

1 A Genome-Wide Association Study Reveals Genes Associated with Fusarium Ear Rot
2 Resistance in a Maize Core Diversity Panel

3
4 Charles T. Zila^{*}, L. Fernando Samayoa[§], Rogelio Santiago[§], Ana Butrón[§], and James B.
5 Holland^{*†1}

6
7 ^{*}Department of Crop Science, North Carolina State University, Raleigh, North Carolina 27695,

8 [§]Misión Biológica de Galicia, CSIC, Pontevedra, Spain, 36080, and [†]U.S. Department of
9 Agriculture—Agricultural Research Service, Plant Science Research Unit, Raleigh, North
10 Carolina, 27695

11

12 GWAS for Fusarium Ear Rot Resistance in Maize

13 KEYWORDS: association analysis, disease resistance, genotype by environment interaction,
14 maize, quantitative trait

15 ¹Corresponding author: USDA-ARS and Department of Crop Science, Campus Box 7620,
16 Raleigh, NC, 27695-7616. Phone: (919) 513-4198. E-mail: james_holland@ncsu.edu.

17

Abstract

18

19 Fusarium ear rot is a common disease of maize that affects food and feed quality
20 globally. Resistance to the disease is highly quantitative, and maize breeders have difficulty
21 incorporating polygenic resistance alleles from unadapted donor sources into elite breeding
22 populations without having a negative impact on agronomic performance. Identification of
23 specific allele variants contributing to improved resistance may be useful to breeders by allowing
24 selection of resistance alleles in coupling phase linkage with favorable agronomic characteristics.
25 We report the results of a genome-wide association study (GWAS) to detect allele variants
26 associated with increased resistance to Fusarium ear rot in a maize core diversity panel of 267
27 inbred lines evaluated in two sets of environments. We performed association tests with 47,445
28 SNPs while controlling for background genomic relationships with a mixed model and identified
29 three marker loci significantly associated with disease resistance in at least one subset of
30 environments. Each associated SNP locus had relatively small additive effects on disease
31 resistance ($\pm 1.1\%$ on a 0-100% scale), but nevertheless were associated with 3 to 12% of the
32 genotypic variation within or across environment subsets. Two of three identified SNPs
33 colocalized with genes that have been implicated with programmed cell death and were
34 expressed at highest levels during the onset of disease symptoms. An analysis of associated allele
35 frequencies within the major maize subpopulations revealed enrichment for resistance alleles in
36 the tropical/subtropical and popcorn subpopulations compared to other temperate breeding pools.

37

Introduction

38 The hemibiotrophic fungus *Fusarium verticillioides* (Sacc) Nirenberg is endemic in most
39 maize fields in the United States and is present in many arable regions of the world (VAN
40 EGMOND *et al.* 2007). This fungus causes Fusarium ear rot disease of maize, especially in low
41 rainfall high-humidity environments, such as the southern United States and some lowland
42 tropics (MILLER and TRENHOLM 1994). Infection by *F. verticillioides* can result in decreased
43 grain yields, poor grain quality, and contamination by the mycotoxin fumonisin, a suspected
44 carcinogen associated with various diseases in livestock and humans (MILLER and TRENHOLM
45 1994; MARASAS 1996; PRESELLO *et al.* 2008).

46 The best strategy for controlling Fusarium ear rot and reducing the incidence of
47 fumonisin contamination of grain is the development and deployment of maize hybrids with
48 genetic resistance. Fusarium ear rot resistance is under polygenic control and strongly influenced
49 by environmental factors; no fully immune genotypes have been discovered (KING and SCOTT
50 1981; NANKAM and PATAKY 1996; CLEMENTS *et al.* 2004). The complexity of this resistance
51 trait has impeded breeding, such that most commercial maize hybrids have lower levels of
52 resistance than are desirable (BUSH *et al.* 2004). Linkage-based mapping studies in biparental
53 populations have shown that Fusarium ear rot resistance QTL have relatively small effects and
54 are not consistent between populations (PÉREZ-BRITO *et al.* 2001; ROBERTSON-HOYT *et al.* 2006;
55 DING *et al.* 2008; MESTERHÁZY *et al.* 2012).

56 Despite the genetic complexity of resistance to Fusarium ear rot and fumonisin
57 accumulation, and despite the very low heritability of resistance measured on individual plants,
58 resistance on the basis of family means from well-replicated studies is moderately to highly
59 heritable (ROBERTSON *et al.* 2006; ELLER *et al.* 2008; BOLDUAN *et al.* 2009). Robertson *et al.*

60 (2006) and Bolduan et al. (2009) reported genotypic correlations between ear rot resistance and
61 fumonisin accumulation of 0.87 in North Carolina and 0.92 in Germany, respectively, indicating
62 that visual selection on Fusarium ear rot resistance should be effective in simultaneously
63 reducing fumonisin contamination. The heritability estimates predict, and empirical selection
64 studies demonstrate, that selection for improved ear rot resistance can be effective (ROBERTSON
65 *et al.* 2006; BOLDUAN *et al.* 2009; ELLER *et al.* 2010). Unfortunately, most sources having high
66 levels of ear rot resistance are older or exotic unadapted inbreds that lack the agronomic
67 performance of modern elite maize lines (CLEMENTS *et al.* 2004; ELLER *et al.* 2008, 2010). Thus,
68 breeders are faced with the difficulty of introducing polygenic resistance alleles of generally
69 small effect linked to inferior polygenic alleles for agronomic performance if they attempt to
70 incorporate improved genetic resistance from unadapted lines into elite breeding gene pools.
71 Identification of specific allelic variants that confer improved resistance would permit maize
72 breeders to select for rare recombinant chromosomes in backcross progeny that have desired
73 target resistance allele sequences in coupling phase with the favorable elite polygenic
74 background, facilitating the improvement of disease resistance without decreasing agronomic
75 performance.

76 Resolving small effect QTL to causal genes for traits that are difficult to accurately
77 measure phenotypically is exceedingly difficult in biparental mapping populations (HOLLAND
78 2007). Compared to traditional linkage-based analyses, association mapping offers higher
79 mapping resolution while eliminating the time and cost associated with developing synthetic
80 mapping populations (FLINT-GARCIA *et al.* 2005; YU and BUCKLER 2006). Historically, a major
81 limitation to association mapping in low linkage disequilibrium (LD) species such as maize has
82 been the large number of genetic markers required to detect marker-trait associations. Limiting

83 the search space to predetermined candidate genes allows for association mapping with a smaller
84 number of markers but requires extensive knowledge of the biochemical pathway contributing to
85 the trait of interest (REMYNTOON and PURUGGANAN 2003). To date, nothing is known about the
86 pathways contributing to Fusarium ear rot resistance in maize. However, the recent availability
87 of the maize 50k SNP genotyping array (GANAL *et al.* 2011) has provided almost 50,000 single
88 nucleotide polymorphism (SNP) markers scored on 279 of the 302 inbred lines of a commonly
89 used maize core diversity panel (FLINT-GARCIA *et al.* 2005; COOK *et al.* 2012). The maize
90 diversity panel captures much of the diversity present in public breeding programs worldwide.
91 The large number of markers available on the diversity panel has enabled genome-wide
92 association studies (GWAS) for several complex traits in maize including kernel composition
93 traits (COOK *et al.* 2012) and the hypersensitive response (OLUKOLU *et al.* 2013). Oluokolu *et al.*
94 (2013) identified SNPs associated with the hypersensitive defense response in or adjacent to five
95 genes not previously known *a priori* to affect disease resistance, but whose predicted gene
96 functions all involved the programmed cell death pathway. In this study, we employed GWAS to
97 identify SNPs associated with Fusarium ear rot resistance in the maize core diversity panel both
98 within and across two contrasting environments – North Carolina, USA and Galicia, Spain.

