Unraveling Genomic Complexity at a Quantitative Disease Resistance Locus in Maize

Tiffany M. Jamann*, Jesse A. Poland§, Judith M. Kolkman†, Laurie G. Smith‡, Rebecca J. Nelson *

*Dept. of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, NY 14853
§ USDA-ARS, Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66502, Dept. of Agronomy, Kansas State University, Manhattan, KS 66506
† Dept. of Plant Breeding, Cornell University, Ithaca, NY 14853
‡ Section of Cell and Developmental Biology, Division of Biological Sciences, University of California at San Diego, La Jolla, CA 92093
Short title: Structural variation and pan1 in disease resistance

Key words: quantitative disease resistance, northern leaf blight, *Setosphaeria turcica*, multiple disease resistance, Stewart’s wilt

Corresponding author:
Rebecca Nelson
303G Plant Science Building
236 Tower Road
Ithaca, NY, 14853
607-254-7475
rjn7@cornell.edu
ABSTRACT

Multiple disease resistance has important implications for plant fitness, given the selection pressure that many pathogens exert directly on natural plant populations and indirectly via variety improvement programs on crop plants. Evidence of a locus conditioning resistance to multiple pathogens was found in bin 1.06 of the maize genome with the allele from inbred line ‘Tx303’ conditioning quantitative resistance to northern leaf blight (NLB) and qualitative resistance to Stewart’s wilt. To dissect the genetic basis of resistance in this region and to refine candidate gene hypotheses, we mapped resistance to the two diseases. Both resistance phenotypes were localized to overlapping regions, with the Stewart’s wilt interval refined to a 95.9-kb segment containing three genes, and the NLB interval to a 3.60-Mb segment containing 117 genes. Regions of the introgression showed little to no recombination, suggesting structural differences between the inbred lines ‘Tx303’ and ‘B73’, the parents of the fine-mapping population. We examined copy number variation across the region using next-generation sequencing data and found large variation in read depth in ‘Tx303’ across the region relative to the reference genome of inbred line ‘B73’. In the fine-mapping region, association mapping for NLB implicated candidate genes, including a putative zinc finger and pan1. We tested mutant alleles and found that pan1 is a susceptibility gene for NLB and Stewart’s wilt. Our data strongly suggest that structural variation plays an important role in resistance conditioned by this region, and pan1, a gene conditioning susceptibility for NLB, may underlie the QTL.
The genes and loci that influence host-pathogen interactions vary in allele effects, specificsities, and linkage relationships. While disease resistance can be conditioned by single genes with large effect (Bent 1996; Jones and Dangl 2006), the emerging model of resistance for many plant diseases is complex in nature, with many genes and loci functioning in concert and each contributing a small proportion of the total phenotypic variation (Kump et al. 2011; Poland et al. 2011; Cook et al. 2012b). Each locus has a unique profile, with some loci contributing broad-spectrum protection against diverse pathogen species and strains. Investigating these intricacies offers the opportunity to understand the diverse ways in which plants defend themselves against microbial assault.

Correlated responses to multiple diseases have been observed in various germplasm panels, implying that there are loci and genes that condition broad-spectrum resistance (Rossi et al. 2006; Gurung et al. 2009; Wisser et al. 2011). At the chromosomal segment level, disease and insect resistance loci co-localize in a non-random fashion (McMullen and Simcox 1995; Williams 2003; Wisser et al. 2005) and loci have been identified that confer resistance to diverse pathogen isolates and taxa (Zwonitzer et al. 2010; Chung et al. 2011; Belcher et al. 2012). There is evidence to suggest that gene clusters can confer resistance to more than one disease. A cluster of germin-like proteins confers resistance to rice blast and sheath blight of rice (Manosalva et al. 2009). Similarly, resistance gene homologs, which are known to co-localize with broad-spectrum disease resistance loci, can cluster in the genome and contribute a diversity of specificsities (Lopex et al. 2003; Ramalingam et al. 2003). Pleiotropy remains uncommon in maize, and correlated responses may be due to linkage or population structure (Wallace et al. 2014),
although in some cases, individual genes have been shown to condition MDR. For example, the putative ABC transporter *Lr34* of wheat provides protection against leaf rust, stripe rust, and powdery mildew (Krattinger et al. 2009). Pattern recognition receptors are able to detect molecular patterns from diverse organisms to confer disease resistance (Zipfel and Rathjen 2008).

While in some cases single genes or alleles common across diverse germplasm confer disease resistance, increasingly, the role of structural variation in plants is being explored and its effects on phenotypic variation recognized (Spinger et al. 2009; Chia et al. 2012; McHale et al. 2012). As quantitative trait loci (QTL) are subjected to fine-mapping, some loci fractionate into many QTL, each conditioned by one or more genes (Steinmetz et al 2002; Studer and Doebley 2011; Johnson et al. 2012). In some cases, the allele effect conditioned by each QTL is small enough that the individual locus cannot be identified in isolation (Poland et al. 2011; Buckler et al. 2009). In other cases, single resistance loci, such as *Rhg1*, are conditioned by multiple genes present in varying copy numbers in different lines (Cook et al. 2012a; Maron et al. 2013). Whole-genome studies have in fact suggested that structural variation is generally associated with disease resistance: structural variation in plants co-localizes with resistance nucleotide-binding proteins, receptor-like proteins, and disease resistance QTL (Lai et al. 2012; McHale et al. 2012; Xu et al. 2012).

