysis increases the utility of information gained from the GWAS and made possible an improved identification of new mechanisms and plant ideotypes that will reduce damage to CEW in maize.

MATERIALS AND METHODS

Germplasm, Environments, and Phenotypic Data

An association mapping panel of 287 maize inbred lines was assembled and characterized by Warburton et al. (2013). These lines were crossed to Va35, a southern-adapted inbred line of the nonstiff stalk heterotic pattern with good general combining ability, and testcrossed hybrids were grown in six locations in 2009 and 2010 in a replicated, randomized complete block design. The randomized complete block design did not guarantee a proper correction of environmental heterogeneity, but was chosen because blocks fell within irrigation zones in the field. Maturity date was recorded as mid silk, or the date that silks had emerged from 50% of the plants in a row. Maize ears were inoculated with *A. flavus* spores 7 d after mid silk. Twenty plants were grown per plot in a single row, and at least 10 plants per plot were scored for all traits and averaged to obtain the final value per plot. Following maturity, ears were harvested and scored for CEW damage under natural infestation. The CEW damage was determined by measuring the length (cm) from the tip of the ear to the deepest penetration according to Widstrom (1967). Variation ranged from 1 cm (close to the tip) to the length of feeding location furthest from the tip of the ear. Data generated on this panel also included SNP data [as per Elshire et al. (2011)]; kinship, population substructure, and linkage disequilibrium as reported by Warburton et al. (2013). Broad-sense heritability estimates were calculated on per-plot and genotype-mean bases within and across all six environments according to Holland et al. (2003).

Genome-Wide Association Study

Genotyping of the 287 inbred line entries in the panel was done via GBS according to Elshire et al. (2011) and reported in Warburton et al. (2015). Briefly, SNPs were extracted from raw GBS data with the Java pipeline for GBS Bioinformatics (Glaubitz et al., 2014). The GBS marker dataset was imputed and filtered by removing SNPs with a minor allele frequency of <5% and any loci with >2 alleles. A total of 260,550 SNPs were used for GWAS. A subset consisting of 2000 SNPs with a low missing data rate (<7.5%) and low frequency imbalance between the two alleles (minor allele frequency > 25%) was extracted for calculation of kinship (K matrix), population substructure (Q matrix), and linkage

disequilibrium as reported by Warburton et al. (2013). The software package TASSEL version 3.0 (Bradbury et al., 2007) was used to perform the GWAS using the plot mean CEW scores from six environments plus the average scores over all environments. The mixed linear model (Yu et al., 2006) was run with six subpopulations (Q matrix) and the K matrix.

Single Nucleotide Polymorphism-to-Gene Algorithm for the Pathway Analysis

In fitting the model, TASSEL calculated three SNP-trait association values that were used in the SNP-to-gene algorithm for the metabolic pathway analysis. They were: significance (p), correlation statistic (R^2 or the proportion of the phenotypic variation accounted for), and effect value (a relative measure of the effect a SNP had on the trait, with positive values reflecting an increase and negative values a decrease in CEW damage). Estimates of linkage disequilibrium (D' , r^2 , and p) between each marker and 50 of the closest neighboring SNPs in the up- and downstream directions were also provided by TASSEL and used in the SNP-to-gene algorithm for the pathway analysis. The SNP-to-gene algorithm used for the pathways analysis was developed by Tang et al. (2015) to identify SNP linkage blocks based on plots of linkage disequilibrium values $-\log(p)$ against r^2 ; linkage was defined as $r^2 > 0.8$. A decision tree based on the SNP-trait association p and R^2 values was implemented to evaluate which SNP in a block of linked SNPs would serve as the tagSNP; this was generally the one with the largest effect on the trait within the linkage block and holding the majority sign (positive or negative) for the effect value. Once the tagSNP was identified, the search window for the causative gene was set to ± 1 kb, which was based on a histogram of distances between linked SNPs. The SNP-trait association values of the tagSNP were then assigned to the gene. The SNP-to-gene algorithm was run with SNP-trait associations for the phenotype from the average environment only.

