nature
genetics

Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population

Kristen L Kump[1], Peter J Bradbury[2,3], Randall J Wisser[4], Edward S Buckler[2,3], Araby R Belcher[5],
Marco A Oropeza-Rosas[1], John C Zwonitzer[5], Stephen Kresovich[3], Michael D McMullen[6], Doreen Ware[3],
Peter J Balint-Kurti[5,7] & James B Holland[1,7]

Nested association mapping (NAM) offers power to resolve complex, quantitative traits to their causal loci. The maize NAM population, consisting of 5,000 recombinant inbred lines (RILs) from 25 families representing the global diversity of maize, was evaluated for resistance to southern leaf blight (SLB) disease. Joint-linkage analysis identified 32 quantitative trait loci (QTLs) with predominantly small, additive effects on SLB resistance. Genome-wide association tests of maize HapMap SNPs were conducted by imputing founder SNP genotypes onto the NAM RILs. SNPs both within and outside of QTL intervals were associated with variation for SLB resistance. Many of these SNPs were within or near sequences homologous to genes previously shown to be involved in plant disease resistance. Limited linkage disequilibrium was observed around some SNPs associated with SLB resistance, indicating that the maize NAM population enables high-resolution mapping of some genome regions.

Southern leaf blight disease, caused by the fungus *Cochliobolus heterostrophus*, constitutes a considerable threat to corn production worldwide. No known genes confer complete immunity to this disease; instead, maize breeders rely on polygenic, quantitative resistance to SLB[1]. Although substantial progress has been made in elucidating the genetic basis of plant disease immunity conditioned by 'R genes'[2], little is known about the genes underlying quantitative disease resistance. The few genes underlying QTLs for plant disease resistance that have been cloned represent a diverse array of gene functions[3–8].

The maize NAM panel, a set of 5,000 RILs derived from crosses between the reference inbred line B73 and 25 other founder inbreds, captures a substantial proportion of the global genetic diversity of maize inbred lines. Maize NAM represents 135,000 recombination events and has been genotyped at 1,106 SNP markers[9]. The high allele

diversity and large sample size provide power for detection and resolution of QTLs[10,11]. With the recent release of the first-generation maize HapMap[12], based on the NAM founder lines, haplotypes can be imputed onto the full set of NAM RILs for genome-wide association analysis[13].

We evaluated the NAM population for quantitative resistance to SLB across three environments and resolved most of the genetic control for resistance to QTLs with joint linkage analysis. High-resolution genome-wide association analysis was conducted using 1.6 million HapMap SNPs, which were identified among the founder lines and imputed onto the complete NAM panel, to identify SNPs significantly associated with SLB resistance. Many of these SNPs are within or adjacent to candidate genes that warrant further investigation as putative causal genes underlying quantitative disease-resistance QTLs.
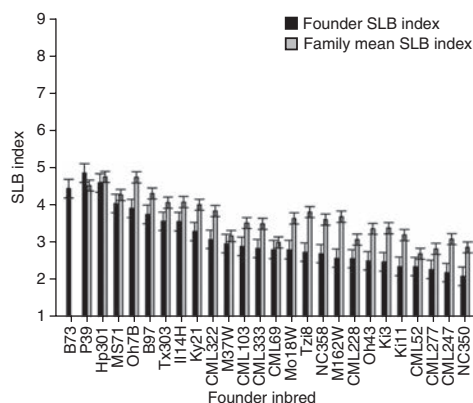
We measured quantitative resistance to SLB using a nine-point rating scale (**Supplementary Fig. 1**) twice on each experimental plot in each of three environments. SLB index values (representing the mean of SLB resistance measured across time points and environments) varied widely among B73 and the 25 other founder inbreds (**Fig. 1**). The most resistant and susceptible founder lines differed by about three points, a difference that can be observed visually (**Supplementary Fig. 1**). The reference founder line, B73, was among the least resistant lines (**Fig. 1**). NAM RIL families also differed in mean resistance, with family mean SLB index values highly correlated with the values of their diverse founder parents ($r = 0.88$, $P < 0.0001$; **Fig. 1**). Heritability of the average SLB index score across time points was 87% (**Supplementary Table 1**), indicating the potential for accurate mapping of SLB-resistance genes.

We identified 32 SLB-resistance QTLs at $P < 0.0001$ in a joint linkage analysis across all families (**Fig. 2** and **Supplementary Tables 2 and 3**). The median length of the support intervals for QTL positions was 6.2 cM, and values ranged from 0.6 cM to 22.0 cM. These 32 QTLs jointly explained 80% of the phenotypic and 93% of the

Figure 1 Parental and RIL family mean values ± s.e.m. of SLB index scores. Black bars, parental SLB best linear unbiased predictor (BLUP) index value; Gray bars, average index value for all RILs derived from the cross of B73 and that parent.

genotypic variation for SLB resistance. We detected no digenic epistatic interactions between QTLs with additive effects; however, we detected one interaction between loci without significant main effects. The QTL model based solely on analysis of the NAM RILs predicted 79% of the variation in resistance of the NAM founders (**Supplementary Fig. 2**).

