A GIANT CHIMERIC NLR GENE CONFERS BROAD RESISTANCE TO PHYTOPHTHORA ROOT AND STEM ROT OF SOYBEAN

by

Weidong Wang

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Agronomy

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Jianxin Ma, Chair
Department of Agronomy

Dr. Mitchell R. Tuinstra
Department of Agronomy

Dr. Steven R. Scofield
Department of Agronomy & USDA-ARS

Dr. Tesfaye Mengiste
Department of Botany and Plant Pathology

Dr. Rajat Aggarwal
Research and Development, Corteva Agriscience

Approved by:
Dr. Ronald Turco

To All who have supported and helped me

致所有在我海外求学之路上提供过帮助的人

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Jianxin Ma, for giving me the opportunity to obtain my doctoral training in his lab. The Ma Lab is an excellent platform for both study and research, with experts in different areas, and I have learned a lot from everyone. From Dr. Ma in particular, I have learned how to think through questions critically, solve problems in a timely manner, and present my results more efficiently. I will definitely benefit from these problem-solving and presentation skills throughout my life.

My special thanks also go to my advisory committee members, Dr. Mitchell R. Tuinstra, Dr. Steven R. Scofield, Dr. Tesfaye Mengiste, Dr. Rajat Aggarwal as well as former committee member Dr. Oswald Crasta, for their generosity in serving on my committee and their insightful suggestions and valuable feedback throughout my course and research.

I also would like to thank all of my lab mates, Dr. Lianjun Sun, Dr. Dajian Zhang, Dr. Zhenyan Miao, Dr. Bo Ren, Dr. Xutong Wang, Dr. Jingbo Duan, Liyang Chen, Chance Clark, Gabriel Fear, for their tremendous assistance in the past six years, and thank Tomara Fleury and Dr. Guohong Cai for their help in resistance screening. I also would like to thank Jim Beaty and his team for the help in field work at ACRE.

I would like to thank Corteva Agriscience™ (Dow AgroSciences) for funding this project, and thank Jon Massman and John Woodward, for their leadership in the Purdue University-Corteva Agriscience project collaboration, thank Kevin Fengler, Joy Bolar, Victor Llaca , Jon Myrvold, David Oneal, Daleen van Dyk, Ashley Hudson, Jesse Munkvold, Andy Baumgarten, Jeff Thompson, Ajit Nott, Lyudmila Sidorenko, Abhijit Sanyal, Jon Allen and Tyler Engelhart, for their technical support in genotyping, transformation, sequence analysis, resistance screening and greenhouse coordination.

Last but not least, I would especially like to thank my family. My parents and brother deserve special thanks for their continued support and encouragement. My wife, Fan Zuo, has been

extremely supportive of me since we got married. My daughter, Yimi Wang, and my upcoming son have brought me endless happiness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Phytophthora root and stem rot is the most destructive soybean soil-borne disease worldwide and can be managed using soybean cultivars with genes conferring resistance to *Phytophthora sojae*. Here we show that soybean *Rps11* is an ~27-kb nucleotide-binding site leucine-rich repeat (NLR) gene that confers broad-spectrum resistance to the pathogen. This giant gene is located in a genomic region containing 12 unusually large NLR genes of a single origin and was formed by rounds of intergenic/intragenic unequal recombination that involves the promoter regions and the LRR regions. Comparison of the genomic region in the Rps11 donor line with its corresponding regions in 29 diverse soybean genomes revealed drastic regional diversification including NLR copy number variation ranging from 5 to 23, and absence of allelic copy of *Rps11* in all 29 genomes. This study highlights innovative evolution and complexity of an NLR cluster and enables precise selection of *Rps11* for cultivar improvement.