99

100

Materials & Methods

101 *Genotypes and experimental design*

102 The maize core diversity panel (sometimes referred to as the “Goodman” association
103 panel, because the seed stocks were originally assembled by Major Goodman at North Carolina
104 State University (FLINT-GARCIA *et al.* 2005)) was evaluated in several years in both North
105 Carolina, USA and Galicia, Spain. Only the 279 inbred lines with available genotypic data were

106 considered in this study. In the Galicia experiment, a subset of 270 inbred lines from the maize
107 diversity panel was evaluated for Fusarium ear rot resistance in a randomized 15×18 α -lattice
108 block design with two replicates in 2010 and 2011. Nine lines with insufficient seed were
109 dropped from the Galicia experiment before randomization. In the North Carolina experiment,
110 the maize diversity panel was part of an evaluation of the entire USDA maize seed bank
111 collection of inbred lines in 2010 (ROMAY *et al.* 2013) and subsets of that collection evaluated in
112 2011 and 2012. The genotypic data on the maize seed bank collection reported by Romay *et al.*
113 (2013) were not available at the time of analysis. The 2010 seed bank collection evaluation
114 included 2572 inbred line entries and was arranged in an augmented single replicate design.
115 Experimental entries were divided into 18 sets of differing sizes based on maturity and field
116 assignment. Each block within each set was augmented with a B73 check plot in a random
117 assignment, and five other checks (IL14H, Ki11, P39, SA24, and Tx303) were included once per
118 set in a random position. Romay *et al.* (2013) reported flowering time evaluations of the entire
119 collection evaluated at three locations in 2010, including North Carolina. Here we include data
120 only from North Carolina because it was the only environment used for Fusarium ear rot
121 evaluation. In 2011 and 2012, the maize core diversity panel was part of a larger sample of
122 inbreds evaluated. The larger population consisted of 771 diverse entries divided into eight sets
123 based on maturity and replicated across years. Although disease measurements were collected on
124 all experimental entries in both years, genotypic data were not available on inbreds outside of the
125 core diversity panel at the time of analysis. Sets were randomized within the field, and each set
126 was blocked using an α -lattice design. As with the seed bank collection evaluation, each block
127 was augmented by a randomly assigned B73 check plot, and five other checks (GE440, NC358,
128 NK794, PHB47, and Tx303) were included once per set.

129 The three North Carolina environments were artificially inoculated with local toxigenic
130 *Fusarium verticillioides* isolates using the toothpick method (CLEMENTS *et al.* 2003).
131 Approximately one week after flowering, a toothpick containing *F. verticillioides* spores was
132 inserted directly into the primary ear of five plants in each plot. At maturity, inoculated ears were
133 harvested and visually scored for Fusarium ear rot symptoms. Scores were assigned to each ear
134 in increments of 5% from 0% to 100% diseased based on the percentage of the ear presenting
135 disease symptoms (ROBERTSON *et al.* 2006; Figure S1). In Galicia, between seven and 14 days
136 after flowering, five primary ears per plot were inoculated with 2 mL of a spore suspension of
137 the local toxigenic isolate of *F. verticillioides*. The spore suspension contained 10^6 spores mL⁻¹
138 and was prepared following the protocol established by Reid *et al.* (1996) with some
139 modifications. Inoculum was injected into the center of the ear using a four-needle vaccinator
140 which perforated the husks and injured three to four kernels. Ears from each plot were collected
141 two months after inoculation and were individually rated for Fusarium ear rot symptoms using a
142 seven-point scale (1=no visible disease symptoms, 2=1-3%, 3=4-10%, 4=11-25%, 5=26-50%,
143 6=51-75%, and 7=76-100% of kernels exhibiting visual symptoms of infection, respectively)
144 devised by Reid and Zhu (2005). Phenotypic data on the seven-point scale from the Galicia
145 environments were transformed to the 0-100% scale used in North Carolina in the analyses.
146 Reliable data could not be obtained for some line-environment combinations because seed set for
147 some plots was limited due to poor adaptation. Raw data are provided in supplemental dataset
148 File S1. Climate data from on-farm weather stations were obtained from
149 <http://www.climate.ncsu.edu> and <http://www.mbg.csic.es/eng/index.php>.

150
151 *Genotypic data*

152 The genotypic data were 47,445 SNPs from the Illumina maize 50k genotyping array
153 filtered by Olukolu et al. (2013). The original array consists of 49,585 SNPs designed by Ganal
154 et al. (2011). Olukolu et al. (2013) filtered the data set to include only those SNP markers that
155 mapped to defined single locations in the maize genome and had <20% missing data
156 (<http://www.genetics.org/content/suppl/2012/12/05/genetics.112.147595.DC1/genetics.112.1475>
157 [95-3.txt](http://www.genetics.org/content/suppl/2012/12/05/genetics.112.147595.DC1/genetics.112.147595-3.txt)).

158

159 *Statistical Analyses*

160 *Estimation of least square means and heritabilities*

161 The Galicia and North Carolina experiments were analyzed separately and then combined
162 in a single multi-environment analysis. Each year of data within each experiment was first
163 analyzed separately by fitting a mixed linear model including line as a fixed effect, silking date
164 as a fixed linear covariate, and replication (Galicia only), block within replication (Galicia only),
165 set (North Carolina only), and block within set (North Carolina only) as random effects. The
166 mixed linear model for the Galicia experiment across years included line as a fixed effect, silking
167 date as a fixed linear covariate, and year, line×year interaction, replication within year, and block
168 within replication as random effects. The North Carolina experiment was analyzed across years
169 with a model including line as a fixed effect, silking date as a fixed linear covariate, and year,
170 line×year interaction, set within year, and block within set as random effects. In the combined
171 experiment analysis, each combination of location and year was considered an environment. The
172 combined analysis model included a fixed line effect, silking date as a fixed covariate nested
173 within environment, a random line×environment interaction effect, and nested random
174 experimental design effects (replication within environment and block within replication at

175 Galicia and set within environment and block within set at North Carolina). All analyses were
176 weighted by the number of ears scored within each plot and utilized a heterogeneous error
177 variance structure. In both experiments, larger predicted ear rot values were associated with
178 larger residuals, so a natural logarithmic transformation of raw ear rot scores (which largely
179 eliminated the relationship between residual variance and predicted values) was used for all
180 analyses. Least square means were estimated for 267 inbred lines within each experiment and
181 across experiments (File S2) using ASReml version 3 software (GILMOUR *et al.* 2009). Means
182 for twelve lines were not estimable due to missing phenotypic observations in all environments
183 (generally due to poor seed production).

184 We conducted a second analysis treating inbred lines as random effects for the purposes
185 of estimating heritability for Fusarium ear rot resistance in the diversity panel. The same models
186 as above were used except lines were treated as random effects to obtain estimates of genetic
187 variance. Line mean-basis heritability was estimated as

$$\hat{H}_c = 1 - \frac{\sigma_{PPE}^2}{2\hat{\sigma}_G^2}$$

188 where σ_{PPE}^2 is the average prediction error variance for all pairwise comparisons of lines and $\hat{\sigma}_G^2$
189 is the estimated genetic variance (CULLIS *et al.* 2006). We estimated line mean-basis
190 heritabilities for each environment individually, across the North Carolina environments, across
191 the Galicia environments, and we also estimated line mean-basis heritability for the combined
192 data set across all environments. The model used to estimate line mean-basis heritability in the
193 combined data set was further modified by nesting the random line effect within environment
194 and modeling the genotype-environment effect (**G**) matrix as unstructured, thereby allowing
195 estimation of unique genetic variance within each environment and a unique genetic correlation
196 between each pair of environments. For the purpose of estimating heritability, the average of the

197 ten pair-wise covariance estimates between environments (which are expected to equal the
198 genotypic variance) was used in the denominator of the above equation.

199 Silking date heritabilities were also calculated for each environment and across
200 environments. The same models used to compute ear rot heritabilities were used to estimate
201 silking date heritabilities, but silking date was treated as the dependent variable instead of as a
202 fixed linear covariate.

203

204 *Association analyses*

205 A genetic kinship matrix (**K**; File S3) based on observed allele frequencies (VANRADEN
206 2008; method 1) was created using R software version 3.0.0 (R CORE TEAM 2013). A subset of
207 4000 SNP markers were used to estimate **K**. Markers were uniformly distributed across the
208 genome (at least 60 kbp between adjacent markers) and had no missing data after excluding
209 heterozygous genotypes. Olukolu et al. (2013) used a kinship matrix produced by Tassel
210 software (BRADBURY *et al.* 2007), which is appropriate for population structure correction for
211 GWAS. In addition to population structure correction, we also wanted to estimate the polygenic
212 background genetic variance component, so we estimated a new **K** matrix that is scaled
213 appropriately to represent realized genomic average identity by descent relationships among the
214 lines (VANRADEN 2008).

