the plant journal

The Plant Journal (2020) **103,** 1246–1255 doi: 10.1111/tpj.14787

RESOURCE

Genotypic and phenotypic characterization of a large, diverse population of maize near-isogenic lines

Laura Morales[1], A. C. Repka[1], Kelly L. Swarts[2,3], William C. Stafstrom[1], Yijian He[4], Shannon M. Sermons[4], Qin Yang[4], Luis O. Lopez-Zuniga[4], Elizabeth Rucker[5], Wade E. Thomason[5], Rebecca J. Nelson[1] and Peter J. Balint-Kurti[4,6,*]

[1]*School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA,*
[2]*Gregor Mendel Institute, Austrian Academy of Sciences, Vienna BioCenter, 1030 Vienna, Austria,*
[3]*Max Perutz Labs, University of Vienna, Vienna BioCenter, 1030 Vienna, Austria,*
[4]*Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA,*
[5]*School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA 24061, USA, and*
[6]*Plant Science Research Unit, United States Department of Agriculture–Agricultural Research Service, Raleigh, NC 27695, USA*

Received 12 February 2020; revised 13 April 2020; accepted 22 April 2020; published online 29 April 2020.
*For correspondence (e-mail Peter.Balint-Kurti@USDA.GOV).

SUMMARY

Genome-wide association (GWA) studies can identify quantitative trait loci (QTL) putatively underlying traits of interest, and nested association mapping (NAM) can further assess allelic series. Near-isogenic lines (NILs) can be used to characterize, dissect and validate QTL, but the development of NILs is costly. Previous studies have utilized limited numbers of NILs and introgression donors. We characterized a panel of 1270 maize NILs derived from crosses between 18 diverse inbred lines and the recurrent inbred parent B73, referred to as the nested NILs (nNILs). The nNILs were phenotyped for flowering time, height and resistance to three foliar diseases, and genotyped with genotyping-by-sequencing. Across traits, broad-sense heritability (0.4–0.8) was relatively high. The 896 genotyped nNILs contain 2638 introgressions, which span the entire genome with substantial overlap within and among allele donors. GWA with the whole panel identified 29 QTL for height and disease resistance with allelic variation across donors. To date, this is the largest and most diverse publicly available panel of maize NILs to be phenotypically and genotypically characterized. The nNILs are a valuable resource for the maize community, providing an extensive collection of introgressions from the founders of the maize NAM population in a B73 background combined with data on six agronomically important traits and from genotyping-by-sequencing. We demonstrate that the nNILs can be used for QTL mapping and allelic testing. The majority of nNILs had four or fewer introgressions, and could readily be used for future fine mapping studies.

Keywords: near-isogenic lines, genotyping-by-sequencing, genome-wide association, allelic analysis, maize, disease resistance, flowering time, plant height, quantitative trait loci, genetics, *Zea mays*.

INTRODUCTION

Over the past decade, nearly 100 genome-wide association (GWA) studies in maize inbred diversity panels and nested association mapping (NAM) populations have been published (Xiao *et al.*, 2017). These studies have implicated hundreds of single-nucleotide polymorphisms (SNPs) and quantitative trait loci (QTL) in the control of dozens of complex traits (Xiao *et al.*, 2017). Candidate loci can be followed up with fine mapping and functional studies.

Near-isogenic lines (NILs) have been used extensively to characterize, validate and dissect QTL in maize (Szalma *et al.*, 2007; Pea *et al.*, 2009; Chung *et al.*, 2010; Eichten *et al.*, 2011; Mideros *et al.*, 2014; Peiffer *et al.*, 2014; Benson *et al.*, 2015; Lennon *et al.*, 2016, 2017; Xiao *et al.*, 2016; Kolkman *et al.*, 2019; Martins *et al.*, 2019). However, the development of NILs is time-consuming and costly, as it requires several generations of backcrossing, selfing and genotyping. As such, NIL populations have been limited in

size and introgression donor diversity in the public sector (Szalma *et al.*, 2007; Pea *et al.*, 2009; Chung *et al.*, 2010; Eichten *et al.*, 2011; Lennon *et al.*, 2016). Via the International Maize and Wheat Improvement Center (CIMMYT), Syngenta AG (Basel, Switzerland) has made public a large panel of maize NILs derived from crosses between 25 diverse inbred lines and B73, henceforth referred to as the Syngenta panel (Gandhi *et al.*, 2008). Kolkman *et al.* (2019) recently used a subset of 412 'nested' NILs (nNILs) from the Syngenta panel, which were chosen for having introgressions surrounding disease resistance QTL, for GWA and candidate gene identification of resistance to northern leaf blight (NLB).

Here we characterize a population of 1270 nNILs from the Syngenta panel containing randomly positioned introgressions from 18 inbred lines in a B73 background. The parents of the nNILs, B73 and the 18 donor parents, represented 19 of the 27 diverse inbred founders of the maize NAM population: B73, CML103, CML228, CML247, CML277, CML322, CML333, CML52, CML69, Ki11, Ki3, M162W, Mo17, Mo18W, NC350, NC358, Oh43, Tx303 and Tzi8 (Yu *et al.*, 2008). We phenotyped the nNILs for six quantitatively inherited morphophysiological and disease resistance traits that had been previously characterized with the maize NAM population: days to anthesis (DTA), ear (EHT) and plant (PHT) height, and resistance to the fungal foliar diseases gray leaf spot (GLS), NLB and southern leaf blight (SLB; Buckler *et al.*, 2009; Kump *et al.*, 2011; Poland *et al.*, 2011; Peiffer *et al.*, 2014; Benson *et al.*, 2015). Using genotyping-by-sequencing (GBS; Elshire *et al.*, 2011), we successfully genotyped 896 of the 1270 nNILs. We used GWA to identify QTL for these traits, and conducted allelic analysis on associated QTL.

RESULTS

Phenotypic and genotypic variation in flowering time, height and disease resistance

We evaluated the 1270 BC_5F_4 nNILs for flowering time (DTA), height (EHT, PHT) and resistance to three foliar diseases (GLS, NLB, SLB) across four–six environments each (Additional file 1). Each disease was assessed separately. We used mixed models to assess genotype and environmental effects on the six traits in the nNILs. Trait variation relative to B73 was small, as would be expected for NILs (Table 1). Broad-sense heritability (H^2) for DTA was low (0.4), which may have been due to the combination of low phenotypic variation (less than half a day relative to B73) and relatively imprecise phenotyping (scores taken approximately 2–3 times per week) of DTA (Table 1). H^2 was relatively high for height and disease resistance (0.61–0.82; Table 1). The interaction between genotype and environment (GxE) explained < 1% of the total variance in flowering time and height (Table 1). GxE also accounted for a small proportion of the total variance in the three disease resistance traits (2.5-8%; Table 1). The evaluation of GLS resistance under artificially inoculated and naturally infected conditions may have contributed to the relatively higher GxE on GLS resistance (8%) than on resistance to NLB and SLB (2.5%; Table 1).

