RESEARCH

Genome-Wide Association Mapping of *Aspergillus flavus* and Aflatoxin Accumulation Resistance in Maize

Marilyn L. Warburton,* Juliet D. Tang, Gary L. Windham, Leigh K. Hawkins, Seth C. Murray, Wenwei Xu, Debbie Boykin, Andy Perkins, and W. Paul Williams

ABSTRACT

Contamination of maize (*Zea mays* L.) with aflatoxin, produced by the fungus *Aspergillus flavus* Link, has severe health and economic consequences. Efforts to reduce aflatoxin accumulation in maize have focused on identifying and selecting germplasm with natural host resistance factors, and several maize lines with significantly reduced aflatoxin accumulation have been identified. Past linkage mapping studies have identified quantitative trait loci (QTL) that consistently reduce aflatoxin levels in maize. In addition, an association mapping panel of 300 maize inbred lines was previously created specifically for the dissection of aflatoxin accumulation resistance. Here we report the results of a genome-wide association study (GWAS) using this panel of testcrossed maize hybrids. Each of the inbred parents of the testcrossed hybrids was genotyped by sequencing to generate 261,184 robust single nucleotide polymorphisms (SNPs), and the entire panel was phenotyped for aflatoxin accumulation following inoculation with *A. flavus* in multilocation, replicated field trials. Results uncovered 107 SNPs associated with aflatoxin accumulation in one or more environments in the association panel at a probability level between 9.78×10^{-6} and 2.87×10^{-10}. Eight SNP trait associations were found with a false discovery rate (FDR) of less than 10% ($p < 3.83 \times 10^{-7}$). These SNPs occur within the sequence of three uncharacterized genes. Variants in 25 other genomic regions showing high association values over more than one environment are also presented. These genomic regions are undergoing validation studies and will be of use to dissect the resistance to aflatoxin accumulation and improve this trait.

M.L. Warburton, J.D. Tang, G.L. Windham, L.K. Hawkins, and W.P. Williams, USDA–ARS, Corn Host Plant Resistance Research Unit, Mississippi State, MS. Seth C. Murray, Dep. of Soil and Crop Sciences, Texas A&M Univ., College Station, TX; W.W. Xu, Dep. of Soil and Crop Sciences, Texas A&M Univ., Lubbock, TX; D. Boykin, USDA–ARS, Jamie Whitten Delta States Research Center, Stoneville, MS; A. Perkins, Dep. of Computer Science and Engineering, Mississippi State University, MS. Received 12 June 2014. Accepted 20 Mar. 2015. *Corresponding author (marilyn.warburton@ars.usda.gov).

Abbreviations: FDR, false discovery rate; GBS, genotyping-by-sequencing; GLM, generalized linear model; GWAS, genome-wide association study; LSMean, least-squared mean; MAF, minor allele frequency; MLM, mixed linear model; MT, Miscellaneous Temperate; Q-Q, quantile–quantile; QTL, quantitative trait loci; SNP, single nucleotide polymorphism.

The infection of maize, peanut (*Arachis hypogaea* L.), cotton (*Gossypium hirsutum* L.), and nut crops with *A. flavus* is problematic at many levels. *Aspergillus flavus* causes ear rot of maize, a staple food and animal feed and can lead to significant aflatoxin accumulation. This is a particular problem in hot growing areas and is further exacerbated by other biotic and abiotic stress (Payne, 1992; van Egmond et al., 2007). Consumption of aflatoxin-contaminated maize has caused human and animal death, and long-term exposure to low levels of aflatoxin causes chronic reduced immune system function and liver cancer (Groopman et al., 2008; Liu and Wu, 2010; Wild and Gong, 2010). Aflatoxin-related immune system depression in the developing world may have increased the severity of epidemics of AIDs, malaria, tuberculosis, and other diseases, as well as impaired child development

Published in Crop Sci. 55:1–11 (2015).
doi: 10.2135/cropsci2014.06.0424
Freely available online through the author-supported open-access option.
© Crop Science Society of America | 5585 Guilford Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

(Gong et al., 2008; Jiang et al., 2008; Pornsri et al., 2011). Furthermore, since many countries regulate or prohibit the sale of aflatoxin-contaminated maize, farmers with infected grain will typically suffer severe economic losses. Because of the high cost of testing, other countries are unable to test at all and cannot identify aflatoxin outbreaks in time to prevent devastating health consequences.

Efforts to reduce aflatoxin accumulation in maize have focused on identifying and selecting different maize lines with natural variation for reduced susceptibility including host resistance factors (Zhang et al., 1999; Mideros et al., 2013). Lower aflatoxin levels may result from a combination of host-plant resistance and agronomic traits (such as flowering time or husk tightness) that help the plant avoid the fungus. In addition, lower aflatoxin levels may be the result of reduced fungal colonization and growth or reduced production of aflatoxin by the fungus, perhaps in response to cues from the host plant. Some past studies have not been structured to determine if avoidance or true resistance is in play. Nevertheless, published results have identified several natural sources of resistance that exhibit significantly reduced aflatoxin accumulation (Campbell and White, 1995; Scott and Zummo, 1988; 1990; 1992; Robertson-Hoyt et al., 2006; Mayfield et al., 2012; Williams and Windham, 2001; 2006; 2012; Warburton et al., 2013). However, the quantitative nature of the trait and the high genotype-by-environment interactions has limited the success in transferring resistance to commercial hybrids (Hamblin and White, 2000). Reported heritability estimates have ranged from 0 to 0.6, with a median near 0.3 (Hamblin and White, 2000; Busboom and White, 2004; Brooks et al., 2005). Furthermore, aflatoxin is quite laborious and expensive to measure, which is further compounded by the need for multiple replicates and environments in accurate screening. Identification of genes and closely linked markers could greatly speed the development of resistant maize lines adapted to different growing environments.

The identification of genetic factors that may increase resistance to aflatoxin accumulation in maize has been accomplished by various methods. These include QTL mapping (reviewed in Mideros et al., 2013), protein studies (Chen et al., 2007; Pechanova et al., 2011), and gene expression analysis (Kelley et al., 2009). However, biparental mapping populations miss alleles that do not segregate between the parents of the population and many of the loci of smaller effect. Association analysis is an alternative method for identifying DNA sequence variants that contribute to quantitative traits, including disease resistance (Kump et al., 2011; Rafalski, 2011; Yan et al., 2011). Genome-wide association study methodologies have been optimized to enable the identification of associations from tens of thousands of polymorphisms without the false positive results due to population structure (Yu et al., 2006; Myles et al., 2009; Elshire et al., 2011). Relatively rapid linkage disequilibrium breakdown in diverse maize panels means that genes can be mapped to a very fine scale, provided that enough polymorphic markers are available to cover the entire genome. The availability of next-generation sequencing and genotyping-by-sequencing (GBS) techniques (Elshire et al., 2011) ensures that this coverage is now possible.

Association mapping for *A. flavus* resistance has not been feasible because previous association panels generally did not have any known sources of aflatoxin resistance. To address this gap, an association mapping population was created specifically for the study of aflatoxin accumulation and *A. flavus* resistance in maize (Warburton et al., 2013). This panel was enriched for diverse maize inbred lines that will grow in the southern United States, where aflatoxin can accumulate at high levels in susceptible lines and spans the spectrum of susceptibility and resistance to aflatoxin accumulation. Many new resistant lines were identified during the process of phenotyping the test-crossed hybrids (Warburton et al., 2013). The previous study presented the germplasm including the phenotypic data and population structure and relationships and the initial pedigree source of resistance found in most of the lines of the panel. The objectives of the current study were (i) to identify associations between polymorphisms found via GBS of the large aflatoxin association mapping panel and aflatoxin accumulation resistance in maize; (ii) to characterize the sequences in which these polymorphisms are found using existing linkage disequilibrium, gene annotation, expression, genetic mapping, and other publicly available data; and (iii) to develop SNP assays for each significant association identified here. These will be useful for independent validation of phenotypic effect with the ultimate goal of marker-assisted improvement of resistance in maize.

