

A Genome-Wide Association Study for Partial Resistance to Maize Common Rust

Bode A. Olukolu, William F. Tracy, Randall Wisser, Brian De Vries, and Peter J. Balint-Kurti

First author: Department of Plant Pathology and Department of Horticulture, North Carolina State University, Raleigh 27695; second and fourth authors: Department of Agronomy, University of Wisconsin-Madison, Madison 53706; third author: Department of Plant & Soil Sciences, University of Delaware, Newark 19716; and fifth author: Department of Plant Pathology, North Carolina State University, and United States Department of Agriculture–Agricultural Research Service Plant Science Research Unit, Raleigh, NC 27695.
Accepted for publication 9 March 2016.

ABSTRACT

Olukolu, B. A., Tracy, W. F., Wisser, R., De Vries, B., and Balint-Kurti, P. J. 2016. A genome-wide association study for partial resistance to maize common rust. *Phytopathology* 106:745-751.

Quantitative resistance to maize common rust (causal agent *Puccinia sorghi*) was assessed in an association mapping population of 274 diverse inbred lines. Resistance to common rust was found to be moderately correlated with resistance to three other diseases and with the severity of the hypersensitive defense response previously assessed in the same population. Using a mixed linear model accounting for the confounding

effects of population structure and flowering time, genome-wide association tests were performed based at 246,497 single-nucleotide polymorphism loci. Three loci associated with maize common rust resistance were identified. Candidate genes at each locus had predicted roles, mainly in cell wall modification. Other candidate genes included a resistance gene and a gene with a predicted role in regulating accumulation of reactive oxygen species.

Additional keywords: association analysis, disease resistance, linkage disequilibrium, QTL.

Maize common rust (CR; causal agent *Puccinia sorghi* Schwein.) is a major disease of field corn and is a particular problem in sweet corn and popcorn production (Groth et al. 1983; Randle et al. 1984). CR symptoms are most obvious after flowering and can result in as much as 49% yield loss in the most susceptible maize cultivars (Groth et al. 1983). The utility of chemical fungicides is limited due to their adverse environmental effects, cost, and limited effectiveness. Genetic resistance is the best, most sustainable alternative. Over 20 resistance major (*R*) genes, known as *Rp* genes, conferring race-specific resistance to CR have been mapped to complex loci on chromosomes 3, 4, and 10 (Delaney et al. 1998; Hooker 1985). These *R* genes mediate the recognition of the corresponding pathogen race, and trigger a rapid localized cell death known as the hypersensitive response (HR) that prevents the spread of the pathogen.

Complete resistance conferred by major *R* genes is generally overcome eventually (Hooker 1985; Vanderplank 1984). Partial or quantitative disease resistance (QDR) reduces but does not entirely prevent pathogen colonization of plant tissue and tends to be more durable. QDR is usually controlled by several genes, each with small effects. QDR can interfere with the infection process in a number of ways (e.g., by altering infection frequency, spore production, incubation period, latent period, and speed of growth within the leaf, among other processes) (Davis et al. 1988; Gingera et al. 1995).

A few studies have mapped quantitative trait loci (QTL) underlying QDR to maize CR. Lübberstedt et al. (1998) detected 20 QTL across four independent mapping populations derived from European maize flint lines with at least one QTL on each of the 10 maize chromosomes. Kerns et al. (1999) identified 16 QTL associated with CR resistance in a population derived from a cross between the temperate maize lines FRMo17 and BS11(FR)c7. Brown et al. (2001) identified two consistently detected CR resistance QTL in $F_{2,3}$ families derived from a cross between the sweetcorn lines IL731a and W6786.

Genome-wide association studies (GWAS) using diverse lines can provide a much higher level of resolution than traditional QTL mapping, potentially to the gene level (Flint-Garcia et al. 2005). This technique is now routinely employed in plant systems to precisely locate QTL for quantitative traits and to identify the underlying candidate genes. Among several association populations that have been used for GWAS in maize (Liu et al. 2011; Lu et al. 2010; Yan et al. 2011; Yu et al. 2011), a panel containing 279 inbred lines (hereafter known as the maize diversity panel) that captures the global diversity in public-sector breeding populations (Flint-Garcia et al. 2005) has been used most frequently (Cook et al. 2012; Harjes et al. 2008; Olukolu et al. 2013; Szalma et al. 2005; Thornsberry et al. 2001; Wilson et al. 2004; Zila et al. 2013).

We previously assessed the diversity panel for three diseases (Wisser et al. 2011): northern leaf blight (NLB; causal agent *Exserohilum turcicum*), southern leaf blight (SLB; causal agent *Bipolaris maydis*), and gray leaf spot (GLS; causal agent *Cercospora zae-maydis*). We also assessed the diversity panel for the strength of the autoactive HR associated with the mutant allele *Rp1-D21* (Olukolu et al. 2013).

In this study, we assessed the diversity panel for resistance to CR and derive genetic correlations between resistance to CR and resistance to SLB, NLB, and GLS and with HR severity in this population. Furthermore, we used GWAS to identify three loci associated with partial resistance to maize CR.

Corresponding authors: B. Olukolu; E-mail address: baolukol@ncsu.edu; and P. Balint-Kurti; E-mail address: peter.balint-kurti@ars.usda.gov

*The e-Xtra logo stands for “electronic extra” and indicates that six supplementary figures, one supplementary table, and five supplementary files are published online.

<http://dx.doi.org/10.1094/PHYTO-11-15-0305-R>

This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source. The American Phytopathological Society, 2016.