215 Tassel version 4.1.24 was used for the genome-wide association analyses based on a
216 mixed linear model (BRADBURY *et al.* 2007). The least square means for inbred lines were used
217 as the input phenotypes, and each set of means (North Carolina, Galicia and combined) was
218 analyzed separately (File S2). The mixed linear model implemented by Tassel was

$$y = X\beta + Zu + e$$

219 where \mathbf{y} is the vector of ear rot least square means (on the natural-log scale), $\boldsymbol{\beta}$ is a vector of
220 fixed effects including SNP marker effects, \mathbf{u} is a vector of random additive genetic effects from
221 background QTL for lines, \mathbf{X} and \mathbf{Z} are design matrices, and \mathbf{e} is a vector of random residuals.
222 The variance of the \mathbf{u} vector was modeled as

$$\text{Var}(\mathbf{u}) = \mathbf{K}\sigma_a^2$$

223 where \mathbf{K} is the $n \times n$ matrix of pairwise kinship coefficients ranging 0—2 and σ_a^2 is the estimated
224 additive genetic variance (YU *et al.* 2006).

225 Restricted maximum likelihood estimates of variance components were obtained using
226 the optimum compression level and population parameters previously determined (P3D) options
227 in Tassel (ZHANG *et al.* 2010). The optimum compression level option reduces the
228 dimensionality of \mathbf{K} by clustering n lines into s groups, thereby reducing computational time and
229 potentially improving model fit. The P -values for each of the 47,445 tests of associations
230 between one SNP and ear rot resistance within each analysis were used to estimate the false
231 positive discovery rate (FDR) using the QVALUE version 1.0 package in R (STOREY and
232 TIBSHIRANI 2003). SNPs significant at $q < 0.10$ in the initial GWAS scan for a particular
233 environment set were then included together in a joint SNP association model together using the
234 GLM procedure in SAS software version 9.2 (SAS INSTITUTE INC 2010) to estimate the total
235 amount of variation explained by the SNPs together and to re-estimate their effects jointly.
236 Candidate genes either containing or located adjacent to associated SNPs were identified using
237 the MaizeGDB genome browser (ANDORF *et al.* 2010).

238

239 *Allele frequency analysis*

240 Lines were grouped into one of five major maize subpopulations (stiff stalk, non-stiff
241 stalk, tropical/subtropical, popcorn, and sweet corn) based on the population structure analysis of
242 the maize core diversity panel reported by Flint-Garcia et al. (2005;
243 http://panzea.org/db/gateway?file_id=pop_structure_xls). Lines of mixed ancestry (the result of
244 admixture among the subpopulations) were dropped from the analysis. Based on the results of
245 the association analyses, the frequencies of alleles that reduced disease severity at significant
246 SNPs were estimated within each subpopulation using the FREQ procedure using SAS software
247 version 9.2 (SAS INSTITUTE INC 2010). At each SNP locus, a Fisher's exact test in R software
248 version 3.0.0 (R CORE TEAM 2013) was used to test the null hypothesis that frequency of the
249 allele conferring increased disease resistance was the same across all five subpopulations.

250

251

Results

Line means and heritability

252
253 Significant ($P < 0.001$) genotypic variation for ear rot resistance was observed in both the
254 North Carolina and Galicia experiments. The mean ear rot observed among 267 inbred lines of
255 the association panel ranged from 4.4% to 100% with an overall mean of 41.1% in North
256 Carolina and from 0% to 89.3% with an overall mean of 7.4% in Galicia (File S2; Table S1). In
257 the combined analysis, mean ear rot ranged from 1.6% to 79.6% with an overall mean of 22.1%.
258 The silking date covariate was highly heritable ($\hat{H}_c = 0.98$ in the combined analysis) and was
259 significantly associated with ear rot resistance in the North Carolina and combined analyses ($P <$
260 0.001), but not in the Galicia analysis ($P = 0.099$; Table S1).

261

262 A significant ($P < 0.001$) line \times environment interaction was detected in the combined
analysis. Results of the mixed model analysis that estimate unique genotypic covariances for

263 each pair of environments indicated that the two Galicia environments had a much stronger
264 genotypic correlation ($r = 0.93$; Table 1 and Figure S2) than did any other pair of environments
265 (range, $r = 0.28$ to 0.51 ; Table 1 and Figure S2). Thus, there was little genotype \times environment
266 interaction between the two Galicia environments, and the heritability of line means across the
267 two years in Galicia was 0.71 . In contrast, pair-wise genotypic correlations were much lower
268 among the North Carolina environments and between North Carolina and Galicia environments
269 (Table 1 and Figure S2), generating much of the observed genotype \times environment interaction in
270 the combined analysis. Despite the strong genotype \times environment interaction among North
271 Carolina environments, heritability of genotype means across the three years in North Carolina
272 (0.73) was higher than within any single North Carolina environment (Table S1). In addition,
273 heritability of line means across all five environments was 0.75 , higher than within any single
274 environment or group of environments (Table S1). Therefore, we conducted separate association
275 analyses on three different sets of genotypic mean values for ear rot: (1) means from three North
276 Carolina environments, (2) means from two Galicia environments, and (3) means from the
277 combined analysis of all five environments.

278

279 *Association mapping of Fusarium ear rot resistance*

280 The optimum compression option in Tassel clustered the 267 lines into 229 groups in the
281 Galicia analysis and 197 groups in the North Carolina and combined analyses (Table 2).
282 Background genetic effects modeled by **K** accounted for 31% of the total variation among line
283 means in the North Carolina analysis, 57% of the total phenotypic variation in the Galicia
284 analysis, and 48% of the total phenotypic variation in the combined analysis (Table 2). In the
285 analysis of means from North Carolina environments, two SNPs were identified as significantly

286 associated with ear rot resistance at $q \leq 0.05$ (raw P -value = 2.4×10^{-7}), and one additional SNP
287 was identified at $q \leq 0.10$ (Table 3 and Figure 1). In the combined analysis, one SNP was
288 identified as significantly associated with ear rot resistance at $q \leq 0.05$ and coincided with one of
289 the SNPs identified in the North Carolina analysis. No SNPs significant at $q \leq 0.10$ were
290 identified in the Galicia analysis, where the minimum raw P -value among SNP association tests
291 was 2.1×10^{-4} .

292

293 *Candidate genes colocalized with associated SNPs*

294 Genes containing or nearby SNPs significantly associated with ear rot resistance were
295 characterized using the filtered predicted gene set from the annotated B73 reference maize
296 genome (SCHNABLE *et al.* 2009). Two of the three genes identified in the North Carolina analysis
297 have predicted functions that have been implicated in disease response pathways in other plant
298 species (TSUNEZUKA *et al.* 2005; HÉMATY *et al.* 2009). The SNP at physical position
299 151,295,233 bp on chromosome 9, which was identified in both the North Carolina and
300 combined analyses, is located in an intronic region of a cellulose synthase-like family A/mannan
301 synthase gene (Table 3). Mean LD r^2 between the chromosome 9 SNP and other SNPs dropped
302 below 0.1 within approximately 100 kbp (Figure 2). The other two SNPs identified in the North
303 Carolina analysis on chromosomes 1 and 5 were located inside of a gene of unknown function
304 and nearby a heat-shock 60-kDa protein (HSP60), respectively. Mean LD r^2 between the
305 chromosome 1 and chromosome 5 SNPs and other SNPs dropped below 0.1 within
306 approximately 10 kbp and 100 kbp, respectively (Figure 2). Although the chromosome 1 and 9
307 SNPs were not significantly associated with ear rot resistance in Galicia, the allele effects at
308 these loci were consistent between North Carolina and Galicia (Table 3). However, the allele

309 effect at the chromosome 5 SNP locus showed a change in direction between North Carolina
310 (+1.149%, Table 3) and Galicia (-0.017%).

311

312 *Allele frequencies at candidate genes*

313 We estimated the allele frequency at the three SNPs significantly associated with ear rot
314 resistance in five of the major maize subpopulations – stiff stalk temperate (SS), non-stiff stalk
315 temperate (NSS), tropical/sub-tropical (TS), popcorn (PC), and sweet corn (SC) (FLINT-GARCIA
316 *et al.* 2005). European flint types are poorly represented in this maize core diversity panel and
317 thus were not considered. Popcorn and sweet corn types were considered in the analysis, but
318 comparisons to either of these two subpopulations may be less reliable than comparisons to other
319 subpopulations due to smaller sample size (Table 4). The allele that reduced disease severity at
320 the chromosome 1 SNP locus is only present in the NSS and TS subpopulations but not at high
321 enough frequencies to be considered significantly different from the other three subpopulations
322 ($P=0.15$, Table 4). The allele with reduced disease severity at the chromosome 5 SNP locus is
323 significantly ($P=6.2\times 10^{-6}$) over-represented in TS and PC lines relative to other temperate (SS,
324 NSS, and SC) lines. At the chromosome 9 SNP locus, the allele associated with reduced disease
325 severity is significantly ($P=3.846\times 10^{-4}$) overrepresented in PC lines compared to the other four
326 subpopulations (Table 4). Averaging least square means from the combined analysis across
327 members of each subpopulation, the SS, NSS, TS, PC, and SC subpopulations had average ear
328 rot scores of 24.0%, 24.3%, 14.6%, 17.9%, and 46.5% respectively (Table 4).