MATERIALS AND METHODS

Association Panel Description and Phenotyping

The association mapping panel of 300 inbred lines was assembled and characterized as reported in Warburton et al. (2013). Briefly, the 300 diverse inbred lines were chosen to represent most of the publicly available sources of *A. flavus* and aflatoxin accumulation resistance known at the time that the panel was assembled, as well as many well-characterized inbred lines from previously published association mapping studies (Flint-Garcia et al., 2005; Setter et al., 2011). These lines were crossed to Va35, a susceptible, southern-adapted inbred line of the nonstiff stalk heterotic pattern with pedigree (C103 × T8)T8,(Gerdes et al., 1993) and phenotyped in seven environments in replicated field trials. Maize ears were inoculated with *Aspergillus flavus* strain NRRL 3357 (ATCC #200026) 7 d after 50% of the plants in each plot had reached silking. Following maturity, ears of inoculated plants were harvested, dried, ground, and aflatoxin levels were measured via the Vicam AflaTest (VICAM). Aflatoxin values were logarithmically transformed to increase normality. Least-squared means (LSMeans) were calculated for transformed

aflatoxin levels for each environment and averaged across environments, as reported in Warburton et al. (2013).

Genotyping-by-Sequencing

Genotyping of the 300 inbred line entries in the panel was done via GBS according to Elshire et al. (2011). Single nucleotide polymorphisms were extracted from the raw GBS data using the Java pipeline for GBS Bioinformatics (Glaubitz et al., 2014). The GBS marker dataset (partially imputed genotypes version December 18, 2013; including data from Romay et al. [2013], which can be found at http://www.panzea.org/db/gateway?file_id=Romay_etal_2013_imputed_geno_data) was stored in HapMap format in separate files for each chromosome. A series of custom scripts and a web interface were developed to query SNPs by Gramene gene identifier or chromosome position. The web interface provided functionality for querying the entire list of SNPs with filtering by missing data rate and minor allele frequency (MAF). Repeatability, as tested by duplicate genotyping of nine individuals, was high, as less than 0.1% of genotyping calls did not agree in the two genotyping runs (data not shown). Overall missing data rate per marker in the imputed data was 16% (18% average per line). After removal of monomorphic SNPs and those with the highest missing data rate, an initial data set containing 405,745 SNPs from the GBS pipeline was filtered for MAF greater than 5.0%, leaving a total of 261,184 for GWAS. A subset consisting of 2000 SNPs with a low missing data rate (<7.5%) and low frequency imbalance between the two alleles (MAF > 25%) was extracted for calculation of kinship, population substructure, and linkage disequilibrium as reported in Warburton et al. (2013). The GBS data were unreadable for four lines, and nine lines were repeated entries to ensure data quality, thus only 287 entries were fully characterized and included in this study (Supplemental Table S1).

Genome-Wide Association Study

The software package TASSEL 3.0 (Bradbury et al., 2007) was used to perform the GWAS using 261,184 SNPs and the aflatoxin LSMeans from seven environments plus the average aflatoxin levels over all environments. The mixed linear model (MLM; Yu et al., 2006) was used with six subpopulations (Q matrix) and pair-wise kinship values (K matrix), both from Warburton et al. (2013) and calculated with the subset of 2000 SNPs. The JMP Genomics software package version 6.1 (SAS Institute, 2012) was used to analyze the same data set with the same parameters; because the results of the two analyses were in good agreement (data not shown), only the results of the TASSEL analysis are presented. The R^2 correlation statistic of the association between any given SNP and aflatoxin accumulation was used to assess the amount of phenotypic variation explained by the model. The distribution of p-values obtained from the SNP association analysis was then corrected for the FDR as calculated by the QVALUE version 1.0 package in R (Storey and Tibshirani, 2003). Linkage disequilibrium was estimated by the r^2 correlation between each marker and 50 of the closest neighboring SNPs in up- and downstream directions. Custom perl scripts were then used to locate genes that contained the marker, contained the linked SNP, or were closest to the marker or linked SNP and lying in the linkage disequilibrium window. Values of r^2 less than 0.8 were considered unlinked. The SNPs and locations of genes from the filtered gene set were from the B73 reference genome assembly V2 (Schnable et al., 2009). Gene identifiers correspond to those used by the MaizeGDB database (Schaeffer et al., 2011).

In a previous characterization of this association mapping panel (Warburton et al., 2013), significant population substructure was found and could be accounted for using six clusters identified by STRUCTURE software (Pritchard et al., 2000). One of the clusters corresponded to germplasm related to B73 and included 23 individuals. Because all lines were phenotyped as test crosses with Va35, which is related to Mo17 (a heterotic partner to B73), it was noted that the mean aflatoxin levels in these 23 lines were lower than average and may have been due to a heterotic effect with the tester (Warburton et al., 2013). Using the Q matrix (population substructure matrix) should account for this effect in the MLM and reduce the chance of false–positive associations. To be sure, however, the MLM analysis was performed with and without the 23 lines to compare the resulting associations.

Allele Frequency Analysis

Lines were grouped into one of six maize subpopulations (Miscellaneous Temperate [MT], Lancaster, Tropical, SC76, Northern US, and B73) on the basis of the population structure analysis according to Warburton et al. (2013). Lines of mixed ancestry (the result of admixture among the subpopulations) were dropped from the allele frequency analysis. On the basis of the results of the association analyses, the frequencies of alleles that reduced disease severity at SNPs with significant trait associations (exceeding 10% FDR) were estimated within each subpopulation. At each SNP locus, a two-tailed Fisher's exact test run with the free online software GraphPad QuickCalcs 2015 (http://graphpad.com/QuickCalcs/) was used to test the null hypothesis that frequency of the allele conferring increased disease resistance was the same between subpopulations.

RESULTS AND DISCUSSION

Loci Associated with Aflatoxin Resistance

The phenotypic values, variation, and heritabilities measured in the test-crossed hybrids of the aflatoxin association panel were reported in Warburton et al. (2013). These trials had little missing data (<5%) and exhibited moderate to high correlation between repetitions within environments (r = 0.47–0.70). High levels of variation among test-crossed lines and high mean-basis heritability for aflatoxin levels (0.87) were indicators that this data would be suitable for GWAS analysis.