MATERIALS AND METHODS

Plant materials and field trials. A maize diversity panel comprising 274 inbred lines was used for this GWAS. The 274 lines (Supplementary File S1) were drawn from a well-characterized association mapping panel of 279 inbred lines (here termed the diversity panel) representing the global diversity of maize available in the public sector of corn-breeding programs (Flint-Garcia et al. 2005). The two trials were planted at West Madison Agricultural Station in a plano silt loam soil on 20 May 2009 and 8 May 2013. A completely randomized block design with two replications in each year was employed. A row of a constant genotype was planted around the edges of the field to eliminate border-row effects. Standard fertilizer, herbicide, and pesticide regimes were applied during the trial, while overhead irrigation was applied as required to ensure normal plant growth. In total, 12 kernels were sown in 3.5-m-long rows, interrow spacing was 0.76 m, and a 0.91-m alley was included at the end of each plot.

Evaluation of resistance to maize CR. The inoculum used in each case was a mix of races of *P. sorghi* spores collected during the previous season in Madison, WI and stored in the freezer until use. A suspension of approximately 15 mg of *P. sorghi* spores was added to 1 liter of water with 5 drops of Tween 20 surfactant. The spore suspension was sprayed from a Solo backpack sprayer into the whorl of the plants when they were at the V8 to V10 growth stage. Maize CR was scored once after flowering on a scale of 0-to-9, with a score of 0 indicating “no rust” and 9 indicating “most severe” disease symptom. The flowering dates, days to anthesis (DTA), and days to silking were scored and used as covariates (DTA) in the genome-wide analysis to account for the confounding effect of flowering.

Statistical analyses in association population. For phenotypic data, to compute inbred line least square mean (LSmean) values adjusted for environmental effects, the rust data were analyzed with a mixed model fitted with inbred lines as fixed effects, while environment, replications nested within environment, line–environment interaction, and DTA were fitted as random effects using the proc mixed procedure in SAS (v9.3) (SAS Institute, Inc. 2000-2004). The significance of the random factors in the model was determined based on the Wald’s Z-test statistic implemented with the restricted maximum likelihood (REML) method (Littell et al. 2006). The LSmeans were used as input phenotype data for the genome-wide association analysis (GWAA). To estimate heritability for CR, a mixed model was fitted as described above and was calculated as previously described (Holland et al. 2002).

For correlation analysis, genotypic correlations were estimated based on REML in ASReml software v3 (Gilmour et al. 2005), following procedures outlined by Wissler et al. (2011) and Holland (2006). Disease scores for SLB, GLS, and NLB were obtained from Wissler et al. (2011). HR scores were obtained from Olukolu et al. (2013).

Genotypic data. Genotypic data from 246,497 single-nucleotide polymorphism (SNP) markers were used for this study. These data were derived from two independent data sets: an Illumina maize 50k array (Ganal et al. 2011) and a genotyping-by-sequencing (GBS) dataset (Romay et al. 2013). Only SNP with <20% missing genotypes

were used. Based on the criterion above, the Illumina maize 50k array data set was filtered from 57,838 to 45,519 SNP, while the GBS SNP data set was filtered from 681,257 to 200,978 SNP.

GWAA was performed while taking account of the confounding effects of population structure (Hirschhorn and Daly 2005) using the kinship matrix (genetic relationship between all pairs of lines), which had been previously estimated in the diversity panel as described by Olukolu et al. (2013). GWAA was performed using the Tassel software (version 4.1.8) with a mixed linear model (MLM). The vector of phenotypes (y) was modeled as $y = X\beta + Zu + e$, where β = vector consisting of fixed effects, including each of the 246,495 SNP being tested; u = vector of the random additive genetic background effects associated with the lines; e = vector of residual effects, and X and Z are the incidence matrices that relate y to β and y to u . The variances of the random effects are modeled as $\text{Var}(u) = 2KV_g$, where $K = n \times n$ matrix of pairwise kinship coefficients that define the degree of genetic covariance between inbred lines and V_g = genetic variance (Yu et al. 2005).

The REML estimates of variance components were estimated following an efficient mixed-model association algorithm method (Kang et al. 2008; Zhang et al. 2010), with the optimum compression MLM and the P3D options. This clusters the lines into groups in order to increase statistical power and computational speed (Zhang et al. 2010).

P values from all the association tests were used as inputs to compute the corresponding positive false discovery rate (FDR) values using the QVALUE software (version 1.0) (Storey and Tibshirani 2003) in the R package (R Development Core Team 2008). With high-density SNP data, background linkage disequilibrium (LD) between SNP is useful for GWAS but can be problematic for multiple testing. Use of P values from SNP in LD violates the assumption of the FDR test because SNP should be completely independent. This can lead to an inflated positive rate (type I error) under nominal significance thresholds (e.g., $\alpha = 0.05$), hence resulting in loss of power (Benjamini and Yekutieli 2001; Nyholt 2004). We have reduced the data set prior to FDR tests to a subset of SNP with a minimum of 2-kbp intervals. Additionally, a few highly associated SNP were retained in the data set while maintaining the 2-kbp intervals so as to test the significance of highly associated SNP. P value threshold (3.0×10^{-5} , $-\log_{10}P = 4.5$) was set at a Q value of 0.10 after the FDR test.

In order to derive resample model inclusion probability (RMIP) values, a second round of GWAA was performed based on a subsampling (or subbagging) method (Valdar et al. 2009). Subsampling was performed 1,000 times, with 80% of the inbred lines subsampled each time. In addition to the P value threshold determined by the FDR test, only SNP detected in at least 250 of the 1,000 subsamples were considered significant (RMIP threshold of 0.25).

Candidate gene selection. The maize B73 reference genome assembly v2 on the MaizeGDB genome browser (Andorf et al. 2010) was used to establish the location of the SNP relative to the candidate genes. Annotation of candidate genes was obtained from the Maize Genome database (http://gbrowse.maizegdb.org/gb2/gbrowse/maize_v2/), Blast2G0 (Conesa et al. 2005), and using tools on the National Center for Biotechnology Information website, including BLASTp (Altschul et al. 1997) and conserved domain search (Marchler-Bauer et al. 2005). Additional annotation of the gene homologs was done by reviewing relevant literature.