329

330

330 **Discussion**

331 *Heritability and false discovery rate estimation*

332 The mean ear rot severity observed across experimental entries was 41.1% in North
333 Carolina and 7.4% in Galicia (Table S1). Mean ear rot in North Carolina 2012 was particularly
334 high (55%; Table S1). The very strong genotypic correlation between Galician environments
335 (Table 1 and Figure S2) justified their grouping as one environmental set in the analysis.
336 Genotypic effects were significantly correlated between each pair of North Carolina
337 environments, but at much lower magnitude (Table 1 and Figure S2). Genotypic values in North
338 Carolina 2010 had slightly higher correlations with the genotypic values in Galicia than in other
339 years of North Carolina (Table 1), so grouping the three North Carolina environments has little
340 justification based on genotype-by-environment patterns. Nevertheless, this environment
341 grouping has a natural interpretation in terms of geography and adaptation, and the heritability of
342 line means across these environments was higher than any individual environment, such that
343 analysis of the three years as a group simplified interpretation of results.

344 The relationship between the *F. verticillioides* isolates used in each location is unknown;
345 as such, it is possible that differences in pathogen aggressiveness could have contributed to the
346 disparity in mean ear rot values across environments. In addition, differences in inoculation
347 methods, as well as variation in temperature and precipitation levels, may have allowed for more
348 favorable disease development in North Carolina as compared to Galicia. Although precipitation
349 levels varied across all five environments, average daily temperatures (both pre- and post-
350 flowering) were higher in all three North Carolina environments compared to the two Galicia
351 environments (Table S2).

352 Heritabilities observed across environments in this study ($\hat{H}_c \geq 0.71$) are consistent with
353 estimates from biparental populations (ROBERTSON *et al.* 2006) and a small sample of North
354 American and European public inbred lines (BOLDUAN *et al.* 2009). These heritability estimates

355 were obtained with a model that assumed each line is a random sample from the reference
356 population of global maize inbreds, modeled by a genotypic variance-covariance structure equal
357 to the genotypic variance component multiplied by an identity matrix. For the purpose of
358 controlling population structure in association analysis, adjusted line means from the original
359 model were then used as observations in a mixed model analysis that modeled the genotypic
360 variance-covariance structure as proportional to the realized genomic relationship matrix, thus
361 incorporating the different pairwise relationships among the lines. This mixed model was
362 simplified by the compression method of ZHANG *et al.* (2010), which clusters lines according to
363 genetic similarity and replaces the full pair-wise realized genomic relationship matrix with a
364 reduced matrix of average relationships among the groups. The optimal level of clustering or
365 compression is determined empirically based on model fit to the observed phenotypic data. A
366 compressed relationship matrix can have better model fit than the original matrix when the
367 empirically observed covariance relationships among lines follow the group relationship
368 averages better than the individual pairwise relationships. Typically, this can happen when
369 closely related lines are grouped and estimate of the group phenotypes and their relationships
370 with other group phenotypes are improved. The optimal compression level can vary among
371 phenotypes for the same set of lines, as observed in this study.

372 Among environment groups, the proportion of phenotypic variance explained by
373 background genetic effects (**K**) was much smaller in North Carolina (31%, Table 2) compared to
374 Galicia (57%). Besides the small polygenic additive effects captured by the kinship matrix, rare
375 allele variants (minor allele frequency < 0.05) with larger effects, as well as epistatic
376 interactions, may explain some of the genotypic variation not captured by either **K** or the
377 significantly associated SNPs (MANOLIO *et al.* 2009).

378 Analyzing the Galicia environments separately from the North Carolina environments
379 revealed no significant SNPs, whereas the North Carolina analysis identified three SNPs
380 significantly associated with Fusarium ear rot resistance (Table 3). Examination of the empirical
381 distribution of P -values for the Galicia analysis revealed a slight skew toward higher P -values,
382 whereas the North Carolina and combined analyses exhibited excesses of small P -values (Figure
383 S3, Figure S4, and Figure S5). The Storey and Tibshirani (2003) method used to compute the
384 false discovery rate assumes that the distribution of P -value for truly null tests follows a flat
385 distribution, such that if the observed proportion of very low P -values is lower than expected
386 based on the flat distribution, the false discovery rate will be high even for the lowest P -values,
387 as we observed in the Galicia analysis. Whereas a few significant SNPs were identified in the
388 North Carolina and combined analyses at $q < 0.10$, no SNP had a q -value of less than 0.9 in the
389 Galicia analysis (Figure S3, Figure S4, and Figure S5). The disparity between the two individual
390 experiment analyses highlights the importance of conducting individual environment association
391 analyses in the presence of significant genotype by environment ($G \times E$) interaction. It should be
392 noted, however, that the appropriate threshold proportion of variation due to $G \times E$ interaction to
393 warrant individual location analyses instead of an overall combined analysis is not clear.

394 One possible mode of $G \times E$ interaction is the relative increase or decrease of additive
395 allelic effects among different loci between environments (FALCONER and MACKAY 1996).
396 Comparison of the absolute value of the allele effect at each of the identified SNP loci between
397 North Carolina and Galicia revealed that allele effects were larger in North Carolina across all
398 three loci (Table 3), congruent with the higher mean ear rot values in North Carolina (Table S1).
399 The largest proportion of phenotypic variance explained by **K** was in Galicia (Table 2), and
400 when combined with comparatively smaller allele effects, suggested that more loci may have

401 contributed to ear rot resistance in Galicia than North Carolina, and on average each locus had a
402 smaller additive effect on disease phenotype in Galicia. Collectively, these two points may
403 explain the deficiency of SNPs significantly associated with ear rot resistance in the Galicia
404 analysis.

405

406 *Association analyses*

407 Three SNPs significantly associated with ear rot resistance were identified in the North
408 Carolina analysis (Table 3), and all localized to separate chromosomes. One of these three SNPs,
409 located on chromosome 9, was also identified in the combined analysis. None of the three SNPs
410 localized to any of the linkage map bins containing resistance QTL reported by Robertson et al.
411 (2006) and Ding et al. (2008). However, the proportion of phenotypic variance explained by each
412 SNP is consistent with individual QTL r^2 values reported by each of the two aforementioned
413 mapping studies. The chromosome 9 SNP explained the largest proportion of the variation in line
414 mean values for ear rot resistance ($R^2=11.5\%$ in NC and $R^2=9.6\%$ in the combined analysis,
415 Table 3), while the chromosome 1 and chromosome 5 SNPs explained 8.8% and 9.6% of the
416 variation in line mean values for ear rot resistance in North Carolina, respectively. Modeling all
417 three SNPs together collectively explained 26% of the line mean variation in ear rot resistance in
418 North Carolina.

419 Although all three SNPs explained a relatively large portion of the total variation in line
420 means, each SNP had a relatively small additive effect on ear rot resistance (± 1.1 percentage
421 points ear rot severity on the back-transformed scale, Table 3). Additive genetic variance
422 estimates for each SNP was computed based on allele effects and frequencies (Table 3), and
423 when scaled to the total line mean variance coincided with the SNP R^2 values computed by

424 Tassel (Table 3). In every case, an increase in disease resistance (decrease in ear rot severity)
425 was associated with the rare allele at each locus. Resistance alleles at the chromosome 1 and 5
426 SNP loci were overrepresented in the tropical subpopulation relative to the other temperate
427 subpopulations (Table 3), consistent with enriched disease resistance observed in tropical maize
428 for some foliar diseases of maize (WISSER *et al.* 2011; OLUKOLU *et al.* 2013) and the lower level
429 of ear rot disease observed in tropical lines in this study.

