



Analysis of Extreme Phenotype Bulk Copy Number Variation (XP-CNV) Identified the Association of *rp1* with Resistance to Goss's Wilt of Maize

Ying Hu^{1†}, Jie Ren^{1‡}, Zhao Peng², Arnoldo A. Umana^{1†}, Ha Le¹, Tatiana Danilova¹, Junjie Fu³, Haiyan Wang⁴, Alison Robertson⁵, Scot H. Hulbert⁶, Frank F. White² and Sanzhen Liu^{1*}

OPEN ACCESS

Edited by:

Patrick Xuechun Zhao,
Noble Research Institute, LLC,
United States

Reviewed by:

Sarah Veronica Schiessl,
Justus-Liebig-Universität Gießen,
Germany
Kirankumar Mysore,
Noble Research Institute, LLC,
United States

*Correspondence:

Sanzhen Liu
liu3zhen@ksu.edu

†Present Address:

Arnoldo A. Umana,
Crop Quest, Inc., Dodge City, KS,
United States

‡These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 02 November 2017

Accepted: 19 January 2018

Published: 09 February 2018

Citation:

Hu Y, Ren J, Peng Z, Umana AA,
Le H, Danilova T, Fu J, Wang H,
Robertson A, Hulbert SH, White FF
and Liu S (2018) Analysis of Extreme
Phenotype Bulk Copy Number
Variation (XP-CNV) Identified the
Association of *rp1* with Resistance to
Goss's Wilt of Maize.
Front. Plant Sci. 9:110.
doi: 10.3389/fpls.2018.00110

¹ Department of Plant Pathology, Kansas State University, Manhattan, KS, United States, ² Department of Plant Pathology, University of Florida, Gainesville, FL, United States, ³ Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, ⁴ Department of Statistics, Kansas State University, Manhattan, KS, United States, ⁵ Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, United States, ⁶ Department of Plant Pathology, Washington State University, Pullman, WA, United States

Goss's wilt (GW) of maize is caused by the Gram-positive bacterium *Clavibacter michiganensis* subsp. *nebraskensis* (Cmn) and has spread in recent years throughout the Great Plains, posing a threat to production. The genetic basis of plant resistance is unknown. Here, a simple method for quantifying disease symptoms was developed and used to select cohorts of highly resistant and highly susceptible lines known as extreme phenotypes (XP). Copy number variation (CNV) analyses using whole genome sequences of bulked XP revealed 141 genes containing CNV between the two XP groups. The CNV genes include the previously identified common rust resistant locus *rp1*. Multiple *Rp1* accessions with distinct *rp1* haplotypes in an otherwise susceptible accession exhibited hypersensitive responses upon inoculation. GW provides an excellent system for the genetic dissection of diseases caused by closely related subspecies of *C. michiganensis*. Further work will facilitate breeding strategies to control GW and provide needed insight into the resistance mechanism of important related diseases such as bacterial canker of tomato and bacterial ring rot of potato.

Keywords: goss's wilt, resistance, copy number variation, *rp1*, *Zea mays*

INTRODUCTION

Goss's bacterial wilt and leaf blight of maize, or simply Goss's wilt (GW), was first identified in 1969 (Schuster et al., 1972; Ruhl et al., 2009). In recent years, the disease has emerged as a serious threat to production (Jackson et al., 2007; Wise et al., 2009; Mueller and Wise, 2014). The estimated total yield loss due to GW, from 2012 to 2015, was more than 500 million bushels in the US and Canada (Mueller et al., 2016). *Clavibacter michiganensis* subsp. *nebraskensis* (Cmn), the causal agent of GW, is a Gram-positive bacterium (Phylum Actinobacteria) and a vascular pathogen of maize, causing systemic wilting on young plants and leaf blight at all stages of plant growth (Ruhl et al., 2009). However, the genetic basis of host resistance to actinobacteria, in general, is poorly understood. The lack of genetic resources to control *C. michiganensis*, specifically, has led to extensive quarantine measures, particularly, for ring spot of potato, which is caused by

C. michiganensis subsp. *sepedonicus*. Quarantine measures are also targeted to *C. michiganensis* subsp. *michiganensis*, a related wilt and canker pathogen of tomato. Ectopic expression of genes conferring broad-spectrum antimicrobial activities was reported to enhance basal defenses of tomato against a related wilt pathogen *C. michiganensis* subsp. *michiganensis* (Kabelka et al., 2002; Coaker and Francis, 2004; Balaji and Smart, 2012). Two additional subspecies, *C. michiganensis* subsp. *tesselarius* and *C. michiganensis* subsp. *insidiosus*, cause diseases of wheat and alfalfa, respectively (Francis et al., 2010). A sixth subspecies has been recently proposed, *C. michiganensis* subsp. *phaseoli*, which causes a systemic disease in common bean (González and Trapiello, 2014).

Plant defense genes dynamically co-evolve with diverse pathogens, and are, therefore, highly polymorphic (Brooks et al., 2006). In general, the adaptive value of defense genes may impose fitness costs and, consequently, they may be eliminated without compromising plant health (Tian et al., 2003). Pathogen resistant (R) genes can occur in clusters at specific genomic loci, particularly in cases of common R gene classes (Marone et al., 2013). Genes in clusters may have adaptive advantages derived from rapid evolution due to rearrangement. Rearrangement can generate new specificities, and under low disease pressure, eliminate deleterious genes, thus generating intra-species copy number variation (CNV) that includes presence and absence variation (PAV). CNV has been identified for plant disease defense genes in a range of species (Bakker et al., 2006; Shen et al., 2006; Cook et al., 2012; Xu et al., 2012; González et al., 2013; Golicz et al., 2016). For example, the 31 kb locus *Rhg1* confers resistance to soybean cyst nematode and appears to function due to multiplication of the locus (Cook et al., 2012).

Maize genomes exhibit high levels of genomic structural variation (Schnable et al., 2009; Springer et al., 2009; Beló et al., 2010; Hirsch et al., 2014). Analysis of sequences of 27 diverse maize inbred lines indicated that the B73 reference genome contains ~70% low copy sequences (Gore et al., 2009). Transcriptomic analysis of 503 diverse maize inbred lines identified 8,681 representative transcript assemblies that are absent in the B73 reference genome (Hirsch et al., 2014). Furthermore, genotyping-by-sequencing (GBS) analysis of 14,129 maize inbred lines found 1.1 million PAVs (Lu et al., 2015a). Relative to inbred line B73, genes with reduced copy number in non-B73 lines are enriched in the pathways of stress responses, indicative of high variability in copy number for disease defense genes (Swanson-Wagner et al., 2010). For example, the *rp1* locus is a highly variable genomic complex of maize, conferring race-specific resistance to the common rust fungus (Hulbert, 1997; Smith et al., 2004). The locus carries multiple *rp1* paralogs, which are members of the nucleotide binding site leucine-rich repeat (NLR) family. Unequal crossovers generate gene duplications, gene losses, and genes with new phenotypic characteristics, and, ultimately, yielding lines with diverse haplotypes at the *rp1* region (Bennetzen et al., 1988; Richter et al., 1995). The number of *rp1* paralogs varies in a broad range from one to greater than fifty copies (Smith et al., 2004). A similar example of a resistant gene that varies in copy number is the recently cloned

maize wall-associated kinase (ZmWAK) that confers resistance to head smut. ZmWAK is absent in many modern maize lines, while present in wild relatives. Absence of ZmWAK is highly correlated with high susceptibility to head smut (Zuo et al., 2015). Despite the few examples, the functional consequences of CNV to plant defense are largely unexplored, and the methodology for such studies needs improvements. Here, we established a simple phenotyping method to quantify disease symptoms and identified disease defense associated genes through comparing average copy number between bulked pools of individual maize lines that displayed highly resistant or susceptible phenotypes, which was referred to as extreme phenotype bulk copy number variation (XP-CNV) analysis.