All QTLs had small effects on SLB resistance, with absolute values of significant allelic effects averaging 0.15, and ranging from 0.09 to 0.39 on the nine-point scale (**Fig. 3**). The one epistatic interaction we detected showed effects ranging from −0.04 to 0.08 points. The QTL with the largest effect estimate mapped to bin 3.04, a region previously associated with SLB resistance[14,15]. We expected plants homozygous for the most resistant allele at this QTL (derived from line IL14H) to have scores 0.78 points lower than plants homozygous for the B73 allele at this locus, a visually observable difference. However, the predicted effects of most other allelic substitutions would be difficult to discern visually. Nevertheless, the combined effect of allelic substitutions across many QTLs leads to the substantial differences in quantitative SLB resistance among the founder lines (**Figs. 1** and **3**). Among all the detected QTLs, significant (at a 5% false discovery rate) functional allelic effects segregated in 1 to 15 families, with a median number of 5 families (**Supplementary Fig. 3**). At 37% of the QTLs, diverse parent founders contributed both significantly positive (greater susceptibility) and negative (greater resistance) allele effects relative to the B73 allele, indicating that allelic series of at least three alleles each exist at these QTLs. Compared to B73, diverse line founder alleles with significant effects contributed only to greater resistance at 13 QTLs, and only to greater susceptibility at 7 QTLs (**Fig. 3**). About two-thirds of significant founder alleles conferred greater resistance relative to B73, as we expected on the basis of the greater resistance observed in most founders compared to B73.

Necrotrophic leaf diseases such as SLB tend to occur after anthesis[15], raising the possibility that some disease-resistance QTLs may result

from pleiotropic effects of flowering time genes. We corrected for this effect by measuring days to anthesis (DTA) on the NAM lines in the same experiments used for SLB-resistance evaluations and fitting DTA as a covariate in the analysis models (**Supplementary Table 4**). To test the hypothesis that some flowering-time QTLs cause SLB resistance by pleiotropy even after applying covariate adjustments, we mapped DTA QTLs and compared their positions and effects to the SLB-resistance QTLs. We identified 30 flowering QTLs, explaining 85.2% of the phenotypic variation for DTA (**Supplementary Tables 5** and **6**). Eight pairs of QTLs for SLB and DTA had overlapping support intervals, but the 25 founder allele effects on the two traits were significantly correlated ($r = −0.52$, $P = 0.008$) at only one colocalized pair (in bin 1.05).

We compared SLB-resistance QTLs detected in the NAM to those reported in previous studies of individual biparental populations representing progeny from combinations of eight parental lines[14,16–20]. We placed QTLs from each study on a common map according to the positions of their closest flanking markers or support intervals (**Supplementary Fig. 4**). The median number of QTLs detected in each previous study was 6.5, about five times fewer than detected in NAM. The number of QTLs with significant allelic effects within each NAM family ranged from 2 (B73 × CML52) to 15 (B73 × CML247), with an average of 7.5. Of the 32 SLB QTLs we detected in NAM, 16 overlapped with previously identified QTLs (**Supplementary Fig. 4**). Another 19 SLB QTLs have been reported in previous studies but not detected in NAM.

We controlled for genetic background variation in the initial step of genome-wide association (GWA) tests of 1.6 million maize HapMap SNPs by testing each SNP effect in combination with the joint linkage QTL model, including the DTA covariates but excluding QTLs identified on the same chromosome as the tested SNP (**Supplementary Fig. 5**). We carried out two distinct initial GWA analyses: a single forward regression search on each chromosome separately using the full RIL dataset, and forward regression for each chromosome within each of 100 subsample data sets containing 80% of the RILs in each family[13,21,22]. Using the two methods combined, we identified 245 significantly associated SNPs (**Fig. 4**, **Supplementary Fig. 6** and **Supplementary Table 7**), which we then considered candidates for inclusion in a final joint GWA model. In the final model selection, we aimed at replacing QTLs (with unique
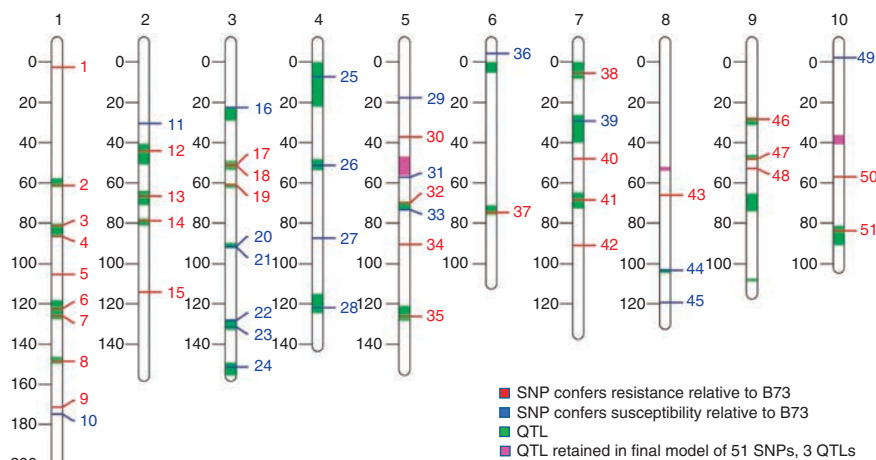

Figure 2 Genetic map of NAM population with support intervals of 32 QTLs and positions of 51 HapMap SNPs retained in final model for SLB resistance. Ten chromosomes of maize genome are shown. Detailed information on these 51 SNPs is in **Table 1**.