The conventional approach of genetic isolation and transgenic complementation remains the gold standard for demonstrating the function of a gene. This approach, however, is proving inadequate for dealing with the complexity underlying some loci, particularly for structural variation. Strong evidence for the importance of copy number variation in explaining trait variation (Cook et al. 2012a; Maron et al. 2013) and the
emerging model of plant defense with many loci each contributing a small effect combine to challenge this paradigm (Kump et al. 2011; Poland et al. 2011; Cook et al. 2012b). There is need for a new approach that can take advantage of whole genome analyses, address presence/absence variation, and examine loci with small effects. This study represents such an approach and provides insights into a genetically complex locus affecting diverse traits.

In maize, chromosomal bin 1.06 has been identified as a key locus for stabilizing yield under adverse conditions, including both biotic and abiotic stress (Landi et al. 2002; Tuberosa et al. 2002; Landi et al. 2010). In additional to plant architectural traits and yield under abiotic stress, resistance to many diseases has been localized to bin 1.06, including northern leaf blight (NLB), Stewart’s wilt, southern leaf blight (SLB), common rust, grey leaf spot (GLS), and ear and stalk rot caused by multiple fungi (Wisser et al. 2006; Chung et al 2010b; Zwonitzer et al. 2010). In a QTL study of the recombinant inbred line (RIL) population Ki14 x B73 evaluated for three foliar fungal diseases, NLB, GLS, and SLB, a 33 Mb region spanning bins 1.05 and 1.06 was the only locus identified that conferred resistance to all three diseases (Zwonitzer et al. 2010). A number of QTL studies for NLB resistance in maize have identified QTL at bin 1.06, ranging in physical size from three to 30 Mb (Freymark et al. 1993; Welz et al. 1999; Wisser et al. 2006; Chung et al. 2010b; Van Esbroeck et al. 2010; Chung et al. 2011; Poland et al. 2011). Additionally, bin 1.06 harbors the dominant Stewart’s wilt resistance gene Sw1 (Ming et al. 1999).

Both NLB, caused by the fungus *Setosphaeria turcica*, and Stewart’s wilt, caused by the bacterium *Pantoea stewartii*, are foliar, hemibiotrophic diseases important to maize production. Both pathogens spread through the vascular tissue, causing wilted lesions by
plugging xylem vessels (Jennings and Ullstrup 1957; Roper et al. 2011). The importance of genes localized to maize bin 1.06 in resistance to both NLB and Stewart’s wilt has been described in multiple mapping populations. Using a population of Tx303 x B73 introgression lines (Szalma et al. 2007), Chung et al. (2010) showed that the NLB resistance QTL at 1.06 protects against fungal penetration.

To explore the genomic complexity of this important region, we constructed high resolution mapping populations at this locus and evaluated NLB and Stewart’s wilt resistance using a set of Tx303 x B73 near-isogenic lines (NILs) (Szalma et al. 2007; Chung et al. 2010). Fine-mapping allowed us to dissect the linkage relationship between the major-effect Stewart’s wilt QTL and the minor-effect NLB QTL and to identify candidate genes. Using association mapping, we further refined the list of candidate genes for NLB resistance and using mutants confirmed a role for the receptor-like kinase, pan1, in plant defense. Furthermore, multiple lines of evidence indicated a lack of genomic stability at the region, including reduced recombination across portions of the fine-mapping region in the NIL population and indicators of copy number variation.

MATERIALS AND METHODS

**Plant Materials:** NILs used for fine-mapping were derived from the TBBC3 (Tx303 x B73 Backcross 3) population, a set of chromosomal segment substitution lines with Tx303 introgressions in a B73 background (Szalma et al. 2007; Chung et al. 2010). Chung et al. (2010) identified families TBBC3-38 and TBBC3-39, both with introgressions in 1.06, as significantly more resistant than B73. Selected families developed from these lines were chosen for fine-mapping: TBBC3-38_19E, TBBC3-38_15G, and TBBC3-38_17A (Chung et al.
The details of population development and evaluation are shown in Figure S1.

Briefly, a population was developed by crossing TBBC3-38_19E to B73. In the F2 generation, 435 individuals were screened for recombinants with flanking markers snp_01_0042 (180,394,924 AGP_V2) and snp_01_0005 (195,557,990 bp AGP_V2). A total of 113 recombinant plants were identified, but seed was available from only 100 plants. Seed from 15 heterozygous F2 individuals was advanced to the F3 generation to screen for additional recombinants. A population of 4,080 F3 seeds was planted and 2,929 plants were screened with flanking markers snp_01_0042 (180,394,924 AGP_V2) and snp_01_0005 (195,557,990 bp AGP_V2). Positions are based on B73 genome sequence release AGP_V2 (Schnable et al. 2009). Individual F3 plants (n=874) were identified as recombinants from the F3 population. Recombinant plants were self-pollinated and homozygous recombinants identified. Homozygous recombinants were increased and evaluated for disease resistance. Subsequently, a population of 1,546 F3 plants was screened from snp_01_0059 (184,633,349 bp AGP_V2) to snp_01_0083 (189,352,206 bp AGP_V2), yielding an additional 156 recombinants. Homozygous plants were identified, self-pollinated and screened for NLB and Stewart’s wilt (Figure S1).