Correlations among flowering time, height and disease resistance

For trait correlation analysis, genotype best linear unbiased predictors (BLUPs) were extracted from mixed models (Additional file 2). As expected given their shared genetic control (Peiffer *et al.*, 2014), EHT and PHT were strongly positively correlated ($r = 0.76$). Height was negatively correlated with resistance to GLS ($r = -[0.25–0.26]$) and SLB ($r = -[0.07–0.11]$; Table 2). The correlation of EHT with GLS and SLB resistance may have resulted from the scoring method, in which disease scores were taken on the ear leaf. For all disease experiments, inoculum was introduced into the maize whorl, from which new leaves emerged, and the disease then progressed upwards on the plant. As such, plants with higher ear leaves may have had slightly reduced GLS and SLB symptom severity scores. In contrast, NLB diseased leaf area (DLA) was measured on a whole plant basis, and no correlation was observed between NLB resistance and height (Table 2). SLB and GLS resistance were positively correlated with each other ($r = 0.38$) and with NLB resistance ($r = 0.23–0.26$; Table 2), suggesting that some of the loci segregating in the nNILs may confer multiple disease resistance.

Characteristics of introgressions present in the nNILs

We successfully genotyped 896 nNILs using GBS (Elshire *et al.*, 2011; Glaubitz *et al.*, 2014) and imputed missing SNPs with FILLIN (Swarts *et al.*, 2014). Gandhi *et al.* (2008) previously estimated an average of 5% donor genome per BC_5 line in the Syngenta panel. After four generations of selfing, the nNILs had a lower proportion of heterozygous SNPs ($\mu = 0.7\%$; $\sigma = 1\%$; range: 0–13%) than expected (approximately 6%), which may have been due to the low coverage of the nNIL GBS data ($\mu_{raw} = 0.36$; $\sigma_{raw} = 0.06$; $\mu_{imputed} = 0.87$; $\sigma_{imputed} = 0.09$) and the undercalling of heterozygous sites by FILLIN (Swarts *et al.*, 2014).

With the TASSEL 5 GenosToABH plugin (Bradbury *et al.*, 2007) and the ABHgenotypeR package (Reuscher and Furuta, 2016), we identified 241 455 'ABH' SNPs that were homozygous and polymorphic between B73 and at least one of the 18 non-B73 parents (Additional file 3). To account for the variation in sequencing depth and quality across samples as well as the genetic variation found across seed stocks of identically named genetic lines (Liang and Schnable, 2016), we utilized several seed sources per parent to reliably call parental genotypes. The ABH SNPs were then used to determine the positions and genotypes of the introgressions in the nNILs (Additional

Table 1 Genotype and environment effects on flowering time, height and disease resistance traits from mixed models

		Proportion of total variance						
Trait	*N*	Genotype	Env.	GxE	B[E]	Error	H^2	Variation relative to B73
DTA	5401	0.020	0.752	0.000	0.079	0.149	0.40	0.431
EHT	5398	0.138	0.294	0.010	0.171	0.387	0.68	4.334
PHT	5393	0.115	0.470	0.006	0.134	0.275	0.71	6.247
GLS	4829	0.132	0.474	0.079	0.079	0.237	0.64	0.173
NLB	4455	0.070	0.467	0.024	0.281	0.158	0.61	1.313
SLB	4333	0.385	0.051	0.026	0.205	0.333	0.82	0.341

N, total number of observations; Env., environment; GxE, interaction between genotype and environment; B[E], block nested within environment; H^2, broad-sense heritability; trait variation relative to B73 was estimated as the standard deviation of the genotype BLUPs, with the following units: days to anthesis (DTA); ear (EHT) and plant (PHT) height in cm; resistance to gray leaf spot (GLS) and southern lead blight (SLB) on a 1 (most resistant) to 9 (most susceptible) scale; percent DLA for northern leaf blight (NLB).

Table 2 Pairwise Pearson correlations between genotype BLUPs of flowering time, height and disease resistance traits

DTA	SLB	NLB	GLS	PHT	EHT
	-0.06^{*}	-0.07^{*}	-0.04^{ns}	-0.01^{ns}	-0.01^{ns}
EHT	-0.11^{**}	0.04^{ns}	-0.26^{**}	0.76^{**}	—
PHT	-0.07^{*}	0.05^{ns}	-0.25^{**}	—	—
GLS	0.38^{**}	0.26^{**}	—	—	—
NLB	0.23^{**}	—	—	—	—

Pearson correlation coefficients are shown in each cell, with significance denoted as: $^{ns}P > 0.05$, $*0.05 > P \geq 0.0001$, $**P < 0.0001$; BLUP, best linear unbiased predictor; DTA, days to anthesis; EHT, ear height; PHT, plant height; GLS, gray leaf spot; NLB, northern leaf blight; SLB, southern leaf blight.

file 4). We found 2638 introgressions across the nNILs. Less than 1% of ABH sites were missing or heterozygous, indicating that the introgression genotype calls matched the introgression donors assigned by Syngenta.

The size, number and distribution of introgressions contained within the nNILs here were similar to the findings of Kolkman *et al.* (2019). The introgressions spanned the entire genome (Figure 1). The distribution of individual introgression lengths was highly positively skewed, with a mean physical length of 13.2 Mbp (σ = 22 Mbp; median = 4.8 Mbp) ranging between 0.2 and 191.6 Mbp, and a mean genetic length of 12.2 cM (σ = 12.2 cM; median = 8.0 cM) ranging between 0 and 96.8 cM. The number of introgressions per nNIL ranged from 1 to 10, with a mean of 2.9 (σ = 1.5). Seventy-five percent of nNILs had four or fewer introgressions, and we identified 138 nNILs with a single introgression. Introgression length decreased with increasing distance from the centromere, as would be expected given the increased probability of recombination events at the ends of the chromosomes and the suppression of cross-overs within pericentromeric regions (Rodgers-Melnick *et al.*, 2015; Figure 1).

We found substantial overlap among introgressions from different nNILs (Figure 1). On average, 45.9 nNILs had introgressions at any given locus (σ = 24.4; range: 4–211), with each donor allele represented by 3.4 nNILs (σ = 2.0; range: 0–17). Pairs of introgressions with overlapping positions had a mean physical overlapping interval of 19.1 Mbp (σ = 29.5 Mbp; median = 6.5 Mbp) and a mean genetic overlapping interval of 18.6 cM (σ = 16.5 cM; median = 14.0 cM). As such, GWA should be feasible with the nNIL panel.

On average, the introgression lengths in each nNIL summed to 38.7 Mbp (σ = 41 Mbp; median = 24.2 Mbp; range: 0.2–283.3 Mbp), approximately 2% of the genome. To investigate whether the size of non-B73 genome (sum of physical introgression lengths per nNIL) contributed to phenotypic variation, we correlated total physical introgression length versus the absolute values of each of the six traits. We found that total introgression length was weakly, but significantly, correlated with EHT, PHT and SLB (r = 0.13, $P < 0.001$), but was not significantly correlated ($P > 0.05$) with the other traits.