Figure 3 Heat map of additive effect estimates of 25 founder inbred alleles at QTLs for SLB resistance relative to B73 on the nine-point SLB index scale. QTLs are indicated by their chromosome and bin numbers (rows) and the allelic effect estimates for each founder allele (columns) are coded by color according to 0.05 score increments (see legend). Squares surrounded by bold lines, allelic effects significantly different from zero at 5% false discovery rate.

effects modeled for each of 25 NAM families) with SNPs (each with only 2 instead of 25 allelic effects across all families, and with effect segregation limited to families in which the SNP segregated) that were simultaneously significant ($P < 0.0001$) in a final joint model (**Supplementary Fig. 5**). This reduced the initial set of 245 SNPs to 51, in part by selecting only one or two SNPs out of localized clusters of linked SNPs (**Fig. 2**, **Table 1** and **Supplementary Figs. 6** and **7**). The final model, containing family mean effects, three QTLs and 51 biallelic SNP effects (**Supplementary Tables 8** and **9**), explained 74% of the phenotypic variation for SLB resistance. Without the family term, the 51 SNPs plus three QTLs explained 69% of the total phenotypic variation. The final model predicted 79% of the variation for SLB resistance among the founder lines (**Supplementary Fig. 8**). SNPs within QTL support intervals were significantly enriched in this final set (**Fig. 2** and **Supplementary Fig. 6**), accounting for 33 (65%) of the 51 SNPs in the final model, compared with 15% in the complete set of 1.6 million HapMap SNPs. Of the 32 SLB-resistance QTL intervals, 5 contained no SNPs, 21 contained a single SNP and 6 contained two SNPs each in the final model (**Fig. 2** and **Supplementary Figs. 6** and **7**).

Using linkage analysis, we found that variation in resistance to SLB in the NAM panel is controlled by at least 32 QTLs, each with relatively small effects. About half of the QTLs detected in previous biparental populations were not detected in NAM (**Supplementary Fig. 4**), possibly because they segregate at low frequency or not at all in NAM. Considering that the NAM founders were chosen to capture maize's diversity, this suggests that there is substantial heterogeneity in the genetic architecture for SLB within maize. The accuracy of prediction of founder line SLB index values on the basis of the additive QTL model was high ($r^2 = 0.79$; **Supplementary Fig. 2**) and only one significant epistatic interaction was identified. Thus, the genetic architecture of SLB resistance in the NAM panel is characterized by largely additive gene action with a minor role for epistasis, and relatively small allelic effects. This genetic architecture is markedly similar to that of flowering time in the NAM[10].

Of the 32 QTL intervals, 5 did not include SNPs maintained in the final model (**Fig. 2**), but this was not related to the magnitude of QTL effects. Discrepancies between QTL and SNP positions can be caused by differences in power of fitting biallelic SNPs compared with
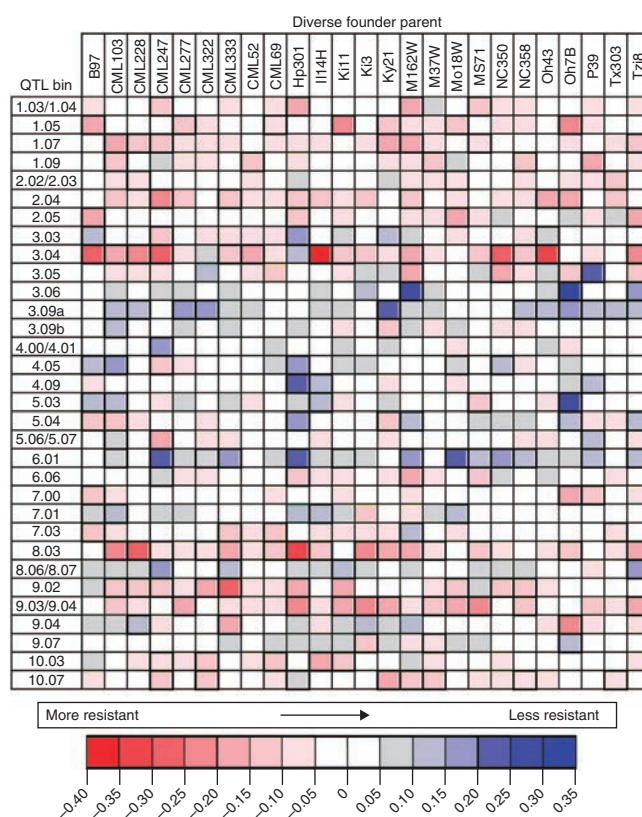
multiallelic QTLs for the detection of causal variants that are truly segregating for differing numbers of alleles in NAM, and to insufficient SNP density for adequate tagging of all functional variants. GWA in NAM has been estimated to require ten times more SNPs than the current HapMap of 1.6M SNPs[12] to ensure complete genome saturation. Thus, we may have missed causal polymorphisms owing to their lack of linkage disequilibrium (LD) with tested SNPs. Conversely, 18 of 51 final model SNPs mapped outside of QTL support interval (**Fig. 2**). These may represent false positives owing to long-range LD within chromosomes or associations with causal genes of small effect that were not detected as QTLs.