CHAPTER 1. INTRODUCTION

Soybean (*Glycine max*) is the most important legume crop, providing 69% of world protein meal and 28% of vegetable oil (www.soystats.com). However, Phytophthora root and stem rot (PRSR), caused by the soil-borne pathogen *Phytophthora sojae*, threatens global soybean production, with annual losses of $200 million in the United States and 1-2 billion worldwide (Dorrance et al., 2007; Tyler, 2007). Deployment of soybean varieties with genes conferring durable and broad-spectrum resistance is the most effective and environmentally friendly strategy to prevent PRSR (Dorrance and Schmitthenner, 2000; Hartman et al., 1999). *Rps1-k* has been widely used in commercial cultivars since 1982 because of its excellent resistance to known races of *P. sojae* (Gordon et al., 2006). However, an increasing number of isolates of *P. sojae*, such as Race 25 and Race 31, have been reported in subsequent years to be virulent to *Rps1-k* (Abney et al., 1997). Studies suggested that *P. sojae* populations could shift rapidly under strong selection pressure (Kaitany et al., 2001; Tooley and Grau, 1982), and individual Rps genes could remain effective for only 8 to 15 years of deployment in cultivars (Schmitthenner, 1985). Therefore, it is highly desired to identify novel Rps genes with durable and broad-spectrum resistance to *P. sojae* for effective management of PRSR.

To this end, more than 30 loci/alleles resistant to *P. sojae* (Rps) have been identified in past decades, most of which were mapped to genomic regions enriched in nucleotide-binding site leucine-rich repeat (NLR) genes (www.soybase.org). Unfortunately, the complex genomic variations found at these loci, including structural and copy number variations, have hindered understanding of the molecular mechanisms underlying the resistance and restricted their application in soybean breeding. First, the high sequence similarity among the different NLR genes often makes it difficult to find unique markers tightly linked to the resistant gene within an NLR gene cluster. Second, marker-assisted introgression is not only time-consuming but also leads to the introduction of an entire cluster of NLR genes, putting the resistant gene at high risk of being disrupted by unequal recombination between different NLR genes. Therefore, precise isolation of the resistant gene from NLR gene clusters is highly desirable, as it will allow more accurate, efficient, and stable introduction of resistance genes into target soybean cultivars.

CHAPTER 2. LITERATURE REVIEW

2.1 Plant defense strategies and known Rps loci

Plant defense includes both constitutive and induced strategies to protect plant from pathogen invasion (Anderson et al., 2010; Doughari, 2015). Constitutive defense includes physical barriers that can avoid the pathogens from entering and spreading in the plant cell and tissues, and biochemical reactions with products that can inhibit or kill the pathogens (Osbourn, 1996; Underwood, 2012). Induced defense also consists of two layers of protection. The first layer is known as pathogen associated molecular pattern (PAMP) triggered immunity (PTI) that recognizes conserved molecules or structure of pathogens and is usually not race specific; The second layer is effector triggered immunity (ETI) that recognizes the effectors released by pathogens and is usually race specific (Jones and Dangl, 2006). Soybean defense against *Phytophthora sojae* is usually ETI controlled by Resistant-to-*P.sojae* (Rps) genes.

For effective management of PRSR, about 34 Rps loci/alleles have been identified onto nine soybean chromosomes in past decades (Table 2.1). Among these loci/alleles, *Rps1a*, *Rps1b*, *Rps1c*, *Rps1d*, *Rps1k*, *Rps7*, *Rps9*, *RpsYu25*, *RpsZheng*, *RpsYD29*, *RpsX*, *RpsHC18*, *RpsQ*, *RpsHN*, *RpsUN1*, *RpsWY* and *RpsGZ* were mapped in a genome region on chromosome 3, forming the largest cluster of Rps loci (Table 2.1). The second largest cluster of Rps loci is located on chromosome 18, including *Rps4*, *Rps5*, *Rps6*, *Rps12* and *RpsJS* (Table 2.1). Chromosome 13 has the third largest cluster of Rps loci, harboring *Rps3a*, *Rps3b*, *Rps3c* and *RpsSN10* (Table 2.1). Besides these Rps cluster, *Rps2* and *RpsUN2* are located on chromosome 16; *RpsZS18* is located on chromosome 2; *Rps11* is located on chromosome 7; *Rps8* is located on chromosome 8; *RpsSu* is located on chromosome 10; *Rps10* is located on chromosome 17; *RpsYB30* is located on chromosome 19 (Table 2.1). Interestingly, most of these Rps loci/alleles were mapped into genome regions enriched with NLR genes. Since these loci are identified from various donor lines, whether they are controlled by different alleles of the same NLR gene or different NLR genes remain unclear.