MATERIALS AND METHODS

Genetic Materials

Six hundred and fifteen maize accessions (largely inbred lines) that were subjected to disease phenotyping, including lines from the Maize 282 Association Panel (Flint-Garcia et al., 2005), were ordered from North Central Regional Plant Introduction Station (NCRPIS). These accessions were used to identify GW resistant and susceptible lines. To validate the association between the *rp1* locus and GW resistance, *Rp1*, *Rp3*, and *Rp5* accessions were used. These accessions were maintained and introgressed to the GW susceptible inbred line H95.

Quantitative Phenotyping of Disease Symptoms

A method was established for quantification of GW disease development. A virulence strain CMN06-1, isolated from Iowa maize field in 2006 by Dr. Charlie Block, was cultured on nutrient broth yeast extraction (NBY) medium at 28°C for 2–3 days. Maize plants were grown in the greenhouse at 28°C with a 14 h photoperiod. The third leaf of three-leaf maize seedlings was inoculated by cutting at 2 cm from the tip with scissors dipped in bacterial inoculum of optical density of 0.55–0.60 at 600 nm. Lengths of lesion were measured from the cut surface at the tip to the distal-most position on the leaf that exhibits a gray, chlorotic or water-soaked lesion at 13 days post-inoculation (DPI). Common lines were used in different batches to examine batch effects.

Selection of Highly Resistant (R) Lines and Highly Susceptible (S) Lines

The lesion phenotyping protocol was used to inoculate the third leaf of three-leaf seedlings. An exception was made for seedlings (335 out of 2,958) that had not developed to the desirable stage on the inoculation day. For those cases, the second or fourth leaves were inoculated instead. From phenotyping data, a statistically significant but weak association was found between seedling heights, from soil surface to the top node of plants, and lesion lengths. To generate comparable lesion phenotypes among maize genotypes, a linear model to obtain best linear unbiased estimation (BLUE) was applied. The model used raw lesion lengths measured at 13 DPI and genotype as response variable and explanatory variable, respectively. Seedling height,

inoculation leaf, and batch were used as other covariates in the model. The BLUE values of lesion lengths were used to represent host resistance levels. Maize genotypes with estimated lesions less than 9 cm were considered as highly resistant (R) genotypes, while those with lesions higher than 22 cm or those with lesions higher than 20 cm and 80% inoculation leaves showing whole leaf wilting were considered as highly susceptible (S) genotypes. Based on genetic relations inferred from genotypes, some highly R and S lines were removed to reduce the genetic redundancies from closely related lines. Selected highly R ($N = 37$) or highly S lines ($N = 44$) were used for the XP-CNV analysis.

WGS Sequencing of Highly R and S Pools

Seeds were germinated and grown in the greenhouse at 28°C with a 14 h photoperiod. Fresh leaves of seedlings at the 2–3 leaf-stage were harvested. Tissues of highly R lines and highly S lines were pooled to form highly R pool and highly S pool, respectively. Pooled tissues were frozen in liquid nitrogen, and homogenized with liquid nitrogen to fine powder. Nuclei were extracted to reduce the proportion of DNAs from organelle genomes, followed by using the Qiagen DNeasy Plant Mini Kit protocol to extract nucleus DNA (Zhang et al., 2012; Liu et al., 2017). Nucleus genomic DNAs were used for TruSeq PCR-free library preparation. Two biologically replicated samples for each of highly R and highly S pools were prepared and subjected to WGS sequencing with one sample per lane using a HiSeq2500. Paired-end reads (2×126 bp) were generated. Sequencing was conducted at Macrogen, Inc., South Korea.

Trimming and Alignment of WGS Sequences

The software Trimmomatic (version 0.32) was used to trim adaptor sequences and low quality sequences (Bolger et al., 2014) with the parameter of “ILLUMINACLIP:<adaptor>;3:20:10:1:true LEADING:3 TRAILING:3 SLIDINGWINDOW:4:13 MINLEN:50.” The adaptor sequence was from <https://github.com/timflutre/trimmomatic/blob/master/adapters/TruSeq3-PE.fa>. Reads retained after trimming were aligned to the B73 reference genome (B73Ref2; Schnable et al., 2009) using BWA (0.7.15-r1140) with the “mem” module (Li and Durbin, 2009). The minimal mapping score of 40 was required and an in-house script (github.com/liu3zhenlab/scripts/blob/master/bwa.filter/samparser.bwa.pl) was used to filter alignments to ensure that reads were uniquely aligned with high confidence. Specifically, each alignment was required to have the insert between 150 and 800 bp, at least 50 bp alignment length, at least 96% identity, and at most 4% unaligned percentage of a read length. The samtools software (version 1.1; Li et al., 2009a) was used to convert alignments in the SAM format to the BAM format.

Statistical Analysis of CNV between R and S Pools

HTSeq (0.6.1p1) was used to count read depth for each gene (filtered gene set, 5b) using the “union” mode. The generalized linear model, implemented in the DESeq2 package (version 1.4.5), for read counts of genes with at least 20 total reads from

all four samples was used for testing the null hypothesis that no difference in read depth between highly R pools and highly S pools (Love et al., 2014). A false discovery rate (FDR) approach was used to account for multiple tests (Benjamini and Hochberg, 1995). The FDR 10% was used as the cutoff for declaration of significant CNVs.

Gene Annotation

Gene annotation was downloaded from the annotation at Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html#!bulk?org=Org_Zmay) and the B73Ref2 annotation at maizesequence.org.

GO Enrichment Test of CNV Genes

Enrichment analysis was performed for determining if a certain GO term is over-represented in Up-CNV genes (greater copies in R pools) vs. Dn-CNV genes (fewer copies in R pools), and vice versa. The randomly resampling method ($N = 10,000$) in the GOSec enrichment test (Young et al., 2010) was employed. GO terms with p-values smaller than 0.05 were considered to be over-represented in a group.

Validation of CNV Genes Using WGS Sequencing Data of the Maize 282 Association Panel

WGS data of the Maize 282 Association Panel were downloaded from iplant shared from the Panzea research group. Among all WGS sequenced lines, 239 maize lines were phenotyped in this study. Sequencing reads were trimmed and aligned to the B73 reference genome (B73Ref2; Schnable et al., 2009) with the same procedures as used for highly R and S WGS data. HTSeq (0.6.1p1) was applied to count read depth for each gene (filtered gene set, 5b) using the “union” mode. Total pairs of reads after trimming were used to determine RPM (read pairs per million of total read pairs) of each gene as normalized values. The correlation was determined between normalized read counts and phenotypic values for each examined gene.

Quantitative Real-Time PCR (qPCR)

Genomic DNA was isolated using fresh leaves of seedlings at the 2–3 leaf-stage. qPCR was performed in 10 μ L reactions containing 4.2 μ L g DNA, 0.4 μ L 10 mM of each primer, and 5 μ L 2x SYBR Green PCR Master Mix (Bio-Rad) on the CFX96™ real time system (Bio-Rad). Primer efficiencies were measured and relative copy number was calculated using the comparative Ct method (Livak and Schmittgen, 2001). The *actin* gene was used as the endogenous control. A gene specific primer pair was designed to amplify GRMZM2G005134. Primers for qPCR were listed in the Supplementary Table 1.