Pathways

Pathway analysis was performed as described by Tang et al. (2015) with the functional annotations and gene-to-pathway mappings of MaizeCyc (Monaco et al., 2013). Gene-set enrichment calculations were based on the study by Subramanian et al. (2005). Briefly, genes were ranked by their effect values from negative to positive and a weighted Kolmogorov-Smirnov running sum statistic was calculated. The enrichment score (ES) for the pathway was the maximum deviation of the running sum statistic from zero. Only pathways with five or more mapped genes (298 pathways) were considered in the analysis. The significance of a pathway was evaluated by making 1000 permutations of the gene effect values, which generated a null distribution for the ES. The mean (μ) and SD of the null distribution were then used to normalize the ES for a pathway, $(ES - \mu) \div SD$, and pathway p -values were computed using the `pnorm` function in R

(R Core Team, 2013). The false discovery rate of a pathway *p*-value was calculated by the QVALUE package in R (Storey and Tibshirani, 2003).

Data Sources and Analysis Tools

The *Z. mays* reference sequence (B73 RefGen version 2 assembly), canonical gene coordinates (ZmB73_5b_FGS_info.txt), and gene functional annotations (ZmB73_5a_gene_descriptors.txt) were obtained from ftp.maizesequence.org (accessed 6 Nov. 2017). The SNP marker dataset of lines was specified by Warburton et al. (2013) and was extracted from the parent file downloaded from panzea.org (accessed 6 Nov. 2017) (Ame-sUSInbreds_AllZeaGBSv1.0_imputed-130508 by Romay et al., 2013). The MaizeCyc version 2.1 pathway genome database (Monaco et al., 2013) was accessed from www.maizegdb.org (accessed 6 Nov. 2017). Unless otherwise specified, scripts to perform analyses were written in Perl 5 version 16 (www.perl.org, accessed 6 Nov. 2017). Graphing and statistical analyses were done in R version 3.0.2 (R Core Team, 2013).

RESULTS AND DISCUSSION

Corn Earworm Damage Data

Corn earworm damage data collected in Starkville, MS (two sites in 2010 and one in 2009); Lubbock, TX (in 2009 and 2010); and College Station, TX (in 2009); are summarized in Table 1; data for each entry are listed in full in Supplemental Table S1. The per-plot heritability across all six environments was 0.1055 and the genotype-mean heritability was 0.5894. There was low CEW infestation in the two Mississippi sites in 2010 and in College Station in 2009, and no significant effect caused by genotype in these three environments (data not shown), and lower overall heritability (Table 1). The other environments had larger heritabilities. In general, the mean rating of CEW damage per plot was much lower in 2010 than in 2009, suggesting that the natural infection was much lower in 2010 than in 2009. The CEW damage scores were correlated for all entries in the six environments. Corn earworm damage was negatively correlated with maturity ($R^2 = -0.46$; data not shown), as later maturing varieties tended to be less damaged. It is unlikely that this was caused by lower insect pressure later in the season, as

CEW can have multiple generations per year. Rather, we suggest it may be simply that the later maturing varieties were more resistant to CEW because they had a higher percentage of tropical germplasm in their pedigrees, which is known to carry greater resistance to CEW than temperate germplasm (Xu et al., 2003). It is hoped that this resistance can be moved to temperate backgrounds.

Although none of the lines in this panel had been selected specifically for CEW resistance, three (Mp705, Mp707, and Mp714) had been selected for leaf feeding caterpillar resistance, and Mp705 and Mp707 were in the least damaged third of the entries. The line with the lowest CEW damage levels in the panel, Mp313E, was bred and selected for resistance to *A. flavus* infection and aflatoxin accumulation (Scott and Zummo, 1990). It is unknown whether the same mechanisms cause resistance to both CEW and *A. flavus* resistance in Mp313E. Mp313E is known to have three QTLs for aflatoxin accumulation resistance that colocalize closely with chitinase genes (Hawkins et al., 2015). Both insects and fungi have chitin and this may thus be a common mechanism of resistance. Mp313E also has very tight husks, which can thwart both *A. flavus* and CEW, but this is not a desirable trait for farmers because of difficulty with dry-down, as well as husking during harvest. The top 10 resistant lines were all from subtropical and tropical breeding programs including the USDA Corn Host Plant Resistance Research Unit, CIMMYT, IITA, and Thai maize breeding programs (Supplemental Table S1). The most resistant lines in this study (the top 40 from Supplemental Table S1) were all of tropical or semi-tropical background, and many were early selections from landrace varieties, showing the importance of non-elite sources of resistance to CEW and other biotic stresses.