Loci associated with height and disease resistance

We used the 241 455 ABH SNPs underlying the introgressions contained within the nNILs for GWA on each of the six disease resistance and morphophysiological traits (Additional files 5 and 6). Multiple testing correction of the GWA results was conducted with the false discovery rate (FDR) method, from which SNP-trait associations with FDR-corrected *P*-values below 0.05 were considered significant (Additional files 5 and 6). GWA identified three QTL for EHT, five for PHT, four for NLB resistance, and 17 for SLB resistance (Table 3). No SNPs were significantly associated with DTA or GLS. The quantile-quantile plots of the actual versus the expected *P*-values from GWA on DTA and GLS further indicated that the nNILs did not have sufficient power to detect QTL for DTA and GLS (Additional file 5).

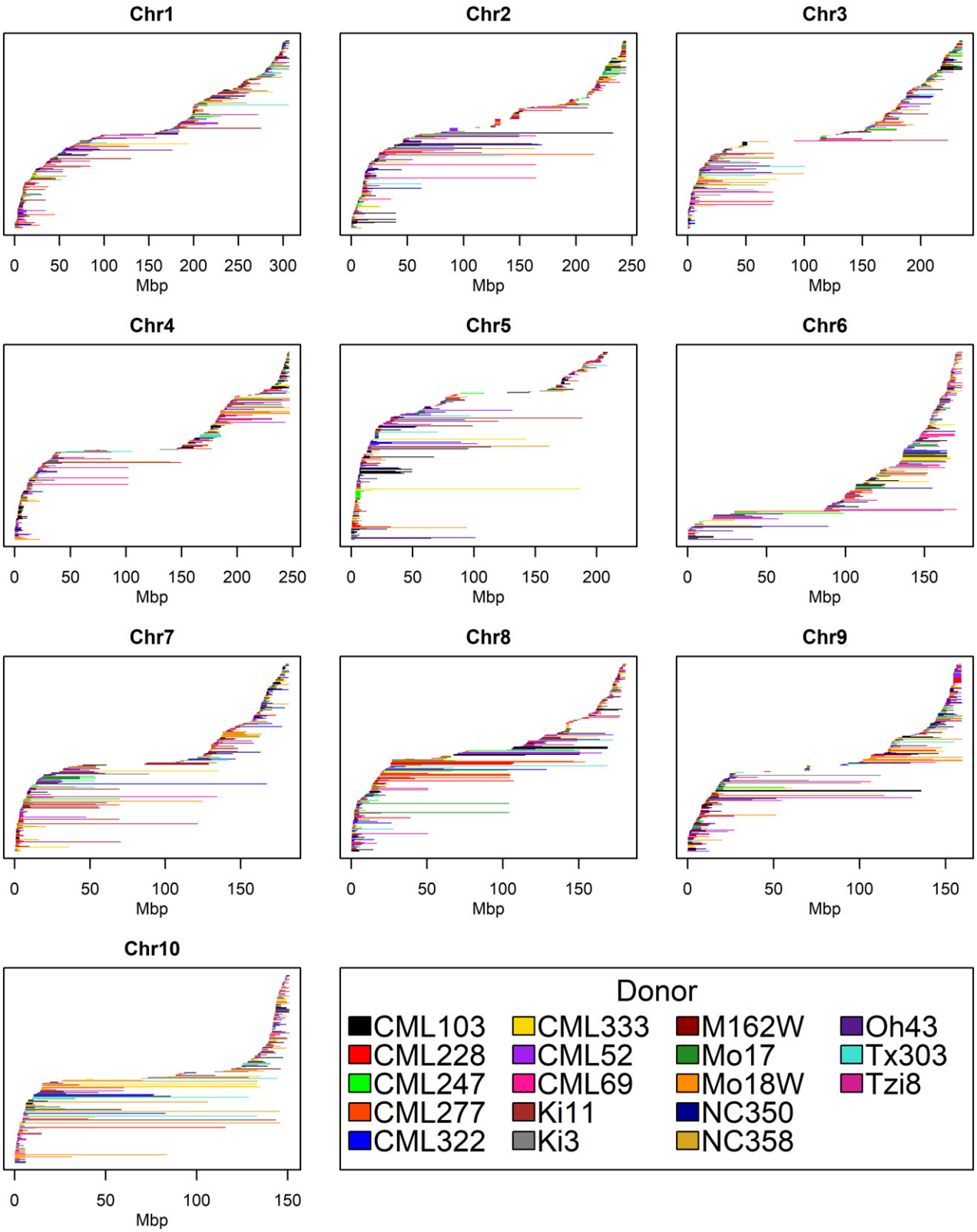

Figure 1. Physical positions and donors of introgressions present in nested near-isogenic lines (nNILs).
Introgressions are ordered along the *y*-axis based on start site position, independent of the nNIL in which they were found, and colored by donor source.

Approximately 40% of QTL in the nNILs (12/29) overlapped with QTL previously reported in the NAM population (Poland *et al.*, 2011; Peiffer *et al.*, 2014). PHT QTL in bins 4.10, 9.04 and 9.04–9.05, and an EHT QTL in bin 1.00 (Table 3) colocalized with NAM QTL for PHT and EHT (Peiffer *et al.*, 2014). SLB resistance QTL in bins 3.02–3.04, 5.03, 6.06 and 7.00 (Table 3) coincided with NAM QTL for SLB resistance (Kump *et al.*, 2011). NAM QTL for NLB resistance (Poland *et al.*, 2011) colocalized with nNIL NLB resistance QTL in bins 6.05, 8.05–8.06 and 9.01–9.02 (Table 3). Both the nNILs here and the smaller nNIL panel used by Kolkman *et al.* (2019) identified QTL associated with NLB resistance in bins 6.05 and 8.05. Differences in mapping methods (single-marker association versus stepwise regression), *Setosphaeria turcica* isolates (mixture of races versus race 0) and environments (North Carolina versus New York) may have led to the detection of distinct NLB loci here and by Kolkman *et al.* (2019).

Diverse allelic variation at quantitative trait loci associated with height and disease resistance

Although the nNILs did not have sufficient allelic replication (three nNILs per donor per site on average) for NAM (Yu *et al.*, 2008), we sought to test introgression donor effects relative to the B73 allele at loci identified by GWA. For each locus, we regressed the associated trait against the 19 possible donor alleles. Donor effects ranged from −1.6 to 0.8 for SLB resistance, which was scored on a 1 (most resistant) to 9 (most susceptible) scale (Figure 2). NLB resistance donor effects ranged from −2.7% to 3.4% DLA. Donor effects on EHT and PHT ranged from −20.5 to 3.9 cm and from −14.2 to 11.4 cm, respectively (Figure 2).

Although EHT and PHT were strongly correlated in the nNILs ($r = 0.8$), only one EHT and one PHT colocalized on chromosome 3 at 0.5–1.5 Mbp (Table 3). This shared QTL may be pleiotropic, as demonstrated by the high positive correlation between donor effects on EHT and on PHT ($r = 0.92$, $P < 0.0001$). NLB and SLB resistance were positively phenotypically correlated ($r = 0.23$), but we did not find any evidence of multiple disease resistance at the QTL level.