GWA relies on LD between SNPs and sequence polymorphisms responsible for phenotypic variation. Such LD can arise even without physical linkage owing to population stratification, greatly complicating the interpretation of GWA results in diverse population samples[23,24]. Random chromosome assortment during NAM population development eliminates interchromosomal LD among its
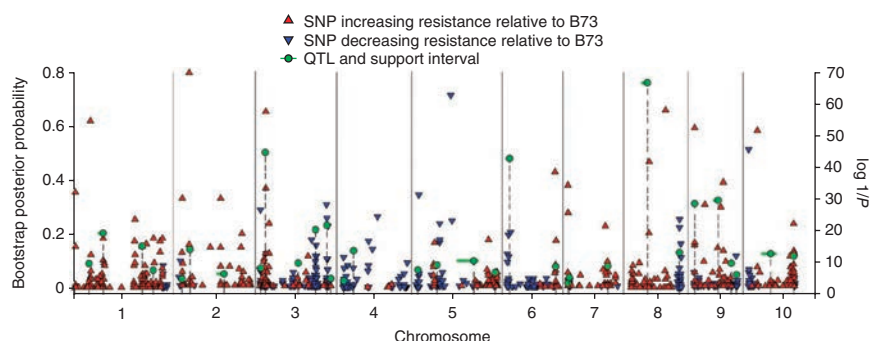
Figure 4 Results of GWA subsampling analysis. All SNPs detected as significant in at least one subsample are triangles relative to their physical sequence position. Red upward triangles, SNPs at which the diverse parent allele increases SLB resistance relative to the reference B73 allele. Blue downward triangles, SNPs at which the diverse parent allele decreases SLB resistance relative to the reference B73 allele. Vertical positions of triangles represent bootstrap posterior probability (BPP) of the SNP. SLB QTLs, green circles whose vertical positions represent their F-test $\log(1/P)$ in the final joint linkage QTL model. Horizontal green bars, QTL support intervals.

Table 1 Physical and genetic map positions of 51 SNPs significantly associated with SLB resistance and predicted function or homology of adjacent candidate genes

| SNP	Chr.	Physical position (AGPv1 bp)	NAM map position (cM)	Mean effect	P	Candidate genes or proteins
1	1	3,345,032	2.4	−0.11	5.0×10^{-9}	
2	1	45,565,372	61.5	−0.07	4.1×10^{-8}	Mitochondrial carrier protein (programmed cell death[32]); Ran GTPase (plant defense response[33])
3	1	80,360,348	81.3	−0.09	2.4×10^{-12}	Glutathione S-transferase (plant defense[34])
4	1	91,029,509	86.8	−0.06	1.9×10^{-14}	Lipid transfer protein (detection of pathogens[35])
5	1	191,462,090	105.2	−0.05	6.2×10^{-17}	BTB/POZ domain gene (interacts with NPR1[36])
6	1	210,676,683	122.4	−0.11	2.0×10^{-8}	LRR receptor kinase (disease resistance[25])
7	1	215,011,781	126.1	−0.06	1.6×10^{-26}	TUBBY domain (upregulated in defense response[37])
8	1	250,552,770	148.6	−0.08	2.7×10^{-13}	
9	1	278,662,769	171.4	−0.15	3.3×10^{-11}	
10	1	280,344,954	174.9	0.10	2.3×10^{-11}	
11	2	10,687,858	30.6	0.05	3.0×10^{-9}	AP2 transcription factor (disease resistance[38])
12	2	16,645,095	44.1	−0.05	8.6×10^{-12}	Cyclin (programmed cell death[39]), multi antimicrobial extrusion protein MatE[40]
13	2	38,763,561	66.6	−0.10	1.8×10^{-40}	
14	2	137,994,984	79.0	−0.10	9.4×10^{-15}	
15	2	205,642,264	114.2	−0.13	3.0×10^{-9}	Armadillo proteins (cell death[41])
16	3	5,756,385	22.5	0.12	1.3×10^{-11}	
17	3	22,604,327	51.2	−0.07	6.0×10^{-99}	NRR (defense response[31,42])
18	3	22,723,018	51.3	−0.19	4.7×10^{-94}	
19	3	129,035,412	61.1	−0.09	7.0×10^{-9}	Cytochrome P450 (phytoalexin production, other defense responses[43,44])
20	3	179,919,456	91.9	0.06	1.7×10^{-19}	Serine-threonine protein kinase (plant defense response[45,46])
21	3	179,938,179	92.0	0.14	4.3×10^{-23}	
22	3	214,769,000	128.3	0.07	5.9×10^{-28}	
23	3	216,041,726	131.4	0.07	2.6×10^{-27}	Serine-threonine protein kinase (plant defense response[45,46])
24	3	227,067,958	151.4	0.06	2.0×10^{-5}	Pectate lyase[47]
25	4	2,078,072	7.2	0.07	3.0×10^{-9}	
26	4	32,640,236	51.4	0.06	3.3×10^{-15}	
27	4	176,435,821	87.6	0.10	3.0×10^{-8}	
28	4	240,050,394	121.9	0.12	3.7×10^{-12}	Pti4, Pti5 and Pti6 ERF transcription factors; ABC transporter (disease resistance[8,48])
29	5	5,052,284	17.7	0.05	1.6×10^{-8}	
30	5	11,846,180	37.1	−0.07	3.2×10^{-8}	
31	5	31,569,318	57.3	0.07	1.1×10^{-18}	
32	5	100,876,057	70.1	−0.08	2.2×10^{-6}	
33	5	151,200,478	73.4	0.07	8.3×10^{-5}	LRR gene (pathogen recognition[2])
34	5	181,340,536	90.8	−0.09	1.7×10^{-10}	
35	5	207,876,608	126.2	−0.09	6.0×10^{-9}	
36	6	5,645,390	−4.4	0.10	1.7×10^{-48}	
37	6	156,545,570	74.8	0.07	2.2×10^{-10}	AP2 transcription factor (disease resistance[38])
38	7	2,159,584	5.4	−0.09	1.3×10^{-12}	Cyclins (programmed cell death[39])
39	7	6,952,875	29.3	0.07	4.4×10^{-9}	Lipases (disease resistance[49])
40	7	35,156,174	48.1	−0.05	5.2×10^{-5}	BTB-POZ domain gene (interacts with NPR1[36])
41	7	127,402,718	68.5	−0.08	2.7×10^{-10}	Basic leucine zipper in TGA subclade (interacts with NPR[29])
42	7	152,488,251	91.2	−0.11	1.5×10^{-9}	TUBBY domains (upregulated in defense response[37])
43	8	122,144,059	66.2	−0.09	2.7×10^{-15}	
44	8	165,649,948	103.2	0.06	8.6×10^{-16}	P450 gene (phytoalexin production and other defense responses[43,44])
45	8	170,075,212	119.4	0.05	2.0×10^{-9}	Cyclins (programmed cell death[39])
46	9	16,221,990	28.5	−0.10	1.4×10^{-33}	
47	9	93,746,250	48.2	−0.08	8.8×10^{-20}	
48	9	106,915,104	53.0	−0.11	7.0×10^{-20}	
49	10	1,221,166	−2.3	0.13	4.0×10^{-9}	NPR1 (disease resistance[27,50])
50	10	133,028,937	57.1	−0.05	3.0×10^{-14}	
51	10	144,732,895	83.9	−0.08	6×10^{-20}	