Genome-Wide Association Study

The 260,550 SNPs filtered from the SNP data for the 287 inbred maize lines in the panel included all biallelic SNPs with a minor allele frequency greater than 5% and less than 20% missing data. Of these, 45.8% of the SNP allele calls were imputed from the regional haplotype, rather than called directly from the GBS procedure. Though no SNPs were associated with CEW damage more significantly than a significance cutoff of $p < 2.74 \times 10^{-8}$ [$= 0.05 \div (260,550 \text{ markers} \times 7 \text{ phenotypes})$], a total of 281 unique SNPs were associated with CEW damage in one or

Table 1. Mean, minimum, and maximum corn earworm (CEW) damage ratings for 287 testcrossed maize lines in six environments, the average over all environments with the SD of the overall average, and the per-plot ($\hat{H}_{plot}$) and genotype mean ($\hat{H}_G$) broad-sense heritability.

CEW damage	Starkville 2010 A	Starkville 2010 B	Lubbock 2010	Starkville 2009	College Station 2009	Lubbock 2009	Avg.	SD
Mean	2.48	2.52	4.57	4.03	4.33	4.68	3.76	1.01
Min.	0.87	0.44	2.77	2.51	2.11	2.79	1.91	1.01
Max.	4.85	4.53	7.97	6.15	8.89	8.67	6.84	1.93
$\hat{H}_{plot}$	0	0.07	0.70	0.56	0.02	0.59	0.11	—
$\hat{H}_G$	0	0.19	0.87	0.80	0.07	0.81	0.59	—

more environments at an arbitrary threshold of $p < 9.99 \times 10^{-5}$ (Supplemental Table S2). Of these, 51 SNPs were associated with CEW damage over more than one environment but not including the average over all environments. Many of these same SNP–trait associations were identified in the most similar environments (Lubbock and College Station in 2009). Few SNPs were associated even at this level with CEW damage in the two Starkville locations in 2010 because of low CEW population pressure, a common problem in damage rating traits that rely on natural infestations. In contrast, 92 SNPs were associated with CEW in Lubbock in 2009, and 45 SNPs on average were associated per environment. The 281 SNPs fell into 186 linkage blocks. In this case, a linkage block was solely defined as SNPs that were less than 20 kb apart. The average length of these linkage blocks was 318 bp. The range of SNP–trait association p and R^2 values were 1.5×10^{-7} to 9.99×10^{-5} and 0.052 to 0.217, respectively.

The associated SNPs fell within 10 kb of 286 genes and two microRNAs (Supplemental Table 2). A majority of the genes encoded DNA-binding proteins or transcription factors; protein-binding or modifying enzymes (including kinases); proteins that bind to, modify, or create RNA, plant hormone-responsive proteins (including those binding abscisic acid, gibberellins, ethylene, or auxin), and proteins involved in signaling, (including potential defense signaling). Taken together, these suggest that a response to biotic stress is occurring in the resistant plants at the transcriptional, translational, and protein interaction or modification levels. In addition, several of the genes are hydrolases, including four

glycosylhydrolases or transferases; three are pectin modifying enzymes; and four are xyloglucanases. All of these enzymes are involved in cell wall construction or degradation, suggesting that resistance may be caused by a strong outer structural defense. Some of the hydrolases may also have chitin-degrading properties, which would be important against fungi and insects. There are also many general stress response genes in Supplemental Table S2 (including multiple cytochrome P450s and heat shock proteins, a drought-induced protein, and a damaged DNA repair protein). However, only a very few have been recognized as insect or biotic stress resistance genes, according to annotation in MaizeGDB (http://www.maizegdb.org/gene_center/gene/, accessed 6 Nov. 2017). These include a RAR1 disease resistance protein homolog, an NB-ARC domain-containing disease resistance protein, a wound-responsive phytochrome signaling receptor kinase, and a flavonol synthase (in the same pathway as the maysin biosynthesis pathway).

Some of the identified genes may be associated simply because of their proximity to another gene of interest, because 76 of the SNPs were close to two or more genes (within a window of ± 10 kb). Some may be false positives because the threshold of selection ($p < 9.99 \times 10^{-5}$) for the genes in Supplemental Table S2 was not highly stringent. Many of the genes will have alleles causing real but only minor effects, or display context dependency or background epistasis, which dilutes the effects of these alleles when screened across such a diverse set of germplasm. Similarly, it is often observed that alleles that are important in inbred lines of maize are masked by hybrid vigor in

Table 2. Pathways affecting corn earworm damage ratings with enrichment scores better than $p < 0.02$, including pathways that contributed to an increase in resistance or an increase in susceptibility.