Of the 261,184 SNPs called from the GBS dataset from the 287 maize inbred lines (MAF > 5%), GWAS identified 107 SNPs associated with aflatoxin accumulation in one or more environments at a probability level between 9.78×10^{-6} and 2.87×10^{-10} (Supplemental Table S2). Fifty-six percent of the significant associations were found in the Starkville 2010 environment. This environment had a good correlation between replications and exhibited a low rate of

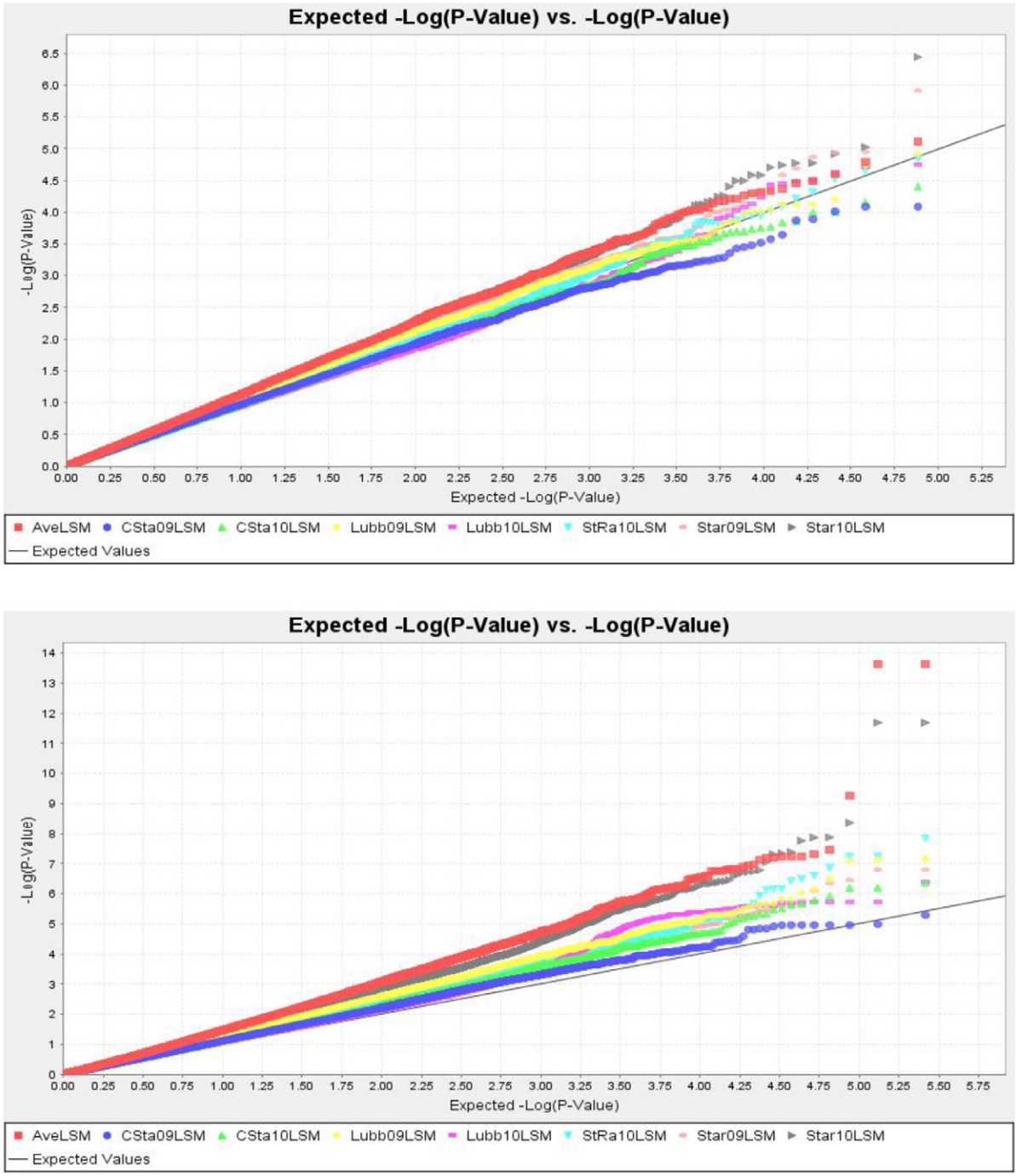

Figure 1. Quantile–quantile (Q-Q) graph for aflatoxin accumulation for each environment and averaged over all environments as calculated with TASSEL using the mixed linear model (including kinship and population substructure data in the model, top graph) and the general linear model (including population substructure data only, bottom graph).

missing data. Overall mean aflatoxin levels were fairly low in this field because the distribution was skewed toward lower levels; the overall range was as high as in other environments, however, and highly susceptible lines were identified.

Visual observation of the quantile–quantile (Q-Q) plots gives an idea of the accuracy of the model used to analyze the data. The Q-Q plots are the plots of observed −log10 *p*-values vs. expected −log10 *p*-values under the null hypothesis that there is no association between marker and the phenotype. As shown in the Q-Q plot for aflatoxin accumulation for each environment and averaged over all environments (Fig. 1), associations were found after employing the MLM to account for population structure and familial relatedness (Yu et al., 2006; Zhang et al., 2010). Some data points fell under the diagonal for the two College Station environments, which may indicate some overfitting of the model for these locations. These two environments were found to form a separate megaenvironment from all other environments in the phenotyping trial (Warburton et al., 2013). These

Table 1. Eight single nucleotide polymorphism (SNP) trait associations that had a probability level exceeding 10% false discovery rate levels. Alleles of each SNP listed with resistant form second. Number of individuals in the panel containing the resistant allele is shown. Genes contained the significant SNPs, but any other genes linked to these SNPs (and thus possibly causing the association) are listed in Table 2.

Environment	SNPs	*p*-value	R^2	Alleles	No. with resist alleles	Allele effect†	Genes containing or linked to SNP
Starkville 2010	2_187449158	2.83×10^{-07}	0.121	A/G	16	1.3223	‡
Lubbock 2009	3_217808740	3.72×10^{-07}	0.113	A/T	18	1.1336	GRMZM2G052991
Lubbock 2009	3_217808747	3.83×10^{-07}	0.112	T/C	18	1.1315	GRMZM2G052991
Starkville 2010	9_117048726	2.24×10^{-09}	0.144	G/A	31	1.2407	GRMZM2G108619
Starkville 2010	9_117048731	2.24×10^{-09}	0.144	C/T	31	1.2407	GRMZM2G108619
Average	9_117048726	2.87×10^{-10}	0.162	G/A	31	0.7867	GRMZM2G108619
Average	9_117048731	2.87×10^{-10}	0.162	C/T	31	0.7867	GRMZM2G108619
Starkville 2010	10_91956547	2.50×10^{-07}	0.099	G/T	28	1.0457	GRMZM2G158141

† Allele effect calculated based on least square mean squared natural log transformed value of aflatoxin levels, which ranged from 0.231 to 8.43.

‡ Nearest gene, GRMZM2G178229, is 5.6 kb from SNP.

environments were reanalyzed using a generalized linear model (GLM), which takes population structure but not kinship into account and lead to a Q-Q graph in which all but two data points fell on or above the diagonal. All associations found via the MLM analysis were also found at very low probability with the GLM analysis (data not shown), lending credibility to the MLM results.

Eight SNP trait associations involving six SNPs and three environments were found with a *q*-value, or FDR, less than 10% FDR levels ($p < 3.83 \times 10^{-7}$) (Table 1). None of these six SNPs had been imputed. Five of the significantly associated SNPs fell within the sequence of three genes, and one associated SNP (S2_187449158) fell far from the known sequence of any gene. The latter SNP, associated with aflatoxin accumulation in one environment, was 1.4 million bases away from multiple SNPs also significantly associated with aflatoxin levels. The closest gene to the single SNP, at 5.6 kilobases (kb) distance upstream of the SNP, was GRMZM2G178229, or CYCD4, a nuclear localized cyclin associated with cell replication. Such a generalized function at such a long physical distance does not lend itself to a hypothesis of a specific mechanism for improving resistance, and this SNP will not be discussed further. The three genes in which the remaining five significantly associated SNPs were characterized for gene ontology from MaizeGDB (www.maizegdb.org), Gramene (www.gramene.org), and Phytozome (http://www.phytozome.net/) and expression data from publicly available databases including MaizeGDB (Sekhon et al., 2011) and qTeller (Schnable and Thompson, 2014). The expression data from Sekhon et al. (2011) represent a comprehensive atlas of global transcription profiles across developmental stages and plant organs measured using a NimbleGen microarray. The data from Schnable and Thompson (2014) summarize several previously published maize expression datasets of different tissues measured by high-throughput sequencing of RNA. None of the maize tissues in these studies had been inoculated with *A. flavus*.