RESULTS

Evaluation of the resistance to maize CR. Maize CR was scored in the diversity panel across two environments, with two replications in each environment. Variance components, test of normality, and heritability estimates were calculated (Table 1). Line effect and line–environment (G–E) interaction were significant. This is consistent with the high phenotypic correlation coefficients (Supplementary Table S1) between replicates in the same

TABLE 1. Significance of factors in mixed model, test of normality, and heritability estimates for resistance to maize common rust (CR) scores in the maize diversity panel^a

| Trait | Line | Rep (Env) | Line × Env | SW ^b | Heritability basis | |
|-------|--------------|-----------|--------------|-----------------|--------------------|------|
| | | | | | Line-mean | Plot |
| CR | $P < 0.0001$ | ns | $P < 0.0001$ | 0.98 | 0.75 | 0.36 |

^a Env = environment, Rep = replicates, and ns = not significant.

^b Shapiro-Wilk (SW) parameter is measured between 0 and 1, and being close to 1 indicates normality. Skewness and Kurtosis for Rust are 0.19 and 1.18, respectively.

environment (0.63 to 0.70), the moderate phenotypic correlations between replicates from the two different environments (0.34 to 0.47), and the heritability estimates across the environments (line mean-basis of 0.75 and plot-basis of 0.36). LSmean values for each line accounting for confounding effects due to variable flowering time (i.e., DTA) were calculated. The Shapiro-Wilk test (Table 1) and diagnostic plot showed that the LSmeans were normally distributed (Supplementary Fig. S1).

After accounting for the effects of population structure and coancestry (Olukolu et al. 2013; Wissler et al. 2011), CR was positively correlated with resistance to the other maize diseases tested (GLS, SLB, and NLB) and with severity of HR defense response (0.18) (Table 2). The strongest correlation with CR was observed for NLB (0.39).

GWAA mapping of maize CR resistance loci. DTA made very little contribution (0.13%) to total variance. In fact, GWAA with and without DTA as a covariate revealed very little difference in results (i.e., very slight changes in P values). The Q-Q plot (Supplementary Fig. S2) for GWAA based on the LSmeans

TABLE 2. Genotypic correlations between resistance to common rust (CR) and resistance to other maize diseases and severity of the hypersensitive defense response (HR) in the maize diversity panel^a

| | Disease scores | | | HR |
|----|---------------------|------------------------|---------------------|---------------------|
| | SLB _{inv} | NLB _{inv} | GLS _{inv} | |
| CR | 0.19* ($n = 244$) | 0.39**** ($n = 240$) | 0.18* ($n = 236$) | 0.18* ($n = 202$) |

^a HR severity due to *Rp1D21* mutant allele, which causes autoactive HR based on the mutant to wild type height ratio (Olukolu et al. 2013); SLB = southern leaf blight resistance score; GLS = gray leaf spot resistance score; NLB = northern leaf blight resistance score; inv indicates that lesion/disease rating was inverted, so that all scores were in the same “direction” (score indicates higher susceptibility/ milder HR in every case); and n = sample size. Significance of correlation coefficients are shown as ****, ***, **, and * = $P < 0.0001$, 0.001, 0.01, and 0.05, respectively. Positive correlations indicate that increasing CR resistance is associated with increasing resistance to the other diseases and with increasing strength of the HR response. Scoring scales: Rust = 0 (resistant) to 9 (susceptible), SLB = 0 (susceptible) to 9 (resistant), NLB = 0% (susceptible) to 100% (resistant), and GLS = 0 (susceptible) to 9 (resistant). HR was scored as the ratio of the height of mutant plants carrying the *Rp1-D21* gene to isogenic wild-type plants, so that the most severe phenotypes approached a value of 0, while lines in which *Rp1-D21* has a very mild effect have values close to 1 (Olukolu et al. 2013).

revealed that the associations were well controlled for population structure (i.e., low false-positive rate). It also implied that we can detect strong associations, reflected in the deviations at the top end and left side of the diagonal. Similar observations were shown in the single-environment analyses, except for the 2009 trial, where the false-positive rate was high (Supplementary Figs. S3 and S4). Following the FDR test and RMIP analysis, the P value threshold was set at $-\log_{10}P$ of 4.5 (corresponding to an FDR Q value of 0.1) and RMIP threshold of 0.25.

GWAA identified six SNP and four candidate genes significantly associated with CR based on the P value and RMIP threshold mentioned above. One of these SNP was located on chromosome 2, four on chromosome 3, and one on chromosome 8 (Fig. 1; Table 3; Supplementary Files S2 to S4). Three of the significantly associated SNP on chromosome 3 at 21,262,214, 21,262,226, and 21,262,243 bp were within 29 bp of each other and cosegregated in the diversity panel so that their RMIP and P values were identical.

The SNP loci were all fitted into a single full model in order to estimate the proportion of total phenotypic variation (R^2) explained. After stepwise selection, three of the four loci (all except the chromosome 8 SNP locus) were significant and explained 27.2% of the total phenotypic variation. GWAA of CR values from the single-environment data also detected associated SNP and candidate genes based on a $-\log_{10}P$ threshold of 7.2 (corresponding to an FDR Q value of 0.002) and 5.9 (corresponding to an FDR Q value of 0.05) for 2009 and 2013 trials, respectively. Based on the P value thresholds and an RMIP threshold set at 0.25, one and two SNP were detected as significantly associated in the 2009 and 2013 trials, respectively.

Candidate genes and colocalization with maize CR QTL.