430 Using the same association panel and marker set as this study, Olukolu *et al.* (2013)
431 reported that LD in the maize core diversity panel is variable across chromosomes and
432 subpopulations. The authors also reported that marker pairs separated by more than 10 kbp had
433 $r^2 < 0.1$ on average, which is consistent with estimates of $r^2 < 0.1$ between marker pairs separated
434 by 5-10 kbp on average in tropical subpopulations and 10-100 kbp on average in temperature
435 subpopulations (LU *et al.* 2011). Increased marker coverage, such as the genotype-by-sequencing
436 (GBS) data (ELSHIRE *et al.* 2011) used in Romay *et al.* (2013), in conjunction with a larger
437 association panel, may be able to uncover more SNPs in higher LD with ear rot resistance loci.
438 Assuming an association panel of between 350 and 400 inbred lines, Van Inghelandt *et al.* (2011)
439 indicated that as few as 4,000 markers would be necessary in a GWAS to detect individual QTL
440 explaining greater than 10% of the total phenotypic variation for a complex trait within the stiff
441 stalk subpopulation, whereas 65,000 markers would be required to detect QTL at the same
442 threshold within European flint types. In a sample of 2,815 inbred lines from the National Plant
443 Germplasm System (USA) representing the same heterotic groups described in this study,
444 Romay *et al.* (2013) reported that the utilization of over 680,000 GBS markers was sufficient to
445 detect most known candidate genes associated with flowering time in maize. Even so,
446 polymorphisms that strongly associated with the lower LD tropical/subtropical subpopulation

447 (such as *ZmCTT*) were more difficult to detect compared to polymorphisms that more frequently
448 associated with higher LD temperate subpopulations (such as *Vgt1*). The results of Romay et al.
449 (2013) indicate that although increased marker coverage and association panel size can improve
450 the power of a GWAS, special care needs to be given to ensure that lower LD subpopulations,
451 such as the tropical/subtropical subpopulation, are adequately represented in an association panel
452 in order to capture rare allele variants associated with those subpopulations.

453

454 *Candidate genes*

455 We used the B73 maize genome reference sequence to identify genes that either included
456 or were nearby SNPs significantly associated with ear rot resistance. The chromosome 9 gene
457 (GRMZM2G178880) that was identified in both the North Carolina and combined analyses
458 belongs to the cellulose synthase-like family A (*Cs1A*) protein family. Given that the associated
459 SNP localized to an intron segment within this gene, it is likely that this SNP is in LD with the
460 causal variant and not the causal variant itself. The expression of this gene is highest in the
461 endosperm of the developing seed kernel between 20 and 24 days after flowering during the
462 growing season (SEKHON *et al.* 2011; <http://www.plexdb.org>). Peak expression of this gene
463 coincides with the initial onset of Fusarium ear rot symptoms, which occurs approximately 28
464 days after flowering (BUSH *et al.* 2004). Genes in the *Cs1A* family encode for non-cellulose
465 polysaccharides (such as mannan polymers) that form part of the wall matrix in plant cells
466 (DHUGGA 2005; LIEPMAN *et al.* 2005). In the model grass species *Brachypodium distachyon*,
467 mannan polymers make up a significant portion of the seed endosperm (GUILLON *et al.* 2011).
468 Dismantling of mannan-rich cell walls may play an important role in programmed cell death
469 (PCD) in host-pathogen interactions (GADJEV *et al.* 2008; RODRÍGUEZ-GACIO *et al.* 2012).

470 Although the interaction between *Fusarium verticillioides* and maize is complex, cell wall
471 structure and PCD may play a role in quantitative resistant to the disease (CHIVASA *et al.* 2005).

472 The SNP on chromosome 5 is located downstream of an HSP60 gene
473 (GRMZM2G111477). Expression levels of this gene are highest in the developing endosperm 12
474 days after flowering (SEKHON *et al.* 2011; <http://www.plexdb.org>). HSP60s are chaperonins that
475 are involved with protein folding under plant stress primarily in the mitochondria and
476 chloroplasts (WANG *et al.* 2004). The role of HSP60s in programmed cell death has been
477 demonstrated in mutants of both rice and Arabidopsis (ISHIKAWA *et al.* 2003; TSUNEZUKA *et al.*
478 2005). The SNP on chromosome 1 is contained within the coding region of GRMZM2G703598.
479 Unfortunately, this gene has no predicted function and has no sequence orthology with related
480 grass species.

481 In conclusion, we have utilized a GWAS approach to identify three novel loci associated
482 with improved resistance to Fusarium ear rot in maize. The identified loci each explain a
483 relatively small proportion of the overall phenotypic variance for ear rot, and each locus has a
484 very small additive genetic effect on resistance, consistent with the highly quantitative nature of
485 the *F. verticillioides*-maize pathosystem. The large amount of variation captured by the kinship
486 matrix, in combination with high false discovery rates, suggests that additive polygenic variation
487 across many loci underlies resistance to Fusarium ear rot. Given the rapid decay of LD along the
488 chromosomes in the maize core diversity panel (OLUKOLU *et al.* 2013), future studies employing
489 increased marker density and larger association panels may be able to identify other novel loci
490 associated with ear rot resistance. Maize breeders can employ targeted allele selection for these
491 three resistance alleles, but may need to also select for recombinations near them as they are
492 introgressed into elite maize from unadapted or undesirable genotypes (such as the tropical

493 maize or popcorn germplasm pools that appear to be enriched for resistance alleles). In addition,
494 given the substantial additive polygenic variation for ear rot resistance, phenotypic and genomic
495 selection approaches should be effective as long as high quality phenotypic evaluations of
496 resistance can be performed to permit direct selection or provide training data for genomic
497 selection models.

498

499 **ACKNOWLEDGEMENTS**

500 This research was supported by the National Science Foundation (projects DBI-0321467, IOS-
501 0820619 and IOS-0604923), USDA-ARS, the National Plan for Research and Development of
502 Spain (AGL2009-12770), and the Excma. Diputación Provincial de Pontevedra. USDA-ARS
503 supported the graduate fellowship of C.T. Zila. R. Santiago acknowledges postdoctoral contract
504 “Isidro Parga Pondal” supported by the Autonomous Government of Galicia and the European
505 Social Fund. The authors wish to thank Jason Brewer, Steve Pigozzo, David Horne, Brittany
506 Scott, Sarah Davidson-Dyer, and Colt Jackson for technical assistance and Bode Olukolu, Funda
507 Ogut, and Shang Xue for assistance with the analyses.

508

509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528

References

ANDORF C. M., LAWRENCE C. J., HARPER L. C., SCHAEFFER M. L., CAMPBELL D. A., *et al.*, 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.

BERNARDO R., 2002 Chapter 5: Phenotypic and Genetic Variances. In: *Breeding for Quantitative Traits in Plants*, Stemma Press, Woodbury, MN, pp. 89–116.

BOLDUAN C., MIEDANER T., SCHIPPRACK W., DHILLON B. S., MELCHINGER A. E., 2009 Genetic variation for resistance to ear rots and mycotoxins contamination in early European maize inbred lines. *Crop Sci.* **49**: 2019–2028.

BRADBURY P. J., ZHANG Z., KROON D. E., CASSTEVENS T. M., RAMDOSS Y., *et al.*, 2007 TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635.

BUSH B. J., CARSON M. L., CUBETA M. A., HAGLER W. M., PAYNE G. A., 2004 Infection and fumonisin production by *Fusarium verticillioides* in developing maize kernels. *Phytopathology* **94**: 88–93.

CHIVASA S., SIMON W. J., YU X.-L., YALPANI N., SLABAS A. R., 2005 Pathogen elicitor-induced changes in the maize extracellular matrix proteome. *Proteomics* **5**: 4894–4904.

CLEMENTS M. J., KLEINSCHMIDT C. E., MARAGOS C. M., PATAKY J. K., WHITE D. G., 2003 Evaluation of inoculation techniques for *Fusarium* ear rot and fumonisin contamination of corn. *Plant Dis.* **87**: 147–153.

CLEMENTS M. J., MARAGOS C. M., PATAKY J. K., WHITE D. G., 2004 Sources of resistance to fumonisin accumulation in grain and *Fusarium* ear and kernel rot of corn. *Phytopathology* **94**: 251–260.