Table 2.1 Summary of known Rps loci

Rps	Donor	Chromosome	Reference
RpsZS18	Zaoshu18	2	(Zhong et al., 2018b)
Rps1a	Mukden	3	(Bernard et al., 1957)
Rps1b	FC 31745	3	(Hartwig et al., 1968)
Rps1c	Arksoy	3	(Mueller et al., 1978)
Rps1d	PI 103091	3	(Buzzell and Anderson, 1992)
Rps1k	Kingwa	3	(Bernard and Cremeens, 1981)
Rps7	Harosoy	3	(Anderson and Buzzell, 1992)
Rps9	Ludou 4	3	(WU et al., 2011a)
RpsYu25	Yudou 25	3	(Sun et al., 2011)
RpsZheng	Zheng 97196	3	(ZHANG et al., 2020)
RpsYD29	Yudou 2	3	(Zhang et al., 2013a)
RpsX	Xiu94-11	3	(Zhong et al., 2019)
RpsHC18	Huachun 18	3	(Zhong et al., 2018a)
RpsQ	Qichadou 1	3	(Li et al., 2017)
RpsHN	Meng8206	3	(Niu et al., 2017)
RpsUN1	PI 567139 B	3	(Lin et al., 2013)
RpsWY	Wayao	3	(Cheng et al., 2017)
RpsGZ	Guizao1	3	(Jiang et al., 2020)
Rps11	PI 594527	7	(Ping et al., 2016)
Rps8	PI 399073	8	(Burnham et al., 2003)
RpsSu	Su88-M21	10	(Wu et al., 2011b)
Rps3a	Mukden	13	(Mueller *et al.*, 1978)
Rps3b	PI 84637	13	(Mueller *et al.*, 1978)
Rps3c	PI 54615-1	13	(Mueller *et al.*, 1978)
RpsSN10	Suinong 10	13	(Yu et al., 2010)
Rps2	CNS	16	(Kilen et al., 1974)
RpsUN2	PI 567139 B	16	(Lin *et al.*, 2013)
Rps10	Wandou 15	17	(Zhang et al., 2013b)
Rps4	PI 86050	18	(Athow et al., 1980)
Rps5	L62-904	18	(Buzzell, 1981)
Rps6	Altona	18	(Athow and Laviolette, 1982)
Rps12	PI 399036	18	(Sahoo et al., 2017)
RpsJS	Nannong 10-1	18	(Sun et al., 2014)
RpsYB30	Youbian 30	19	(Zhendong et al., 2010)

2.2 Summary of released high-quality soybean genomes

Soybean (*Glycine max*) is proposed to be domesticated from its wild relative, *Glycine soja*, ~6,000-9,000 years ago (Carter et al., 2004; Kim et al., 2012). Two rounds of whole genome duplication have occurred at approximately 59 and 13 million years ago, respectively, resulting in a highly duplicated genome with nearly 3/4 of the genes present in multiple copies (Schmutz et al., 2010). In 2010, Schmutz *et al*. reported the first chromosome-level reference genome of a soybean cultivar, Williams 82 (Schmutz *et al.*, 2010). However, a lot of sequencing gaps are still present in this reference genome, especially in genome regions enriched of nucleotide-binding leucine-rich repeat genes (NLR), due to the limitation in sequencing technology.