Contributes	Pathway ID	Pathway name	p	q
Resistance	PWY-1081	Homogalacturonan degradation	0.000	0.051
Resistance	PWY-5885	Wax ester biosynthesis II	0.002	0.240
Resistance	PWY-6118	Glycerol-3-phosphate shuttle	0.005	0.443
Susceptibility	NONMEVIPP-PWY	Methylerythritol phosphate pathway	0.005	0.440
Susceptibility	ANAGLYCOLYSIS-PWY	Glycolysis III	0.005	0.440
Susceptibility	PWY-5868	Simple coumarin biosynthesis	0.006	0.440
Resistance	PWY-5053	Superpathway of gibberellic acid biosynthesis	0.006	0.443
Susceptibility	GLYCOLYSIS	Glycolysis I	0.008	0.440
Resistance	PWY-5032	Ent-kaurene biosynthesis I	0.008	0.463
Susceptibility	PWY-5901	2,3-Dihydroxybenzoate biosynthesis	0.009	0.440
Susceptibility	PWY-5098	Chlorophyll a degradation	0.010	0.440
Susceptibility	PWY-1042	Glycolysis IV (plant cytosol)	0.012	0.440
Resistance	PWY-5084	2-Ketoglutarate dehydrogenase complex	0.012	0.544
Susceptibility	PWY-6040	Chlorogenic acid biosynthesis II	0.013	0.440
Susceptibility	PWY-5121	Superpathway of geranylgeranyldiphosphate biosynthesis II (via the methylerythritol phosphate pathway)	0.013	0.440
Resistance	PWY-1353	Succinate to cytochrome bd oxidase electron transfer	0.015	0.544
Resistance	PWY-3861	Mannitol degradation II	0.019	0.544

hybrid combinations. Most genes identified by this GWAS analysis, therefore, do not currently make tempting plant breeding targets of selection, especially in the Va35 hybrid background, but could indeed be valuable in some genetic backgrounds.

Most Significant Pathways

To use the cumulative effect of multiple genes in linkage blocks, regardless of the magnitude of effect of individual genes, pathway analysis was run on all associations. In total, 25,483 unique tagSNPs were used to locate as many genes in this study. Of the 25,483 gene associations found, 2880 of the genes mapped to the 298 MaizeCyc pathways with five or more genes. Analyzing pathways with fewer genes leads to results with low statistical power because of the small sample size. Of the 298 pathways, genes with a negative effect on the CEW damage score (i.e., those that made the plant more resistant) were enriched from eight pathways (Table 2). These pathways contained between 11 and 154 genes and all had an ES better than expected at $p < 0.02$. Genes assigned to a positive effect on the CEW damage score (i.e., those that were associated with susceptibility and increased CEW damage) were enriched from nine pathways containing a range of 7 to 183 genes, with an ES better than $p < 0.02$ (Table 2).

Possible CEW Resistance Mechanisms

Only one pathway (homogalacturonan degradation, with an effect on reducing CEW damage) had an ES that was significant at a false discovery rate of <0.20 (Table 2). Homogalacturonan is a polysaccharide that forms one type of backbone found in the pectin of cell walls. The wax esters biosynthesis II pathway was the second most significant; it is involved in the formation of epicuticular wax, which is a physical barrier to water loss and insect predation. It is made of cutin and wax esters, both derived from fatty acids. It is interesting to note that the top two significant pathways are involved in creating a strong physical barrier to the exterior and may protect maize ears from insect herbivory in this manner. Two other significant and linked pathways were gibberellic acid biosynthesis, and *ent*-kaurene biosynthesis I. *Ent*-kaurene is converted into gibberellic acid, and gibberellins are plant hormones involved in almost all stages of plant growth but especially promote flowering. They are associated with biotic and abiotic stress resistance in some cases, and may act in defense by promoting growth and morphological changes resulting in physical defense strategies, such as the creation of strong cell walls with thick epicuticular wax. In addition, on an extended list of pathways (those significant at $p < 0.1$; Supplemental Table S3), five pathways support the production of fatty acids, which are necessary to build the cutins and wax esters that make up the epicuticular wax layer. Also on the extended pathway list is the cellulose biosynthesis pathway, which also points to stronger cell walls. Over one-third of the 31 pathways on the extended list of resistant

pathways appear to be involved in an increased physical barrier at the cell wall (Supplemental Table S3).