DISCUSSION

The large population of 1270 nNILs is an important resource for the maize genetics community. Here we publicly provide data on six distinct phenotypes, as well as raw sequencing and SNP data from GBS. The introgressions present in the 896 genotyped nNILs not only span the entire maize genome but are also derived from 18 diverse inbred lines, which were founders of the well-established maize NAM population (Yu *et al.*, 2008). The diversity of the nNIL introgressions allows for the evaluation of allelic series relative to B73, the recurrent parent of both the nNILs and the NAM population (Yu *et al.*, 2008), and the inbred line that has been used as the maize reference genome for over a decade (Schnable *et al.*, 2009; Chia *et al.*, 2012; Law *et al.*, 2015; Jiao *et al.*, 2017). For future applications, researchers could use nNILs carrying introgressions at loci of interest for QTL characterization, validation, fine mapping and allelic testing. The six traits measured on the nNILs exhibited relatively high heritability (0.4–0.8), indicating that the introgressions within the nNILs contribute to phenotypic variation.

We have demonstrated that the nNILs have sufficient allelic replication for GWA. We identified 29 QTL for height and disease resistance, 12 of which had been previously mapped in the NAM population (Kump *et al.*, 2011; Poland *et al.*, 2011; Peiffer *et al.*, 2014). These results indicate that the low number of introgressions (three on average) and small proportion of non-B73 genome (approximately 2% of the genome) per nNIL can allow for the evaluation of QTL with minimal confounding effects from the genetic background. Differences in mapping methods and genomic features may partially explain the relatively high number of novel QTL detected (i.e. QTL detected here but not in the NAM population). While we used single marker regression for GWA and allelic analysis, previous NAM studies employed joint linkage mapping approaches that modeled multiple markers with nested effects by family (Kump *et al.*, 2011; Poland *et al.*, 2011; Peiffer *et al.*, 2014; Benson *et al.*, 2015). Although the introgression lengths (12.2 ± 12.2 cM) and introgression overlapping intervals (18.6 ± 16.5 cM) in the nNILs are larger than the recombination breakpoints across the genomes of the NAM recombinant inbred lines (0.9 ± 1.3 cM; Li *et al.*, 2015), QTL confidence intervals in the nNILs (4.0 ± 6.8 cM) were comparable to those previously described in the NAM population (6.5 ± 4.7 cM; Kump *et al.*, 2011; Poland *et al.*, 2011). Donor allele effects on NLB resistance (−3 to 3) were similar in the nNILs and the NAM (Poland *et al.*, 2011). SLB resistance donor allele effects were relatively larger in the nNILs (−1.6 to 0.8) than in the NAM (−0.4 to 0.4).

EXPERIMENTAL PROCEDURES

Population development

We requested 1270 nNILs from the greater Syngenta AG (Basel, Switzerland) panel, which had been derived from crosses between 18 diverse inbred lines and the recurrent inbred parent B73 (Gandhi *et al.*, 2008), from CIMMYT. The nNILs had been backcrossed for five generations (BC_5), and Syngenta had previously estimated an average of 5% donor genome per BC_5 seed stock (Gandhi *et al.*, 2008). The BC_5 nNILs were then self-pollinated for four generations (BC_5F_4). The 1270 BC_5F_4 nNILs each contained introgressions from one of the 18 donor lines. nNILs with introgressions from the same donor were considered a subpopulation, with each subpopulation comprising 50–108 BC_5F_4 nNILs (Table 4). The introgression donors were of tropical, temperate non-stiff stalk, and mixed origin, while B73 is a temperate stiff

Table 3 QTL significantly associated with resistance to NLB and SLB, and with EHT and PHT

QTL	Chr.	Physical position (bp)	Bin(s)	Effect size	*P*-value	Number of nNILs with introgressions at QTL	NAM QTL colocalization
NLB1	6	122 704 858–124 071 097	6.05	−1.07	9×10^{-4}	17	Poland *et al.* (2011)
NLB2	8	106 931 964–115 490 705	8.03–8.04	−1.17	2×10^{-4}	18	
NLB3	8	124 043 604–161 476 885	8.05–8.06	−0.47	8×10^{-3}	58	Poland *et al.* (2011)
NLB4	9	8 145 673–12 433 407	9.01–9.02	0.59	3×10^{-3}	44	Poland *et al.* (2011)
SLB1	3	6 383 526–28 243 869	3.02–3.04	−0.2	1×10^{-6}	72	Kump *et al.* (2011)
SLB2	3	31 701 560–38 013 747	3.04	−0.19	1×10^{-3}	35	
SLB3	3	72 227 366	3.04	−0.39	6×10^{-4}	9	
SLB4	3	151 815 973–152 447 840	3.05	0.52	7×10^{-4}	5	
SLB5	3	218 139 425–224 230 787	3.09	0.18	2×10^{-3}	34	
SLB6	5	28 911 521–31 726 291	5.03	0.21	9×10^{-4}	31	Kump *et al.* (2011)
SLB7	5	32 818 953–32 945 804	5.03	0.2	9×10^{-4}	33	
SLB8	5	34 841 144–40 164 346	5.03	0.18	2×10^{-3}	34	
SLB9	6	153 427 535–155 606 473	6.06	−0.19	2×10^{-4}	47	Kump *et al.* (2011)
SLB10	7	1 395 966–2 868 330	7.00	−0.2	1×10^{-4}	45	Kump *et al.* (2011)
SLB11	8	60 724 302–73 563 802	8.03	−0.21	1×10^{-3}	27	
SLB12	8	102 416 404–105 682 245	8.03	−0.23	2×10^{-3}	22	
SLB13	9	60 693 136–68 800 792	9.03	−0.36	3×10^{-4}	12	
SLB14	9	69 430 265–70 796 565	9.03	−0.37	2×10^{-5}	16	
SLB15	9	71 749 971–88 688 648	9.03	−0.44	6×10^{-4}	7	
SLB16	9	100 696 675	9.03	−0.38	3×10^{-4}	11	
SLB17	10	139 318 740–140 281 980	9.06	−0.23	1×10^{-3}	24	
EHT1	1	56 073	1.00	−9.36	2×10^{-4}	3	Peiffer *et al.* (2014)
EHT2	3	472 473–1 2 05 491	3.00	−3.24	2×10^{-5}	33	
EHT3	7	55 319 100–59 929 944	7.02	−3.45	8×10^{-4}	18	
PHT1	3	472 473–1 457 941	3.00	−3.26	2×10^{-3}	35	
PHT2	4	238 089 470–238 092 172	4.10	3.74	1×10^{-3}	30	Peiffer *et al.* (2014)
PHT3	9	102 115 269–106 003 663	9.04	6.87	8×10^{-5}	13	Peiffer *et al.* (2014)
PHT4	9	106 480 828–119 380 527	9.04	3.78	3×10^{-4}	37	Peiffer *et al.* (2014)
PHT5	9	120 343 929–130 365 828	9.04–9.05	3.11	4×10^{-3}	34	Peiffer *et al.* (2014)

Physical positions are in AGPv4 coordinates; bins divide the maize genetic map into 100 approximately equal segments; effect sizes are in the following units: SLB (scale of 1–9, from most resistant to most susceptible), NLB (% DLA), EHT and PHT (cm); QTL that colocalized with QTL previously reported in the NAM population for the same trait are indicated by the corresponding NAM publication.
EHT, ear height; NAM, nesting association mapping; NLB, northern leaf blight; nNIL, nested near-isogenic line; PHT, plant height; QTL, quantitative trait loci; SLB, southern leaf blight.

stalk line (Romay *et al.*, 2013; Table 4). Seed for the 1270 nNILs can be requested from CIMMYT for future studies.