Candidate genes that might underlie the detected effects were identified based on proximity of the associated SNP with predicted genes (File S4). The candidate genes were defined and annotated based on the filtered gene set of the B73 reference maize genome, version 2 (http://ensembl.gramene.org/Zea_mays/Info/Index) (Schnable et al. 2009). The SNP were localized within the promoter, an intron, an exon, and 22 kb downstream of the stop codon of xylanase, powdery-mildew-resistant protein 5 (PMR5), methionine sulfoxide reductase B2 (MSRB2), and aroenate dehydratase 6 (ADT6) genes (Table 3). In the 2009 trial, the associated SNP colocalized within the exon of an R gene (RGA2 homolog). Two SNP, which colocalize within the exons of a hypothetical protein

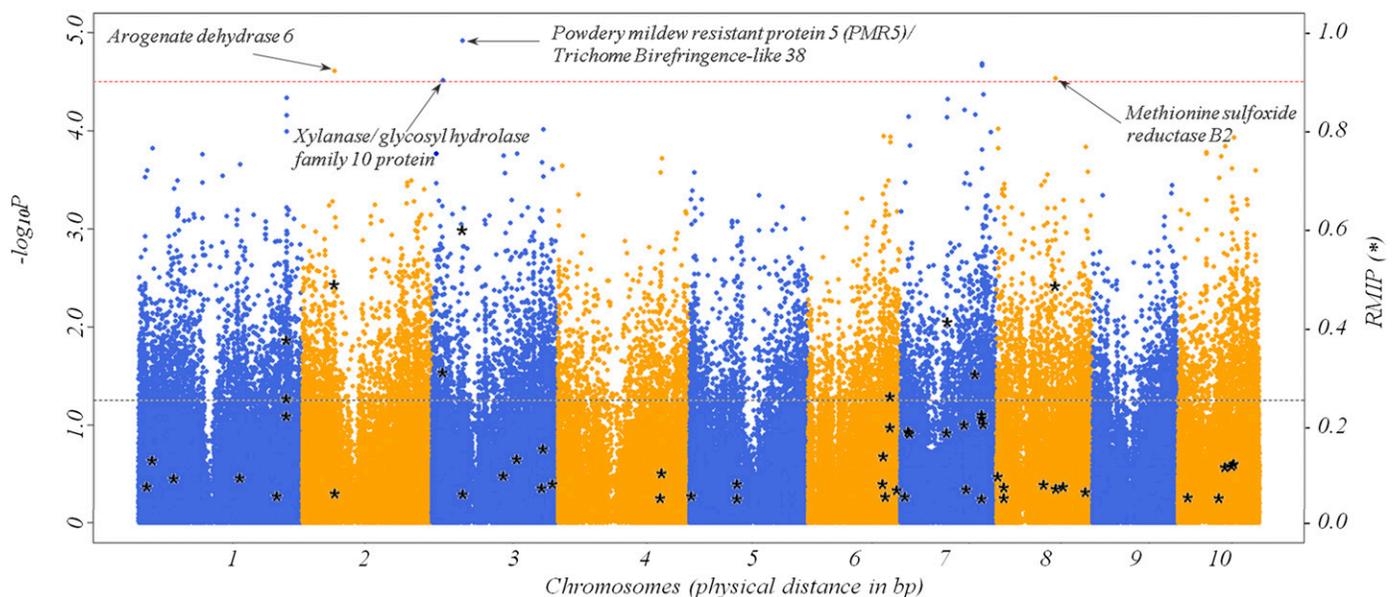


Fig. 1. Results of genome-wide association analysis in the maize diversity panel showing the significantly associated single-nucleotide polymorphisms (SNP) above the P value ($-\log_{10}P = 4.5$) and resample model inclusion probability (RMIP) (0.25) thresholds. The left vertical axis indicates the $-\log_{10}$ of P values and the right vertical axis indicates the RMIP values, while the horizontal axis indicates chromosomes and physical map positions of SNP. The lower dashed line indicates the RMIP threshold and the upper dashed line indicates the P value threshold. Dots represent the P values of each SNP and the stars represent their RMIP values.

and a nodulin MtN21/EamA-like transporter family protein, were detected in the 2013 trial (Supplementary Figs. S5 and S6).

DISCUSSION

Correlation analyses. In a previous study, we compared the resistance levels of the diversity panel to the three foliar fungal maize diseases SLB, GLS, and NLB (Wisser et al. 2011). We showed that resistances between these diseases were highly genetically correlated, suggesting that their genetic bases were controlled, in part, pleiotropically (i.e., some alleles conferred resistance to two or all three of these diseases). We also previously assessed the severity of the response observed when an H95 maize line carrying the mutant *Rp1-D21* allele that causes autoactive HR to resistance is crossed to each line of the diversity panel (Olukolu et al. 2013). The genetic correlations between HR and SLB, GLS, and NLB were either not significant or were much lower (approximately 0.1) than those between the diseases (0.62 to 0.67).

All the previous diseases studied were, to a greater or lesser extent, necrotrophic. Among them, NLB is considered the most biotrophic because it can penetrate living cells and, although it does eventually kill cells, it can grow for periods of several days in living plant tissue (Carson 1995; Hilu and Hooker 1964). CR, by contrast, is a strictly biotrophic disease. Therefore, it is of note that the strongest genetic correlation between the CR scores and those of the other diseases was between CR and NLB (0.39) (Table 2), and it is not surprising that the correlations between CR and the other three diseases are lower than those previously calculated between the

other three diseases. The fact that the correlations are positive and significant nevertheless indicates that resistance to CR and the other three diseases (especially with respect to NLB) may be controlled, to some extent, pleiotropically.

HR is a resistance mechanism that is primarily effective against biotrophic diseases; however, the correlations between the CR and HR scores (0.18) are not substantially higher than those observed between HR and the necrotrophic diseases (Olukolu et al. 2013). This is surprising but suggests that partial resistance to CR observed in this diversity panel may primarily be based on other mechanisms, rather than HR. It should also be noted that, for logistical reasons, HR was scored in F_1 hybrids with the line H95 while disease resistance was scored in the *per se* lines. Had HR been scored in *per se* lines, we might have expected higher correlations.