Gene GRMZM2G052991, in which two SNPs significantly associated with aflatoxin levels in one environment were identified, appears to encode an uncharacterized protein with a methyl transferase domain and a reverse transcriptase domain; this second domain may actually be homologous to a retrotransposon. This gene, located in bin 3.09, was not present in the MaizeGDB expression data set and expressed at very low levels in all tissues studied in qTeller. The second gene with two SNPs associated with aflatoxin level in one environment and the average of all environments was GRMZM2G108619, located in bin 9.04. The deduced protein shared homology only to other uncharacterized proteins in other grass species. Transcript levels were low in nearly every tissue examined and higher in the leaves and the developing ear according to the expression datasets of both MaizeGDB and qTeller. The third gene in which one SNP was associated with aflatoxin level in one environment was GRMZM2G158141, which encodes an uncharacterized protein with a Zn binding or Zn finger DNA binding motif located in bin 10.04. It was expressed at low levels in all tissue studied.

Including all associations less than 10^{-4} (5242 associations in total) and looking at an expanded Manhattan graph generated in Excel allows genomic regions with multiple associations (more than one environment or more than one linked SNP) to be easily visualized. Including multiple associations with at least one of $p < 1.00 \times 10^{-6}$ that occur in at least two environments may highlight genes contributing to stable resistance. The expanded Manhattan graph uncovered 27 genomic regions, including two of the four regions of highest significance (those less than 10% FDR), and they can be found in Supplemental Fig. S1. The 27 genomic locations where multiple SNP–environment associations converged included 63 SNPs, which are listed in Table 2, along with the most probable candidate gene (or genes) in the region. Many of the 27 regions were represented by both imputed and nonimputed SNPs, lending confidence to the imputed data. This list must be considered incomplete, as only SNPs matching the B73 reference

Table 2. Genes linked to singe nucleotide polymorphisms (SNPs) associated with grain aflatoxin levels ($p < 10^{-4}$) in seven different environments or the average of all environments (column labeled Environment) and potentially causing the association. Bin location of the SNP and association statistics including *F*, *p*, and R^2 (from Supplemental Table S2) are given. The proximity in kilobase (kb) pairs and in linkage disequilibrium (LD) r^2 values is shown. Gene names correspond to the Gramene and MaizeGDB databases and locations of SNPs and distances to genes were taken from version 2 of the maize reference sequence.

Region	Environment	SNPs	BIN	*F*	*p*-value	R^2	Proximity†	Gene
1	Starkville2010	S1_272220177		23.12	2.78×10^{-06}	0.099	Gene contains marker	GRMZM2G049349
	Starkville2010	S1_272220142	1.10	16.68	5.92×10^{-05}	0.064	Gene contains linked SNP 55 kb away, LD > 0.3	GRMZM2G137891
	Average	S1_272220177		16.01	8.55×10^{-05}	0.066	Gene is 34 kb from marker and lies in LD window > 0.35	GRMZM2G061442
	Starkville2010	S1_272220852		15.73	9.50×10^{-05}	0.058		
	Starkville2010	S1_272272944		16.13	8.09×10^{-05}	0.066	Gene is 2 kb from marker and lies in LD window > 0.36	GRMZM2G137891
	Starkville2010	S1_272272952		20.08	1.19×10^{-05}	0.083	Gene is 15 kb from marker and lies in LD window > 0.3	GRMZM2G061442
	Starkville2010	S1_272273063		16.52	6.62×10^{-05}	0.068		
2	Starkville2010	S1_280635905	1.10	25.10	1.14×10^{-06}	0.101	Gene contains linked SNP 55 kb away, LD > 0.8	GRMZM2G009958
	CollStation2009	S1_280635905		15.98	8.83×10^{-05}	0.072	Gene is 47 kb from marker and is outside LD window	GRMZM2G164580
	Starkville2010	S1_280635931		25.10	1.14×10^{-06}	0.101	Gene is 45 kb from marker and lies in LD window > 0.67	GRMZM2G312239
	CollStation2009	S1_280635931		15.98	8.83×10^{-05}	0.072	Gene contains linked SNP 56 kb away, LD > 0.7	GRMZM2G009944
	Starkville2010	S1_280635934		25.10	1.14×10^{-06}	0.101		
	CollStation2009	S1_280635934		15.98	8.83×10^{-05}	0.072		
	Starkville2010	S1_280635950		25.10	1.14×10^{-06}	0.101		
	CollStation2009	S1_280635950		15.98	8.83×10^{-05}	0.072		
	Starkville2010	S1_280635958		25.10	1.14×10^{-06}	0.101		
	CollStation2009	S1_280635958		15.98	8.83×10^{-05}	0.072		
	Starkville2010	S1_280635967		19.52	1.61×10^{-05}	0.080		
	CollStation2009	S1_280635967		16.33	7.48×10^{-05}	0.077		
3	Average	S2_22947761	2.03	22.05	4.24×10^{-06}	0.076	Gene contains marker	GRMZM2G003784
	Starkville2009	S2_22947761		18.15	2.82×10^{-05}	0.063	Gene is 2.4 kb from marker but no nearby LD info	GRMZM2G003769
4	Starkville2010	S2_153128978	2.06	20.98	8.53×10^{-06}	0.111	Gene contains marker	GRMZM2G155437
	Average	S2_153128978		16.47	7.30×10^{-05}	0.088		
5	Starkville2010	S2_183190432	2.06	23.27	3.80×10^{-06}	0.138	Gene is 5 kb from marker and lies in LD window = 0.5	GRMZM2G037574
	Average	S2_183190432		16.22	9.45×10^{-05}	0.116	Gene contains linked SNP 325.6 kb away, LD > 0.3	GRMZM2G019901
							Gene is 7 kb from marker but no nearby LD info	GRMZM2G057852
6	Starkville2009	S2_188872774	2.07	21.35	6.08×10^{-06}	0.075	Gene contains marker	GRMZM2G026065
	Starkville2010	S2_188872774		17.59	3.78×10^{-05}	0.066		
	Starkville2010	S2_188872782		17.59	3.78×10^{-05}	0.066		
	Average	S2_188872911		19.55	1.47×10^{-05}	0.078		
	Lubbock2009	S2_188872911		17.14	4.76×10^{-05}	0.067		
7	Starkville2010	S2_205035222	2.08	23.26	2.38×10^{-06}	0.081	Gene contains marker	GRMZM2G166337
	Average	S2_205035222		16.17	7.53×10^{-05}	0.057		
8	Starkville2010	S3_127577411	3.05	21.89	4.85×10^{-06}	0.092	Gene is 18 kb from marker but not in LD	GRMZM2G333619
	Average	S3_127577411		16.43	6.84×10^{-05}	0.063	Gene contains linked SNP 36 kb away, LD = 0.32	GRMZM2G160840
9	Starkville2010	S3_217358368	3.09	23.73	1.85×10^{-06}	0.080	Gene contains marker	GRMZM2G089525
	Average	S3_217358368		19.78	1.25×10^{-05}	0.067		

(cont'd)

genome could be used in the analysis. Genes present in other maize lines will not have been sampled, and these will probably include resistance genes, as B73 is one of the more susceptible maize varieties in the association panel.