Cytogenetic Analysis

Maize somatic chromosome preparations using the drop technique, direct probe labeling by nick translation and the fluorescence *in situ* hybridization (FISH) were performed as described previously (Kato et al., 2004, 2006). The clone containing ~1.7 kb part of *rp1* gene was provided by Dr. James A. Birchler. The insert was sequenced by Sanger sequencing

and found 100% similar to the first 1,521 bp of the coding region of the *Rp1-D* allele (Genbank accessions: AY581258 and AF107293) (Supplementary Figure 1), corresponding to the N-terminal NBS region of the *rp1* gene, which is the most conservative part of the gene (Ramakrishna et al., 2002). To make the FISH probe, the insert was amplified with standard primers M13; PCR products were purified with Invitrogen PCR purification kit (Life Technologies, USA) and labeled with Texas red-5-dCTP (PerkinElmer, USA). Oligonucleotide probes labeling centromeric repeats CentC CCTAAAGTAGTGGATTG GGCATGTTTCG and 5S ribosomal DNA TAGTAAAAATGGGT GACCGTTCTCGTGTTA were synthesized by Integrated DNA Technologies with 6-FAM attached to the 5' end. For nucleolus organizing region (NOR) probe, wheat clone pTa71 (Gerlach and Bedbrook, 1979) was labeled with Fluorescein-12-dUTP (PerkinElmer, USA). Chromosome preparations were mounted and counterstained with 4',6-diamidino-2-phenylindole solution (DAPI) in Vectashield (Vector Laboratories, USA). Images were captured with a Zeiss Axioplan 2 microscope using a cooled charge-coupled device camera CoolSNAP HQ2 (Photometrics, USA) and AxioVision 4.8 software (Zeiss). Images were processed using the Adobe Photoshop software.

Data Access

WGS Illumina sequencing data of the pools of highly R and S lines have been deposited at Sequence Read Archive (SRP100278).

RESULTS

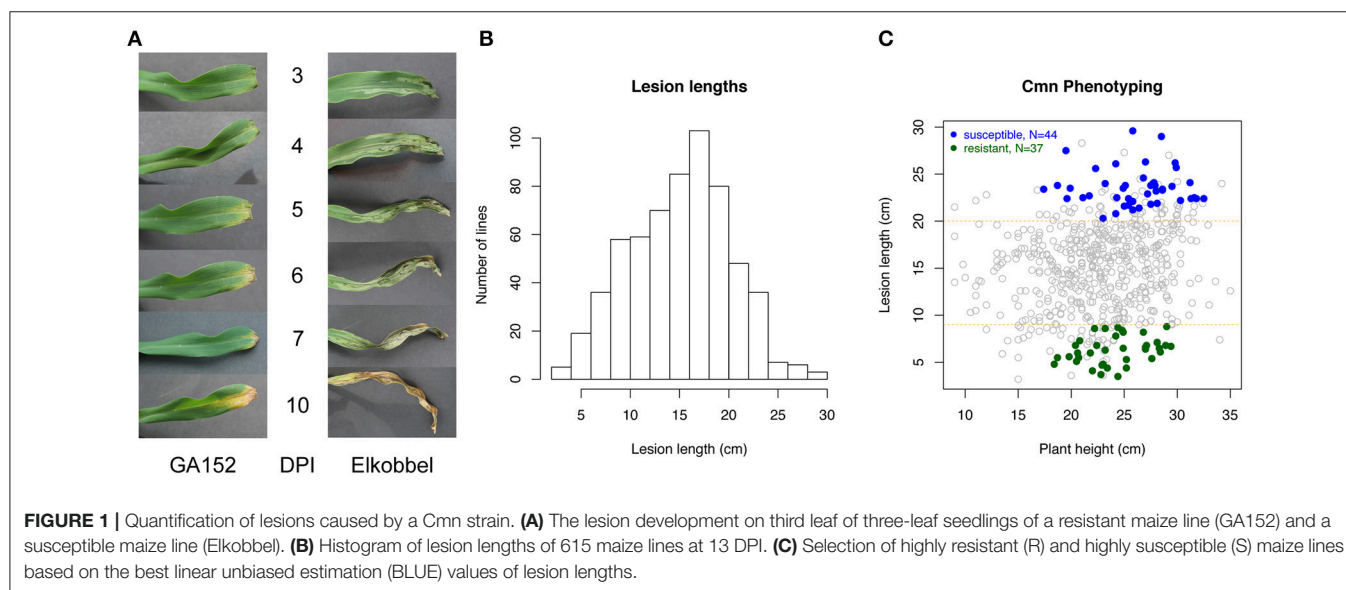
A Simple Inoculation Method for GW Disease Quantification

To establish a rapid and reproducible method for quantification of GW, the effects of growth stages and leaf positions on lesion expansion were first tested on seedlings of a maize inbred line Mo17. The results indicated that three-leaf seedlings

overall exhibited longer lesion expansion compared to four-leaf seedlings in the same time period, allowing for greater potential variation in symptoms (Supplementary Figure 2A). Three-leaf seedlings of 25 maize inbred lines were inoculated on the second or third leaf, and the lesion lengths were measured every 2 days from 3 to 13 DPI. Two-way ANOVA analysis on the factors of inoculation leaf and maize genotype found the greater mean expansion of lesions on third leaves vs. second leaves of diverse maize lines. Also high variation in lesion length between maize lines and low variation within individuals of a given line were observed on the third leaf compared to the second leaf (Supplementary Figure 2B), suggesting that the third leaf at the three-leaf seedling stage is optimal for lesion phenotyping. Twenty-four NAM parents and Mo17 showed a wide range of among-genotype variation in lesion length using this phenotyping method at 13 DPI (Supplementary Figure 3). The method was used to measure lesion lengths at 13 DPI to represent host resistance levels for 615 maize lines. Overall, maize lines showed a wide range of variation in lesion length from 2.8 to 32.1 cm at 13 DPI (Figures 1A,B, Supplementary Table 2; a time-lapse video of a highly susceptible line provided in Supplementary Movie 1). To account for potential variation from plant growth habits of the individual lines, a linear regression model was fitted for lesion lengths based on plant heights. Based on corrected lesion lengths, 37 so-called highly resistant (R) genotypes and 44 highly susceptible (S) genotypes were selected from 615 lines as extreme phenotypes (XP) (see Methods) (Figure 1C).

Differential Gene Copy Number in R and S Pools

Two biologically replicated pools of R lines and pools S lines selected from 615 lines were subjected to whole genome shotgun (WGS) sequencing. In total, ~50.3x (125.7 Gb) and 49.6x (123.9 Gb) sequences were generated for the R and S pools, respectively.



WGS sequencing data were aligned to the B73 reference genome and uniquely mapped reads were used for determining read depths per gene. We employed a novel approach, XP-CNV, to examine CNV between the R and S pools through comparing genic read depths. Each of the genes ($N = 37,483$) with at least 20 total reads from all the four samples (two biological replicates of R and S pools) was subjected to a statistical test for the null hypothesis of no difference in sequencing depths between the R and S pools (Love et al., 2014). As a result, 141 genes exhibited significantly differential read depths between R and S pools. These genes were considered as candidate CNV genes showing distinct copies in the R and S lines. Among the candidate CNV genes, 90 and 51 genes were displayed more copies (Up-CNV) and fewer copies (Dn-CNV) in R vs. S, respectively (Supplementary Table 3).