In 2018, Shen *et al*. reported the first high-quality soybean reference genome, Zhonghuang 13, assembled based on a combination of SMRT, Hi-C and optical mapping data (Shen et al., 2018) and most of the NLR gene clusters were well assembled (Table 2.2). In 2019, Xie et al. reported the first high-quality reference-grade genome of wild soybean (Table 2.2), W05, and a few structural variations has been identified compared to *G. max* genomes (Xie et al., 2019). Another milestone of soybean genome sequencing is the pan-genome of wild and cultivated soybeans, release by Liu *et al.* in 2020. The authors reported the high-quality genomes of 26 soybean accessions, including 3 wild soybeans, 9 landraces and 14 elite cultivars, which are selected as representatives of 2,898 accessions (Table 2.2). The breakthrough sequencing and assembly technologies have allowed us to decode complete complex NLR gene regions that are highly repetitive and variable in gene copy number and structure. For example, the physical distance of the NLR gene cluster in the *Rps11* region ranges from 422 kb in SoyW02 to 1,206 kb in SoyC10 (Table 2.2), suggesting complicate structural and copy number variations. Besides the 29 released soybean genomes, ~ 4,000 soybean accessions have been re-sequenced and several million single nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) have been identified (Liu et al., 2020; Valliyodan et al., 2021; Zhou et al., 2015), and over 20,000 soybean germplasm have been genotyped using the Illumina Infinium SoySNP6k BeadChip (Illumina, San Diego, Calif. USA) (Song et al., 2015), which have highly accelerated the genetic study and breeding of soybeans.

Table 2.2 Summary of released high-quality soybean genomes

Soybean Variety	Accession Name	Genome Size (Mb)	Length of *Rps11* Region (Kb)
SoyW01	PI 562565	1008.7	479.6
SoyW02	PI 549046	1005.1	422.1
SoyW03	PI 578357	1013.3	632.7
SoyL01	Zhutwinning2	996.9	628.5
SoyL02	Zi Hua No.4	1004.1	474.2
SoyL03	Tong Shan Tian E Dan	1033.9	453.8
SoyL04	58-161	1000.7	602.5
SoyL05	PI 398296	1051.5	459.5
SoyL06	Zhang Chun Man Cang Jin	996.9	521.3
SoyL07	Feng Di Huang	1003.4	531.7
SoyL08	Tie Jia Si Li Huang	998.5	532.8
SoyL09	Shi Sheng Chang Ye	1025.4	404.7
SoyC01	Xu Dou No.1	1000.2	834.5
SoyC02	Tie Feng No.18	1005.1	685.7
SoyC03	Ju Xuan No.23	999.5	713.3
SoyC04	Wan Dou No.28	995.4	517.5
SoyC05	Amsoy	988.3	642.5
SoyC06	Yu Dou No.22	1000.5	595.1
SoyC07	Jin Dou No.23	1006.0	817.6
SoyC08	Qi Huang No.34	995.6	517.0
SoyC09	Han Dou No.5	993.1	611.2
SoyC10	PI 548362	998.3	1206.6
SoyC11	Ji Dou No.17	1019.0	972.7
SoyC12	Dong Nong No.50	1015.6	521.0
SoyC13	Hei He No.43	1007.8	516.5
SoyC14	Ke Shan No.1	1003.9	530.4
ZH13	Zhong Huang 13	1011.2	471.9
W05	W05	1013.2	603.3
Williams 82	PI 518671	978.4	517.6

2.3 *Phytophthora soaje* and virulence genes

Phytophthora is a genus of plant-damaging oomycetes with more than 170 identified species (Goheen and Frankel, 2009), many of which are plant pathogens of considerable economic importance such as *Phytophthora infestans* that caused the Irish potato famine and subsequent

diaspora (Nowicki et al., 2012), and *Phytophthora sojae* that causes the Phytophthora stem and root Rot (PRSR) of soybean, which has been ranked as the second most destructive soybean disease since it was first found in Indiana in 1948 (Kaufmann and Gerdemann, 1958; Wrather and Koenning, 2009). The zoospores produced by *P. sojae* can be attracted by soybean roots in wet conditions, and then infect plant tissues. Infected soybean plants usually develop lesions on root and stem and the entire plants are eventually killed.

P. sojae is a diploid organism with a genome size ~95 Mb (Tyler et al., 2006) and the pathogenic mechanisms of *P. soja* have been extensively explored. Several hundreds of effector proteins have been identified to date. Among these effector proteins, a conserved RXLR motif has been proposed to play crucial role in virulence by *P. sojae* (Dou et al., 2008; Jiang et al., 2008). In addition, several dozen of *P. sojae* islates have been re-sequenced, which will benefit the studies of the natural diversity of P. sojae populations and guide better deployment of Rps genes towards effective management of PRSR.