These 27 regions will be considered of interest to aflatoxin resistance breeding, but nonsignificant *p*-values do not lend sufficient confidence to these regions for in-depth speculation as to role in resistance mechanisms. Gene model, expression, and linkage disequilibrium information of the 66 genes that are closely linked to these 27 genomic locations are presented in Supplemental Table S3 for the information of the reader. Ontology of some of the

Table 2. Continued.

Region	Environment	SNPs	BIN	*F*	*p*-value	R^2	Proximity†	Gene
10	Lubbock2009	S3_217808747	3.09	27.33	3.83×10^{-07}	0.112	Gene contains marker	GRMZM2G052991
	Starkville2010	S3_217808272		17.40	4.19×10^{-05}	0.070	Gene contains linked SNP 12 kb away, LD > 0.46	GRMZM2G053047
	Lubbock2009	S3_217808302		20.76	8.14×10^{-06}	0.081		
	Lubbock2009	S3_217808740		27.42	3.72×10^{-07}	0.113		
	Average	S3_217808740		16.85	5.62×10^{-05}	0.071	Gene is 2.6 kb from marker but no nearby LD info	GRMZM2G377609
	Average	S3_217808747		16.71	6.00×10^{-05}	0.070		
11	CollStation2009	S3_217820604	3.09	15.67	9.66×10^{-05}	0.054	Gene contains marker	GRMZM2G053047
	Starkville2010	S3_217820604		15.63	9.84×10^{-05}	0.055	Gene contains linked SNP 12 kb away, LD > 0.46	GRMZM2G052991
12	Starkville2010	S3_217832603	3.09	21.56	5.65×10^{-06}	0.081	Gene contains marker	GRMZM2G053140
	Average	S3_217832603		18.25	2.80×10^{-05}	0.072	Gene contains linked SNP 44 kb away, LD > 0.3	GRMZM2G148532
	StarkvilleR2010	S3_217832603		16.05	8.25×10^{-05}	0.063	Gene is 1.3 kb from marker but no nearby LD info	GRMZM2G053047
13	Average	S4_26406913	4.04	16.74	5.63×10^{-05}	0.056	Gene contains marker	GRMZM2G003814
							Gene is 6.3 kb from marker but no nearby LD info	GRMZM2G003642
14	StarkvilleR2010	S4_26653796	4.04	24.31	1.52×10^{-06}	0.090	Gene contains marker	GRMZM2G111261
							Gene is 1.7 kb from marker but not in LD	GRMZM2G111300
							Gene is 1.1 kb from marker and lies in LD window > 0.3	GRMZM2G111228
15	Starkville2010	S5_20311281	5.03	23.04	3.04×10^{-06}	0.108	Gene contains marker	GRMZM2G545615
	Average	S5_20311281		22.49	3.94×10^{-06}	0.101		
	Lubbock2009	S5_20311281		17.25	4.79×10^{-05}	0.077		
16	Average	S5_185650835	5.05	23.31	2.63×10^{-06}	0.105	Gene is 5 kb from marker and in LD = 0.3	GRMZM2G028736
	Starkville2010	S5_185650835		21.96	4.97×10^{-06}	0.102	Gene is 4 kb from marker but no nearby LD info	GRMZM2G028830
17	Starkville2010	S5_206795119	5.07	22.94	3.58×10^{-06}	0.131	Gene contains marker	GRMZM2G105874
	Starkville2010	S5_206795116		21.73	6.13×10^{-06}	0.121	Gene contains linked SNP 45 kb away, LD > 0.31	GRMZM2G105844
	Average	S5_206795116		18.12	3.35×10^{-05}	0.087	Gene contains linked SNP 50.1 kb away, LD > 0.6	GRMZM2G162233
	StarkvilleR2010	S5_206795116		18.08	3.41×10^{-05}	0.093	Gene contains linked SNP 160 kb away, LD = 0.3	GRMZM2G127632
	StarkvilleR2010	S5_206795119		18.92	2.32×10^{-05}	0.099		
	Average	S5_206795119		17.55	4.46×10^{-05}	0.088		
	Starkville2010	S5_206797714		18.66	2.42×10^{-05}	0.081		
18	Starkville2010	S6_74967491	6.01	23.12	2.73×10^{-06}	0.095	Gene contains marker	GRMZM2G005499
	Average	S6_74967491		16.31	7.31×10^{-05}	0.063	Gene is 48 kb from marker and lies in LD window > 0.37	AC186592.3_FGT003
	Starkville2010	S6_74967515		18.56	2.42×10^{-05}	0.076		
	Starkville2010	S6_74967516		18.56	2.42×10^{-05}	0.076		
	Starkville2010	S6_74967517		18.56	2.42×10^{-05}	0.076		
	Starkville2010	S6_74967532		18.56	2.42×10^{-05}	0.076		
19	Starkville2010	S6_121311711	6.05	20.68	8.26×10^{-06}	0.075	Gene contains marker	GRMZM5G841142
	Average	S6_121311711		19.82	1.25×10^{-05}	0.073	Gene contains linked SNP 59.5 kb away, LD > 0.55	GRMZM2G136800
	Starkville2009	S6_121311711		16.93	5.20×10^{-05}	0.062	Gene is 19 kb from marker and lies in LD window > 0.55	AC231604.1_FGT001
							Gene contains linked SNP 95 kb away, LD = 0.3	GRMZM5G841142
							Gene is 58.7 kb from marker and lies in LD window = 0.3	GRMZM2G149700
20	Starkville2010	S7_155754021	7.03	21.89	7.09×10^{-06}	0.155	Gene is 8.6 kb from marker and lies in LD window > 0.5	GRMZM2G444075
	Average	S7_155752575		19.66	1.52×10^{-05}	0.094	Gene contains linked SNP 207 kb away, LD = 0.31	GRMZM2G173863
	Average	S7_155752599		19.66	1.52×10^{-05}	0.094		

(cont'd)

Table 2. Continued.