Association analysis. We used GWAA to identify SNP associated with resistance to maize CR. GWAA provides much more precise position estimates than the method of QTL mapping in biparental mapping populations that had been previously performed to analyze this trait. The rapid decay of LD in the diversity panel breaks the genome into small LD blocks and provides us with the ability to fine map QTL, oftentimes to the gene level (Olukolu et al. 2013). However, GWAA in association populations has limited statistical power for detecting QTL compared with linkage analysis in biparental mapping populations. This may be why we only detected three QTL loci in this study compared with the 8 to 20 QTL (most of which are small effect QTL) in other studies (Brown et al. 2001; Kerns et al. 1999; Lübberstedt et al. 1998).

TABLE 3. Significantly associated single-nucleotide polymorphism (SNP) hits for resistance to maize common rust (CR) in the maize diversity panel based on genome-wide association analysis and the candidate genes in close proximity of the SNP hits based on maize B73 reference genome assembly v2 (Andorf et al. 2010) and their predicted functions

| Chr ^a | Position (bp) ^b | P^c | RMIP ^d | Allele ^e | Freq. ^f | Effect ^g | R^2 (%) ^h | Gene ID | Gene or ortholog ⁱ | e-Value ^j | Funct ^k |
|------------------|----------------------------|-------|-------------------|---------------------|--------------------|---------------------|------------------------|---------------|-------------------------------|----------------------|--------------------|
| 2 | 59,014,463 | 4.61 | 0.49 | C | 0.977 | 2.0 | 8.3 | GRMZM2G437912 | ADT6 | 0.0 | dfr, phe |
| | ... | ... | ... | T | 0.023 | 0.0 | ... | ... | ... | ... | ... |
| 3 | 21,262,214 ^l | 4.51 | 0.31 | G | 0.967 | 1.63 | 7.5 | GRMZM2G031004 | XYL/GH10 | 1×10^{-171} | dfr, cwm |
| | ... | ... | ... | A | 0.033 | 0.0 | ... | ... | ... | ... | ... |
| 3 | 56,476,524 | 4.92 | 0.60 | A | 0.981 | -2.31 | 10.4 | GRMZM2G409309 | PMR5/TBL38 | 4×10^{-119} | dfr, cwm |
| | ... | ... | ... | G | 0.019 | 0.0 | ... | ... | ... | ... | ... |
| 8 | 107,796,411 | 4.54 | 0.48 | A | 0.869 | -0.91 | 7.5 | GRMZM2G089308 | MSRB2 | 2×10^{-76} | dfr, rh |
| | ... | ... | ... | G | 0.131 | 0.0 | ... | ... | ... | ... | ... |

^a Chromosome.

^b Physical map position based on B73 maize reference genome v2.

^c The $-\log_{10}$ of average P value from GWAA based on the 100 resampling method.

^d Resampling model inclusion probability.

^e Alleles are from homozygote genotypes.

^f Allele frequency is the proportional representation of each allele in the populations.

^g Allele effect is the effect of replacing one allele with the other. A positive effect implies increasing susceptibility.

^h R^2 = proportion of phenotypic variance explained by SNP.

ⁱ Maize gene name or best matching ortholog. ADT6 = argenatase dehydratase 6, XYL/GH10 = xylanase/glycosyl hydrolase 10, PMR5 = powdery-mildew-resistant protein 5, TBL38 = trichome birefringence-like 38, MSRB2 = methionine sulfoxide reductase B2.

^j Based on protein sequence alignment with the best *Arabidopsis* best matching ortholog.

^k Functional annotation: dfr = defense response, phe = phenylpropanoid pathway, cwm = cell wall modification, and rh = redox homeostasis.

^l Three significantly associated SNP on chromosome 3 at 21,262,214, 21,262,226, and 21,262,243 bp were within 29 bp of each other and cosegregated in the diversity panel so that their RMIP and P values were identical. The first of these SNP is reported here.

TABLE 4. Genome-wide association study-associated single-nucleotide polymorphisms (SNP) that overlap with quantitative trait loci (QTL) underlying partial resistance to maize common rust (CR) in three independent mapping populations

| SNP (chr:bp) ^b | Bin position (kb) | R^2 (%) ^a | | | |
|---------------------------|--|-----------------------------|----------------------------|----------------------------|----------------------------|
| | | IL731a × W6786 ^c | KW1265 × D146 ^d | KW1265 × D145 ^d | D145 × KW1292 ^d |
| 2:59,014,463 | Bin2:04 (Chr2:28,771,109-71,742,767) | 16.1 | ... | 25.5 | 14.5 |
| 3:21,262,214 | Bin3:04 (Chr3:13,102,315-126,236,346) | 20.0 | 5.6 | ... | 15.3 |
| 3:56,476,524 | Bin3:04 (Chr3:13,102,315-126,236,346) | 20.0 | 5.6 | ... | 15.3 |
| 8:107,796,411 | Bin8:03 (Chr8:108,690,693-122,478,285) | ... | 4.2 | 12.4 | 15.9 |

^a R^2 = proportion of phenotypic variance explained by QTL.

^b Associated SNP position (chromosome:base pair).

^c Maize CR scored over 2 years (Brown et al. 2001).

^d Maize CR scored over 3 years (Lübberstedt et al. 1998).

It should be noted that three of the four associated SNP loci have low minor allele frequencies (2 to 3%) (Table 3). These loci are likely genuinely associated with CR resistance because they are identified using relatively stringent thresholds (see Materials and Methods) (Fig. 1; Table 3) and their positions coincide with CR resistance QTL identified in previous studies (see below) (Table 4). Furthermore, the distribution of the lines carrying the minor alleles does not coincide with population subgroups (Supplementary File S5); therefore, these SNP associations are not likely an artifact of population structure.

Comparison of GWAS hits with previous CR QTL studies.