Candidate CNV genes were found on all 10 maize chromosomes (Figure 2A). Some genes were physically close. Six loci on chromosomes 1, 2, 6, and 10, each of which harbors multiple Up-CNV genes in relation to R pools, were closely clustered within 250 kb of each other, while four loci were observed close together on chromosomes 2, 6, and 7 for Dn-CNV genes in relation to R pools. Noticeably, a locus at the short arm on chromosome 10 contained four Up-CNV genes (GRMZM2G083246, GRMZM2G349565, GRMZM2G005134, and GRMZM2G143769). Functional annotation revealed that all these four genes are *rp1* family members. The *rp1* gene, GRMZM2G083246, displayed the highest fold change (2.5x) of sequencing depths in R pools vs. S pools among all Up-CNV genes (Figure 2B). *Rp1* is an NLR gene that was previously discovered to confer race-specific resistance to common rust of maize (Smith et al., 2004).

Besides *rp1*, a number of other Up-CNV genes were associated with plant disease responses. For example, a *SEC-14*-related gene (GRMZM2G363377), and one of top Up-CNV genes, displayed a 2.4x fold change in sequencing depths of R pools vs. S pools (Table 1). *SEC-14*-like genes encode phospholipid transfer proteins, which can play an important role in plant innate immune responses (Kiba et al., 2014). Another Up-CNV gene (GRMZM2G362303) encodes a wall-associated kinase (WAK). WAK family members have been demonstrated to be associated with plant defenses, including the rice OsWAK1, *Arabidopsis* WAK10 and WAK22 (Diener and Ausubel, 2005; Li et al., 2009b; Meier et al., 2010), and maize ZmWAK (Zuo et al., 2015). Gene ontology (GO) enrichment analysis of Up-CNV vs. Dn-CNV showed that the GO term transport (GO:0006810) was overrepresented in Dn-CNV ($p < 0.03$). Other GO terms associated with protein kinase, electron carrier, ion binding and heme binding, although not significantly overrepresented, occurred at least three times in Up-CNV but were absent in Dn-CNV.

Confirmation of the Association of the *rp1* Locus with GW Resistance

WGS data of individual lines of 239 maize accessions from the Maize 282 Association Panel was used to corroborate the association of *rp1* with disease resistance. The 239 lines are a subset of 615 lines whose disease resistance levels (lesion lengths) were phenotyped in this study. For each *rp1* gene, the correlation of lesion lengths with normalized read counts of the gene were examined. All four *rp1* genes identified as Up-CNV genes showed negative correlations with susceptibility as determined by lesion length (all four *p*-values smaller than 0.002, correlations

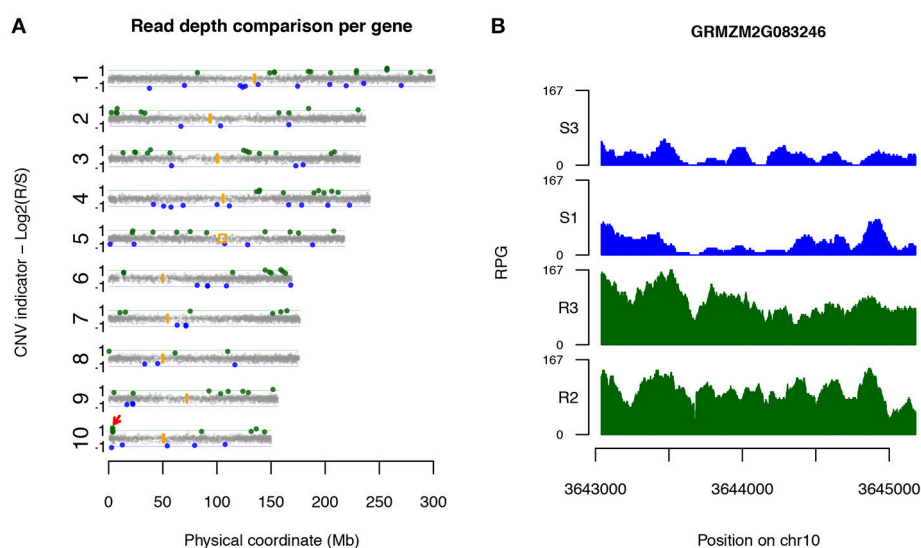


FIGURE 2 | Candidate disease-associated genes identified via XP-CNV. **(A)** A genome-wide view of the distribution of sequencing depth comparison between highly R pools and highly S pools across 10 maize chromosomes. Each dot designates a gene. Y-axis represents log2 values of ratios of read depths of R to S, signifying CNV between R and S pools. Genes with statistically significantly higher and lower read depths of R vs. S pools were interpreted as Up-CNV (green) and Dn-CNV (blue), respectively. Orange boxes indicate centromeric positions. Red arrow points to the *rp1* locus. **(B)** Read depths (reads per gigabit total aligned reads, RPG) across the Up-CNV gene GRMZM2G083246 from the pools of two highly R (R2 and R3) and two highly S (S1 and S3) pools.

TABLE 1 | Information of the top ten Up-CNV and the top 10 Dn-CNV genes.

GeneID	log2FC ^a	p-Value	padj ^b	Chr	Start	End	Description ^c
GRMZM2G083246	1.32	4.05E-09	2.54E-05	10	3643040	3645180	LRR and NB-ARC domains containing disease resistance protein (<i>rp1</i>)
GRMZM2G026176	1.28	1.51E-09	1.42E-05	1	257112452	257127180	Unknown
GRMZM2G363377	1.25	5.04E-11	9.48E-07	2	184717427	184745555	Sec14p-like phosphatidylinositol transfer family protein
GRMZM2G169013	1.25	2.08E-08	1.12E-04	2	7859326	7861362	Unknown
GRMZM2G702514	1.19	7.19E-12	2.70E-07	10	136325395	136339339	Flavanone 3-hydroxylase
GRMZM2G357081	1.19	1.21E-09	1.42E-05	1	257033199	257043603	Shaggy-like kinase 13
GRMZM2G349565	1.17	6.10E-07	1.76E-03	10	3764373	3766967	LRR and NB-ARC domains containing disease resistance protein (<i>rp1</i>)
GRMZM2G009770	1.12	9.94E-07	2.44E-03	9	151759076	151762771	Leucine-rich repeat protein kinase family protein
GRMZM2G127619	1.08	1.16E-06	2.57E-03	6	159004500	159005754	Unknown
GRMZM2G701441	1.08	2.77E-09	2.08E-05	3	38609811	38632417	Unknown
GRMZM2G086935	-0.94	3.38E-05	2.70E-02	7	71388331	71391082	Unknown
GRMZM2G106165	-0.95	2.06E-05	1.94E-02	4	50720961	50724861	Serine-type carboxypeptidase
GRMZM2G049027	-0.96	5.95E-05	3.82E-02	6	91353470	91355242	Cleavage and polyadenylation specificity factor 30
GRMZM2G575323	-0.96	4.89E-05	3.47E-02	1	126269847	126270723	Unknown
GRMZM2G164672	-0.97	3.74E-05	2.87E-02	1	174565242	174567097	Unknown
GRMZM5G829946	-0.98	2.30E-05	2.06E-02	2	66708324	66712872	PLC-like phosphodiesterases superfamily protein
GRMZM2G166695	-1.08	6.98E-08	2.92E-04	4	57590015	57595523	Remorin family protein
GRMZM2G009624	-1.09	1.04E-06	2.44E-03	1	123774374	123777663	RNA-binding (RRM/RBD/RNP motifs) family protein
GRMZM2G320305	-1.15	8.25E-08	3.10E-04	10	2690434	2694516	Ferredoxin-NADP(+)-oxidoreductase 2
GRMZM2G014323	-1.22	3.94E-08	1.85E-04	1	37667524	37685753	Cell division control 2