Not all reductions in CEW damage seen in the panel may be caused by mechanical means. Another pathway with an ES significant at $p < 0.02$ is the phospholipase pathway. Phospholipases are active in response to wounding and initiate production of important defense signaling molecules, such as oxylipins and jasmonates. They also activate the secondary messenger phosphatidic acid, which has been shown to modulate the activity of a variety of proteins involved in defense signaling. The phospholipases, as well as the gibberellin and brassinosteroid pathways from the extended list, are all compounds that activate defense responses following pathogen and insect attack (De Vleeschauwer et al., 2014; Ryu, 2004; Santamaria et al., 2013). A more direct antibiosis defense is also indicated from the pathways identified in the extended list, including 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) glucoside degradation pathway (leading to DIMBOA glycoside), the coumarin biosynthesis pathway, and the anthocyanin biosynthesis pathway. Anthocyanins and coumarins are flavonoids synthesized via the phenylpropanoid pathway. DIMBOA and flavonoids are toxic to insects, and if these were to accumulate in the ear, they would provide a chemical defense against the CEW. Pathways involved in this antibiosis, defense signaling, or general stress resistance are common among the resistance pathways, accounting for another third of the total (Supplemental Table S3). Pathways involved in N uptake, cell growth, or energy production make up the other third.

Pathways Associated with Susceptibility

The nine most significant pathways associated with an increase in susceptibility to CEW damage can also be found in Table 2. It may be that the genes in these pathways are under different selection pressure from those in the pathways boosting resistance, and different types of pathways are indeed seen in this group. Most (six of nine) are involved in general cell growth or energy generation. Three, however, may be directly related to antibiosis; these include the biosynthesis of 2,3-dihydroxybenzoate, chlorogenic acid (another dihydroxybenzene) and simple coumarins. These three compounds are phenolics, which include flavones and serve as defense compounds by repelling feeding herbivores and inhibiting enzymes. In addition, however, chlorogenic acid is an important intermediate in lignin biosynthesis and may thus create stronger cell walls. The extended list of pathways ($p < 0.1$) that increase susceptibility (Supplemental Table S3) are also dominated by growth and energy pathways, which make up over 60% of the total susceptibility-increasing pathways. Pathways involved in the production of cell walls or fatty acids are rare (only one of each out of 45 total susceptibility-increasing pathways; these include the suberin biosynthesis pathway and the saturated fatty acid elongation pathway). Pathways involved in antibiosis, defense signaling, or general stress resistance make up about one-third of the 45 pathways. Growth

and energy related pathways may not be directly related to a resistance mechanism, but in terms of susceptibility, plants that are not vigorous would generally be more prone to infection or predation.

Corn earworm resistance in sweet corn has been reported due to the presence of the C-glycosyl flavone, maysin [2'-O- α -L-rhamnosyl-6-C-(6-deoxy-xylo-hexos-4-ulosyl) luteolin], and chlorogenic acid (Nuessly et al., 2007). Past QTL mapping studies of CEW resistance only found one major resistance pathway: the production of maysin. One major flavonoid related gene was identified by GWAS in this paper (*GRMZM2G152801*, *fls1*, flavonol synthase), and three pathways [phenylpropanoid biosynthesis, anthocyanin biosynthesis (pelargonidin 3-O-glucoside, cyanidin 3-O-glucoside) and chlorogenic acid biosynthesis] were identified in our pathway analysis. These three pathways could all lead to the production of flavonoids. However, most of the genes and pathways appear to be pointing to resistance mechanisms that are not related to flavonoids and thus provide for a new approach to designing an ideotype of a highly resistant maize plant.

CONCLUSIONS

Both the GWAS and the pathway analyses point to strong cell walls as being important for defense against CEW damage and suggest that production and accumulation of compounds that repel and/or harm the feeding larvae are components of CEW resistance. These include flavonoids and phenolics (produced in the pathways of anthocyanin biosynthesis, simple coumarin biosynthesis, coumarin biosynthesis via 2-coumarate, 2,3-dihydroxybenzoate biosynthesis, chlorogenic acid biosynthesis II, and phenylpropanoid biosynthesis I and II and initial reactions). One GWAS-identified gene (flavonol synthase) and one direct pathway (anthocyanin biosynthesis), plus some indirect pathways, confirm past studies that identified maysin as being important in CEW defense. The pathway analysis, however, provided a more focused view of resistance than the GWAS alone, and that genes in each pathway identified in this analysis should be considered as candidate genes for CEW resistance that are in need of further testing. The pathway analysis also uncovered new potential mechanisms for resistance, which could be used in plant breeding programs by providing (i) a physiological mechanism and ideotype for phenotypic selection, (ii) more candidate genes for marker-assisted selection (following validation), and (iii) directions for more basic research into insect and biotic stress resistance.

Supplemental Information

Supplemental Table S1. Corn earworm damage ratings for 287 testcrossed maize lines averaged over three replicates per environment in six environments, and the average over all environments with the SD of the overall average.