**Disease trials**

**Northern leaf blight:** NLB trials were carried out at the Cornell University Robert Musgrave Research Farm in Aurora, NY. In the fine-mapping populations, 194, 80, and 146 homozygous recombinants were screened for NLB in 2010, 2011, and 2012, respectively. Lines carrying mutations in the genes *pan1* and *pan2* were tested in Aurora, NY in 2011, 2012, and 2013, and in Ithaca, NY in 2011. Plants were inoculated with *S. turcica* isolate
StNY001 (race 1) using previously described inoculation procedures (Chung et al. 2010a).

Briefly, cultures of the fungus were grown on lactose casein agar for three to four weeks prior to inoculation or sorghum culturing. A spore suspension was prepared by flooding the cultures with 5 mL sterilized distilled water and conidia were dislodged using a glass rod. The spore suspension was filtered through two layers of cheesecloth and adjusted to a concentration of 4x10³ spores per ml using a haemocytometer. The spore suspension contained a final concentration of 0.02% Tween 20. Sorghum seed cultures were prepared by soaking 900 mL of sorghum grains in 600 mL distilled water overnight in a one-gallon clear milk jug and autoclaved twice for 25 minutes. One mL of unfiltered spore suspension was then introduced to each jug, which was then cultured at room temperature for about three weeks before field inoculations were conducted. Jugs were shaken daily to prevent caking and provide uniform infestation. For field inoculations, 0.50 mL spore suspension and ~1.25 mL sorghum grains colonized by *S. turcica* were placed into the whorl of each plant at the five- to six-leaf stage.

Diseased leaf area (DLA) was rated on a per row basis three times after flowering at an interval of seven to ten days using a percentage scale of 0-100 with increments of one, where 0 indicates a plant with no disease and 100 indicates a completely diseased plant. Area under the disease progress curve (AUDPC) was calculated as described previously (Chung et al. 2010a).

**Stewart’s wilt:** Stewart's wilt trials were conducted at the Musgrave Research Farm in Aurora, NY in 2010, 2011, and 2012 for fine-mapping and 2012 and 2013 for *pan1* and *pan2* mutants. Results from 2013 were not included in the analysis due to flooding. Plants were inoculated with *Pantoea stewartii* strain PsNY003, originally collected in NY in 1991,
at the five- to six-leaf stage, with inoculum prepared and a modified pinprick method used for inoculations as previously described (Chung et al. 2010a). DLA was rated on a per row basis at two to four weeks after inoculation on a row basis using a percentage scale of 0 to 100, with 0 being no disease and 100 being completely diseased.

**SNP Marker Development:** The maize diversity project database http://www.panzea.org (Canaran et al. 2008) was used to locate polymorphisms between the two inbred lines using a number of datasets including markers from the NAM genetic map (McMullen et al. 2009), HapMapV1 (Gore et al. 2009), and HapMapV2 (Chia et al. 2012). Single nucleotide polymorphisms (SNPs) polymorphic between B73 and Tx303 were chosen for this study. SNP markers used for the fine-mapping study are shown in Table S1. In addition, an Illumina MaizeSNP50 Beadchip assay (Illumina) was conducted on pooled DNA from families TBBC3-38_05F and TBBC3-38_19E at the David H. Murdock Research Institute, (Kannapolis, NC, USA) which included genotypic information for 52,686 SNPs.

**DNA Extractions:** Both Sigma ExNAmp (Sigma-Aldrich) and CTAB DNA extractions were used. ExNAmp DNA extractions were used to identify recombinants and homozygous recombinants. CTAB extractions were used for genotype confirmation and breakpoint analysis. To perform ExNAmp extractions, one mm² of plant tissue was collected in a 0.2-mL PCR tube and the tubes were placed on ice. Eight μL of extraction buffer was added to each tube, and tubes were incubated at 95°C for ten minutes. Following the incubation, 8 μL of dilution buffer was added. The resulting DNA was diluted 1:100 with water for KASPar (LGC Genomics, Hoddesdon, Herfordshire, UK) genotyping. CTAB DNA extractions were
performed using about 0.1 mg of fresh tissue as described previously (Doyle et al. 1987; Chung et al. 2010a).

Allele-specific PCR: When using ExN Amp-extracted DNA, 10 μL of 1:100 diluted DNA was dried down in a 384-well KASPar plate (LGC Genomics) and a 4-μL reaction performed. For reactions using CTAB extracted DNA, DNA was quantified and approximately 5 ng/μL of DNA was used per reaction. Reaction conditions were as follows: 1x KASPar reaction mix (LGC Genomics), 0.4 mM MgCl₂, 0.41 mM common reverse primer, and 0.165 μM of each allele-specific primer. Standard oligonucleotides were obtained from IDT (Coralville, IA, USA). PCR thermocycling parameters were as follows: 94°C for 15 minutes, 20 cycles of 94°C for 10 seconds, 57°C for 5 seconds, and 72°C for 10 seconds, followed by 26 cycles of 94°C for 10 seconds, 57°C for 20 seconds, and 72°C for 40 seconds. Results were read using an Applied Biosystems 7900 HT (Life Technologies) and analyzed using SDS v2.1 (Life Technologies).