Field design and inoculation

The panel was evaluated for GLS, NLB and SLB severity across four separate year/field replication environments for each disease. For each experiment, we used an augmented incomplete block design, in which each subpopulation was grown separately with 25-plot blocks. nNIL plots were replicated once per environment and randomized within the subpopulation. One plot of B73, the recurrent parent, was randomized within each 25-plot block. The NLB and SLB experiments were conducted at the Central Crops Research Station in Clayton, NC. NLB was evaluated in 2015 and 2018 with one replication per year, and in 2016 with two replications. SLB was assessed in 2014 with two replications, and in 2015 and 2016 with one replication per year. GLS was screened in Andrews, NC in 2014 with one replication, and at College Farm Research Station in Blacksburg, VA in 2016 with two replications and in 2017 with one replication.

Inoculum for the GLS (Blacksburg, VA only), NLB and SLB experiments was prepared from mixtures of isolates of *Cercospora zeae-maydis*, *S. turcica* (race 0, race 1, race 2,3 and race 2,3,N) and *Cochliobolus heterostrophus* (including 2–16 Bm and Hm540), respectively, as described by Sermons and Balint-Kurti (2018) and Martins *et al.* (2019). The Andrews, NC site has naturally high and consistent levels of *C. zeae-maydis* inoculum present in soil crop residues, which allowed for evaluation of GLS under natural infection (Benson *et al.*, 2015; Lennon *et al.*, 2016; Sermons and Balint-Kurti, 2018; Martins *et al.*, 2019).

Phenotyping

In all SLB experiments and in the 2015 NLB experiment, we took observations on flowering time two–three times per week, and estimated plot anthesis dates. EHT and PHT were measured in each plot in all SLB experiments, and in the 2015 and 2018 NLB experiments. Each plot was visually evaluated for disease severity at two time points per experiment: GLS on July 28 and August 6 in 2014, August 9 and 19 in 2016, and August 9 and 25 in 2017; NLB on July 13 and 22 in 2015, July 17 and 21 in 2016, and July 18 and 26, 2018; SLB on July 14 and 31 in 2014, July 10 and 25 in 2015, and July 11 and 20, 2016. GLS and SLB plots were scored on a 1–9 scale, where 1 corresponds to most resistant and 9

denotes most susceptible (Kump *et al.*, 2011; Wisser *et al.*, 2011; Benson *et al.*, 2015; Sermons and Balint-Kurti, 2018). NLB plots were scored for 0–100% DLA with 5% increments (Poland *et al.*, 2011; Sermons and Balint-Kurti, 2018). The two disease scores were averaged for each plot that was scored twice (Lennon *et al.*, 2016). Five nNILs exhibited lesion-mimic mutant phenotypes (CML333/B73 NIL-1002, CML333/B73 NIL-1007, Ki11/B73 NIL-1103, Ki11/B73 NIL-1104, Ki3/B73 NIL-1258), and thus were not included in further analyses on disease resistance.

Linear models and heritability

We used SAS software, Version 9.4 (SAS Institute Inc, 2011) to fit linear mixed models for all traits measured in this study. For each trait, a mixed model was fit using PROC MIXED with environment, block[environment], genotype, and genotype-by-environment (GxE) as random effects. Variance components were extracted from each model and used to calculate broad-sense heritability (H^2) as

$$H^2 = \frac{\sigma^2_G + \frac{\sigma^2_{GE}}{y_h} + \frac{\sigma^2_e}{p_h}}{\sigma^2_G},$$

where σ^2_G, σ^2_{GE}, and σ^2_e are the genotype, GxE and error variances, respectively, and y_h and p_h are defined as

$$y_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{y_i}}$$

$$p_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{p_i}},$$

where n is the number of genotypes, and y_i and p_i are the number of environments and plots for the ith genotype, respectively (Holland, Nyquist and Cervantes-Martinez, 2003).

DNA extraction and sequencing

Three seeds per nNIL were planted in a greenhouse at the Cornell University Kenneth Post Laboratory. Approximately 1 month after germination, 25–50 mg of seedling leaf tissue was collected from each nNIL. Genomic DNA was extracted from fresh leaf tissue using the DNeasy Plant Mini Kit system (Qiagen, Valencia, CA, USA). GBS (Elshire *et al.*, 2011) was performed on 914 unique nNIL DNA samples by BGI Americas (Cambridge, MA, USA) using 100-base pair paired-end sequencing with the HiSeq 4000 system (Illumina, San Diego, CA, USA).

Genotyping-by-sequencing analysis

For proprietary reasons, the BGI team clipped the barcodes on the raw read fastq files before delivering the data to us. In order to use the TASSEL-GBS pipeline (Glaubitz *et al.*, 2014) to call SNPs on our nNIL panel, we modified the TASSEL 5 (Bradbury *et al.*, 2007) base code to bypass the requirement for barcodes (Additional file 7). In addition, the TASSEL-GBS pipeline (Glaubitz *et al.*, 2014) only accepts single-end reads. In order to meet the single-end read requirement, we generated the reverse complement of the second paired-end fastq file for each set of paired-end fastq files with the FASTX-Toolkit (Hannon, 2010), and then merged the two fastq files.

We made a GBS build from 4603 lines from the USA maize inbred seed bank that were previously analyzed with GBS by the Panzea Maize Diversity Project (Zhao, 2006). We accessed the Panzea Maize Diversity Project GBS fastq files (NCBI BioProject accession #PRJNA200550) from the public NCBI Sequence Read Archive (Leinonen, Sugawara and Shumway, 2011). The Panzea Maize Diversity Project GBS fastq files were processed with the GBS Discovery Pipeline in TASSEL 5 (Bradbury *et al.*, 2007; Glaubitz *et al.*, 2014) with the B73 reference genome 4.0 (NCBI BioProject accession #PRJNA10769), a k-mer length filter of 64 and a minimum quality score of 20. The TASSEL GBS Discovery Pipeline yielded a GBS database, which we then used to call SNPs on the Panzea Maize Diversity Project and nNIL fastq files with the GBS Production Pipeline in TASSEL 5 (Bradbury *et al.*, 2007; Glaubitz *et al.*, 2014) with a k-mer length filter of 64 and a minimum quality score of 20 for the Panzea Maize Diversity Project lines and 10 for the nNILs. This resulted in 564 764 SNPs. We then imputed the GBS SNPs using the FILLIN plugin (Swarts *et al.*, 2014) in TASSEL 5 (Bradbury *et al.*, 2007; Additional file 8).