The four SNP loci identified in our association analysis map to three previously identified major QTL regions (Table 4) (Brown et al. 2001; Lübberstedt et al. 1998). Our analysis suggests that there is one gene underlying the QTL mapped to Bin2:04 and Bin8:03, while the QTL mapped to Bin 3:04 may be caused by the combined effects of two genes separated by approximately 35 MB. The phenomenon of multiple linked genes underlying a single QTL has been identified in many previous studies (Olukolu et al. 2013; Studer and Doebley 2011). Several major genes for resistance to specific races of CR have been identified (Wisser et al. 2006). Because these genes are race specific and because we used a mixed-race inoculum in this study, the presence of major genes would not be expected to have a large effect. The only major CR resistance gene that maps close to any of the association hits is the *Rp3* gene on chromosome 3 at approximately 114 Mb. This is still almost 60 Mb distant from the closest of our association hits on the same chromosome; therefore, major genes are unlikely to underlie any of the CR QTL we have identified in this study.

The single-environment GWAA for 2009 did identify a resistance gene homolog (RGA2 homolog) that had been previously isolated and mapped in a study by Collins et al. (1998) but without any association yet to resistance to a specific pathogen.

Candidate genes and underlying biological processes.

The identification of genes controlling disease resistance can lead to an understanding of the molecular and physiological mechanisms underlying resistance. We identified four candidate genes based on colocalization of the associated SNP with genic sequence (introns and exons) or flanking regulatory sequence (i.e., promoter sequences). Homology searches suggested that they might all play roles in the modification of cell wall components and redox homeostasis (Table 3).

ADT6 is one of the six members of a gene family in *Arabidopsis* involved in the biosynthesis of L-phenylalanine from L-arogenate or prephenate. L-phenylalanine is the substrate that feeds into the phenylpropanoid metabolic pathway required for production of lignin, flavonoids, and salicylic acid (SA). All of these molecules are involved in the plant defense response, with lignin also being an important component of plant cell walls. Although flavonoid compounds play multiple roles, their involvement in defense against pathogens, herbivores, and environmental stress has been well documented (Treutter 2005). They are important for detoxifying damaging reactive oxygen species (ROS) produced during the defense response (Grace and Logan 2000). SA is a signaling molecule that induces rapid systemic-acquired resistance through the SA-dependent defense pathway (Kunkel and Brooks 2002).

The two candidate genes identified within the chromosome 3 QTL (PMR5 and xylanase/glycosyl hydrolase 10) are both predicted to be involved in cell-wall modification. Xylanases are found in both fungal and plant genomes while PMR5 is plant specific. Fungal pathogen xylanases contain cellulose binding domains (Hall et al. 1989) and utilize their hydrolase activity to degrade xylan, a major component of hemicellulose. This weakens the host cell wall and mediates penetration and infection of the host cell. Plant xylanases have been identified in wheat and maize (Banik et al. 1996; Bih et al. 1999) and contain a signal peptide and catalytic domain. They are required to degrade the cell wall of the tapetum cells during seed germination and the cell wall of the stigma for initial pollen tube penetration (Slade et al. 1989). It is possible that CR might also co-opt

host xylanases to weaken host cells to enhance penetration and infection.

Arabidopsis lines with mutations in PMR5 exhibit several phenotypes, including cell walls enriched in pectin with reduced pectin modification, smaller cells, and increased resistance to the powdery mildew pathogens *Erysiphe cichoracearum* and *E. orontii* but not to the bacterial pathogens *Pseudomonas syringae* or *Peronospora parasitica* (Vogel et al. 2004). Other studies, however, suggest that PMR5 might be an important component of nonhost resistance in *Arabidopsis* (Nakao et al. 2011).

One candidate gene (MSRB2) was implicated in redox homeostasis. This is of interest because the oxidative burst and accumulation of ROS at the point of pathogen infection is a downstream response after detection of pathogens by *R* genes, and mediates localized cell death during HR. HR is particularly important for resistance to biotrophic pathogens such as CR. MSRB2 plays a role in the repair of oxidized methionine, the amino acid most sensitive to oxidation by ROS. MSRB2 reverts methionine sulfoxide to methionine and could participate in the scavenging of ROS and reactive nitrogen (Oh et al. 2010). Expression of MSRB2 leads to enhanced disease resistance in pepper, while the loss of function resulted in enhanced susceptibility and accelerated cell death (Oh et al. 2010).

In conclusion, we show here that the genetic control of partial resistance to CR and NLB appears, to some extent, to be pleiotropic but that CR resistance is less well correlated with resistance to SLB and GLS and to severity of the HR response. We report on the mapping of four CR resistance loci and the detection of plausible causal genes underlying partial resistance to maize CR. Their relatively large effects suggest that they might be useful for marker-assisted breeding.

ACKNOWLEDGMENTS

This work was funded by the United States Department of Agriculture–Agricultural Research Service and by the National Science Foundation Plant Genome research program grant number 0822495. We thank M. Goodman and the Maize Genetics Cooperation Stock Center for donating seed; E. Buckler, P. Bradbury, J. Glaubitz, and the Maize Diversity project team for access to the GBS genotype dataset; and J. Holland for helpful discussions.

LITERATURE CITED

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Andorf, C. M., Lawrence, C. J., Harper, L. C., Schaeffer, M. L., Campbell, D. A., and Sen, T. Z. 2010. The Locus Lookup tool at MaizeGDB: Identification of genomic regions in maize by integrating sequence information with physical and genetic maps. *Bioinformatics* 26:434-436.
- Banik, M., Garrett, T. P. J., and Fincher, G. B. 1996. Molecular cloning of cDNAs encoding (1->4)-beta-xylan endohydrolases from the aleurone layer of germinated barley (*Hordeum vulgare*). *Plant Mol. Biol.* 31:1163-1172.
- Benjamini, Y., and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29:1165-1188.
- Bih, F. Y., Wu, S. S. H., Ratnayake, C., Walling, L. L., Nothnagel, E. A., and Huang, A. H. C. 1999. The predominant protein on the surface of maize pollen is an endoxylanase synthesized by a tapetum mRNA with a long 5' leader. *J. Biol. Chem.* 274:22884-22894.
- Brown, A. F., Juvik, J. A., and Pataky, J. K. 2001. Quantitative trait loci in sweet corn associated with partial resistance to Stewart's wilt, northern corn leaf blight, and common rust. *Phytopathology* 91:293-300.
- Carson, M. L. 1995. Inheritance of latent period length in maize infected with *Exserohilum turcicum*. *Plant Dis.* 79:581-585.
- Collins, N. C., Webb, C. A., Seah, S., Ellis, J. G., Hulbert, S. H., and Pryor, A. 1998. The isolation and mapping of disease resistance gene analogs in maize. *Mol. Plant-Microbe Interact.* 11:968-978.
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.
- Cook, J. P., McMullen, M. D., Holland, J. B., Tian, F., Bradbury, P., Ross-Ibarra, J., Buckler, E. S., and Flint-Garcia, S. A. 2012. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* 158:824-834.