Experimental design and statistical analysis: A randomized incomplete block design was used for all fine-mapping field experiments, with three replications for NLB trials and two for Stewart’s wilt trials. NILs carrying B73 and Tx303 alleles across the region were included in each block as check lines. Two rows were planted around the edge of the experiment to reduce border effects. For the breakpoint analysis, best linear unbiased predictors (BLUPs) were calculated using the ‘lmer’ command in the lme4 package in R version 2.14 (R Development Core Team 2013) where line, year, replication within year, and block nested within replication were fitted as random factors in a mixed-effects model for NLB. Similarly, Stewart’s wilt BLUPs were calculated using the ‘lmer’ command including line and year as random effects. Fine-mapping statistical analyses were
completed in R version 2.14 (R Development Core Team 2013) using R/qtl (Broman et al. 2003). First, individuals with fewer than 11 genotyped markers and markers with fewer than 250 individuals genotyped were removed from the analysis. A genetic map was then constructed using the Kosambi mapping function (Kosambi 1943). Single marker regression was conducted (Paterson et al. 1990; Kump et al. 2010) using the ‘scanone’ function in R/qtl. Confidence intervals were calculated on the basis of a 95% Bayes credible interval using the function ‘bayesint’ in R/qtl (Broman et al. 2003). Genotypes and phenotypes can be found in File S1.

**Recombination rate diversity:** Recombination rates were calculated using genotyping-by-sequencing SNPs on the NAM sub-populations, using phased and fully imputed at 1-cM resolution genotypes (AllZea_GBSv2.3) (http://panzea.org/lit/data_sets.html). Genetic maps were constructed using R/qtl with the “est.map” function in R (Broman et al. 2003; R Development Core Team 2013).

**Read depth variation:** Illumina single and paired-end sequencing reads averaging ~4x coverage of the genome generated as part of the HapMapV1 and V2 projects (NCBI accession SRA051245) were used for the read depth variation analysis (Chia et al. 2012). Aligned sequencing reads for B73 and Tx303 (Chia et al. 2012) were downloaded from iPlant (Goff et al. 2011). SAMtools was used to count the number of reads at each nucleotide location (Li et al. 2009). A Perl script was written to divide the interval into 11 bins of equal size (327,181 bp) and to compile the number of reads per bin.

**Candidate gene identification:** All B73 RefGenV2 (AGP_V2) filtered genes between the two flanking markers snp_01_0047 at 185,737,089 bp (AGP_V2) and snp_01_0082 at 189,336,643 bp (AGP_V2) of the narrowed NLB fine-mapping interval were considered as...
candidate genes. SNPs from genome-wide nested association mapping with a bootstrap posterior probability (BPP) >0.01 that fell within the narrowed fine-mapping interval were considered further as candidates (Poland et al. 2011; Chia et al. 2012).

**Association analysis of the 282-line maize diversity panel:** Association mapping was conducted for the NLB fine-mapping interval. BLUPs that included design factors, flowering time, and population structure were used for association analysis (Wisser et al. 2011). A mixed linear model (MLM) was implemented using TASSEL v4 (Bradbury et al. 2007). The markers assayed included 47,445 Illumina MaizeSNP50 SNPs (Cook et al. 2012b) and 425,035 genotyping-by-sequencing SNPs (Romay et al. 2013), filtered to remove sites with >20% missing data (Olukolu et al. 2013). The kinship (K) matrix was constructed in TASSEL using a 5,000-SNP subset of the Illumina MaizeSNP50 dataset that had no missing data (Olukolu et al. 2013). Q values were calculated using qvalue package (Storey 2002). Associations with a false discovery rate of FDR<0.15 were noted.

**Mutant analysis:** Mutants in the pan1 and pan2 genes were evaluated for NLB and Stewart’s wilt reaction. Two mutant alleles of pan1 were evaluated in a B73 background: one mutant generated by ethyl methanesulfonate mutagenesis (pan1-EMS) and one line with a Mutator1 (Mu1) transposon in pan1 (Gallagher and Smith 2000; Cartwright et al. 2009). Two ethyl methanesulfonate alleles of pan2 were evaluated in a B73 background: pan2-0 and pan2-3 (Cartwright et al. 2009). All pan1 and pan2 mutant alleles except pan2-0 can be considered null alleles based on the nature of the mutations and analysis of PAN protein accumulation in mutants, whereas pan2-0 is a missense allele that may encode a partially functional protein (Cartwright et al. 2009). pan1 mutant lines were backcrossed to B73 3-5X and pan2 mutants were backcrossed to B73 3-7X. For mutant analysis a complete
block design was used, with five replications per location for *pan1* and *pan2* mutants. NLB results were analyzed with a mixed linear model in JMP 9.0 (SAS, Cary, NC, USA), with genotype as a fixed effect and replication nested within environment, and environment as random effects. Stewart’s wilt results were analyzed with genotype as a fixed effect and replication as a random effect.

**RT-PCR analysis**: Tissue for RNA extraction was collected from mature leaf tissue of (BC4F3)BC1F5 plants carrying either the B73 (*qNLB1.06B73*) or Tx303 allele (*qNLB1.06Tx303*) at the *qNLB1.06* locus during the summer of 2011 and 2012. RNA was extracted using an RNeasy kit (Qiagen) and first cDNA was prepared from this RNA using a RETROscript First Strand cDNA Synthesis kit (Life Technologies). PCR was carried out using the following primers for amplification of *pan1* (5’-TCGGGATGGAGCTGGAGGAG-3’ and 5’-TGGACAGACGACGGACCC-3’) and actin as a control (5’-TCAGCAGGTCTTCTTGTCTT-3’ and 5’-TCCTTCATATTTTCTTCCGTT-3’) with Q5 Hot Start Taq Polymerase (New England Biolabs). *pan1* and actin PCR products were quantified from gel images using NIH ImageJ v. 1.47g.