Region	Environment	SNPs	BIN	*F*	*p*-value	R^2	Proximity†	Gene
21	Starkville2010	S8_94752247	8.02	25.30	8.98×10^{-07}	0.089	Gene contains marker	GRMZM2G147221
	Average	S8_94752242		20.36	9.58×10^{-06}	0.071	Gene contains linked SNP 118.7 kb away, LD > = 0.3	GRMZM2G083344
	Starkville2010	S8_94752242		19.72	1.30×10^{-05}	0.069	Gene is 3.8 kb from marker but no nearby LD info	GRMZM2G147271
	Average	S8_94752243		20.39	9.42×10^{-06}	0.071		
	Starkville2010	S8_94752243		19.98	1.15×10^{-05}	0.069		
	Average	S8_94752246		21.07	6.77×10^{-06}	0.074		
	Starkville2010	S8_94752246		20.64	8.35×10^{-06}	0.072		
	Average	S8_94752247		23.63	2.00×10^{-06}	0.085		
	Average	S8_94752249		20.11	1.08×10^{-05}	0.070		
	Starkville2010	S8_94752249		19.76	1.28×10^{-05}	0.070		
	Starkville2010	S8_94752251		25.10	9.83×10^{-07}	0.090		
	Average	S8_94752251		23.34	2.27×10^{-06}	0.083		
	Average	S8_94752253		20.61	8.42×10^{-06}	0.072		
	Starkville2010	S8_94752253		19.75	1.28×10^{-05}	0.068		
	Average	S8_94752254		20.11	1.08×10^{-05}	0.070		
	Starkville2010	S8_94752254		19.76	1.28×10^{-05}	0.070		
	Average	S8_94752255		20.11	1.08×10^{-05}	0.070		
	Starkville2010	S8_94752255		19.76	1.28×10^{-05}	0.070		
	Starkville2010	S8_94752257		25.10	9.83×10^{-07}	0.090		
	Average	S8_94752257		23.34	2.27×10^{-06}	0.090		
	Average	S8_94752258		20.11	1.08×10^{-05}	0.070		
	Starkville2010	S8_94752258		19.76	1.28×10^{-05}	0.070		
	Average	S8_94752280		17.09	4.73×10^{-05}	0.058		
22	Average	S9_107333254	9.04	22.52	3.47×10^{-06}	0.085	Gene is 1 kb from marker but no nearby LD info	GRMZM2G331766
	Lubbock2009	S9_107333254		21.87	4.76×10^{-06}	0.085		
23	Average	S9_117048726	9.04	43.12	2.87×10^{-10}	0.162	Gene contains marker	GRMZM2G108619
	Starkville2010	S9_117048726		38.46	2.24×10^{-09}	0.144	Gene is 6.3 kb from marker but not in LD	GRMZM2G409245
	Starkville2009	S9_117048726		22.99	2.78×10^{-06}	0.088	Gene is 4 kb from marker but not in LD	GRMZM2G108668
	Average	S9_117048731		43.12	2.87×10^{-10}	0.162		
	Starkville2010	S9_117048731		38.46	2.24×10^{-09}	0.144		
	Starkville2009	S9_117048731		22.99	2.78×10^{-06}	0.088		
24	Starkville2009	S10_95855766	10.04	23.80	2.04×10^{-06}	0.095	Gene contains linked SNP 38 kb away, LD > 0.3	GRMZM2G407650
	Average	S10_95855766		19.68	1.44×10^{-05}	0.079		
25	Starkville2009	S10_109718061	10.04	20.60	8.86×10^{-06}	0.079	Gene contains marker	GRMZM2G159675
	Average	S10_109718061		19.83	1.28×10^{-05}	0.073	Gene contains linked SNP 4 kb away, LD > 0.4	GRMZM2G159720
	Lubbock2009	S10_109718061		16.28	7.29×10^{-05}	0.063	Gene contains linked SNP 13 kb away, LD > 0.3	GRMZM2G159732
							Gene is 106 kb away ad lies in LD window > 0.38	GRMZM2G130351
							Gene contains linked SNP 338.5 kb away, LD > 0.38	GRMZM2G126194
26	Average	S10_125923329	10.04	24.44	1.34×10^{-06}	0.084	Gene contains marker	GRMZM2G058573
	StarkvilleR2010	S10_125923329		21.63	5.17×10^{-06}	0.079		
27	Starkville2010	S10_139081513	10.06	24.04	1.71×10^{-06}	0.087	Gene contains marker	GRMZM2G148467
	Average	S10_139081513		16.95	5.23×10^{-05}	0.064	Gene is 33.8 kb from marker and lies in LD window > 0.3	GRMZM2G124965
	Starkville2010	S10_139081517		24.04	1.71×10^{-06}	0.087	Gene is 4 kb from marker and lies in LD window > 0.3	GRMZM2G148475
	Average	S10_139081517		16.95	5.23×10^{-05}	0.064		

† LD, linkage disequilibrium as measured by r^2.

genes is highly suggestive, including MYB stress-response factors, one WRKY disease-resistance factor, and one known fungal resistance gene. While they warrant further study, they will not be discussed further in this manuscript.

Single Nucleotide Polymorphisms in Different Clusters

Allele frequencies of candidate genes containing SNP–trait associations exceeding FDR in the overall mapping panel

and in each cluster can be found in Supplemental Table S4; a more detailed analysis of these and the other 25 associations of low probability is presented in Supplemental Table S2. The resistant allele of the two SNPs in GRMZM2G052991 was present in the B73, Lancaster, and Tropical clusters, and allele frequencies were significantly different between the Tropical and the MT clusters, and between the Lancaster and the MT clusters. The resistant alleles of the two SNPs in gene GRMZM2G108619 were present in the SC76, Tropical and MT clusters, and allele distribution was different between SC76 and all other clusters except MT, and between the Tropical and MT clusters. Finally, the resistant allele of the single SNP in gene GRMZM2G158141 was present only in the Northern US and Tropical clusters, and frequencies were significantly different between the Tropical and the MT clusters, and between the Northern US and the MT clusters. Unfortunately, the allele frequencies in the clusters are very unbalanced because of the size of each cluster and the actual allele frequency differences, and this can lead to cryptic false positives difficult to resolve with substructure and kinship mixed-model analysis as has been show for *dwarf8* (Larsson et al., 2013).

The Tropical cluster and the B73 cluster contained lines with significantly less aflatoxin than the other clusters (Warburton et al., 2013). To ensure that the B73 lines did not show reduced aflatoxin due to a heterotic response with the Lancaster tester, Va35, GWAS was performed with and without the B73 cluster. There were no differences found in the results of the GWAS performed with or without the 23 B73 related lines, and no significant SNP associations were lost or gained (data not shown). The kinship data included in the Q + K model seems to have accounted for both the allele frequency differences in different clusters and the possible effects of heterosis causing a bias in the results, or the bias was too small to be noted in the GWAS results. Because no new SNP associations were identified at high probabilities, only the results of the analysis with all lines were presented here.

Genomic-Wide Association Study of Flowering Time

The number of days from planting to midsilk (when 50% of the plants had extruding silks) had been recorded to inoculate all ears at the same physiological stage. A GWAS was performed on days to silking as well, to see if the same SNPs would be associated with both aflatoxin levels and female flowering time. These data can be found in Supplemental Table S5. Two SNP regions associated with female flowering time ($p < 10^{-4}$) on chromosomes 1 and 5 were found within one million base pairs of SNPs associated with aflatoxin resistance, but since linkage disequilibrium values between these SNPs and the aflatoxin-associated SNPs were low (0.133 and 0.008 for chromosomes 1 and 5, respectively), they are probably not influencing that association. One SNP was found associated with female flowering time ($p < 10^{-5}$) and was within 80 kb of an aflatoxin-associated SNP on chromosome 10. Despite being so close, the flowering time and the aflatoxin SNPs were not in linkage disequilibrium (r^2 values were less than 0.013).

Corroborating Evidence for Single Nucleotide Polymorphism Associations

If significantly associated SNPs fall within the region of previously reported QTL identified through biparental linkage mapping, this lends weight to the claim that SNP is causing, or very closely linked to, the trait being mapped. Of the four SNP regions with association probabilities under FDR, only one on chromosome 2 fell within the interval of a significant QTL in previously published mapping populations (as reviewed in Mideros et al., 2013). The associated SNP in bin 2.07 falls within the confidence interval of a QTL in two mapping populations, and the resistant form of the SNP is found in the resistant parent (Mp313E) and the susceptible form of the SNP in the susceptible parents (Va35 and B73). Not finding a SNP within the region of a previously reported QTL interval does not disprove the SNP effect and may indicate only that it was not polymorphic and segregating in the two parents of a given population.