CHAPTER 3. MATERIALS AND METHODS

3.1 Plant materials

The mapping populations were generated from an initial cross between PI 594527 and Williams. In 2015, 2640 F3 plants derived from heterozygous F2 individuals were screened for identification of recombinants. In 2016 and 2017, two additional larger populations, including 7680 and 6730 F4 plants respectively, derived from heterozygous F3 individuals, were screened for more recombinants (Figure 3.1). Only the recombinants with one side heterozygous and the other side homozygous Williams genotype were used for fine mapping, because the two expected phenotypes could be easily and accurately distinguished.

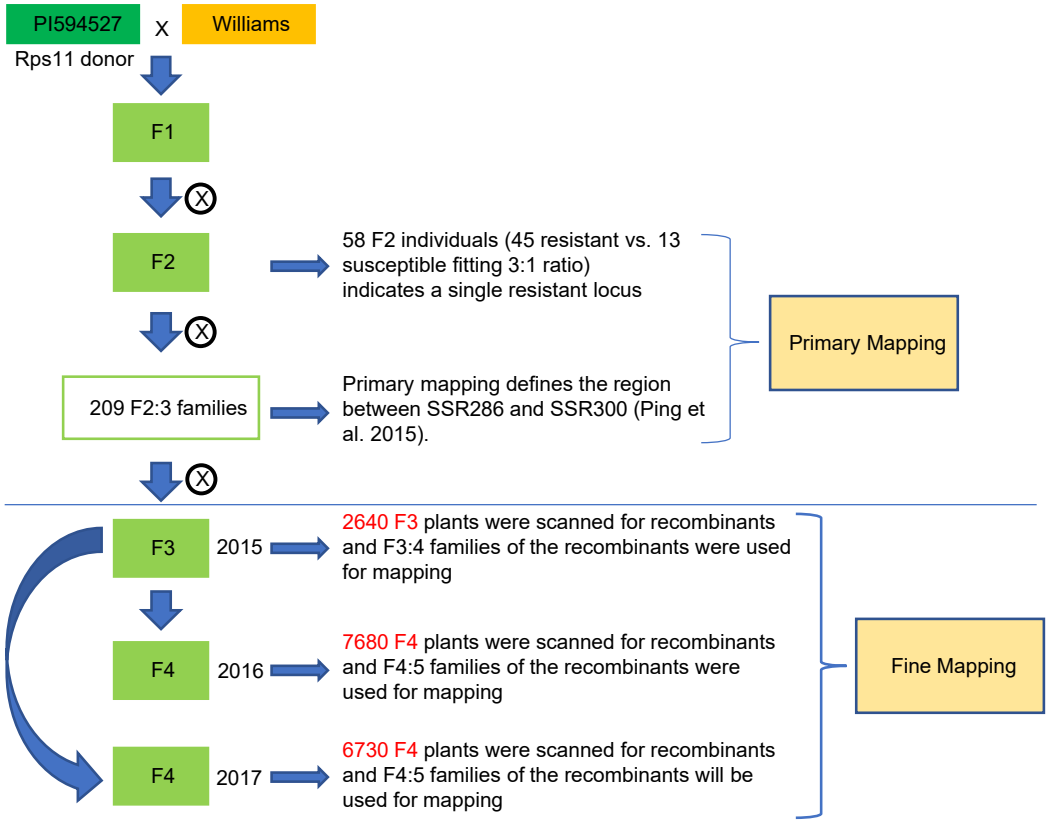


Figure 3.1 Flowchart of the *Rps11* fine mapping process and the plant materials used.