**RESULTS**

**Identification of MDR in bin 1.06**: A number of QTL studies have localized resistance to NLB to maize bin 1.06, with varying resolution (Table S2). These studies have consistently implicated the region spanning from 180 to 205 Mb. Lines carrying a Tx303 introgression at this interval were found to be associated with resistance to NLB and Stewart’s wilt (families TBBC3-38 and TBBC-39 of the TBBC3 population) (Chung et al. 2010b). Based on genetic background and seed availability, TBBC3-38_19E, TBBC3-38_15A,
and TBBC3-38_17G were selected for fine-mapping. The Tx303 introgression in TBBC3-38
spans from ss196428597 (172,877,033 bp) to ss196518155 (196,244,799 bp) (Figure 1).
An interval of 15.16 Mb spanning from snp_01_0042 (180,394,890 bp AGP_V2) to
snp_01_0005 (195,557,990 bp AGP_V2) was targeted for fine-mapping, based on NIL
introgression locations and previous QTL mapping studies (Table S2 and Figure 1).

Fine-mapping of MDR at 1.06: Two markers flanking the qNLB1.06 region,
snp_01_0042 and snp_01_0005, were used to screen 435 F2 and 4,475 F3 plants. We
identified a total of 1,130 recombinants spanning the 15 Mb interval of interest. Plants
were self-pollinated and progeny were assayed for homozygous (fixed) recombinants.
Recombinant plants were selected for phenotypic evaluation based on breakpoint analysis
and seed availability. Fixed recombinants were screened for disease phenotype in a
randomized incomplete block design with two control NILs, one carrying the B73 allele and
one carrying the Tx303 allele in the region of interest. Fixed recombinants were screened
for NLB (n=194, 80, and 146 in 2010, 2011, and 2012, respectively) and Stewart’s wilt
(n=60, 78, and 140 in 2010, 2011 and 2012, respectively) (Figure 1). An additional 17 SNP
markers were assayed on the population to determine the physical position of breakpoints
(Table S1). Five were removed from the breakpoint analysis because of missing genotypes
or a lack of recombination with neighboring markers. The order of the physical map
matched the order of the genetic map. Based on these data, qNLB1.06 was narrowed to a
3.60-Mb region flanked by snp_01_0047 (185,737,089 bp AGP_V2) and snp_01_0082
(189,336,643 bp AGP_V2) (Figure 1). qSw1.06 was narrowed to a 95.9-kb interval flanked
by snp_01_0137 (187,245,104 bp AGP_V2) and snp_01_0139 (187,341,010 bp AGP_V2)
While confidence intervals differed for the diseases, breakpoint analyses for both showed similar profiles.

**Genomic integrity:** We observed a low frequency of recombination across part of the fine-mapping region. To determine whether this was an anomaly only found in the NIL fine-mapping population, we examined the recombination rates across the RIL populations that comprise the NAM population by examining genetic distances in NAM sub-populations, as shown in Figure 2. We found a depressed recombination rate across portions of the target interval in the Tx303 x B73 RIL population, confirming a low rate of recombination found in the current study between B73 and Tx303 in this region. Furthermore, most NAM founder lines showed low recombination rates across parts of the region. Hp301, however, showed slightly higher levels of recombination, suggesting structural similarity between B73 and Hp301. Other lines, such as Ms71 and M37W, showed repressed recombination in other regions of the interval.

Because low recombination rates are hypothesized to be due to structural variation such as inversions, indels, transposable elements, or presence/absence variation (McMullen et al. 2009), we examined Illumina reads from the HapMapV2 project (Chia et al. 2012) to assess read depth variation as a proxy for structural variation. We found variation in the number of Tx303 reads that mapped to the B73 reference sequence, while little variation was observed in the number of B73 reads that mapped to the reference sequence (Figure 3). The region with little to no recombination in the NAM sub-populations showed a reduced number of reads in Tx303. Conversely, the region with significant CNV NAM GWAS associations at 187 Mb had an elevated number of reads that mapped to the reference genome in this location, suggesting possible duplications and genome expansion.
Candidate genes underlying qSw1.06Tx303: The Stewart’s wilt fine-mapping region was narrowed to a 95.9-kb interval that contains three genes in the B73 reference: a putative zinc finger (GRMZM2G445684) and two uncharacterized genes (GRMZM2G445676, and AC213857.4_FG001). The two uncharacterized genes have no known homologs and no annotated domains. Resistance at this locus is conferred by Tx303, and it is therefore plausible that the resistance gene is absent from the B73 reference.