A previous GWAS study, using a panel that contained many of the same lines as in the present study, was conducted in different years but some of the same locations (Barrero et al., 2015). Unfortunately, many fewer SNPs were used in the previous GWAS (only 61K), and many of the most significantly associated SNPs in the current study were not included. Additionally, aflatoxin resistance was not the main focus of the previous study, but grain aflatoxin levels were measured in that panel. No SNP–trait associations with aflatoxin accumulation were found lower than FDR, but three SNPs were associated consistently and with low *p*-value in multiple environments. One of these regions (within 130 bps) was associated with aflatoxin levels in the current study at lower than 9.68×10^{-4}, not a significant value but suggestive of corroboration. In addition, one of the SNPs found significantly associated in the current study (S3_217808272) was associated with aflatoxin levels in two environments and the average across environments in the Barrero et al. (2015) study (at $p = 7.40\ 10^{-4}$).

The frequency of the six significantly associated SNPs was investigated in a much larger set of several thousand maize inbred lines from the US national seed bank analyzed with GBS previously (Romay et al., 2013). There is significant overlap in the two data sets, but many of the inbred lines in the aflatoxin association mapping panel reported in the current study are not contained in the US seed bank study; the majority of the inbreds unique to the current data set are the most aflatoxin accumulation resistant lines. Because some of the SNPs are completely linked to each other, only five haplotypes were segregating in the inbreds

of both data sets. The SNP 2_187449158 (not in any gene) was not segregating in the US seed bank data set (Romay et al., 2013), and the resistant allele occurred only in the aflatoxin association mapping panel at a frequency of 8.1% (data not shown) where it appeared to be enriched particularly in the resistant lines. Many of the resistant alleles from the present study are rare in maize in general. The resistant allele for SNP 3_217808740 and 3_217808747 in gene GRMZM2G052991 was present in 1.9% of the individuals in the US seed bank data set and in 7.2% of the aflatoxin association mapping panel; for SNP 9_117048726 and 9_117048731 in gene GRMZM2G108619, the resistance allele was present in 7.9% of the in the US seed bank data set and 11.0% in the aflatoxin association panel; and for SNP 10_91956547 in gene GRMZM2G158141, the resistant allele was present in 7.0% of the individuals in the US seed bank data set and 9.8% in the aflatoxin association panel. This last SNP was the one where the resistant allele was most frequent in the US seed bank inbred lines out of all significantly associated SNPs and was found across diverse (and even some temperate) germplasm, whereas the resistance alleles at other associated SNPs were mainly found in tropical material (and were always enriched in the most resistant lines).

The six SNPs with *q*-values less than 10% FDR and the SNPs in the 25 additional regions have been converted to KASP genotyping assays (LGC Genomics) and ordering information is available from the author. These can now be used for independent confirmation of phenotypic effect in near isogenic or recombinant inbred lines or in other aflatoxin panels. Eventually, they can be used in marker-assisted backcrossing of these genes into elite temperate germplasm currently lacking sufficient aflatoxin accumulation resistance and possibly to guide a more informed genomic selection (Snelling et al., 2013). Meanwhile, a pathway analysis using all SNP–trait associations from this study is ongoing. By including all the genome-wide SNP associations rather than those exceeding specific *p*-values, we hope to highlight resistance mechanisms that involve metabolic pathways associated with reduced aflatoxin accumulation.

Acknowledgments

The authors wish to thank Jack Haynes for his excellent technical assistance, the Cornell Institute for Genomic Diversity for their help in generating and analyzing the GBS data, and Drs. W. Suwarno and I. Barrero for their critical review of the manuscript. This work was supported by the USDA and the joint USDA–USAID Feed the Future initiative. This paper is a joint contribution of USDA–ARS and the Mississippi Agricultural and Forestry Experiment Station and is published as Journal No. J-12668 of the Miss. Agric. and Forestry Exp. Stn. Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by USDA.

References

Barrero, I.D., G.N. De La Fuente, S.C. Murray, T. Isakeit, P.C. Huang, M.L. Warburton, W.P. Williams, G.L. Windham, and M. Kolomiets. 2015. Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. PLoS ONE 10:e0117737. doi:10.1371/journal.pone.0117737

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, et al. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. doi:10.1093/bioinformatics/btm308

Brooks, T.D., W.P. Williams, G.L. Windham, M.C. Willcox, and H.K. Abbas. 2005. Quantitative trait loci contributing resistance to aflatoxin accumulation in maize inbred Mp313E. Crop Sci. 45:171–174.

Busboom, K.N., and D.G. White. 2004. Inheritance of resistance to aflatoxin production and Aspergillus ear rot of corn from the cross of inbreds B73 and Oh516. Phytopathology 94:1101–1115. doi:10.1094/PHYTO.2004.94.10.1107

Campbell, K.W., and D.G. White. 1995. Evaluation of corn genotypes for resistance to Aspergillus ear rot, kernel infection and aflatoxin production. Plant Dis. 79:1039–1045. doi:10.1094/PD-79-1039

Chen, Z.Y., R.L. Brown, K.E. Damann, and T.E. Cleveland. 2007. Identification of maize kernel endosperm proteins associated with resistance to aflatoxin contamination by *Aspergillus flavus*. Phytopathology 97:1094–1103. doi:10.1094/PHYTO-97-9-1094

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. doi:10.1371/journal.pone.0019379

Flint-Garcia, S.A., A.C. Thuillet, J.M. Yu, G. Pressoir, S.M. Romero, S.E. Mitchell, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2005. Maize association population: A high resolution platform for quantitative trait locus dissection. Plant J. 44:1054–1064. doi:10.1111/j.1365-313X.2005.02591.x

Gerdes, J.T., C.F. Behr, J.G. Coors, W.F. Tracy, J.L. Geadelmann, and M.K. Viney. 1993. Compilation of North American maize breeding germplasm. CSSA, Madison, WI.

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346. doi:10.1371/journal.pone.0090346

Gong, Y.Y., P.C. Turner, A.J. Hall, and C.P. Wild. 2008. Aflatoxin exposure and impaired child growth in West Africa: An unexplored international public health burden? In: J.F. Leslie, R. Bandyopadhyay, and A. Visconti, editors, Mycotoxins. Detection methods, management, public health and agricultural trade. CABI, Wallingford, Oxfordshire. p. 53–65.