- Davis, D. W., Randle, W., and Groth, J. V. 1988. Some sources of partial resistance to common leaf rust (*Puccinia sorghi*) in maize and strategy for screening. *Maydica* 33:1-13.
- Delaney, D. E., Webb, C. A., and Hulbert, S. H. 1998. A novel rust resistance gene in maize showing overdominance. *Mol. Plant-Microbe Interact.* 11:242-245.
- Flint-Garcia, S. A., ThUILlet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., Doebley, J., Kresovich, S., Goodman, M. M., and Buckler, E. S. 2005. Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J.* 44:1054-1064.
- Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., Clarke, J. D., Graner, E.-M., Hansen, M., Joets, J., Le Paslier, M.-C., McMullen, M. D., Montalent, P., Rose, M., Schön, C.-C., Sun, Q., Walter, H., Martin, O. C., and Falque, M. 2011. A large maize (*Zea mays* L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS One* 6:e28334.
- Gilmour, A. R., Cullis, B., Gogel, B., Welham, S. J., and Thompson, R. 2005. *ASReml User's Guide*. Release 2.0. VSN International, Ltd., Hemel Hempstead, UK.
- Gingera, G. R., Davis, D. W., and Groth, J. V. 1995. Identification and inheritance of delayed first pustule appearance to common leaf rust in sweet corn. *J. Am. Soc. Hortic. Sci.* 120:667-672.
- Grace, S. C., and Logan, B. A. 2000. Energy dissipation and radical scavenging by the plant phenylpropanoid pathway. *Phil. Trans. Royal Soc. London Series B-Biol. Sci.* 355:1499-1510.
- Groth, J. V., Zeyen, R. J., Davis, D. W., and Christ, B. J. 1983. Yield and quality losses caused by common rust (*Puccinia sorghi* Schw.) in sweet corn (*Zea mays*) hybrids. *Crop Prot.* 2:105-111.
- Hall, J., Hazlewood, G. P., Huskisson, N. S., Durrant, A. J., and Gilbert, H. J. 1989. Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subspecies *cellulosa*—Internal signal sequence and unusual protein processing. *Mol. Microbiol.* 3:1211-1219.
- Harjes, C. E., Rocheford, T. R., Bai, L., Brutnell, T. P., Kandianis, C. B., Sowinski, S. G., Stapleton, A. E., Vallabhaneni, R., Williams, M., Wurtzel, E. T., Yan, J., and Buckler, E. S. 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319:330-333.
- Hilu, H. M., and Hooker, A. L. 1964. Host-pathogen relationship of *Helminthosporium turcicum* in resistant and susceptible corn seedlings. *Phytopathology* 54:570-575.
- Hirschhorn, J. N., and Daly, M. J. 2005. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6:95-108.
- Holland, J. B. 2006. Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. *Crop Sci.* 46:642-654.
- Holland, J. B., Nyquist, W. E., and Cervantes-Martinez, C. T. 2002. Estimating and interpreting heritability for plant breeding: An update. In: *Plant Breeding Reviews*, Vol. 22. J. Janick, ed. John Wiley & Sons, Inc., Oxford, UK.
- Hooker, A. L. 1985. Corn and sorghum rusts. Pages 207-236 in: *The Cereal Rusts: Diseases, Distribution, Epidemiology, and Control*. A. P. Roelfs and W. R. Bushnell, eds. Academic Press, New York.
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., and Eskin, E. 2008. Efficient control of population structure in model organism association mapping. *Genetics* 178:1709-1723.
- Kerns, M. R., Dudley, J. W., and Rufener, G. K. 1999. QTL for resistance to common rust and smut in maize. *Maydica* 44:37-45.
- Kunkel, B. N., and Brooks, D. M. 2002. Cross talk between signaling pathways in pathogen defense. *Curr. Opin. Plant Biol.* 5:325-331.
- Littell, R. C., Milliken, G. A., Stroup, W. A., Wolfinger, R. D., and Schabenberger, O. 2006. *SAS System for Mixed Models*, 2nd ed. SAS Institute, Inc., Cary, NC.
- Liu, W., Gowda, M., Steinhoff, J., Maurer, H., Würschum, T., Longin, C., Cossic, F., and Reif, J. 2011. Association mapping in an elite maize breeding population. *Theor. Appl. Genet.* 123:847-858.
- Lu, Y. L., Zhang, S. H., Shah, T., Xie, C. X., Hao, Z. F., Li, X. H., Farkhari, M., Ribaut, J. M., Cao, M. J., Rong, T. Z., and Xu, Y. B. 2010. Joint linkage-linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proc. Nat. Acad. Sci. USA* 107:19585-19590.
- Lübberstedt, T., Klein, D., and Melchinger, A. E. 1998. Comparative quantitative trait loci mapping of partial resistance to *Puccinia sorghi* across four populations of European flint maize. *Phytopathology* 88:1324-1329.
- Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., and Ke, Z. 2005. CDD: A conserved domain database for protein classification. *Nucleic Acids Res.* 33(Database Issue):D192-D196.
- Nakao, M., Nakamura, R., Kita, K., Inukai, R., and Ishikawa, A. 2011. Non-host resistance to penetration and hyphal growth of *Magnaporthe oryzae* in *Arabidopsis*. *Sci. Rep.* 1:Article 171. doi:10.1038/srep00171
- Nyholt, D. R. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* 74:765-769.
- Oh, S. K., Baek, K. H., Seong, E. S., Joung, Y. H., Choi, G. J., Park, J. M., Cho, H. S., Kim, E. A., Lee, S., and Choi, D. 2010. CaMsrb2, pepper methionine sulfoxide reductase B2, is a novel defense regulator against oxidative stress and pathogen attack. *Plant Physiol.* 154:245-261.