Introgression calling

In order to determine introgression positions, we first used the GenosToABH plugin in TASSEL 5 (Bradbury *et al.*, 2007) to convert SNPs from the nucleotide-based format (e.g. A, C, G, T, etc.) to a parent-based format (A, B, H, N) for each subpopulation separately. In the parent-based format, 'A' was assigned to SNPs that matched the B73 allele, 'B' was assigned to SNPs that matched the non-B73 introgression donor of the subpopulation, 'H' was assigned to heterozygous SNPs, and 'N' was assigned to missing sites. The Panzea Maize Diversity Project (Zhao, 2006) previously GBS-sequenced several seed sources per NAM parent (35 B73 DNA samples and seven per non-B73 line). To call ABH genotypes on each subpopulation with the TASSEL 5 GenosToABH plugin, we utilized all samples of B73 and of the respective non-B73 parent.

We then imputed ABH genotypes with the ABHgenotypeR package (Reuscher and Furuta, 2016) in R version 3.3.1 (R Core Team, 2019). Heterozygous and missing genotypes accounted for a small proportion of all ABH sites (< 1%), and were usually adjacent to or within long stretches of B sites. As such, we attributed H and N calls to low site coverage and converted them to B genotypes. Next, our script converted A sites within 5 Mbp of a B site to B, as A genotypes found within stretches of B genotypes were likely due to sequencing error. To remove small sequencing artifacts, we converted short stretches of < 50 consecutive B sites to A. Similar to the findings of Kolkman *et al.* (2019) in a parallel panel of nNILs, we observed a large sequencing artifact at the end of chromosome 5 (approximately 210–224 Mbp). Eighteen nNILs had a single introgression in the chromosome 5 artifact region and were thus not included in further analysis. For the remaining 896 nNILs, introgressions within the chromosome 5 artifact region were removed before further analysis. This resulted in 241 455 SNPs that were polymorphic between B73 and at least one donor. We used the GenomicRanges R package (Aboyoun *et al.*, 2013) to find overlapping regions between introgressions in different nNILs.

We sought to compare introgression lengths and QTL confidence intervals found in the nNILs with recombination breakpoints and QTL confidence intervals from the NAM population, which have largely been reported as genetic distances (cM). As such, we converted the physical positions of the introgressions in the nNILs to genetic positions using a genetic map of 7386 SNPs scored on the NAM population (Olukolu *et al.*, 2014; Additional file 4).

Genome-wide association

For GWA, the 241 455 SNPs described above were converted to numerical genotypes, where 0 = B73 allele and 1 = non-B73 allele. We used the generalized linear model (GLM) plugin in TASSEL 5 (Bradbury *et al.*, 2007) for single marker regression on each trait.

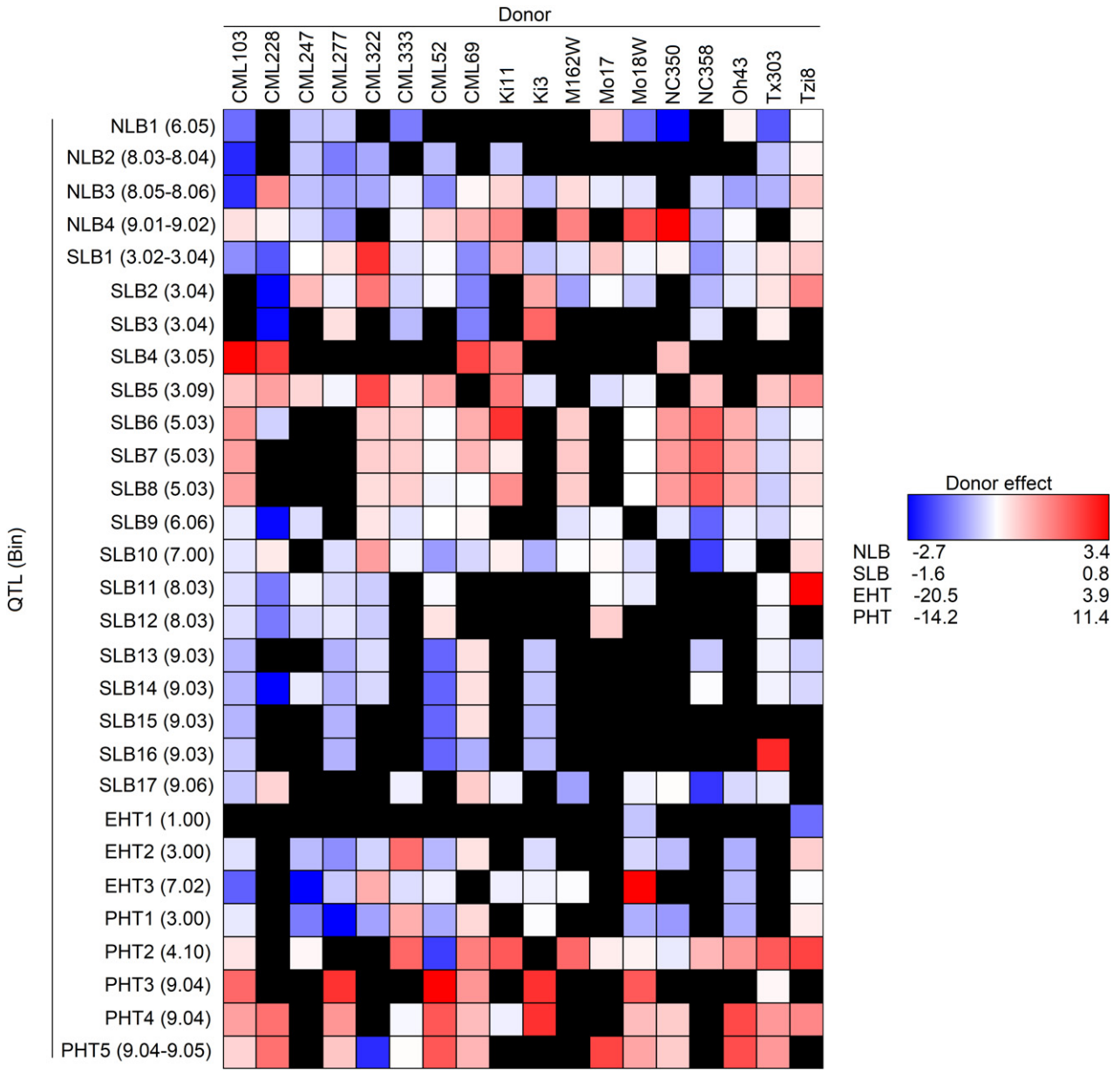

Figure 2. Introgression donor allele effects on disease resistance and height at significantly associated quantitative trait loci (QTL).
The bin location of each QTL is noted in parentheses next to the respective QTL. Bins divide the maize genetic map into 100 approximately equal segments. For each QTL (rows), introgression donor allele effect estimates (columns) on the trait associated with the QTL are displayed. Donor effects significantly ($P < 0.05$) associated with the trait are colored with a heat map, scaled with blue as the minimum and red as the maximum. Significantly associated ($P < 0.05$) donor effects are outlined in black. Donor effects that could not evaluated (NA) are black. Allele effects are in the following units: gray leaf spot (GLS) and southern leaf blight (SLB; scale of 1 to 9, from most resistant to most susceptible); northern leaf blight (NLB; % DLA); ear height (EHT) and plant height (PHT; cm).

Briefly, a GLM was fit for every trait-SNP pair, where the trait was the response and the SNP was a fixed effect.