- 529 COOK J. P., MCMULLEN M. D., HOLLAND J. B., TIAN F., BRADBURY P., *et al.*, 2012 Genetic architecture of maize
530 kernel composition in the nested association mapping and inbred association panels. *Plant*
531 *Physiol.* **158**: 824–834.
- 532 CULLIS B., SMITH A., COOMBES N., 2006 On the design of early generation variety trials with correlated
533 data. *J. Agric. Biol. Environ. Stat.* **11**: 381–393.
- 534 DHUGGA K. S., 2005 Plant Golgi cell wall synthesis: From genes to enzyme activities. *Proc. Natl. Acad. Sci.*
535 *U. S. A.* **102**: 1815–1816.
- 536 DING J.-Q., WANG X.-M., CHANDER S., YAN J.-B., LI J.-S., 2008 QTL mapping of resistance to *Fusarium* ear
537 rot using a RIL population in maize. *Mol. Breed.* **22**: 395–403.
- 538 ELLER M. S., HOLLAND J. B., PAYNE G. A., 2008 Breeding for improved resistance to fumonisin
539 contamination in maize. *Toxin Rev.* **27**: 371–389.
- 540 ELLER M. S., PAYNE G. A., HOLLAND J. B., 2010 Selection for reduced *Fusarium* ear rot and fumonisin
541 content in advanced backcross maize lines and their topcross hybrids. *Crop Sci.* **50**: 2249–2260.
- 542 ELSHIRE R. J., GLAUBITZ J. C., SUN Q., POLAND J. A., KAWAMOTO K., *et al.*, 2011 A robust, simple genotyping-
543 by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**: e19379.
- 544 FALCONER D. S., MACKAY T. F. C., 1996 Chapter 19: Correlated Characters. In: *Introduction to Quantitative*
545 *Genetics*, Longman, Essex, UK, pp. 312–334.
- 546 FLINT-GARCIA S. A., THUILLET A., YU J., PRESSOIR G., ROMERO S. M., *et al.*, 2005 Maize association population:
547 a high-resolution platform for quantitative trait locus dissection. *Plant J.* **44**: 1054–1064.

548 GADJEV I., STONE J. M., GECEV T. S., 2008 Programmed cell death in plants: New insights into redox
549 regulation and the role of hydrogen peroxide. *Int. Rev. Cell Mol. Biol.* **270**: 87–144.

550 GANAL M. W., DURSTEWITZ G., POLLEY A., BÉRARD A., BUCKLER E. S., *et al.*, 2011 A large maize (*Zea mays* L.)
551 SNP genotyping array: development and germplasm genotyping, and genetic mapping to
552 compare with the B73 reference genome. *PLoS ONE* **6**: e28334.

553 GILMOUR A. R., GOGEL B. J., CULLIS B. R., THOMPSON R., 2009 *ASReml User Guide Release 3.0*. VSN
554 International Ltd, Hemel Hempstead, HP1 1ES, UK.

555 GUILLON F., BOUCHET B., JAMME F., ROBERT P., QUÉMÉNER B., *et al.*, 2011 *Brachypodium distachyon* grain:
556 characterization of endosperm cell walls. *J. Exp. Bot.* **62**: 1001–1015.

557 HÉMATY K., CHERK C., SOMERVILLE S., 2009 Host–pathogen warfare at the plant cell wall. *Curr. Opin. Plant*
558 *Biol.* **12**: 406–413.

559 HOLLAND J. B., 2007 Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* **10**: 156–161.

560 ISHIKAWA A., TANAKA H., NAKAI M., ASAHI T., 2003 Deletion of a chaperonin 60 β gene leads to cell death in
561 the *Arabidopsis lesion initiation 1* mutant. *Plant Cell Physiol.* **44**: 255–261.

562 KING S. B., SCOTT G. E., 1981 Genotypic differences in maize to kernel infection by *Fusarium moniliforme*.
563 *Phytopathology* **71**: 1245–1247.

564 LIEPMAN A. H., WILKERSON C. G., KEEGSTRA K., 2005 Expression of cellulose synthase-like (*Csl*) genes in
565 insect cells reveals that *CsIA* family members encode mannan synthases. *Proc. Natl. Acad. Sci. U.*
566 *S. A.* **102**: 2221–2226.

567 LU Y., SHAH T., HAO Z., TABA S., ZHANG S., *et al.*, 2011 Comparative SNP and haplotype analysis reveals a
568 higher genetic diversity and rapider LD decay in tropical than temperate germplasm in maize.
569 PLoS ONE **6**: e24861.

570 MANOLIO T. A., COLLINS F. S., COX N. J., GOLDSTEIN D. B., HINDORFF L. A., *et al.*, 2009 Finding the missing
571 heritability of complex diseases. Nature **461**: 747–753.

572 MARASAS W. F., 1996 Fumonisin: history, world-wide occurrence and impact. Adv. Exp. Med. Biol. **392**:
573 1–17.

574 MESTERHÁZY Á., LEMMENS M., REID L. M., 2012 Breeding for resistance to ear rots caused by *Fusarium* spp.
575 in maize – a review. Plant Breed. **131**: 1–19.

576 MILLER J. D., TRENHOLM H. L. (Eds.), 1994 The biosynthetic diversity of secondary metabolites;
577 Epidemiology of *Fusarium* ear diseases of cereals. In: *Mycotoxins in Grain: Compounds Other*
578 *Than Aflatoxin*, Eagan Press, St. Paul, pp. 3–36.

579 NANKAM C., PATAKY J. K., 1996 Resistance to kernel infection by *Fusarium moniliforme* in the sweet corn
580 inbred IL 125b. Plant Dis. **80**: 593–598.

581 OLUKOLU B. A., NEGERI A., DHAWAN R., VENKATA B. P., SHARMA P., *et al.*, 2013 A connected set of genes
582 associated with programmed cell death implicated in controlling the hypersensitive response in
583 maize. Genetics **193**: 609–620.

584 PÉREZ-BRITO D., JEFFERS D., GONZÁLEZ-DE-LEÓN D., KHAIRALLAH M., CORTÉS-CRUZ M., *et al.*, 2001 QTL mapping
585 of *Fusarium moniliforme* ear rot resistance in highland maize, Mexico. Agrociencia **35**: 181–196.

586 PRESELLO D. A., BOTTA G., IGLESIAS J., EYHÉRABIDE G. H., 2008 Effect of disease severity on yield and grain
587 fumonisin concentration of maize hybrids inoculated with *Fusarium verticillioides*. *Crop Prot.* **27**:
588 572–576.

589 R CORE TEAM, 2013 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
590 Computing, Vienna, Austria.

591 REID L. M., HAMILTON R. I., MATHER D. E., 1996 *Screening Maize for Resistance to Gibberella Ear Rot*.
592 Agriculture and Agri-Food Canada, Ottawa, ON, Canada.

593 REID L. M., ZHU X., 2005 *Screening Corn for Resistance to Common Diseases in Canada*. Agriculture and
594 Agri-Food Canada, Ottawa, ON, Canada.

595 REMINGTON D. L., PURUGGANAN M. D., 2003 Candidate genes, quantitative trait loci, and functional trait
596 evolution in plants. *Int. J. Plant Sci.* **164**: S7–S20.

597 ROBERTSON L. A., KLEINSCHMIDT C. E., WHITE D. G., PAYNE G. A., MARAGOS C. M., *et al.*, 2006 Heritabilities
598 and correlations of *Fusarium* ear rot resistance and fumonisin contamination resistance in two
599 maize populations. *Crop Sci.* **46**: 353–361.

600 ROBERTSON-HOYT L. A., JINES M. P., BALINT-KURTI P. J., KLEINSCHMIDT C. E., WHITE D. G., *et al.*, 2006 QTL
601 mapping for *Fusarium* ear rot and fumonisin contamination resistance in two maize populations.
602 *Crop Sci.* **46**: 1734–1743.

603 RODRÍGUEZ-GACIO M. del C., IGLESIAS-FERNÁNDEZ R., CARBONERO P., MATILLA Á. J., 2012 Softening-up
604 mannan-rich cell walls. *J. Exp. Bot.* **63**: 3976–3988.

605 ROMAY M. C., MILLARD M. J., GLAUBITZ J. C., PEIFFER J. A., SWARTS K. L., *et al.*, 2013 Comprehensive
606 genotyping of the USA national maize inbred seed bank. *Genome Biol.* **14**: R55.

607 SAS INSTITUTE INC, 2010 *SAS® 9.2 Intelligence Platform: System Administration Guide*. SAS Institute Inc,
608 Cary, NC.

609 SCHNABLE P. S., WARE D., FULTON R. S., STEIN J. C., WEI F., *et al.*, 2009 The B73 maize genome: complexity,
610 diversity, and dynamics. *Science* **326**: 1112–1115.