SNP number refers to positions in **Figure 2**. Chr., chromosome.

founders. Furthermore, recombination erodes founder intrachromosomal LD as a function of the genetic distance between loci, with LD decaying toward zero as recombination frequency approaches 0.5. Thus, undertaking GWA in the framework of NAM eliminates spurious unlinked associations and reduces the occurrence of linked associations[13]. The NAM design allows calculation of expected LD among HapMap SNPs by interpolation of the observed LD among NAM map markers into the genetic map intervals where the HapMap SNP genetic positions were projected. We inspected LD between each of 16 representative SNPs in the final GWA model and all other SNPs on the same chromosome in the NAM founders and in the progeny RIL generation (**Fig. 5** and **Supplementary Fig. 9**). We observed relatively high levels of founder LD between these sites and distant sites, but recombination during NAM population development often dissipated LD to much lower levels (**Fig. 5a**). In other cases, recombination in the NAM over a 5 cM interval surrounding the SNP was not sufficient to erode LD below $r^2 = 0.5$ (for example, chromosome 1, 191.5 Mb; **Fig. 5b**). Thus, GWA resolution varies among genome regions in NAM but in some cases seems sufficient for identification of one or a few probable causal gene candidates.

A diverse range of genes are probably involved in natural variation for plant disease resistance[3–8]. We identified several genes which may function in known plant disease-resistance pathways among the genes containing or immediately adjacent to the 51 SNPs identified herein (**Table 1** and **Supplementary Table 7**). For example, we identified two genes with leucine-rich repeat (LRR) domains, which are important in plant responses to a variety of external stimuli including pathogens[25,26] (**Table 1**). We also identified a gene with strong similarity to *NPR1*, which is important for the defense response[27,28]. *NPR1* interacts with the *TGA* family of basic leucine zipper transcription factors[29,30]; we also identified a member of the TGA gene family adjacent to a final model SNP (**Table 1**). Another SNP was adjacent to a homolog of the rice gene *NRR* (negative regulator of resistance) (**Table 1**), which encodes a protein that interacts with the *NPR1* protein during the defense response[31]. Several other SNPs were within or adjacent to genes containing domains homologous to various other gene classes that seem to function in plant disease resistance (**Table 1**).

Future research will focus on validating the effects of these genes, understanding how genes engender resistance to SLB in maize

Figure 5 Linkage disequilibrium. (**a**,**b**) LD (r^2) between significant SNPs on chromosome 7 (**a**), physical position 127 Mb, and chromosome 1 (**b**), physical position 191 Mb, and all other SNPs on the same chromosome. Position of the significant SNP to which all others are compared, dashed vertical line (this SNP has $r^2 = 1.0$ with itself on the line). Red circles, observed r^2 values in the 25 NAM founders. Black circles, expected r^2 values in the NAM RILs. Shaded vertical bars, all QTL support intervals on a chromosome. Top, LD across an entire chromosome. Bottom, LD across an entire QTL support interval (if the significant SNP is inside of a QTL support interval) or a length of chromosome representing 5 cM of genetic

distance to either side of the significant SNP (if the SNP is outside of a QTL support interval). Discrete bands of founder LD in **b** result from the high allele frequency of this SNP (it segregates in 23 of 25 NAM families) and the limited sample size of 25 founders.
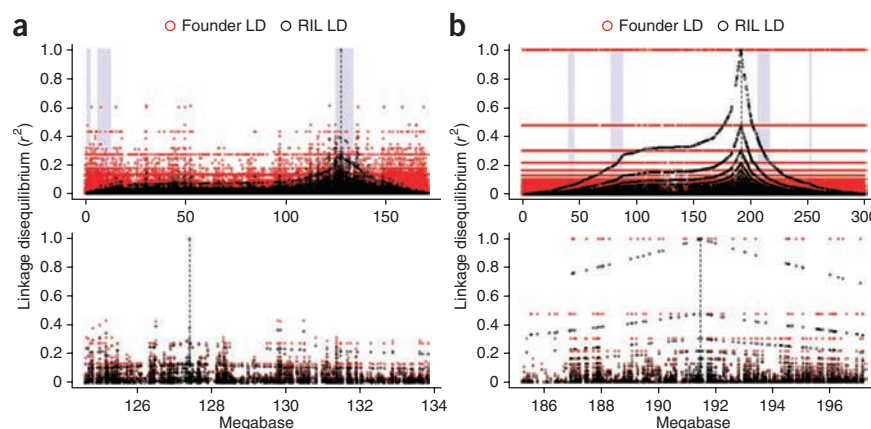
and integrating this knowledge to explain what mechanisms underlie quantitative resistance to necrotrophic pathogens in higher plants.