^aLog2-fold change in read depth between R and S pools.^bAdjusted p-values.^cGene functional annotation from Phytozome or maizegenome.org.

were from -0.346 to -0.203) (**Figure 3A**, Supplementary Figure 4). Quantitative PCR (qPCR) of one of *rp1* genes, GRMZM2G005134, on B73, 22 R lines and 22 S lines, which randomly selected from highly R and S lines, using a specific primer pair of GRMZM2G005134, showed a high correlation between qPCR signals and sequencing depths of this gene (correlation = 0.831, $p = 3.04 \times 10^{-6}$) (Supplementary Figure 5). Of 22 R lines and 22 S lines, 11 R and 17 S lines displayed extremely low qPCR signals, indicating that GRMZM2G005134 probably was absent in these lines (Supplementary Figure 6). The *t*-test comparing the R lines and B73 with S lines resulted in a significant association between qPCR signals of GRMZM2G005134 and GW resistance ($p < 0.05$) (**Figure 3B**). Eight R and S lines were randomly selected for fluorescent *in situ* hybridization (FISH) using a 1.7kb *rp1* probe to examine the overall *rp1* FISH signals (**Figure 3C**, Supplementary Figure 7). Resistant and susceptible lines showed strong or weak *rp1* FISH signals, depending on the line. For examples, A441-5, a susceptible line, and NC306, a resistant line, had strong *rp1* signals, whereas K14758, a susceptible line, and GA152, a resistant line had weak *rp1* signals (**Figure 3C**). Thus, if *rp1* is responsible for the resistance, specific combinations or specific *rp1* paralogs rather than the total number of *rp1* paralogs are associated with resistance.

Multiple *Rp1* haplotypes, containing various *rp1* copies and different *rp1* genes (Collins et al., 1999), were collected and introgressed to an inbred line H95 that is highly susceptible to both common rust and GW. All *Rp1* haplotypes confer race-specific resistance to common rust. We phenotyped the

introgression *Rp1* lines to assess GW resistance. As additional controls, *Rp3* and *Rp5* introgression lines were also tested. Of eleven *Rp1* accessions, three lines, containing *RpG* (an *rp1* haplotype), *Rp1-IG* or *Rp1-JC13*, showed high levels of resistance in comparison to H95, with 61.6% ($p = 2.4 \times 10^{-6}$), 64.9% ($p = 1.4 \times 10^{-11}$), and 47.5% ($p = 2.8 \times 10^{-10}$) reduction in lesion length, respectively (**Figure 4A**). All three lines showed necrotic and/or chlorotic symptoms at inoculation sites, resembling an HR reaction (**Figure 4B**). Some *Rp1* accessions exhibited moderate resistance to GW in comparison to H95, including accessions *Rp1-JF69*, *Rp1-J*, and *Rp1-JD46* with 36.8% ($p = 3.2 \times 10^{-8}$), 33.8% ($p = 7.3 \times 10^{-6}$), and 29.4% ($p = 1.2 \times 10^{-3}$) reduction in lesion length, respectively. Both *Rp3* and *Rp5* accessions did not show enhanced resistance. Quantification of disease resistance of these *rp1* haplotypes corroborated the association between the *rp1* locus and GW resistance.

DISCUSSION

The genetic basis of host/pathogen interactions to *Clavibacter*, in any system, is poorly understood. This study examined the genetic basis of the emerging maize Gram-positive bacterial disease GW. We applied XP-CNV to identify GW-associated CNV and discovered an associated locus in the region of the R locus *rp1* to *Puccinia sorghi*, the causal agent of common leaf rust (Hulbert, 1997; Smith et al., 2004). The co-localization was corroborated by the observation of effective resistance of multiple but not all *Rp1* accessions. Resistance to rust and possibly

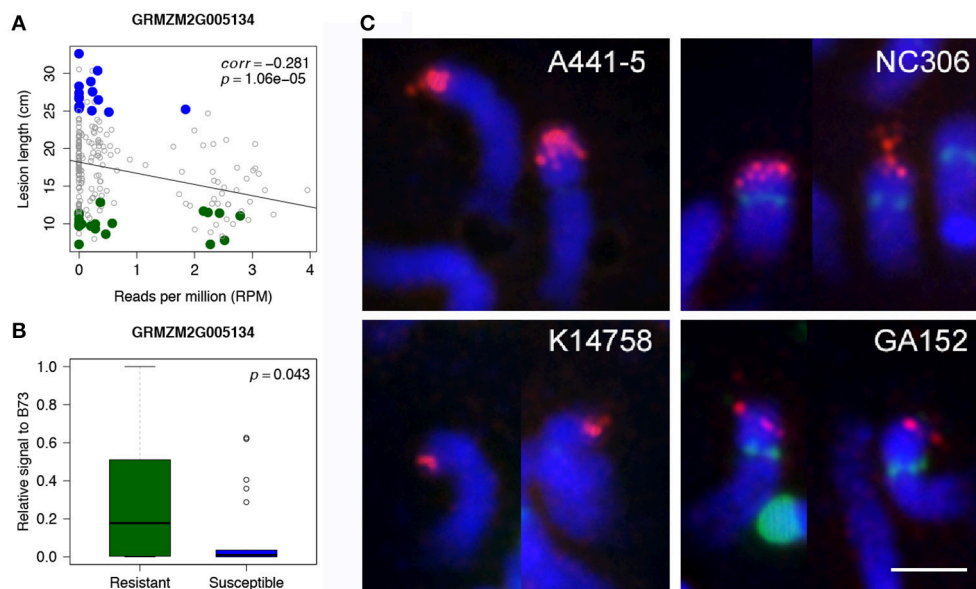


FIGURE 3 | Association of an *rp1* gene with disease resistance. **(A)** Correlations between read depths of an *rp1* gene (GRMZM2G005134) and lesion lengths of 239 lines from the Maize 282 Association Panel. Highly resistant and susceptible lines that are overlapped with the Maize 282 Association Panel are labeled in green and blue, respectively. **(B)** Boxplots of qPCR signals of GRMZM2G005134 in the group of selected R lines and B73 as well as the group of S lines. **(C)** FISH of the 1.7kb *rp1* probe on somatic late prophase of chromosome 10 of four maize lines. *Rp1* signals are in red; CentC, NOR, and 5S rDNA repeats are in green; chromosomes counterstained with DAPI are in blue. Bar corresponds to 5 μ m.