3.2 Isolates of *P. sojae* and resistance evaluation

Fourteen isolates of *P. sojae* (Table 3.1) were used for inoculation of a progeny populations, including 14 lines with *Rps11/Rps11* (PI 594527) genotype and 14 lines with *rps11/rps11* (Williams) genotype derived from the mapping population; 158 isolates collected from fields across Indiana (Figure 3.2 and Table 3.2) were used to further explore the resistance spectrum of the *Rps11* locus. Race 1 was used for the inoculation of all the recombinants.

For inoculating the recombinants, about 30 seedlings from each line were inoculated with *P. sojae* Race 1 using a protocol previously described (Dorrance *et al.*, 2007; Lin *et al.*, 2013). Lines with less than 25% of progenies surviving after inoculation were classified as susceptible (S); lines with more than 25% of progenies surviving were classified as segregating (H). Evaluation was repeated twice for each line. For resistance spectrum tests, lines with less than 25% of progenies surviving after inoculation were classified as susceptible (S), lines with 25% to 75% of progenies surviving after inoculation were classified as partially resistant (H), and lines with more than 75% of progenies surviving after inoculation were classified as completely resistant (R).

Table 3.1 List of P. sojae isolates used to inoculate segregating progeny population.

Isolates of P. sojae
Race 1
Race 3
Race 4
Race 7
Race 25
OH001
OHC2S1
OH003
MINI2004.03.01
MINI2004.01.01
MINI2002.01.05
MINI2002.05.01
MINI2005.07.02
MINI2002.05.05