URLs. NCBI, http://www.ncbi.nlm.nih.gov/; MaizeSequence, http://www.maizesequence.org/; INTERPROSCAN, http://www.ebi.ac.uk/Tools/InterProScan.

METHODS
Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

AUTHOR CONTRIBUTIONS
E.S.B., M.D.M., S.K. and J.B.H. developed the NAM population. K.L.K., R.J.W., A.R.B., M.A.O.-R., J.C.Z., P.J.B.-K. and J.B.H. conducted the field experiments. E.S.B. and D.W. created HapMap data. P.J.B. and E.S.B. developed GWAS methods and software. K.L.K. and J.B.H. analyzed the data. K.L.K., R.J.W., P.J.B.-K. and J.B.H. wrote the paper.

1. Holley, R.N. & Goodman, M.M. New sources of resistance to southern corn leaf blight from tropical hybrid maize derivatives. *Plant Dis.* **73**, 562–564 (1989).
2. Bent, A.F. & Mackey, D. Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol.* **45**, 399–436 (2007).
3. Broglie, K. *et al.* Polynucleotides and methods for making plants resistant to fungal pathogens. United States Patent number 7,619,133 (2009).
4. Fukuoka, S. *et al.* Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* **325**, 998–1001 (2009).
5. Fu, D. *et al.* A kinase-start gene confers temperature-dependent resistance to wheat stripe rust. *Science* **323**, 1357–1360 (2009).
6. Manosalva, P.M. *et al.* A germin-like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. *Plant Physiol.* **149**, 286–296 (2009).
7. Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C. & Nelson, R.J. Shades of gray: The world of quantitative disease resistance. *Trends Plant Sci.* **14**, 21–29 (2009).
8. Krattinger, S.G. *et al.* A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* **323**, 1360–1363 (2009).
9. McMullen, M.D. *et al.* Genetic properties of the maize nested association mapping population. *Science* **325**, 737–740 (2009).
10. Buckler, E.S. *et al.* The genetic architecture of maize flowering time. *Science* **325**, 714–718 (2009).
11. Yu, J., Holland, J.B., McMullen, M. & Buckler, E.S. Genetic design and statistical power of nested association mapping in maize. *Genetics* **178**, 539–551 (2008).
12. Gore, M.A. *et al.* A first-generation haplotype map of maize. *Science* **326**, 1115–1117 (2009).
13. Tian, F. *et al.* Genome wide association identifies genes that have contributed to maize improvement. *Nat. Genet.* advance online publication, doi:10.1038/ng.746 (9 January 2010).
14. Balint-Kurti, P.J. *et al.* Precise mapping of quantitative trait loci for resistance to southern leaf blight, caused by Cochliobolus heterostrophus race O, and flowering time using advanced intercross maize lines. *Genetics* **176**, 645–657 (2007).
15. Wisser, R. The genetic architecture of disease resistance in maize: a synthesis of published studies. *Phytopathology* **96**, 120–129 (2006).
16. Balint-Kurti, P.J. & Carson, M.L. Analysis of quantitative trait loci for resistance to southern leaf blight in juvenile maize. *Phytopathology* **96**, 221–225 (2006).
17. Balint-Kurti, P.J. *et al.* Identification of quantitative trait loci for resistance to southern leaf blight and days to anthesis in a maize recombinant inbred line population. *Phytopathology* **96**, 1067–1071 (2006).
18. Balint-Kurti, P.J. *et al.* Identification of quantitative trait loci for resistance to southern leaf blight and days to anthesis in two maize recombinant inbred line populations. *Phytopathology* **98**, 315–320 (2008).
19. Zwonitzer, J. *et al.* Use of selection with recurrent backcrossing and QTL mapping to identify loci contributing to southern leaf blight resistance in a highly resistant maize line. *Theor. Appl. Genet.* **118**, 911–925 (2009).
20. Zwonitzer, J.C. *et al.* Mapping resistance quantitative trait loci for three foliar diseases in a maize recombinant inbred line population-evidence for multiple disease resistance? *Phytopathology* **100**, 72–79 (2010).
21. Valdar, W., Holmes, C.C., Mott, R. & Flint, J. Mapping in structured populations by resample model averaging. *Genetics* **182**, 1263–1277 (2009).
22. Valdar, W. *et al.* Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet.* **38**, 879–887 (2006).
23. Atwell, S. *et al.* Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627–631 (2010).
24. Brachi, B. *et al.* Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* **6**, e1000940 (2010).
25. Morris, E.R. & Walker, J.C. Receptor-like protein kinases: The keys to response. *Curr. Opin. Plant Biol.* **6**, 339–342 (2003).
26. Nicaise, V., Roux, M. & Zipfel, C. Recent advances in PAMP-triggered immunity against bacteria: Pattern recognition receptors watch over and raise the alarm. *Plant Physiol.* **150**, 1638–1647 (2009).
27. Dong, X. NPR1, all things considered. *Curr. Opin. Plant Biol.* **7**, 547–552 (2004).
28. Cao, H., Glazebrook, J., Clarke, J.D., Volko, S. & Dong, X.N. The *Arabidopsis* NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* **88**, 57–63 (1997).
29. Despres, C., DeLong, C., Glaze, S., Liu, E. & Fobert, P.-R. The *Arabidopsis* NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. *Plant Cell* **12**, 279–290 (2000).
30. Chern, M., Fitzgerald, H.A., Canlas, P.E., Navarre, D.A. & Ronald, P.C. Overexpression of a rice NPR1 homolog leads to constitutive activation of defense response and hypersensitivity to light. *Mol. Plant Microbe Interact.* **18**, 511–520 (2005).