Supplemental Table S2. SNP-trait associations significant at $p < 9.99 \times 10^{-5}$ and all genes within a window

of ± 10 bp of the SNP according to the B73 reference genome (version 2).

Supplemental Table S3. Pathways affecting corn earworm damage ratings with enrichment scores better than $p < 0.1$, including pathways that contributed to an increase in resistance or an increase in susceptibility.

Conflict of Interest Disclosure

The authors declare that there is no conflict of interest.

Acknowledgements

The authors thank Jack Haynes and Boo Matthews for their excellent technical help, and Dr. Julio Bernal and Dr. Michael Brewer for their thoughtful review and suggestions. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

References

- Archer, T.L., B.R. Wiseman, and A.J. Bockholt. 1994. Factors affecting corn earworm (Lepidoptera: Noctuidae) resistance in food corn. *J. Agric. Entomol.* 11:9–16.
- Baute, J., D. Herman, F. Coppens, J. De Block, B. Slabbinck, M. Dell'Aquila, et al. 2016. Combining phenotyping and transcriptomics in maize reveals a robust growth regulatory network. *Plant Phys.* 170(3):1848–1867. doi:10.1104/pp.15.01883
- Bhatnagar-Mathur, P., S. Sunkara, M. Bhatnagar-Panwar, F. Waliyar, and K.K. Sharma. 2015. Biotechnological advances for combating *Aspergillus flavus* and aflatoxin contamination in crops. *Plant Sci.* 234:119–132. doi:10.1016/j.plantsci.2015.02.009
- Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19):2633–2635. doi:10.1093/bioinformatics/btm308
- Butron, A., and N.W. Widstrom. 2001. Mass selection for agronomic performance and resistance to ear-feeding insects in three corn populations. *Maydica* 46:207–212.
- Byrne, P.F., M.D. McMullen, B.R. Wiseman, M.E. Snook, T.A. Musket, J.M. Theuri, et al. 1997. Identification of maize chromosome regions associated with antibiosis to corn earworm (Lepidoptera: Noctuidae) larvae. *J. Econ. Entomol.* 90:1039–1045. doi:10.1093/jee/90.4.1039
- Byrne, P.F., M.D. McMullen, M.E. Snook, T.A. Musket, J.M. Theuri, N.W. Widstrom, et al. 1996. Quantitative trait loci and metabolic pathways: Genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. *Proc. Natl. Acad. Sci. USA* 20:8820–8825. doi:10.1073/pnas.93.17.8820
- Chen, G., C.R. Hooks, T.W. Patton, R. Kratochvil, and G. Dively. 2016. Tolerance to stalk and ear-invading worms and yield performance of Bt and conventional corn hybrids. *Agron. J.* 108:73–84. doi:10.2134/agronj15.0139
- Cortés-Cruz, M., M. Snook, and M.D. McMullen. 2003. The genetic basis of C-glycosyl flavone B-ring modification in maize (*Zea mays* L.) silks. *Genome* 46:182–194. doi:10.1139/g02-119
- De Vleeschauwer, D., J. Xu, and M. Höfte. 2014. Making sense of hormone-mediated defense networking: From rice to Arabidopsis. *Front. Plant Sci.* 5:611. doi:10.3389/fpls.2014.00611
- Dively, G.P., P.D. Venugopal, and C. Finkenbinder. 2016. Field-evolved resistance in corn earworm to Cry proteins expressed by transgenic sweet corn. *PLoS One* 11:e0169115. doi:10.1371/journal.pone.0169115
- Dowd, P.F. 1998. Involvement of arthropods in the establishment of mycotoxigenic fungi under field conditions. Marcel Dekker: New York.
- Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi:10.1371/journal.pone.0019379
- Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, et al. 2014. TASSEL-GBS: A high capacity genotyping by sequencing