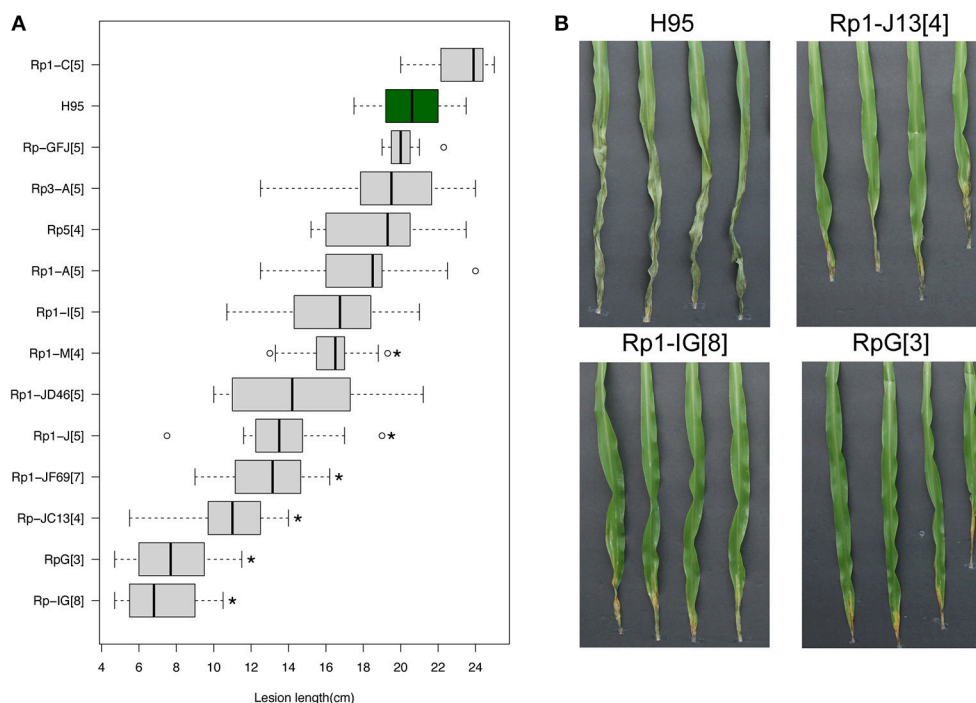


FIGURE 4 | Effects of multiple *Rp1* accessions on disease resistance. **(A)** Distribution of lesion lengths at 13 DPLs in multiple *Rp1* accessions of which the *rp1* locus was introgressed into the inbred line H95. In the boxplot, the vertical line within each box indicates a median of lesion lengths. Asterisks indicate statistically significant differences between H95 and an *Rp1* accession (*t*-test, $*p < 0.001$). **(B)** Goss's wilt phenotype of H95 and three *Rp1* accessions, *RpG* (an *Rp1* haplotype), *Rp1-IG*, and *Rp1-JC13*, at 13 DPLs with a Cmn strain. Numbers in square brackets indicate the times of backcross generations.

CMN is due to either the unique combination of *rp1* genes (Hu et al., 1997) or intragenic recombination events generating novel *rp1* genes at the *rp1* locus (Smith et al., 2010). Some of these haplotypes mediate spontaneous defense (lesion mimic) responses even in the absence of rust infection (Hu et al., 1997; Chintamanani et al., 2010; Smith et al., 2010). Lesion mimic phenotypes have not been observed in most of the GW resistant lines identified in this study. The GW resistance identified in these lines is thus probably due to a specific response to the bacterial pathogen and not a non-specific induction of defense responses. In the simplest application of the results, manipulation of known resistant *rp1* haplotypes in elite germplasm may provide enhanced resistance to GW.

Many plant disease resistance genes are dispensable to the general fitness of the plant (Collins et al., 2003; McHale et al., 2012; Yao et al., 2015; Hardigan et al., 2016). Therefore, in theory, some CNV such as PAV or tandem duplication should be detectable by SNP markers that exhibit high linkage disequilibrium with CNV locations. A human study indicated that 77% of CNVs were effectively tagged by SNPs, suggesting GWAS can capture the majority of CNV (Conrad et al., 2010). Genotyping by CNV of multiple unlinked homologous family members may obfuscate SNP analysis. Unlike SNPs or small insertions or deletions (InDels), PAV or other CNV have not been well established. In this study, XP-CNV was used with bulked WGS sequencing data for the purpose of reducing sequencing cost. Sequencing depths per gene were determined to infer average gene copy numbers of individuals in each GW R and S pools. The statistical comparison between the R and S pools resulted in a set of genes with differential average gene CNVs between the two groups, and many Up-CNV genes in R pools seemed to be defense-related. The strategy of XP-CNV should be efficient for direct identification of causal CNV genes.

We have utilized sequencing data of the Maize 282 Association Panel to verify the GW association with *rp1* genes, suggesting that the XP-CNV result is largely repeatable. However, XP-CNV analyses are subject to false discovery. First, organelle sequences homologous to nuclear chromosomes could result in high variation in read depths (Lough et al., 2015). The number of organelles, such as mitochondria or chloroplasts, varies in different tissues and the percentage of organelle genomes in DNA samples is subject to DNA extraction procedures. Therefore, falsely discovered CNV can result from organelle sequences. Here, two biological replicates of each R and S pool were done to minimize the effect of organelle variation. Second, the method is influenced by the population structure of selected lines, which could lead to spurious association (Balding, 2006). One solution is to generate sequences of individual lines that can be used to assess population structure, and to establish phenotype-CNV association with the control of population structure.

Genomic structural variation (SV) includes CNV and other SV types such as inversions and rearrangements, and only CNV was examined in this study. It is still challenging to reliably genotype SV other than CNV in a population scale. However, the advance of longer read sequencing and bioinformatics algorithms has improved the calling accuracy and reliability (Pirooznia et al., 2015; Dong et al., 2016; Huddleston et al., 2016; Peng et al., 2016). In future, with the continuous decrease of sequencing

cost and improvement of sequencing technologies, the majority of plant germplasms will be sequenced in a decent depth. Phenotypic association of CNV and perhaps other genomic structural variation would be a regular analysis in addition to regular GWAS.

Through XP-CNV, we identified the association between the *rp1* genes and disease resistance to GW caused by the Gram-positive bacterium *Cmn*. HR-like resistant responses were observed for multiple *Rp1* haplotypes. It is unknown whether one or multiple separate genes with separate elicitors occur at the locus. The *Rp1* locus encodes a number of nucleotide binding site/leucine rich repeat proteins, which are known to participate in the recognition of pathogen effectors (Deyoung and Innes, 2006). In Gram-negative bacteria, many of the known effectors are delivered to plant cells through the type III systems (T3SS). However, no secretion systems equivalent to T3SS has been identified in *Cmn* (Bentley et al., 2008; Gartemann et al., 2008; Lu et al., 2015b). The known virulence factors of related *Clavibacter* pathogens are secreted by the type II secretion system (T2SS). Alternatively, toxins or other non-proteinaceous compounds may trigger the HR. The revelation of the association between the *rp1* locus and GW provides a system to elucidate molecular interactions between hosts and bacterial species of *C. michiganensis*.

AUTHOR CONTRIBUTIONS

YH, JR, FW, AR, and SL designed experiments; YH, HW, JF, and SL analyzed data. YH, JR, ZP, AU, HL, and TD performed wet-lab experiments; SH maintained and provided *rp1* lines; YH, FW, and SL wrote the manuscript with the input from ZP, AR, TD, JF, SH, and HW. All authors read and approved the final manuscript.

FUNDING

We thank the support from the Kansas Agricultural Experiment Station of Kansas State University and the U.S. National Science Foundation (Award No. 1238189). This is contribution number 17-274-J from the Kansas Agricultural Experiment Station.