Candidate genes underlying qNLB1.06: The fine-mapping region falling within the Bayes 95% confidence interval for NLB resistance, 185.7 Mb to 189.3 Mb of maize chromosome 1, comprises 117 annotated coding genes, 30 pseudogenes, and 39 transposable elements in the B73 genome sequence (Table S3). A number of genes within the narrowed fine-mapping interval are credible candidates based on the involvement of those gene classes in plant defense as documented in the scientific literature. These include three putative leucine rich repeat-encoding genes, three putative protein kinases, two putative wall-associated receptor kinases, and one putative lipoxygenase sharing homology with Arabidopsis thaliana LOX2. Furthermore, genome-wide nested association mapping conducted by Chia et al. (2012) identified three significant associations clustered within the 40-kb region spanning 187.23 to 187.27 Mb AGP_V1. The most significant hit was an intergenic 10 kb read depth variation with a bootstrap posterior probability (BPP) of 44, which was 30 kb upstream of a putative serine-threonine protein kinase and 129 kb downstream of a putative zinc-finger encoding gene (GRMZM2G441903). Two additional polymorphisms, one intronic SNP (BPP=3) and one intergenic copy number variation (BPP=1) were within 40 kb of the significant association with BPP=44.
Association mapping: Association analysis of this region using the 282-line Goodman diversity panel (Flint-Garcia et al. 2005) evaluated for NLB (Wisser et al. 2011) identified an association between 185.7 Mb to 189.3 Mb (Figure 4). For significant intergenic associations, adjacent genes were considered as well as genes implicated by long-distance linkage disequilibrium. The SNP implicated by association analysis was located at 188,018,070 bp (AGP_V2) \( (p\)-value=1.72x10^{-4}, \( q\)-value=0.136) 260 bp downstream of a SpoU methylase (GRMZM5G854901) and 13.54 Kb upstream of an uncharacterized gene with a helix-loop-helix DNA-binding domain (GRMZM5G879527).

The receptor-like kinase \textit{pan1} (GRMZM5G836190) is located 37.555 kb from the significant diversity panel association, a distance close enough to link the SNP to \textit{pan1} by long-range LD (Chia et al. 2012) or the SNP could implicate a downstream element regulating the expression of \textit{pan1}.

\textbf{\textit{pan1} is a susceptibility gene for NLB:} \textit{pan1} (GRMZM5G836190 at position 187,978,007 - 187,980,232 on chromosome 1) was initially a candidate gene for disease resistance at 1.06 based on its location inside the fine-mapping interval, its proximity to an NLB association from the diversity panel analysis, and its identity as a receptor-like kinase (Cartwright et al. 2009) (this class of proteins is known to detect microbe-associated molecular patterns; Zipfel 2008)). Phenotypically similar, but unlinked, \textit{pan2} mutants (Zhang et al. 2012) were also assessed to test of the hypothesis that the \textit{pan} genes influence the disease response through their known effect on stomatal morphology or related pathways. We tested \textit{pan1} and \textit{pan2} mutants for NLB and Stewart’s wilt in replicated, multi-year trials and found a significant genotype effect \( (p\)-value <0.0001) for both diseases. Both mutants with null alleles of \textit{pan1} were found to be significantly more
resistant to both diseases than B73, indicating *pan1* is a susceptibility gene for NLB and Stewart’s wilt, while *pan2* mutants showed no significant difference from B73 (Figure 5).

For NLB, *pan1-Mu* had a 41% decrease in AUDPC, as compared with B73, and was generally more resistant than *pan1-ems*, which had a 27% decrease in AUDPC, as compared to B73.

For Stewart’s wilt, *pan1* mutants were nearly immune.

These results suggest that partial or complete loss of *pan1* may contribute to the increased resistance to NLB seen for *qNLB1.06Tx303* compared to *qNLB1.06B73*. We examined the gene action of the QTL, and in two different environments found that the QTL acted in a dominant fashion, where the heterozygote NIL fell into the same phenotypic class as the resistant NIL carrying two copies of the Tx303 allele. Preliminary data from one environment showed that the *pan1-ems* mutation was also dominant for resistance.

Further data is needed to confirm this observation. To further test whether *pan1* underlies *qNLB1.06*, we compared *pan1* gene expression levels in mature leaves of these two lines via RT-PCR. Consistent with reduced *pan1* function in *qNLB1.06Tx303*, we found *pan1* to control actin signal ratios were decreased from 0.915 +/- 0.168 in *qNLB1.06B73* to 0.553 +/- 0.035 in *qNLB1.06Tx303* (+/- standard errors, p<0.05 using Student’s *t*-test). However, further work will be needed to determine whether the reduction in *pan1* expression level in *qNLB1.06Tx303* is causally related to the increase in NLB resistance seen in this line.

**DISCUSSION**

Loci that underlie a number of traits present an opportunity to investigate the complex relationship between variation for traits, genome structure, recombination, and causative genes. Maize bin 1.06 is one such locus that is associated with effects on diverse
traits. This chromosomal region of interest has been described as a yield-stabilizing locus associated with effects on resistance to several diseases, root architecture, plant height, flowering time and yield across different soil moisture levels and genetic backgrounds (Tuberosa et al. 2002; Landi et al. 2002; Wisser et al. 2006; Landi et al. 2010). This locus is a relatively QTL-dense segment of the maize genome, with more than double the average number of QTL (35 QTL v. an average of 15 QTL/bin), but of average genetic size based on the NAM genetic map and based on physical size (McMullen et al. 2009; Andorf et al. 2010). The elevated number of QTL in this bin, coupled with an average gene content and genetic size, indicate that this bin is important for maize breeding across a broad set of traits. Hence, it is of interest not only to identify the genes underlying those traits, including multiple disease resistance, but also to investigate the genome dynamics shaping the region.