- analysis pipeline. *PLoS One*. 9(2):e90346. doi:10.1371/journal.pone.0090346
- Goergen, G., P.L. Kumar, S.B. Sankung, A. Togola, and M. Tamò. 2016. First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (JE Smith) (Lepidoptera, Noctuidae), A new alien invasive pest in West and Central Africa. *PLoS One* 11(10):e0165632. doi:10.1371/journal.pone.0165632
- Guo, B.Z., Z.J. Zhang, A. Butron, N.W. Widstrom, M.E. Snook, R.E. Lynch, et al. 2004. Lost P1 allele in sh2 sweet corn: Quantitative effects of p1 and a1 genes on concentrations of maysin, apimaysin, methoxymaysin, and chlorogenic acid in maize silk. *J. Econ. Entomol.* 97(6):2117–2126. doi:10.1093/jee/97.6.2117
- Hawkins, L.K., D. Alves, J.S. Smith, S. Ozkan, G.W. Windham, W.P. Williams, et al. 2015. Characterization of the chitinase gene family and the effect on *A. flavus* and aflatoxin resistance in maize. *PLoS One* 10(6):E0126185.
- Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding. *Plant Breed. Rev.* 22:9–112.
- Lee, E.A., P.F. Byrne, M.D. McMullen, M.E. Snook, B.R. Wiseman, N.W. Widstrom, et al. 1998. Genetic mechanisms underlying apimaysin and maysin synthesis and corn earworm antibiosis in maize (*Zea mays* L.). *Genetics* 149(4):1997–2006.
- McMillian, W.W., D.M. Wilson, and N.W. Widstrom. 1978. Insect damage, *Aspergillus flavus* ear mold, and aflatoxin contamination in south Georgia corn fields in 1977. *J. Environ. Qual.* 7:564–566. doi:10.2134/jeq1978.00472425000700040019x
- Monaco, M.K., T.Z. Sen, P.D. Dharmawardhana, L. Ren, M. Schaeffer, S. Naithani, et al. 2013. Maize metabolic network construction and transcriptome analysis. *Plant Genome* 6(1). doi:10.3835/plantgenome2012.09.0025
- Musungu, B., D. Bhatnagar, R.L. Brown, A.M. Fakhoury, and M. Geisler. 2015. A predicted protein interactome identifies conserved global networks and disease resistance subnetworks in maize. *Front. Genet.* 6:201.
- Nuessly, G.S., B.T. Scully, M.G. Hentz, R. Beiriger, M.E. Snook, and N.W. Widstrom. 2007. Resistance to *Spodoptera frugiperda* (Lepidoptera: Noctuidae) and *Euxesta stigmatias* (Diptera: Ulidiidae) in sweet corn derived from exogenous and endogenous genetic systems. *J. Econ. Entomol.* 100:1887–1895. doi:10.1093/jee/100.6.1887
- Owens, B.F., A.E. Lipka, M. Magallanes-Lundback, T. Tiede, C.H. Diepenbrock, C.B. Kandianis, et al. 2014. A foundation for provitamin A biofortification of maize: Genome-wide association and genomic prediction models of carotenoid levels. *Genetics* 198:1699–1716. doi:10.1534/genetics.114.169979
- R Core Team. R: A language and environment for statistical computing. 2013. <http://www.R-project.org>, accessed 6 Nov. 2017.
- Ranum, P., J.P. Peña-Rosas, and M.N. Garcia-Casal. 2014. Global maize production, utilization, and consumption. *Ann. N. Y. Acad. Sci.* 1312(1):105–112. doi:10.1111/nyas.12396
- Romay, M.C., M.J. Millard, J.C. Glaubit, J.A. Peiffer, K.L. Swarts, T.M. Casstevens, et al. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14(6):R55.
- Ryu, S.B. 2004. Phospholipid-derived signaling mediated by phospholipase A in plants. *Trends Plant Sci.* 9(5):229–235. doi:10.1016/j.tplants.2004.03.004
- Santamaria, M.E., M. Martínez, I. Cambra, V. Grbic, and I. Diaz. 2013. Understanding plant defense responses against herbivore attacks: An essential first step towards the development of sustainable resistance against pests. *Transgenic Res.* 22:697–708.
- Scott, G.E., and N. Zummo. 1990. Registration of Mp313E parental line of maize. *Crop Sci.* 30:1378. doi:10.2135/cropsci1990.0011183X0030000600080x
- Snook, M.E., B.R. Wiseman, N.W. Widstrom, and R.L. Wilson. 1994. Chemicals associated with maize resistance to corn earworm and fall armyworm. In: J.A. Mihm, editor, *Insect Resistant Maize: Recent Advances and Utilization*, Proceedings of an International Symposium held at the International Maize and Wheat Improvement Center (CIMMYT), Mexico D.F. 27 Nov.–3 Dec. 1994. CIMMYT: Mexico D.F., Mexico. p. 37–45.