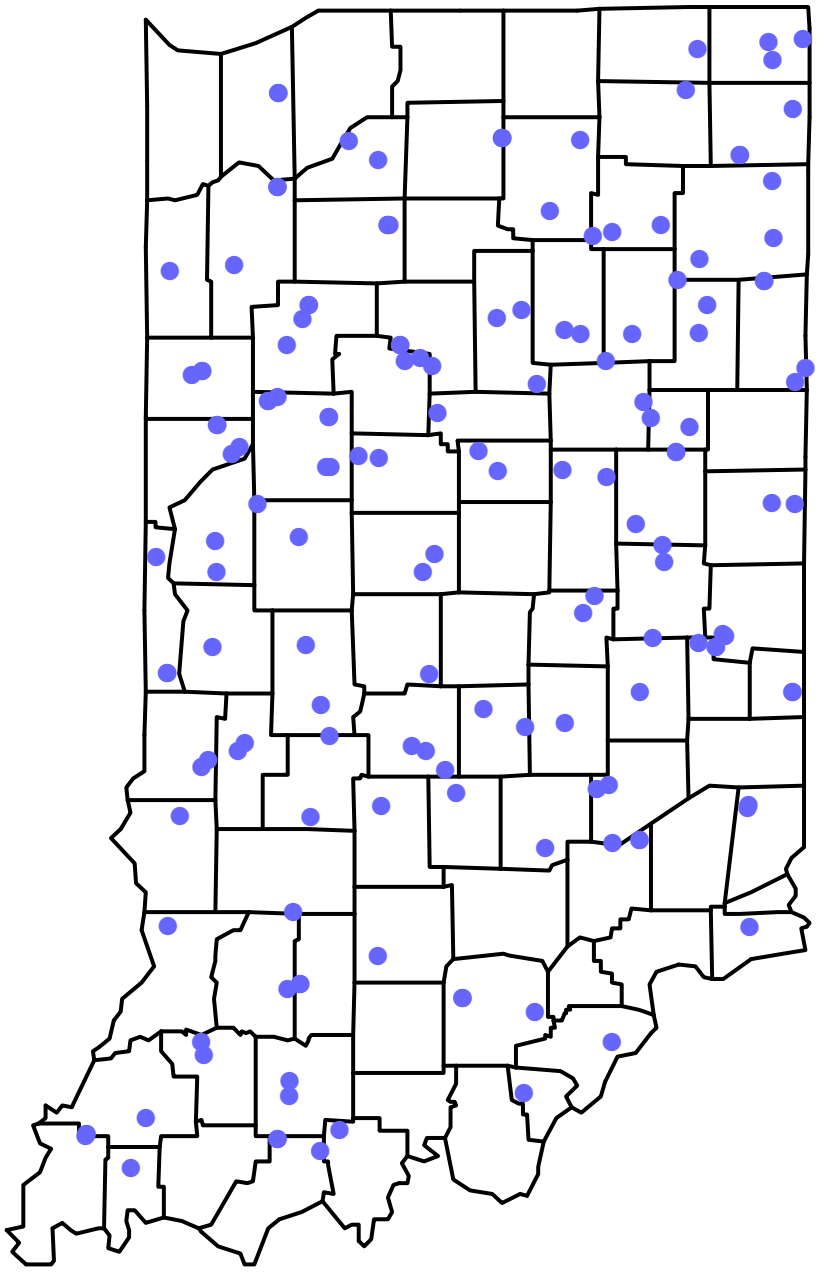


Figure 3.2 Geographic distrubution of the 158 isolates collected from Indiana.

Table 3.2 Informaiton of the 158 isolates collected from indiana.

Isolates	Latitude	Longitude	City	County
1 A-1	N 40.90496	W 084.98903	Decatur	Adams
2 A-1	N 40.90496	W 084.98903	NA	NA
3 B-1	N 40.58507	W 084.86088	Geneva	Adams
4 E-2	N 40.63035	W 084.81773	Berne	Adams
5 A-1	N 41.22283	W 084.95517	Grabill	Allen
6 E-2	N 41.04152	W 084.95053	New Haven	Allen
7 E-2	N 40.97488	W 085.25479	Roanoke	Allen
8 A-1	N 40.91072	W 085.34553	Roanoke	Huntington
9 A-1	N 40.82905	W 086.86231	Monon	White
9 A-1	N 40.82905	W 086.86231	Monon	White
10 B-1	N 40.78548	W 086.88977	Reynolds	White
11 B-1	N 40.70295	W 086.95348	Reynolds	White
12 E-2	N 40.93737	W 087.43598	Morocco	Newton
14 A-1	N 41.34976	W 086.69878	Brems	Starke
15 A-1	N 41.28823	W 086.57769	Knox	Starke
16 C-3	N 41.08433	W 086.53154	Monteray	Pulaski
16b C-2	N 41.08380	W 086.53840	Monteray	Pulaski
16b C-2	N 41.08380	W 086.53840	Monteray	Pulaski
16b H-5	N 41.08380	W 086.53840	Monteray	Pulaski
17 A-1	N 39.42980	W 086.43896	Martinsville	Morgan
18 H-3	N 39.41669	W 086.38110	Martinsville	Morgan
19 A-1	N 39.35630	W 086.30166	Morgantown	Morgan
20 B-2	N 39.28340	W 086.25600	Beanblossom	Brown
21 A-1	N 39.30665	W 085.62824	Greensburg	Decatur
22 A-1	N 39.29596	W 085.67756	Greensburg	Decatur
25 C-2	N 39.10662	W 085.89056	Waynesville	Bartholomew
29 A-2	N 39.13217	W 085.50211	North Vernon	Jennings
29 E-2	N 39.13217	W 085.50211	North Vernon	Jennings
30 C-1	N 39.12446	W 085.61291	North Vernon	Jennings
32 B-1	N 40.66336	W 086.40500	Deer Creek	Carroll
33 B-1	N 40.65193	W 086.46801	Deer Creek	Carroll
35 A-1	N 40.70462	W 086.48701	Clymers	Cass
35 F-3	N 40.70462	W 086.48701	Clymers	Cass
36 A-1	N 41.36051	W 086.06642	Bourbon	Marshall
36 A-1	N 41.36051	W 086.06642	Bourbon	Marshall
36 B-2	N 41.36051	W 086.06642	Bourbon	Marshall
37 E-3	N 41.35224	W 085.74662	Syracuse	Kosciusko
38 G-4	N 41.04783	W 085.69312	North Manchester	Kosciusko