- Olukolu, B. A., Negeri, A., Dhawan, R., Venkata, B. P., Sharma, P., Garg, A., Gachomo, E., Marla, S., Chu, K., Hasan, A., Ji, J., Chintamani, S., Green, J., Shyu, C.-R., Wissner, R. J., Holland, J. B., Johal, G., and Balint-Kurti, P. 2013. A connected set of genes associated with programmed cell death implicated in controlling the hypersensitive response in maize. *Genetics* 193:609-620.
- Randle, W. M., Davis, D. W., and Groth, J. V. 1984. Improvement and genetic-control of partial resistance in sweet corn to corn leaf rust. *J. Am. Soc. Hortic. Sci.* 109:777-781.
- R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Romay, M., Millard, M., Glaubitz, J., Peiffer, J., Swarts, K., Casstevens, T., Elshire, R., Acharya, C., Mitchell, S., Flint-Garcia, S., McMullen, M., Holland, J., Buckler, E., and Gardner, C. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14:R55.
- SAS Institute, Inc. 2000-2004. *SAS 9.2 Help and Documentation*. SAS Institute, Inc., Cary, NC.
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reilly, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., Belter, E., Du, F., Kim, K., Abbott, R. M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S. M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M. J., McMahan, L., Van Buren, P., Vaughn, M. W., Ying, K., Yeh, C. T., Emrich, S. J., Jia, Y., Kalyanaraman, A., Hsia, A. P., Barbazuk, W. B., Baucom, R. S., Brutnell, T. P., Carpita, N. C., Chaparro, C., Chia, J. M., Deragon, J. M., Estill, J. C., Fu, Y., Jeddalo, J. A., Han, Y., Lee, H., Li, P., Lisch, D. R., Liu, S., Liu, Z., Nagel, D. H., McCann, M. C., SanMiguel, P., Myers, A. M., Nettleton, D., Nguyen, J., Penning, B. W., Ponnala, L., Schneider, K. L., Schwartz, D. C., Sharma, A., Soderlund, C., Springer, N. M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T. K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J. L., Dawe, R. K., Jiang, J., Jiang, N., Presting, G. G., Wessler, S. R., Aluru, S., Martienssen, R. A., Clifton, S. W., McCombie, W. R., Wing, R. A., and Wilson, R. K. 2009. The B73 maize genome: Complexity, diversity, and dynamics. *Science* 326:1112-1115.
- Slade, A. M., Hoj, P. B., Morrice, N. A., and Fincher, G. B. 1989. Purification and characterization of 3 (1-14)-beta-D-xylan endohydrolases from germinated barley. *Eur. J. Biochem.* 185:533-539.
- Storey, J. D., and Tibshirani, R. 2003. Statistical significance for genomewide studies. *Proc. Nat. Acad. Sci. USA.* 100:9440-9445.
- Studer, A. J., and Doebley, J. F. 2011. Do large effect QTL fractionate? A case study at the maize domestication QTL teosinte branched1. *Genetics* 188: 673-681.
- Szalma, S., Buckler, E., Snook, M., and McMullen, M. 2005. Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theor. Appl. Genet.* 110:1324-1333.
- Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D., and Buckler, E. S. 2001. Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28:286-289.
- Treutter, D. 2005. Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol.* 7:581-591.
- Valdar, W., Holmes, C. C., Mott, R., and Flint, J. 2009. Mapping in structured populations by resample model averaging. *Genetics* 182:1263-1277.
- Vanderplank, J. E. 1984. *Disease Resistance in Plants*, 2nd ed. Academic Press, New York.
- Vogel, J. P., Raab, T. K., Somerville, C. R., and Somerville, S. C. 2004. Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. *Plant J.* 40:968-978.
- Wilson, L. M., Whitt, S. R., Ibanez, A. M., Rocheford, T. R., Goodman, M. M., and Buckler, E. S. 2004. Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* 16:2719-2733.
- Wisser, R. J., Balint-Kurti, P. J., and Nelson, R. J. 2006. The genetic architecture of disease resistance in maize: A synthesis of published studies. *Phytopathology* 96:120-129.
- Wisser, R. J., Kolkman, J. M., Patzoldt, M. E., Holland, J. B., Yu, J., Krakowsky, M., Nelson, R. J., and Balint-Kurti, P. J. 2011. Multivariate

- analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a GST gene. *Proc. Nat. Acad. Sci.* 108:7339-7344.
- Yan, J., Warburton, M., and Crouch, J. 2011. Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. *Crop Sci.* 51:433-449.
- Yu, J. M., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., and Buckler, E. S. 2005. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38:203-208.
- Yu, L.-X., Lorenz, A., Rutkoski, J., Singh, R., Bhavani, S., Huerta-Espino, J., and Sorrells, M. 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theor. Appl. Genet.* 123:1257-1268.
- Zhang, Z., Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., and Buckler, E. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42:355-360.
- Zila, C. T., Samayoa, L. F., Santiago, R., Butrón, A., and Holland, J. B. 2013. A genome-wide association study reveals genes associated with fusarium ear rot resistance in a maize core diversity panel. *G3: Genes Genomes Genet.* 3:2095-2104.