- Storey, J.D., and R. Tibshirani. 2003. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* 100(16):9440–9445. doi:10.1073/pnas.1530509100
- Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, et al. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102:15545–15550.
- Tang, J.D., A. Perkins, W.P. Williams, and M.L. Warburton. 2015. Using genome-wide associations to identify metabolic pathways involved in maize aflatoxin accumulation resistance. *BMC Genomics* 16(1):673.
- Tay, W.T., M.F. Soria, T. Walsh, D. Thomazoni, P. Silvie, G.T. Behere, et al. 2013. A brave new world for an old world pest: *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Brazil. *PLoS One* 8(11):e80134.
- van Heerwaarden, J., M.B. Hufford, and J. Ross-Ibarra. 2012. Historical genomics of North American maize. *Proc. Natl. Acad. Sci. USA* 109(31):12420–12425. doi:10.1073/pnas.1209275109
- Walley, J.W., R.C. Sartor, Z. Shen, R.J. Schmitz, K.J. Wu, M.A. Urich, et al. 2016. Integration of omic networks in a developmental atlas of maize. *Science* 353(6301):814–818. doi:10.1126/science.aag1125
- Warburton, M.L., W.P. Williams, G. Windham, S. Murray, W. Xu, L. Hawkins, et al. 2013. Phenotypic and genetic characterization of a maize association mapping panel developed for the identification of new sources of resistance to *Aspergillus flavus* and aflatoxin accumulation. *Crop Sci.* 53:2374–2383. doi:10.2135/cropsci2012.10.0616
- Warburton, M.L., J. Tang, G.L. Windham, L.K. Hawkins, S.C. Murray, W. Xu, et al. 2015. Genome-wide association mapping of *Aspergillus flavus* and aflatoxin accumulation resistance in maize. *Crop Sci.* 55:1857–1867. doi:10.2135/cropsci2014.06.0424
- Wang, K., M. Li, and M. Bucan. 2007. Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.* 81:1278–1283. doi:10.1086/522374
- Widstrom, N.W., A. Butron, B.Z. Guo, D.M. Wilson, M.E. Snook, T.E. Cleveland, et al. 2003. Control of preharvest aflatoxin contamination in maize by pyramiding QTL involved in resistance to ear-feeding insects and invasion by *Aspergillus* spp. *Eur. J. Agron.* 19(4):563–572. doi:10.1016/S1161-0301(03)00004-2
- Widstrom, N.W., and M.E. Snook. 2001. Registration of EPM6 and SIM6 maize germplasm, high silk-maysin sources of resistance to corn earworm. *Crop Sci.* 41:2009–2010. doi:10.2135/cropsci2001.2009
- Widstrom, N.W., B.R. Wiseman, and W.W. McMillian. 1988. Registration of six corn earworm resistant germplasm lines of maize. *Crop Sci.* 28:202. doi:10.2135/cropsci1988.202a.rgp
- Widstrom, N.W. 1967. An evaluation of methods for measuring corn earworm injury. *J. Econ. Entomol.* 60:791–794. doi:10.1093/jee/60.3.791
- Wilson, R.L., B.R. Wiseman, and G.L. Reed. 1991. Evaluation of J.C. Eldredge popcorn collection for resistance to corn earworm, fall armyworm (Lepidoptera: Noctuidae), and European corn borer (Lepidoptera: Pyralidae). *J. Econ. Entomol.* 84:693–698. doi:10.1093/jee/84.2.693
- Wiseman, B.R., M.E. Snook, R.L. Wilson, and D.J. Isenhour. 1992. Allelochemical content of selected popcorn silks: Effects on growth of corn earworm larvae (Lepidoptera: Noctuidae). *J. Econ. Entomol.* 85:2500–2504. doi:10.1093/jee/85.6.2500
- Wiseman, B.R., and N.W. Widstrom. 1992. Resistance of dent corn inbreds to larvae of the corn earworm (Lepidoptera: Noctuidae). *J. Econ. Entomol.* 85:289–292. doi:10.1093/jee/85.1.289
- Wong, A.Y., and A.W. Chan. 2016. Genetically modified foods in China and the United States: A primer of regulation and intellectual property protection. *Food Sci. Human Wellness.* 5:124–140. doi:10.1016/j.fshw.2016.03.002
- Xu, W., L. Pollak, and E.D. Bynum, Jr. 2003. Tropical × temperate germplasm resistant to corn earworm (Lepidoptera: Noctuidae). *Crop Prot.* 22:859–864. doi:10.1016/S0261-2194(03)00069-3
- Yu, J., G. Pressoir, W.H. Briggs, I.V. Bi, M. Yamasaki, J.F. Doebley, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38(2):203–208. doi:10.1038/ng1702