Consistent with the observation of high QTL density, variations in the targeted interval were shown to be under selection during domestication and subsequent varietal improvement. Candidate genes for domestication syndrome in this region, identified as genes lying in extended regions with allele frequency differentiation between landraces and *Zea mays* ssp. *parviglumis*, include genes such as a putative lipoxygenase, putative frataxin, and a putative zinc finger among others, while “improvement candidates” (those that contrast for improved lines versus landraces) include putative protein kinases, putative EF-hand proteins, a putative alcohol dehydrogenase, and a putative antifreeze protein, among others (Hufford et al. 2012). The observed patterns of selection may be related to disease resistance, as the transcriptional rewiring of the maize transcriptome during domestication suggests that genes related to biotic stress are overrepresented.
among the group of genes up-regulated during domestication (Swanson-Wagner et al. 2012). Indeed, some of these candidate genes, such as the lipoxygenase, serine/threonine protein kinase, and the antifreeze protein, could be involved in one of the many disease resistances conditioned by this locus.

This region harboring numerous QTL for diverse traits shows signs of high genome complexity and plasticity. A low recombination rate in the fine-mapping interval was observed in the NILs, with a limited number of recombinants identified between 184.6 Mb and 187.6 Mb, but an average gene density as compared to the maize genome as a whole (Schnable et al. 2009). The fine-mapping interval, including the region of low recombination, co-localizes with the yield-stabilizing QTL reported by Landi et al. (2010). A reduced recombination rate in this interval was observed in many of the NAM subpopulations, providing support for the hypothesis that B73 has a lack of synteny with other maize lines at this region, with the exception of Hp301, which recombines with B73 in this region. The low recombination could be due to small inversions, indels, transposon insertion, or presence/absence variation (McMullen et al. 2009). Such differences that suppress recombination may be selected upon to conserve the yield-stabilizing haplotype located at this region.

Increasingly, copy number variation (CNV) has been found to underlie trait variation, including biotic and abiotic stress tolerance (Cook et al 2012a; Maron et al. 2013). A CNV polymorphism was significantly associated with NLB in the 1.06 interval based on the NAM GWAS. Together with the lack of recombination in part of the fine-mapping population, this suggests that genome content variation across diverse maize germplasm may underlie the differences in disease response. To test this hypothesis,
structural differences were explored by examining read depth variation across the region. Pronounced variation in the number of Tx303 reads that mapped to the reference was found, which can be interpreted as evidence for duplications or genome expansion at this region in Tx303. This is not uncommon, as the maize genome is highly plastic, with read depth variation in 90% of the genome (Chia et al. 2012) and presence/absence variation thought to be a major driver of phenotypic variation in maize (Wallace et al. 2014).

Given the complex genetic basis of quantitative traits and this region, a fine-mapping approach was taken to refine the genomic region associated with resistance to NLB and Stewart’s wilt, complemented by association mapping to identify candidate genes. The fine-mapping approach allowed for the dissection of the multi-trait nature of this QTL. Resistance to Stewart’s wilt was localized to a 95.9-kb region within the larger 3.60 Mb NLB fine-mapping interval. While the majority of major QTL have been shown not to be pleiotropic in nature (Wallace et al. 2014), a pleiotropic basis of disease resistance cannot be excluded at this locus. The breakpoint analysis for both diseases was similar, although the confidence interval of NLB was calculated to be larger than that for Stewart’s wilt.

The candidate region for Stewart’s wilt contains three genes in the B73 genome sequence: a gene with a putative zinc finger and two uncharacterized genes. Both uncharacterized genes lack homologs and one lacks expression evidence (Dong et al. 2004; Sen et al. 2010). However, the genic content of the region may differ in Tx303. A physical map assembly for Tx303 across the fine-mapping region would clarify this. A number of mapping studies have implicated the region on Chr. 1 between 180 and 190 Mb across diverse populations for resistance to NLB. The $q_{NLB1.06_{Tx303}}$ region has been successfully narrowed to 3.6 Mb. While 117 candidate genes from the B73 genome sequence were
implicated through fine-mapping, it is again likely that Tx303 differs in the genic content of this region. A subset of the NLB candidate genes was particularly credible based on previous knowledge of plant defense, including four putative leucine-rich repeat protein kinases. Other candidates include a putative frataxin, a putative ABC transporter, and a putative lipoxygenase.

Association mapping provides a complementary approach for identifying candidate genes. Significant associations within the qNLB1.06Tx303 fine-mapping interval were detected using both the NAM and Goodman diversity panels (Flint-Garcia et al. 2005; Poland et al. 2011; Wisser et al. 2011; Chia et al. 2012). The most significant association within this region in the NAM was an intergenic copy number variation, which had a BPP value of 44 (p-value = 0.0000737), one of the most highly significant associations from the analysis (Poland et al. 2011; Chia et al. 2012). Among the genes implicated by NAM, the putative A20/AN1 zinc finger was the strongest candidate; a gene domain that is associated with stress tolerance in plants and the immune system in animals (Vij and Tyagi 2008). This polymorphism is close to the Stewart’s wilt fine-mapping interval, lending support to the hypothesis that a region present in Tx303 but not B73 may contain gene(s) for resistance to both diseases. Association analysis using the Goodman diversity panel revealed a significant intergenic SNP within the fine-mapping interval, approximately 800 Kb from the NAM CNV (Flint-Garcia et al. 2005; Wisser et al. 2011). The NLB-associated CNV and SNP polymorphisms may be in linkage disequilibrium with one or more genes in the vicinity. Candidate genes from association mapping can be further investigated through expression analyses, re-sequencing, and testing across different germplasm sets.
Within the fine-mapping interval shown in Figure 1, the significant Goodman panel SNP was 38 kb from *pan1*. We tested the *pan* mutants for both NLB and Stewart’s wilt because the mapping results for the two diseases were similar (Figure 1). The smaller confidence interval for Stewart’s wilt may reflect the stronger phenotype, while the similar profile indicates that there may be multiple genes underlying the QTL for both diseases. Two independent null mutant alleles of this gene conferred resistance phenotypes for NLB and Stewart’s wilt (Figure 5), demonstrating that mutations in *pan1* itself (not a linked gene present in one or other mutant background) increase resistance for the two diseases. This finding suggests that a loss of function allele of *pan1* derived from Tx303 may contribute to the disease resistance phenotype(s) conferred by *qNLB1.06Tx303* and *qSW1.06Tx303*. Preliminary data showed that the QTL and *pan1* mutants showed similar inheritance patterns with dominant resistance in both cases. This is unexpected for a loss of function mutation. Further work is needed to confirm the gene action of the *pan1* mutants and to dissect the relationship between the MDR QTL and *pan1*.