31. Chern, M., Canlas, P.E., Fitzgerald, H.A. & Ronald, P.C. Rice NRR, a negative regulator of disease resistance, interacts with *Arabidopsis* NPR1 and rice NH1. *Plant J.* **43**, 623–635 (2005).
32. Mignotte, B. & Vayssiere, J.-L. Mitochondria and apoptosis. *Eur. J. Biochem.* **252**, 1–15 (1998).
33. Sacco, M.A. *et al.* The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and rangap2-dependent plant cell death. *PLoS Pathog.* **5**, e1000564 (2009).
34. Wisser, R.J., Sun, Q., Hulbert, S.H., Kresovich, S. & Nelson, R.J. Identification and characterization of regions of the rice genome associated with broad-spectrum, quantitative disease resistance. *Genetics* **169**, 2277–2293 (2005).
35. Maldonado, A.M., Doerner, P., Dixon, R.A., Lamb, C.J. & Cameron, R.K. A putative lipid transfer protein involved in systemic resistance signalling in *Arabidopsis*. *Nature* **419**, 399–403 (2002).
36. Boyle, P. *et al.* The BTB/POZ domain of the *Arabidopsis* disease resistance protein NPR1 interacts with the repression domain of TGA2 to negate its function. *Plant Cell* **21**, 3700–3713 (2009).
37. Kou, Y., Qiu, D., Wang, L., Li, X. & Wang, S. Molecular analyses of the rice tubby-like protein gene family and their response to bacterial infection. *Plant Cell Rep.* **28**, 113–121 (2009).
38. Gutterson, N. & Reuber, T.L. Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Biol.* **7**, 465–471 (2004).
39. Taj, G., Kumar, A., Bansal, K.C. & Garg, G.K. Signaling components in programmed cell death: Molecular and functional similarities between animals and plants. in *Recent Advances in Plant Biotechnology and Its Applications* (eds. Kumar, A. & Sopory, S.) 391–409 (I.K. International Publishing, New Dehli, India, 2008).
40. Nawrath, C., Heck, S., Parinthawong, N. & Metraux, J.-P. Eds5, an essential component of salicylic acid-dependent signaling for disease resistance in *Arabidopsis*, is a member of the mate transporter family. *Plant Cell* **14**, 275–286 (2002).
41. Zeng, L.-R. *et al.* Spotted leaf11, a negative regulator of plant cell death and defense, encodes a u-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. *Plant Cell* **16**, 2795–2808 (2004).
42. Chern, M., Fitzgerald, H.A., Canlas, P.E., Navarre, D.A. & Ronald, P.C. Overexpression of a rice NPR1 homolog leads to constitutive activation of defense response and hypersensitivity to light. *Mol. Plant Microbe Interact.* **18**, 511–520 (2005).
43. Zhou, N., Tootle, T.-L. & Glazebrook, J. *Arabidopsis* PAD3, a gene required for camalexin biosynthesis, encodes a putative cytochrome P450 monooxygenase. *Plant Cell* **11**, 2419–2428 (1999).
44. Takemoto, D., Hayashi, M., Doke, N., Nishimura, M. & Kawakita, K. Molecular cloning of a defense-response-related cytochrome P450 gene from tobacco. *Plant Cell Physiol.* **40**, 1232–1242 (1999).
45. Zhou, J., Loh, Y.-T., Bressan, R.A. & Martin, G.B. The tomato gene Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell* **83**, 925–935 (1995).
46. Loh, Y.T. & Martin, G.B. The Pto bacterial resistance gene and the Fen insecticide sensitivity gene encode functional protein kinases with serine/threonine specificity. *Plant Physiol.* **108**, 1735–1739 (1995).
47. Vogel, J.P., Raab, T.K., Schiff, C. & Somerville, S.C. PMR6, a pectate lyase-like gene required for powdery mildew susceptibility in *Arabidopsis*. *Plant Cell* **14**, 2095–2106 (2002).
48. Gu, Y.-Q. *et al.* Tomato transcription factors Pti4, Pti5, and Pti6 activate defense responses when expressed in *Arabidopsis*. *Plant Cell* **14**, 817–831 (2002).
49. Shah, J. Lipids, lipases, and lipid-modifying enzymes in plant disease resistance. *Annu. Rev. Phytopathol.* **43**, 229–260 (2005).
50. Zhang, Y., Fan, W., Kinkema, M., Li, X. & Dong, X. Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. *Proc. Natl. Acad. Sci. USA* **96**, 6523–6528 (1999).

ONLINE METHODS

The NAM population comprises 5,000 $F_{5:6}$ RILs derived from crosses between B73 and 25 diverse inbred lines. B73 selected as the reference line for its ubiquity in maize inbred line development and its sequenced genome[10,11,51]. Other parents were chosen to maximize allelic diversity on the basis of simple sequence repeat marker data[52]. B73 was crossed to each of the 25 inbred lines, and six generations of self-fertilization without selection were used to create RILs. Each RIL descended from a unique F_2 plant[9]. We derived 200 RILs from each family, leading to a population size of 5,000 lines. SNP genotyping later showed that some lines had >8% heterozygosity or contained nonparental alleles; these lines were dropped from data analysis, resulting in 4,699 RILs in the data set.