For multiple testing correction, we extracted *P*-values from the TASSEL GLM output, and used the FDR method to adjust the *P*-values with the p.adjust function in R version 3.3.1 (R Core Team, 2019). FDR tests were conducted for each trait separately. FDR-adjusted *P*-values less than 0.05 were considered significant. Adjacent significant SNPs were considered to be part of the same QTL, and QTL start and end positions were determined by the left-most and rightmost SNPs, respectively.

Quantitative trait loci allelic testing

Single-nucleotide polymorphisms with a numerical genotype of 0 were coded as having a B73 allele. For each nNIL, SNPs with a numerical genotype of 1 were coded as having the allele of the nNIL introgression donor. This resulted in 19 possible alleles at any given SNP. SNPs underlying each QTL were concatenated. For each QTL, a linear model was fit with the lm function in R version 3.3.1 (16), with the QTL-associated trait as the response and allele as a fixed effect.

Table 4 Numbers of BC_5F_4 nNILs phenotyped and genotyped per introgression donor

		Number of BC_5F_4 nNILs	
Subpop donor	Donor origin	Phenotyped	Genotyped
108	Tropical		96
50	Tropical		41
64	Tropical		41
56	Tropical		46
63	Tropical		46
95	Tropical		74
77	Tropical		57
57	Tropical		49
63	Tropical		46
65	Tropical		31
73	Non-stiff stalk		22
65	Non-stiff stalk		29
92	Mixed		74
65	Tropical		48
69	Tropical		41
72	Non-stiff stalk		60
62	Mixed		47
73	Tropical		59

BC_5F_4, backcrossed five generations followed by four generations of self-pollination; nNIL, nested near-isogenic line; Subpop, subpopulation.

ACKNOWLEDGEMENTS

This work was partially funded by NSF PGRP grant #1127076. The authors thank Greg Marshall, Judith Kolkman and Molly Towne for technical and logistic support. The authors also thank Denise Costich and the CIMMYT Germplasm Bank for originally providing the nNIL seed. The authors also acknowledge the staff at the Central Crops Research Station in Clayton, NC for field trial management, the Cornell University Computational Biology Service Unit for providing computational resources, and Detlef Weigel and Hernán Burbano at the Max Planck Institute for Developmental Biology for desk space.

AUTHOR CONTRIBUTIONS

Conceptualization: LM, PJBK, RJN; methodology: LM, PJBK, RJN; software: KLS, LM; validation: ACR, LM, WCS; formal analysis: LM; investigation: ACR, LM, LOLZ, PJBK, QY, SMS, YH; resources: ACR, ER, LM, PJBK, WET; data curation: RJN; writing – original draft preparation: LM; writing – review and editing: LM, PJBK, RJN; visualization: LM; supervision: PJBK, RJN; project administration: ACR, LM, PJBK, RJN; funding acquisition: PJBK, RJN.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Fastq files containing raw sequencing data are available in the NCBI Sequence Read Archive under the BioProject accession PRJNA591863. Seed of the nNILs characterized here are available from the CIMMYT Germplasm Bank.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Additional file 1. Tab-delimited text file with the raw phenotypes and experimental plot information. For each field plot, the columns contain data on the environment ID (Env), location (Loc), year (Year), replication (Rep), block number (Block), plot number (Plot), genotype (Geno), days to anthesis (DTA), ear height in cm (EHT), plant height in cm (PHT), first (GLS1) and second (GLS2) disease scores for GLS, and the mean of the two scores (GLSmean) on a 1–9 scale (1 = most resistant, 9 = most susceptible), first (NLB1) and second (NLB2) disease scores for NLB and the mean of the two scores (NLBmean) based on percent DLA (0–100%), and first (SLB1) and second (SLB2) disease scores for SLB and the mean of the two scores (SLBmean) on a 1–9 scale (1 = most resistant, 9 = most susceptible).

Additional file 2. Tab-delimited text file with NIL genotype BLUPs for DTA, EHT and PHT, GLS, NLB and SLB.

Additional file 3. Compressed tab-delimited text file containing 241 455 ABH SNPs across 896 nNILs in TASSEL numerical genotype format in AGPv4 coordinates.

Additional file 4. Tab-delimited text file containing the physical (bp, AGPv4 coordinates) and genetic (cM, based on NAM genetic map) positions of 2638 introgressions present in 896 nNILs.

Additional file 5. TIF image file showing plots of GWA results for DTA, EHT and PHT, and resistance to GLS, NLB and SLB. Quantile-quantile (QQ) plots show the expected versus the actual −log10(*P*-values) from GWA on each trait. GWA Manhattan plots are plotted by chromosome and trait, with significantly associated SNPs (FDR-adjusted $P < 0.05$) highlighted in red, SNP physical positions in AGPv4 coordinates (Mbp) on the *x*-axes, and −log10 (*P*-values) on the *y*-axes.

Additional file 6. Tab-delimited text file with statistics and FDR-adjusted *P*-values from GWA on DTA, EHT and PHT, GLS, NLB and SLB.

Additional file 7. Executable jar file of the TASSEL 5 program modified for the GBS fastq files used here.

Additional file 8. HDF file containing 564 764 FILLIN-imputed SNPs across 896 nNILs in AGPv4 coordinates.

REFERENCES

Aboyoun, P., Carlson, M., Lawrence, M. *et al.* (2013) Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**(8), e1003118. https://doi.org/10.1371/journal.pcbi.1003118

Benson, J.M., Poland, J.A., Benson, B.M., Stromberg, E.L. and Nelson, R.J. (2015) Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLoS Genet.* **11**, e1005045. https://doi.org/10.1371/journal.pgen.1005045

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633–2635.

Buckler, E.S., Holland, J.B., Bradbury, P.J. *et al.* (2009) The genetic architecture of maize flowering time. *Science*, **325**, 714–718. https://doi.org/10.1126/science.1174276

Chia, J.-M., Song, C., Bradbury, P.J. *et al.* (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* **44**, 803–807. https://doi.org/10.1038/ng.2313

Chung, C., Longfellow, J.M., Walsh, E.K. *et al.* (2010) Resistance loci affecting distinct stages of fungal pathogenesis: use of introgression lines for QTL mapping and characterization in the maize - Setosphaeria turcica

pathosystem. *BMC Plant Biol.* **10**, 103. https://doi.org/10.1186/1471-2229-10-103

Eichten, S.R., Foerster, J.M., de Leon, N. *et al.* (2011) B73-Mo17 near-isogenic lines demonstrate dispersed structural variation in maize. *Plant Physiol.* **156**(4), 1679–1690. https://doi.org/10.1104/pp.111.174748

Elshire, R.J., Glaubitz, J.C., Sun, Q. *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379. https://doi.org/10.1371/journal.pone.0019379

Gandhi, S., Lauchner, M., Williams, T.*et al.* (2008) High throughput allele discovery and incorporation in elite maize germplasm. In *Proceedings for the 5th International Crop Science Congress*.