ACKNOWLEDGMENTS

We thank the North Central Regional Plant Introduction Station (NCRPIS) and the Maize Genetics COOP Stock Center for sharing seeds of maize germplasms. We thank Dr. Charlie Block for his kindness of sharing *Cmn* strains. We thank Patrice Albert and Dr. James Birchler for supplying *rp1* plasmid used for the FISH experiment and Guifang Lin and Pierre Migeon for technical assistance. We thank the Panzea group for sharing sequencing data of the Maize 282 Association Panel. Many thanks are extended to Dr. Bailin Li and Dr. Patrick Schnable for helpful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00110/full#supplementary-material>

REFERENCES

- Bakker, E. G., Toomajian, C., Kreitman, M., and Bergelson, J. (2006). A genome-wide survey of R gene polymorphisms in Arabidopsis. *Plant Cell* 18, 1803–1818. doi: 10.1105/tpc.106.042614
- Balaji, V., and Smart, C. D. (2012). Over-expression of snakin-2 and extensin-like protein genes restricts pathogen invasiveness and enhances tolerance to *Clavibacter michiganensis* subsp. *michiganensis* in transgenic tomato (*Solanum lycopersicum*). *Transgenic Res.* 21, 23–37. doi: 10.1007/s11248-011-9506-x
- Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* 7, 781–791. doi: 10.1038/nrg1916
- Beló, A., Beatty, M. K., Hondred, D., Fengler, K. A., Li, B., and Rafalski, A. (2010). Allelic genome structural variations in maize detected by array comparative genome hybridization. *Theor. Appl. Genet.* 120, 355–367. doi: 10.1007/s00122-009-1128-9
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B-Methodol.* 57, 289–300.
- Bennetzen, J. L., Blevins, W. E., and Ellingboe, A. H. (1988). Cell-autonomous recognition of the rust pathogen determines *rp1*-specified resistance in maize. *Science* 241, 208–210. doi: 10.1126/science.241.4862.208
- Bentley, S. D., Corton, C., Brown, S. E., Barron, A., Clark, L., Doggett, J., et al. (2008). Genome of the actinomycete plant pathogen *Clavibacter michiganensis* subsp. *sepedonicus* suggests recent niche adaptation. *J. Bacteriol.* 190, 2150–2160. doi: 10.1128/JB.01598-07
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Brooks, S. A., Huang, L., Herbel, M. N., Gill, B. S., Brown-Guedira, G., and Fellers, J. P. (2006). Structural variation and evolution of a defense-gene cluster in natural populations of *Aegilops tauschii*. *Theor. Appl. Genet.* 112, 618–626. doi: 10.1007/s00122-005-0160-7
- Chintamanani, S., Hulbert, S. H., Johal, G. S., and Balint-Kurti, P. J. (2010). Identification of a maize locus that modulates the hypersensitive defense response, using mutant-assisted gene identification and characterization. *Genetics* 184, 813–825. doi: 10.1534/genetics.109.111880
- Coaker, G. L., and Francis, D. M. (2004). Mapping, genetic effects, and epistatic interaction of two bacterial canker resistance QTLs from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* 108, 1047–1055. doi: 10.1007/s00122-003-1531-6
- Collins, N., Drake, J., Ayliffe, M., Sun, Q., Ellis, J., Hulbert, S., et al. (1999). Molecular characterization of the maize *Rp1-D* rust resistance haplotype and its mutants. *Plant Cell* 11, 1365–1376. doi: 10.1105/tpc.11.7.1365
- Collins, N. C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J. L., et al. (2003). SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977. doi: 10.1038/nature02076
- Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* 464, 704–712. doi: 10.1038/nature08516
- Cook, D. E., Lee, T. G., Guo, X., Melito, S., Wang, K., Bayless, A. M., et al. (2012). Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338, 1206–1209. doi: 10.1126/science.1228746
- Deyoung, B. J., and Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat. Immunol.* 7, 1243–1249. doi: 10.1038/ni1410
- Diener, A. C., and Ausubel, F. M. (2005). Resistance to *Fusarium Oxysporum* 1, a dominant Arabidopsis disease-resistance gene, is not race specific. *Genetics* 171, 305–321. doi: 10.1534/genetics.105.042218
- Dong, J., Feng, Y., Kumar, D., Zhang, W., Zhu, T., Luo, M. C., et al. (2016). Analysis of tandem gene copies in maize chromosomal regions reconstructed from long sequence reads. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7949–7956. doi: 10.1073/pnas.1608775113
- Flint-Garcia, S. A., Thuillet, A. C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064. doi: 10.1111/j.1365-3113XX.2005.02591.x
- Francis, I., Holsters, M., and Vereecke, D. (2010). The gram-positive side of plant-microbe interactions. *Environ. Microbiol.* 12, 1–12. doi: 10.1111/j.1462-2920.2009.01989.x
- Gartemann, K. H., Abt, B., Bekel, T., Burger, A., Engemann, J., Flügel, M., et al. (2008). The genome sequence of the tomato-pathogenic actinomycete *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382 reveals a large island involved in pathogenicity. *J. Bacteriol.* 190, 2138–2149. doi: 10.1128/JB.01595-07
- Gerlach, W. L., and Bedbrook, J. R. (1979). Cloning and characterization of ribosomal-RNA genes from wheat and barley. *Nucleic Acids Res.* 7, 1869–1885. doi: 10.1093/nar/7.7.1869
- Golicz, A. A., Batley, J., and Edwards, D. (2016). Towards plant pangenomics. *Plant Biotechnol. J.* 14, 1099–1105. doi: 10.1111/pbi.12499
- González, A. J., and Trapiello, E. (2014). *Clavibacter michiganensis* subsp. *phaseoli* subsp. nov., pathogenic in bean. *Int. J. Syst. Evol. Microbiol.* 64, 1752–1755. doi: 10.1099/ijs.0.058099-0
- González, V. M., Aventin, N., Centeno, E., and Puigdomenech, P. (2013). High presence/absence gene variability in defense-related gene clusters of *Cucumis melo*. *BMC Genomics* 14:782. doi: 10.1186/1471-2164-14-782
- Gore, M. A., Chia, J. M., Elshire, R. J., Sun, Q., Ersoz, E. S., Hurwitz, B. L., et al. (2009). A first-generation haplotype map of maize. *Science* 326, 1115–1117. doi: 10.1126/science.1177837
- Hardigan, M. A., Crisovan, E., Hamilton, J. P., Kim, J., Laimbeer, P., Leisner, C. P., et al. (2016). Genome reduction uncovers a large dispensable genome and adaptive role for copy number variation in asexually propagated *Solanum tuberosum*. *Plant Cell* 28, 388–405. doi: 10.1105/tpc.15.00538
- Hirsch, C. N., Foerster, J. M., Johnson, J. M., Sekhon, R. S., Muttoni, G., Vaillancourt, B., et al. (2014). Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* 26, 121–135. doi: 10.1105/tpc.113.119982
- Hu, G., Webb, C. A., and Hulbert, S. H. (1997). Adult plant phenotype of the *Rp1-DJ* compound rust resistance gene in maize. *Phytopathology* 87, 236–241. doi: 10.1094/PHYTO.1997.87.3.236
- Huddleston, J., Chaisson, M. J. P., Steinberg, K. M., Warren, W., Hoekzema, K., Gordon, D., et al. (2016). Discovery and genotyping of structural variation from long-read haploid genome sequence data. *Genome Res.* 27, 677–685. doi: 10.1101/gr.214007.116
- Hulbert, S. H. (1997). Structure and evolution of the *rp1* complex conferring rust resistance in maize. *Annu. Rev. Phytopathol.* 35, 293–310. doi: 10.1146/annurev.phyto.35.1.293
- Jackson, T. A., Harveson, R. M., and Vidaver, A. K. (2007). *Goss's Bacterial Wilt and Leaf Blight of Corn*. NebGuide G1675. Lincoln, NE: Nebraska Institute of Agriculture and Natural Resources.
- Kabelka, E., Franchino, B., and Francis, D. M. (2002). Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology* 92, 504–510. doi: 10.1094/PHYTO.2002.92.5.504
- Kato, A., Lamb, J. C., and Birchler, J. A. (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13554–13559. doi: 10.1073/pnas.0403659101
- Kato, A., Albert, P. S., Vega, J. M., and Birchler, J. A. (2006). Sensitive fluorescence *in situ* hybridization signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. *Biotech. Histochem.* 81, 71–78. doi: 10.1080/10520290600643677
- Kiba, A., Galis, I., Hojo, Y., Ohnishi, K., Yoshioka, H., and Hikichi, Y. (2014). SEC14 phospholipid transfer protein is involved in lipid signaling-mediated plant immune responses in *Nicotiana benthamiana*. *PLoS ONE* 9:e98150. doi: 10.1371/journal.pone.0098150
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009a). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, H., Zhou, S. Y., Zhao, W. S., Su, S. C., and Peng, Y. L. (2009b). A novel wall-associated receptor-like protein kinase gene, OsWAK1, plays important roles in rice blast disease resistance. *Plant Mol. Biol.* 69, 337–346. doi: 10.1007/s11103-008-9430-5
- Liu, S., Zheng, J., Migeon, P., Ren, J., Hu, Y., He, C., et al. (2017). Unbiased k-mer analysis reveals changes in copy number of highly repetitive