The NAM panel was evaluated for SLB resistance in randomized trials across three environments (**Supplementary Note**), two of which coincided with those used for the NAM flowering-time evaluation reported[10]. About 10% of plots were planted to replicate check founder lines to control within-environment spatial variability. The lines were exposed to *C. heterostrophus* naturally in one environment and were inoculated artificially in the other two. After flowering, plots were scored for severity of SLB symptoms twice, through assessment of number and area of lesions on the leaf closest to the ear ('ear leaf'), and the leaf immediately above the ear leaf. Disease scores were given according to a nine-point scale, on which a higher score indicate greater SLB susceptibility[17]. DTA, the time between planting and 50% of the plants in a plot shedding pollen, was also recorded on each plot.

To minimize the effects of environmental variation, BLUPs of each NAM line were used for QTL mapping. We used ASReml version 2 software[53] to implement a multivariate mixed model, treating the disease scores at the two time points on each plot as distinct traits (**Supplementary Note**). We predicted each RIL's genetic effect on SLB score at each time point after removing effects of environments and spatial field effects from this model. Plant maturity also affects SLB[54], therefore DTA was used as a linear and quadratic covariate to minimize confounding effects of maturity on SLB scores. The SLB index score for each RIL was computed as the average of its BLUPs for the two disease ratings.

Genotyping and map construction were done by McMullen *et al.*[9]. Joint linkage mapping was implemented as described[10]. Genome-wide association analysis was carried out by projecting founder SNP genotypes from the maize HapMap[12] on RILs[11,13]. All SNPs on a given chromosome were tested on residual phenotypic values resulting from a model containing all identified QTLs outside of the tested chromosome. Forward regression was conducted on the complete RIL dataset one chromosome at a time (after removing the effects of QTLs mapped on other chromosomes) to identify SNPs significant at $P < 0.0001$. In addition, 100 subsample data sets, each containing 80% of each family, were analyzed in the same way[13,21]. SNPs detected as significant in at least five subsamples (BPP ≥ 0.05) were considered for further testing.

All SNPs detected with either forward regression or subsampling were tested for inclusion in a single final model (**Supplementary Fig. 5**). SNPs outside of QTL support intervals were tested for significance in the context of the complete 32-QTL joint linkage model. Each SNP within a QTL support interval was tested for significance in the context of a 31-QTL joint linkage model that did not include the surrounding QTLs. All QTLs significant at $P < 0.0001$ in these tests plus the five QTLs that did not include significant SNPs were included in a final stepwise regression analysis with significance thresholds

to enter or leave the model of $P = 0.0001$. SNPs were considered as candidate genes for detailed sequence homology analysis if they were included in the final SNP plus QTL regression model. The predictive value of the QTL model was assessed by predicting SLB index values of founders on the basis of the allelic effects estimated in their progeny RILs. Founder values were predicted by summing their respective RIL population mean and their corresponding allelic effects at all 32 allelic effects, multiplying by a factor of two (because founders are homozygous, carrying two identical alleles at each QTL) and adding this sum to the intercept of the joint linkage model. A similar prediction was carried out based on the final SNP plus QTL model. The predictions were then evaluated by regressing the observed founder SLB values on the model predictions.

We computed LD between selected SNPs of greatest interest in the final GWA model and all other SNPs on the same chromosome. LD in the founder generation was computed based on the 25 founder HapMap SNP genotypes. LD expected in the RIL generation was computed by calculating the expected frequency of haplotypes in the NAM population given the haplotype frequencies in the founder generation and the expected recombination frequency for each pair of SNPs. The expected recombination frequency was computed by first estimating the NAM centimorgan map position of each HapMap SNP based on its physical position, using its relative physical distance from its flanking NAM markers. The RIL centimorgan map distance was converted to a single meiosis (F_2 generation) recombination frequency using the Haldane mapping function. The F_2 generation recombination frequency was then used to compute the RIL generation recombination frequency as described[55]. The frequency of the doubly non-B73 SNP allele haplotype (1,1) in the RIL generation was then computed as:

$$\text{freq}(1,1)_{\text{RIL}} = \left(1 - 2r_{\text{RIL}}\right)\left(\text{freq}(1,1)_{\text{P}}\right) + \left(2r_{\text{RIL}}\right)\left(p_1 q_1\right),$$

where r_{RIL} is the expected recombination frequency in the RIL generation, $\text{freq}(1,1)_{\text{P}}$ is the frequency of the (1,1) haplotype in the founders, and p_1 and q_1 are the frequencies of the non-B73 allele at the two SNPs. LD computations and graphics were carried out with R[56].

The coding sequences from which each SNP in the final model originated or from the two predicted genes flanking it (if the SNP is in a noncoding region) were investigated through use of protein-protein BLAST against the NCBI nonredundant protein database (see URLs). Information at MaizeSequence and/or the INTERPROSCAN program was used to identify conserved domains in the candidate protein sequences (see URLs).

51. Schnable, P.S. *et al.* The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009).
52. Liu, K. *et al.* Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics* **165**, 2117–2128 (2003).
53. Gilmour, A.R., Cullis, B., Gogel, B., Welham, S.J. & Thompson, R. *Asreml User's Guide. Release 2.0* (VSN International, Hemel Hempstead, UK, 2005).
54. Hooker, A.L. Southern leaf blight of corn present status and future prospects. *J. Environ. Qual.* **1**, 244–248 (1972).
55. Haldane, J.B.S. & Waddington, C.H. Inbreeding and linkage. *Genetics* **16**, 357–374 (1931).
56. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2010).