611 SEKHON R. S., LIN H., CHILDS K. L., HANSEY C. N., BUELL C. R., *et al.*, 2011 Genome-wide atlas of transcription
612 during maize development. *Plant J.* **66**: 553–563.

613 STOREY J. D., TIBSHIRANI R., 2003 Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci.*
614 **100**: 9440–9445.

615 TSUNEZUKA H., FUJIWARA M., KAWASAKI T., SHIMAMOTO K., 2005 Proteome analysis of programmed cell
616 death and defense signaling using the rice lesion mimic mutant *cdr2*. *Mol. Plant. Microbe*
617 *Interact.* **18**: 52–59.

618 VAN EGMOND H. P., SCHOTHORST R. C., JONKER M. A., 2007 Regulations relating to mycotoxins in food;
619 Perspectives in a global and European context. *Anal. Bioanal. Chem.* **389**: 147–157.

620 VAN INGHELANDT D., REIF J. C., DHILLON B. S., FLAMENT P., MELCHINGER A. E., 2011 Extent and genome-wide
621 distribution of linkage disequilibrium in commercial maize germplasm. *Theor. Appl. Genet.* **123**:
622 11–20.

623 VANRADEN P. M., 2008 Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**: 4414–4423.

624 WANG W., VINOUCUR B., SHOSEYOV O., ALTMAN A., 2004 Role of plant heat-shock proteins and molecular
625 chaperones in the abiotic stress response. *Trends Plant Sci.* **9**: 244–252.

626 WISSER R. J., KOLKMAN J. M., PATZOLDT M. E., HOLLAND J. B., YU J., *et al.*, 2011 Multivariate analysis of maize
627 disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. *Proc. Natl.*
628 *Acad. Sci.* **108**: 7339–7344.

629 YU J., BUCKLER E. S., 2006 Genetic association mapping and genome organization of maize. *Curr. Opin.*
630 *Biotechnol.* **17**: 155–160.

631 YU J., PRESSOIR G., BRIGGS W. H., VROH BI I., YAMASAKI M., *et al.*, 2006 A unified mixed-model method for
632 association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**: 203–208.

633 ZHANG Z., ERSOZ E., LAI C.-Q., TODHUNTER R. J., TIWARI H. K., *et al.*, 2010 Mixed linear model approach
634 adapted for genome-wide association studies. *Nat. Genet.* **42**: 355–360.

635

636

Table 1. Genotypic covariance/variance/correlation matrix for Fusarium ear rot from the combined analysis of a maize diversity panel evaluated in five environments. The diagonal (bold) is an estimate of genetic variance ($\hat{\sigma}_G^2$) plus the genotype by environment interaction ($\hat{\sigma}_{GE}^2$) within each environment. Estimates of genetic variance (covariance between pairs of environments) are shown below the diagonal, and genetic correlations between inbred lines in each pair of environments are shown above the diagonal.

Environment	NC 2010	NC 2011	NC 2012	Galicia 2010	Galicia 2011
NC 2010	0.27	0.42	0.44	0.51	0.44
NC 2011	0.15	0.45	0.38	0.33	0.28
NC 2012	0.19	0.21	0.68	0.36	0.35
Galicia 2010	0.15	0.12	0.17	0.32	0.93
Galicia 2011	0.11	0.09	0.14	0.25	0.23

Table 2. Number of lines, number of groups, compression level, polygenic additive background genetic variance component, residual genotypic variance component, and proportion of total line mean variance explained by additive relationship matrix from the three mixed-linear model (MLM) analyses.

	N ^a	Groups ^b	Compression ^c	$(\hat{\sigma}_G^2)^d$	$(\hat{\sigma}^2)^d$	$\left(\frac{\hat{\sigma}_G^2}{\hat{\sigma}_G^2 + \hat{\sigma}^2}\right)^e$
North Carolina	247	197	1.25	0.09	0.20	0.31
Galicia	254	229	1.11	0.18	0.14	0.57
Combined	267	197	1.36	0.10	0.11	0.48

^a Total number of lines included in the analysis.

^b Number of groups determined by optimum compression.

^c Compression level is the average number of individuals per group.

^d Polygenic additive background genetic variance and residual genotypic variance components are estimated in Tassel by fitting the kinship matrix (**K**) in the mixed linear model without any SNP marker effects.

^e Background genetic variance divided by total phenotypic variance.

Table 3. Chromosome locations (AGP v2 coordinates), allele effect estimates, genes containing or adjacent to SNP, and other summary statistics for the three SNPs significantly associated with Fusarium ear rot resistance in the North Carolina analysis and the single SNP associated with resistance in the combined analysis. Statistics from environments in which the SNPs were not significantly associated with ear rot are also shown for comparison.

Chromosome	SNP physical position (bp)	P-value	q-value	Allele	N ^a	Allele effect (%) ^b	Additive variance estimate ^c	(R ²) ^d	Gene containing or adjacent to SNP
<i>North Carolina analysis</i>									
1	63,540,590	5.5×10 ⁻⁶	0.084	A	224	+0.945	0.036	8.8	GRMZM2G703598
				G	22	0.0			
5	30,997,717	2.2×10 ⁻⁶	0.050	G	225	+1.149	0.042	9.6	GRMZM2G111477
				A	19	0.0			
9	151,295,233	2.4×10 ⁻⁷	0.011	A	67	-0.365	0.041	11.5	GRMZM2G178880
				G	176	0.0			
<i>Galicia analysis</i>									
1	63,540,590	0.826 ^{NS}	1.000	A	231	+0.035	9.55×10 ⁻⁵	1.9×10 ⁻²	GRMZM2G703598
				G	22	0.0	2.49×10 ⁻⁵	4.2×10 ⁻³	GRMZM2G111477
5	30,997,717	0.918 ^{NS}	1.000	G	228	-0.017			
				A	23	0.0			
9	151,295,233	0.198 ^{NS}	1.000	A	71	-0.115	0.003	0.7	GRMZM2G178880
				G	179	0.0			
<i>Combined analysis</i>									
1	63,540,590	4.5×10 ⁻³	0.689	A	244	+0.425	0.010	3.1	GRMZM2G703598
				G	22	0.0			
5	30,997,717	2.6×10 ⁻³	0.689	G	240	+0.428	0.011	3.5	GRMZM2G111477
				A	24	0.0			
9	151,295,233	9.1×10 ⁻⁷	0.042	A	74	-0.292	0.024	9.6	GRMZM2G178880
				G	189	0.0			

^a N, total number of lines with the specific SNP genotype.

^b Allele effects are reported back-transformed to the original 0-100% disease severity scale.

^c Additive variance for an inbred population was computed as two times the product of the separate allele frequencies times the genotypic value from Tassel squared using the formula $2pqa^2$ from Bernardo (2002).

^d R², proportion of total line mean variance explained by SNP as computed by Tassel.

Table 4. Allele frequencies of significantly associated SNPs in the five major maize subpopulations.

Chromo- some	SNP physical position (bp)	Allele increasing resistance	Allele frequency (%)					<i>P</i> -value	<i>N</i> ^b					Ear rot mean (%) ^c				
			<i>SS</i> ^a	<i>NSS</i>	<i>TS</i>	<i>PC</i>	<i>SC</i>		<i>SS</i>	<i>NSS</i>	<i>TS</i>	<i>PC</i>	<i>SC</i>	<i>SS</i>	<i>NSS</i>	<i>TS</i>	<i>PC</i>	<i>SC</i>
1	63,540,590	G	0.0	8.4	15.4	0.0	0.0	0.1488	28	107	65	8	6	24.0	24.3	14.6	17.9	46.5
5	30,997,717	A	0.0	3.8	26.6	37.5	0.0	6.193×10 ⁻⁶	28	106	64	8	6					
9	151,295,233	A	14.3	34.9	26.6	100.0	33.3	3.846×10 ⁻⁴	28	106	64	7	6					

^a *SS*, Stiff Stalk; *NSS*, non-Stiff Stalk; *TS*, tropical/sub-tropical; *PC*, popcorn; *SC*, sweet corn.

^b *N*, total number of lines within each subpopulation.

^c Overall phenotypic ear rot means are the average of least square means from the combined analysis across members of each subpopulation.