Groopman, J.D., T.W. Kensler, and C.P. Wild. 2008. Protective interventions to prevent aflatoxin-induced carcinogenesis in developing countries. Annu. Rev. Public Health 29:187–203. doi:10.1146/annurev.publhealth.29.020907.090859

Hamblin, A.M., and D.G. White. 2000. Inheritance of resistance to Aspergillus ear rot and aflatoxin production of corn from Tex6. Phytopathology 90:292–296. doi:10.1094/PHYTO.2000.90.3.292

Jiang, Y., P.E. Jolly, P. Preko, S.W. Jia, W.O. Ellis, T.D. Phillips, and J.H. Williams. 2008. Aflatoxin-related immune dysfunction in health and in human immunodeficiency virus disease. Clin. Dev. Immunol. 2008:790309. doi:10.1155/2008/790309

Kelley, R.Y., W.P. Williams, J.E. Mylroie, D.L. Boykin, L.K. Hawkins, G.L. Windham, T.D. Brooks, S.M. Bridges, B.E. Scheffler, and J.R. Wilkinson. 2009. Genomic profile of maize response to *Aspergillus flavus* infection. Toxin Rev. 28:129–141. doi:10.1080/15569540903089239

Kump, K.L., P.J. Bradbury, R.J. Wisser, E.S. Buckler, A.R. Belcher, M. Oropeza-Rosas, J.C. Zwonitzer, S. Kresovich, M.D. McMullen, D. Ware, P.J. Balint-Kurti, and J.B. Holland. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat. Genet. 43:163–168. doi:10.1038/ng.747

Larsson, S.J., A.E. Lipka, and E.S. Buckler. 2013. Lessons from *Dwarf8* on the strengths and weaknesses of structured association mapping. PLoS Genet. 9:e1003246 doi:10.1371/journal.pgen.1003246

Liu, Y., and F. Wu. 2010. Global burden of Aflatoxin-induced hepatocellular carcinoma: A risk assessment. Environ. Health Perspect. 118:818–824. doi:10.1289/ehp.0901388

Mayfield, K., F. J. Betrán, T. Isakeit, G. Odvody, S.C. Murray, W.L. Rooney, and J.C. Landivar. 2012. Registration of maize germplasm lines Tx736, Tx739, and Tx740 for reducing preharvest aflatoxin accumulation. J. Plant Registrations 6:88–94.

Mideros, S.X., M.L. Warburton, T.M. Jamann, G.L. Windham, W.P. Williams, and R.J. Nelson. 2013. Quantitative trait loci for resistance to Aspergillus ear rot: Analysis by linkage mapping, characterization of near-isogenic lines and meta-analysis. Crop Sci. 54:127–142. doi:10.2135/cropsci2013.04.0249

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association mapping: Critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202. doi:10.1105/tpc.109.068437

Payne, G.A. 1992. Aflatoxins in maize. Crit. Rev. Plant Sci. 10:423–440. doi:10.1080/07352689209382320

Pechanova, O., T. Pechan, W.P. Williams, and D.S. Luthe. 2011. Proteomic analysis of the maize rachis: Potential roles of constitutive and induced proteins in resistance to *Aspergillus flavus* infection and aflatoxin accumulation. Proteomics 11:114–127. doi:10.1002/pmic.201000368

Pornsri, K., G.S. Shephard, and F. Wu. 2011. Aflatoxins and growth impairment: A review. Crit. Rev. Toxicol. 41:740–755. doi:10.3109/10408444.2011.575766

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rafalski, J.A. 2011. Association genetics in crop improvement. Curr. Opin. Plant Biol. 13:174–180. doi:10.1016/j.pbi.2009.12.004

Robertson Hoyt, L.A., M.P. Jines, P. Balint-Kurti, C.E. Kleinschmidt, D.G. White, G.A. Payne, C.M. Maragos, T.L. Molnar, and J.B. Holland. 2006. QTL mapping for Fusarium ear rot and fuminisin contamination resistance in two maize populations. Crop Sci. 46:1734–1743. doi:10.2135/cropsci2005.12-0450

Romay, M.C., M.J. Millard, J.C. Glaubitz, J.A. Peiffer, K.L. Swarts, T.M. Casstevens, R.J. Elshire, C.B. Acharya, S.E. Mitchell, S.A. Flint-Garcia, M.D. McMullen, J.B. Holland, E.S. Buckler, and C.A. Gardner. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14:R55. doi:10.1186/gb-2013-14-6-r55

SAS Institute. 2012. JMP Genomics, release 6.1. SAS Inst. Inc., Cary, NC.

Scott, G.E., and N. Zummo. 1988. Sources of resistance in maize to kernel infection by *Aspergillus flavus* in the field. Crop Sci. 28:504–507. doi:10.2135/cropsci1988.0011183X002800030016x

Scott, G.E., and N. Zummo. 1990. Registration of Mp313E parental line of maize. Crop Sci. 30:1378. doi:10.2135/cropsci1990.0011183X003000060080x

Scott, G.E., and N. Zummo. 1992. Registration of Mp420 germplasm line of maize. Crop Sci. 32:1296. doi:10.2135/cropsci1992.0011183X003200050054x

Schaeffer, M.L., L.C. Harper, J.M. Gardiner, C.M. Andorf, D.A. Campbell, E.K. Cannon, T.Z. Sen, and C.J. Lawrence. 2011. MaizeGDB: Curation and outreach go hand-in-hand. Database 2011:bar022. doi:10.1093/database/bar022

Schnable, J., and A. Thompson. 2014. qTeller: Simple tool to identify genes under your QTL. http://qteller.com/ (accessed 7 Apr. 2014)

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T.A. Graves, et al. 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112–1115. doi:10.1126/science.1178534

Sekhon, R., K.L. Childs, C.N. Hansey, C.R. Buell, N. de Leon, and S.M. Kaeppler. 2011. Genome-wide atlas of transcription during maize development. Plant J. 66:553–563. doi:10.1111/j.1365-313X.2011.04527.x

Setter, T.L., J. Yan, M. Warburton, J.M. Ribaut, Y. Xu, M. Sawkins, E.S. Buckler, Z. Zhang, and M. Gore. 2011. Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. J. Exp. Bot. 62:701–716. doi:10.1093/jxb/erq308

Snelling, W.M., R.A. Cushman, J.W. Keele, C. Maltecca, M.G. Thomas, M.R.S. Fortes, and A. Reverter. 2013. Networks and pathways to guide genomic selection. J. Anim. Sci. 91:537–552. doi:10.2527/jas.2012-5784

Storey, J.D., and R. Tibshirani. 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA 100:9440–9445. doi:10.1073/pnas.1530509100

van Egmond, H.P., R.C. Schothorst, and M.A. Jonker. 2007. Regulations relating to mycotoxins in food: Perspectives in a global and European context. Anal. Bioanal. Chem. 389:147–157. doi:10.1007/s00216-007-1317-9

Warburton, M.L., W.P. Williams, G. Windham, S. Murray, W. Xu, L. Hawkins, and J. Franco. 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

Williams, W.P., and G.L. Windham. 2001. Registration of maize germplasm line Mp715. Crop Sci. 41:1374–1375. doi:10.2135/cropsci2001.4141374x

Williams, W.P., and G.L. Windham. 2006. Registration of Mp718 and Mp719 germplasm lines of maize. J. Plant Reg. 6:200–202. doi:10.3198/jpr2011.09.0489crg

Williams, W.P., and G.L. Windham. 2012. Registration of maize germplasm line Mp717. Crop Sci. 46:1407–1408. doi:10.2135/cropsci2005.09-0330

Wild, C.P., and Y.Y. Gong. 2010. Mycotoxins and human disease: A largely ignored global health issue. Carcinogenesis 31:71–82 10.1093/carcin/bgp264. doi:10.1093/carcin/bgp264

Yan, J.B., M. Warburton, and J. Crouch. 2011. Association mapping for enhancing maize genetic improvement. Crop Sci. 51:433–449. doi:10.2135/cropsci2010.04.0233

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208. doi:10.1038/ng1702

Zhang, Y., M.S. Kang, and R. Magari. 1999. Genetics of resistance to kernel infection by *Aspergillus flavus* in maize. Plant Breed. 116:146–152. doi:10.1111/j.1439-0523.1997.tb02169.x

Zhang, Z.W., E. Ersoz, C.Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, P.J. Bradbury, J. Yu, D.K. Arnett, J.M. Ordovas, and E.S. Buckler. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. 42:355–360. doi:10.1038/ng.546