*PAN1* has been shown to play a role in promoting features of actin organization that support asymmetric cell division (Cartwright et al. 2009). Interestingly, lines carrying mutations for *pan2*, which also show asymmetric cell division, were not significantly different from B73 for NLB, indicating aberrant stomata found in both mutant lines are not the underlying mechanism of resistance in the *pan1* mutants. Susceptibility conditioned by wild-type *pan1* could be due to a passive mechanism, such as altered anatomical structures, or an active process, such as actin re-organization during pathogen attack. We have successfully refined *q1.06Tx303*, identified candidate genes, and demonstrated a role for *pan1* in multiple disease resistance. NLB and Stewart’s wilt
resistance regions have been narrowed sufficiently that the markers within the NLB and Stewart's wilt intervals can be used for marker-assisted selection. These data strongly suggest that structural variation underlies this locus and *pan1*, a gene in which lowered expression is correlated with higher resistance levels, may underlie the NLB and Stewart's wilt QTL. Loss of susceptible *pan1* alleles could be used to decrease maize susceptibility to diverse pathogens. Through fine-mapping, examining recombination rates and re-sequencing data, and evaluating mutant lines, we were able to dissect a complex locus and identified a role for *pan1* in plant defense. This approach shed light on a locus known for its complexity and quantitative effect.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Santiago Mideros for bioinformatics support, and Drs. Randall Wisser, Bode Olukolu, and Chia-Lin Chung for helpful discussions and technical assistance. We would like to acknowledge Laura Morales, Alyssa Cowles, William Miller, Chris Mancuso, Katharine Constas, Ariel Fialko, and Xingyu Luo for their assistance with fieldwork, and Yeri Park for help with the *pan1* RT-PCR experiments. This work was funded by National Science Foundation award 1127076, the McKnight Foundation, USDA Agricultural Experimental Station (HATCH) Award NYC-153418, and Cornell University.


Landi, P., Sanguineti, M., Darrah, L., Giuhani, M., Salvi, S., *et al.*, 2002 Detection of QTLs for vertical root pulling resistance in maize and overlap with QTLs for root traits in
hydroponics and for grain yield under different water regimes. Maydica 47: 233-243.


Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., et al., 2009 The B73 maize genome: complexity, diversity, and dynamics. Science 326: 1112-1115.


Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., et al., 2009 The B73 maize genome: complexity, diversity, and dynamics. Science 326: 1112-1115.

Springer, N. M., Ying, K., Fu, Y., Ji, T., Yeh, C. T., et al., 2009 Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet 5: e1000734.


Figure Legends:

**Figure 1** Mapping of *qNLB1.06* and breakpoint analyses for *qNLB1.06*.

(A) Location of introgression TBBC3_19E is shown with marker names. (B) Breakpoint analysis for NLB and Stewart’s wilt. (C) Selected representative recombinants and their associated phenotypes. For genotypes, dark shading indicates the Tx303 allele, while white indicates the B73 allele. Light grey shading indicates the region of a recombination event. For phenotypes, blue shading indicates a more resistant line, while red indicates a more susceptible line.

**Figure 2** Recombination rates across fine-mapping interval.

Recombination rates were calculated between nine markers for the NAM sub-populations. NAM founders are shown on the left and the gene counts for the eight sub-intervals across the top. Red indicates regions of high recombination and blue indicates regions of low recombination.

**Figure 3** Read depth variation across *qNLB1.06* fine-mapping region.

Blue bars indicate the number of B73 Illumina reads that align to the given bin, while red bars indicate the number of Tx303 Illumina reads that align to the given bin. Bins are 327 Kb and the start position of the bin in noted on the X-axis.

**Figure 4** Diversity panel association.
Association analysis for NLB in the fine-mapping region using the Goodman diversity panel (Flint-Garcia et al. 2005; Wisser et al. 2011). The significant SNP at 188,018,070 bp ($p$-value=1.72x10^{-4}, $q$-value=0.136) is highlighted in blue.

**Figure 5**  Mutant phenotypes of *pan* mutants.

Plants homozygous for both *pan1-ems* and *pan1-Mu* alleles are significantly different from B73, the background for the mutants, while *pan2* mutants are not. Letters denote significance with Student’s *t*-test ($p$-value < 0.0001). *pan2* mutants are shown in blue, while *pan1* mutants are shown in green. B73 is shown in red.

**File S1**  Genotypic and phenotypic data

Data for fine-mapping study is included in this file.