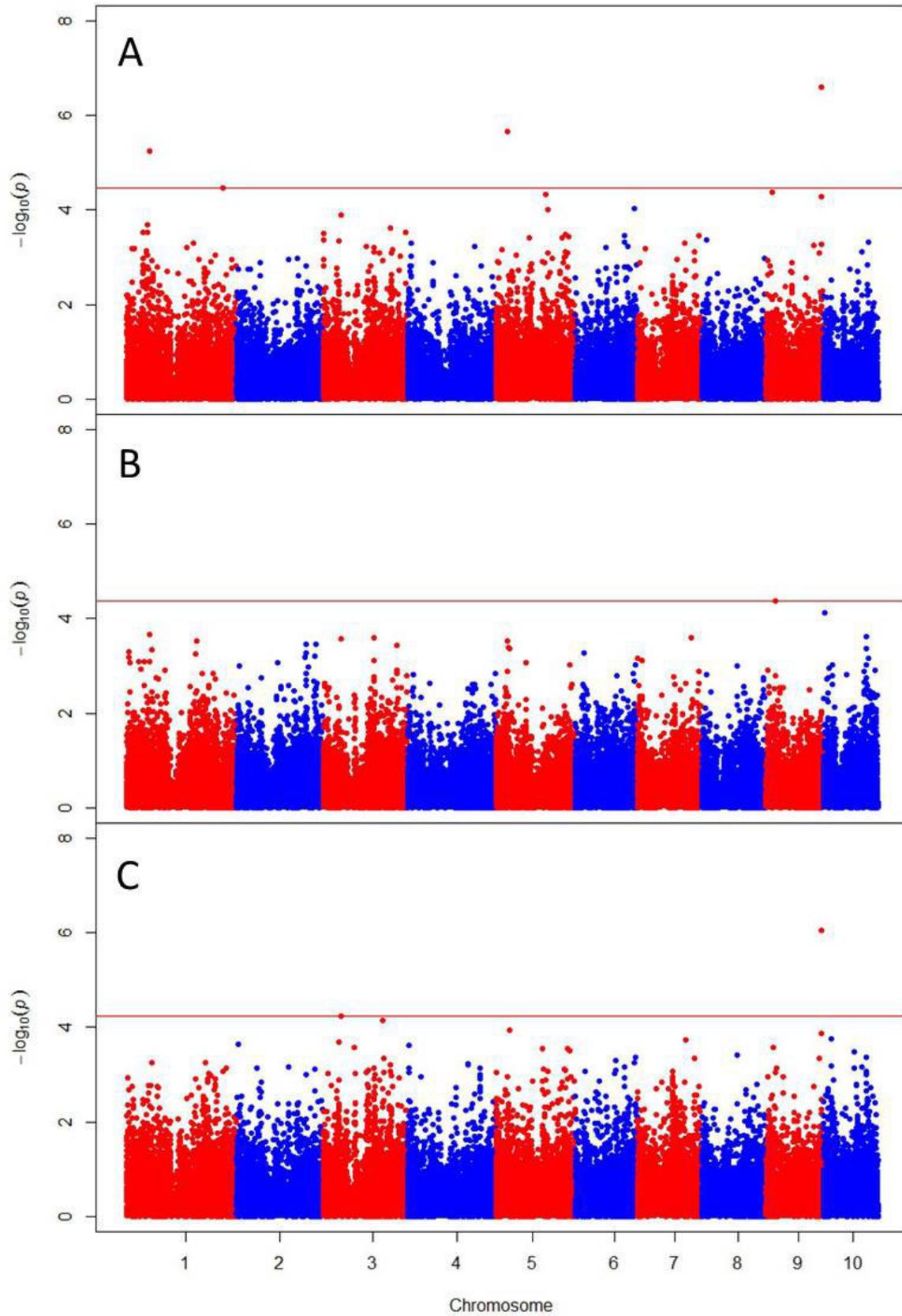


Figure 1. Results of the three GWAS showing significant associations (points above red FDR = 0.10 threshold lines) in the North Carolina (A), Galicia (B), and combined (C) analyses. The vertical axis indicates $-\log_{10}$ of P -value scores, and the horizontal axis indicates chromosomes and physical positions of SNPs.

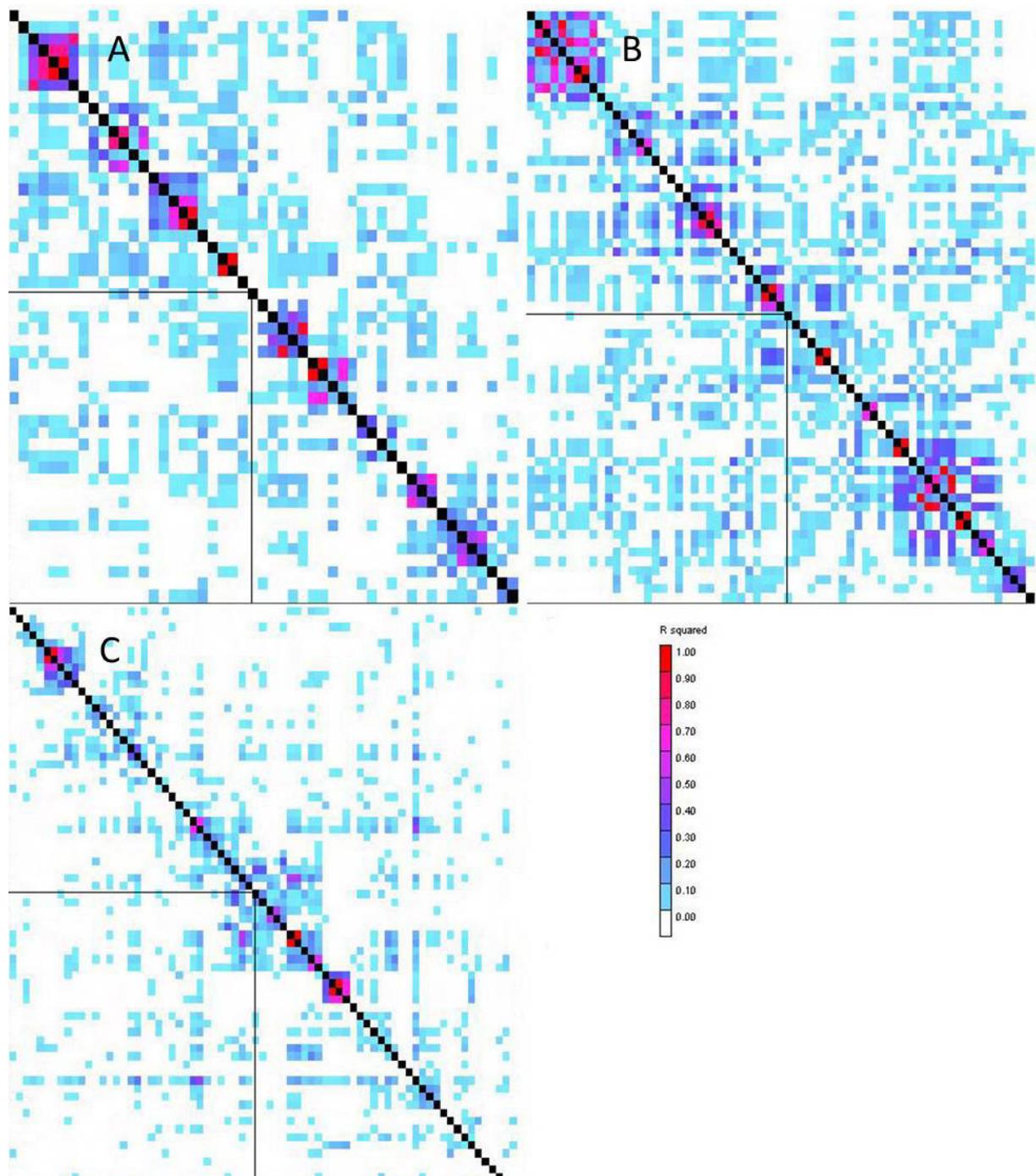


Figure 2. LD heatmaps showing LD measure (r^2) calculated for each pairwise combination of SNPs in an approximately ± 1 Mbp region surrounding each SNP significantly associated with ear rot resistance in the North Carolina analysis. (A) LD around chromosome 1 SNP. (B) LD around chromosome 5 SNP. (C) LD around chromosome 9 SNP. The significant SNP on each chromosome is highlighted by the perpendicular black lines within each heatmap. Colors indicate the magnitude of each pairwise r^2 measure ($r^2=1$ is red to $r^2=0$ is white).