Glaubitz, J.C., Casstevens, T.M., Lu, F. *et al.* (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, **9**, e90346. https://doi.org/10.1371/journal.pone.0090346

Hannon, G.J. (2010) *FASTX-Toolkit*. Available at: http://hannonlab.cshl.edu/fastx_toolkit/

Holland, J.B., Nyquist, W.E. and Cervantes-Martinez, C.T. (2003) Estimating and interpreting heritability for plant breeding: an update. *Plant Breed. Rev.* **22**, 9–112.

Jiao, Y., Peluso, P., Shi, J. *et al.* (2017) Improved maize reference genome with. *Nature*, **546**(7659), 524–527. https://doi.org/10.1038/nature22971

Kolkman, J.M., Strable, J., Harline, K., Kroon, D.E., Wiesner-Hanks, T., Bradbury, P.J. and Nelson, R.J. (2019) Maize nested introgression library provides evidence for the involvement of liguleless1 in resistance to northern leaf blight, bioRxiv. https://doi.org/10.1101/818518

Kump, K.L., Bradbury, P.J., Wisser, R.J. *et al.* (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* **43**, 163–168. https://doi.org/10.1038/ng.747

Law, M., Childs, K.L., Campbell, M.S. *et al.* (2015) Automated update, revision, and quality control of the maize genome annotations USING MAKER-P improves the B73 RefGen_v3 gene models and identifies new genes. *Plant Physiol.* **167**(January), 25–39. https://doi.org/10.1104/pp.114.245027

Leinonen, R., Sugawara, H. and Shumway, M. (2011) The sequence read archive. *Nucleic Acids Res.* **39**(November 2010), 2010–2012. https://doi.org/10.1093/nar/gkq1019

Lennon, J.R., Krakowsky, M., Goodman, M., Flint-Garcia, S. and Balint-Kurti, P.J. (2016) Identification of alleles conferring resistance to gray leaf spot in maize derived from its wild progenitor species teosinte. *Crop Sci.* **56**(1), 209–218. https://doi.org/10.2135/cropsci2014.07.0468

Lennon, J.R., Krakowsky, M., Goodman, M., Flint-Garcia, S. and Balint-Kurti, P.J. (2017) Identification of teosinte alleles for resistance to southern leaf blight in near isogenic maize lines. *Crop Sci.* **57**(4), 1973–1983. https://doi.org/10.2135/cropsci2016.12.0979

Li, C., Li, Y., Bradbury, P.J. *et al.* (2015) Construction of high-quality recombination maps with low-coverage genomic sequencing for joint linkage analysis in maize. *BMC Biol.* **13**, 78. https://doi.org/10.1186/s12915-015-0187-4

Liang, Z. and Schnable, J.C. (2016) RNA-Seq based analysis of population structure within the maize inbred B73. *PLoS ONE*, **11**(6), e0157942. https://doi.org/10.1371/journal.pone.0157942

Martins, L.B., Rucker, E., Thomason, W., Wisser, R.J. and Holland, J.B. (2019) 'Validation and Characterization of maize multiple disease. *Resistance QTL'* **9**(September), 2905–2912. https://doi.org/10.1534/g3.119.400195

Mideros, S.X., Warburton, M.L., Jamann, T.M., Windham, G.L., Williams, W.P. and Nelson, R.J. (2014) Quantitative trait loci influencing mycotoxin contamination of maize: analysis by linkage mapping, characterization of near-isogenic lines, and meta-analysis. *Crop Sci.* **54**(1), 127. https://doi.org/10.2135/cropsci2013.04.0249

Olukolu, B.A., Wang, G.F., Vontimitta, V. *et al.* (2014) A genome-wide association study of the maize hypersensitive defense response identifies genes that cluster in related pathways. *PLoS Genet.* **10**, e1004562. https://doi.org/10.1371/journal.pgen.1004562

Pea, G., Paulstephenraj, P., Canè, M.A. *et al.* (2009) Recombinant near-isogenic lines: a resource for the mendelization of heterotic QTL in maize. *Mol. Genet. Genom.* **281**(4), 447–457. https://doi.org/10.1007/s00438-008-0422-6

Peiffer, J.A., Romay, M.C., Gore, M.A. *et al.* (2014) The genetic architecture of maize height. *Genetics*, **196**, 1337–1356. https://doi.org/10.1534/genetics.113.159152

Poland, J.A., Bradbury, P.J., Buckler, E.S. and Nelson, R.J. (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl Acad. Sci.* **108**, 6893–6898. https://doi.org/10.1073/pnas.1010894108

R Core Team. (2019) R: A Language and Environment for Statistical Computing. Vienna, Austria. Available at: www.R-project.org

Reuscher, S. and Furuta, T. (2016) *Easy visualization of ABH genotypes*. Available at: http://github.com/StefanReuscher/ABHgenotypeR

Rodgers-Melnick, E., Bradbury, P.J., Elshire, R.J. *et al.* (2015) Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl Acad. Sci.* USA, **112**(12), 201413864. https://doi.org/10.1073/pnas.1413864112

Romay, M.C., Millard, M.J., Glaubitz, J.C. *et al.* (2013) Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* **14**, R55. https://doi.org/10.1186/gb-2013-14-6-r55

SAS Institute, Inc. (2011) SAS System for Windows. Cary, NC.

Schnable, P.S., Ware, D., Fulton, R.S. *et al.* (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, **326**(5956), 1112–1115.

Sermons, S.M. and Balint-Kurti, P.J. (2018) Large scale field inoculation and scoring of maize southern leaf blight and other maize foliar fungal diseases. *Bio-protocol*, **8**(5), e2745. https://doi.org/10.21769/BioProtoc.2745

Swarts, K., Li, H., Romero Navarro, J.A. *et al.* (2014) Novel methods to optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants. *The Plant Genome*, **7**, 1–7. https://doi.org/10.3835/plantgenome2014.05.0023

Szalma, S.J., Hostert, B.M., LeDeaux, J.R., Stuber, C.W. and Holland, J.B. (2007) QTL mapping with near-isogenic lines in maize. *Theor. Appl. Genet.* **114**(7), 1211–1228. https://doi.org/10.1007/s00122-007-0512-6

Wisser, R.J., Kolkman, J.M., Patzoldt, M.E., Holland, J.B., Yu, J. and Krakowsky, M. (2011) Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a GST gene. *Proc. Natl Acad. Sci. USA*, **108**, 7339–7344. https://doi.org/10.1073/pnas.1011739108

Xiao, Y., Tong, H., Yang, X. *et al.* (2016) Genome-wide dissection of the maize ear genetic architecture using multiple populations. *New Phytol.* **210**(3), 1095–1106. https://doi.org/10.1111/nph.13814

Xiao, Y., Liu, H., Wu, L., Warburton, M. and Yan, J. (2017) Genome-wide association studies in maize: praise and stargaze. *Molec. Plant* **10**(3), 359–374. https://doi.org/10.1016/j.molp.2016.12.008

Yu, J., Holland, J.B., McMullen, M.D. and Buckler, E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics*, **178**(1), 539–551. https://doi.org/10.1534/genetics.107.074245

Zhao, W. (2006) Panzea: a database and resource for molecular and functional diversity in the maize genome. *Nucleic Acids Res.* **34**, D752–D757. https://doi.org/10.1093/nar/gkj011