- sequences during maize domestication and improvement. *Sci. Rep.* 7:42444. doi: 10.1038/srep42444
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lough, A. N., Faries, K. M., Koo, D. H., Hussain, A., Roark, L. M., Langewisch, T. L., et al. (2015). Cytogenetic and sequence analyses of mitochondrial DNA insertions in nuclear chromosomes of maize. *G3* 5, 2229–2239. doi: 10.1534/g3.115.020677
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8
- Lu, F., Romy, M. C., Glaubitz, J. C., Bradbury, P. J., Elshire, R. J., Wang, T., et al. (2015a). High-resolution genetic mapping of maize pan-genome sequence anchors. *Nat. Commun.* 6:6914. doi: 10.1038/ncomms7914
- Lu, Y., Hatsugai, N., Katagiri, F., Ishimaru, C. A., and Glazebrook, J. (2015b). Putative serine protease effectors of *Clavibacter michiganensis* induce a hypersensitive response in the apoplast of *Nicotiana glauca*. *Mol. Plant Microbe Interact.* 28, 1216–1226. doi: 10.1094/MPMI-02-15-0036-R
- Marone, D., Russo, M. A., Laidò, G., De Leonardi, A. M., and Mastrangelo, A. M. (2013). Plant nucleotide binding site-leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. *Int. J. Mol. Sci.* 14, 7302–7326. doi: 10.3390/ijms14047302
- McHale, L. K., Haun, W. J., Xu, W. W., Bhaskar, P. B., Anderson, J. E., Hyten, D. L., et al. (2012). Structural variants in the soybean genome localize to clusters of biotic stress-response genes. *Plant Physiol.* 159, 1295–1308. doi: 10.1104/pp.112.194605
- Meier, S., Ruzvidzo, O., Morse, M., Donaldson, L., Kwezi, L., and Gehring, C. (2010). The *Arabidopsis* wall associated kinase-like 10 gene encodes a functional guanylyl cyclase and is co-expressed with pathogen defense related genes. *PLoS ONE* 5:e8904. doi: 10.1371/journal.pone.0008904
- Mueller, D. S., and Wise, K. A. (2014). *Corn Disease Loss Estimates from the United States and Ontario, Canada—2013*. West Lafayette, IN: Purdue Extension Publication, BP-96-13-W.
- Mueller, D. S., Wise, K. A., Sisson, A. J., Allen, T. W., Bergstrom, G. C., Bosley, D. B., et al. (2016). Corn yield loss estimates due to diseases in the United States and Ontario, Canada from 2012 to 2015. *Plant Health Prog.* 17, 211–222. doi: 10.1094/PHP-RS-16-0030
- Peng, Z., Hu, Y., Xie, J., Potnis, N., Akhunova, A., Jones, J., et al. (2016). Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics* 17, 21. doi: 10.1186/s12864-015-2348-9
- Pirooznia, M., Goes, F. S., and Zandi, P. P. (2015). Whole-genome CNV analysis: advances in computational approaches. *Front. Genet.* 6:138. doi: 10.3389/fgene.2015.00138
- Ramakrishna, W., Emberton, J., Ogden, M., Sanmiguel, P., and Bennetzen, J. L. (2002). Structural analysis of the maize Rp1 complex reveals numerous sites and unexpected mechanisms of local rearrangement. *Plant Cell* 14, 3213–3223. doi: 10.1105/tpc.006338
- Richter, T. E., Pryor, T. J., Bennetzen, J. L., and Hulbert, S. H. (1995). New rust resistance specificities associated with recombination in the Rp1 complex in maize. *Genetics* 141, 373–381.
- Ruhl, G., Wise, K., Creswell, T., Leonberger, A., and Speers, C. (2009). First report of goss's bacterial wilt and leaf blight on corn caused by *Clavibacter michiganensis* subsp. *nebraskensis* in Indiana. *Plant Dis.* 93, 841–841. doi: 10.1094/PDIS-93-8-0841B
- Schnable, P. S., Ware, P., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534
- Schuster, M. L., Compton, W. A., and Hoff, B. (1972). Reaction of corn inbred lines to the new Nebraska leaf freckles and wilt bacterium. *Plant Dis. Rep.* 56, 863–865.
- Shen, J., Araki, H., Chen, L., Chen, J. Q., and Tian, D. (2006). Unique evolutionary mechanism in R-genes under the presence/absence polymorphism in *Arabidopsis thaliana*. *Genetics* 172, 1243–1250. doi: 10.1534/genetics.105.047290
- Smith, S. M., Pryor, A. J., and Hulbert, S. H. (2004). Allelic and haplotypic diversity at the rp1 rust resistance locus of maize. *Genetics* 167, 1939–1947. doi: 10.1534/genetics.104.029371
- Smith, S. M., Steinau, M., Trick, H. N., and Hulbert, S. H. (2010). Recombinant Rp1 genes confer necrotic or nonspecific resistance phenotypes. *Mol. Genet. Genomics* 283, 591–602. doi: 10.1007/s00438-010-0536-5
- Springer, N. M., Ying, K., Fu, Y., Ji, T., Yeh, C. T., Jia, Y., 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. doi: 10.1371/journal.pgen.1000734
- Swanson-Wagner, R. A., Eichten, S. R., Kumari, S., Tiffin, P., Stein, J. C., Ware, D., et al. (2010). Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome Res.* 20, 1689–1699. doi: 10.1101/gr.109165.110
- Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M., and Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* 423, 74–77. doi: 10.1038/nature01588
- Wise, K., Ruhl, G., Creswell, T., Leonberger, A., and Speers, C. (2009). Occurrence and impact of Goss's bacterial wilt and leaf blight on corn in Indiana. *Phytopathology* 99:S143.
- Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30, U105–U157. doi: 10.1038/nbt.2050
- Yao, W., Li, G., Zhao, H., Wang, G., Lian, X., and Xie, W. (2015). Exploring the rice dispensable genome using a metagenome-like assembly strategy. *Genome Biol.* 16, 187. doi: 10.1186/s13059-015-0757-3
- Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14
- Zhang, M., Zhang, Y., Scheuring, C. F., Wu, C. C., Dong, J. J., and Zhang, H. B. (2012). Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. *Nat. Protoc.* 7, 467–478. doi: 10.1038/nprot.2011.455
- Zuo, W., Chao, Q., Zhang, N., Ye, J., Tan, G., Li, B., et al. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. *Nat. Genet.* 47, 151–157. doi: 10.1038/ng.3170

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer KM and handling Editor declared their shared affiliation.

Copyright © 2018 Hu, Ren, Peng, Umana, Le, Danilova, Fu, Wang, Robertson, Hulbert, White and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.