Table 3.2 continued

39 B-2	N 41.12950	W 085.87054	Claypool	Kosciusko
40 G-4	N 41.51197	W 085.31149	Kendallville	Nobel
42 B-1	N 41.45224	W 084.87048	Butler	Dekalb
43 A-1	N 41.30648	W 085.08826	Auburn	Dekalb
43 B-1	N 41.30648	W 085.08826	Auburn	Dekalb
44 B-1	N 40.81373	W 085.98716	Peru	Miami
45 B-1	N 40.58074	W 085.92374	Converse	Miami
46 C-2	N 40.78996	W 086.08952	Peru	Miami
47 A-1	N 40.34442	W 086.57448	Frankfort	Clinton
48 A-1	N 40.44903	W 087.24055	Pine Village	Warren
48 B-1	N 40.44903	W 087.24055	Pine Village	Warren
49 A-1	N 40.38105	W 087.14914	Otterbein	Warren
50 B-1	N 40.35755	W 087.17988	Attica	Warren
51 M-4	N 40.03183	W 087.49126	Perrysville	Vermillion
52 B-1	N 39.66355	W 087.44700	Clinton	Vermillion
52 B-1	N 39.66355	W 087.44700	Clinton	Vermillion
54 F-4	N 40.65244	W 085.64075	La Fontaine	Grant
55 A-3	N 40.52213	W 085.48476	Upland	Grant
56 C-1	N 41.50157	W 086.98971	Valparasio	Porter
57 E-1	N 41.50291	W 086.98766	Valparasio	Porter
58 J-4	N 41.20449	W 086.99295	Wheatfield	Jasper
58 L-5	N 41.20449	W 086.99295	Wheatfield	Jasper
59 A-1	N 40.95789	W 087.17056	Rensselear	Jasper
61 A-1	N 39.24440	W 085.05331	Sunman	Dearborn
62 B-2	N 39.23475	W 085.05738	Sunman	Dearborn
65 B-1	N 38.32988	W 085.97841	Georgetown	Floyd
67 F-3	N 38.49106	W 085.6161	Charlestown	Clark
69 E-2	N 38.58755	W 085.9325	Salem	Washington
70 A-1	N 38.63133	W 086.2307	Campbellsburg	Washington
70 B-1	N 38.63133	W 086.2307	Campbellsburg	Washington
71 D-1	N 38.85856	W 085.0488	Bennington	Switzerland
72 B-1	N 40.60660	W 087.34538	Fowler	Benton
73 E-2	N 40.62072	W 087.30252	Fowler	Benton
73 E-2	N 40.62072	W 087.30252	Fowler	Benton
75 E-3	N 39.90692	W 085.68655	Greenfield	Hancock
77 F-4	N 39.85382	W 085.73348	Greenfield	Hancock
79 A-1	N 40.35025	W 086.65878	Mulberry	Clinton
82 A-1	N 40.63588	W 086.35467	Camden	Cass
85 A-1	N 40.47463	W 086.78065	Buck Creek	Tippecanoe
85 B-2	N 40.47463	W 086.78065	Buck Creek	Tippecanoe

Table 3.2 continued

86 A-1	N 40.31692	W 086.79200	Lafayette	Tippecanoe
87 A-1	N 40.31547	W 086.77392	Lafayette	Tippecanoe
89 G-2	N 40.52630	W 087.03092	West Lafayette	Tippecanoe
89 I-3	N 40.52630	W 087.03092	West Lafayette	Tippecanoe
90 B-2	N 40.53872	W 086.99163	West Lafayette	Tippecanoe
91 A-1	N 40.36445	W 085.35042	Eaton	Delaware
91 B-1	N 40.36445	W 085.35042	Eaton	Delaware
92 C-2	N 40.44283	W 085.29508	Hartford City	Blackford
93 E-3	N 40.47115	W 085.45500	Upland	Grant
94 E-3	N 39.78147	W 085.14852	Milton	Wayne
95 E-2	N 39.78807	W 085.15885	Milton	Wayne
96 A-1	N 39.74610	W 085.18735	Bentonville	Fayette
96 C-2	N 39.74610	W 085.18735	Bentonville	Fayette
97 A-1	N 39.75875	W 085.25815	Bentonville	Fayette
98 E-4	N 40.01417	W 085.40000	New Castle	Henry
99 E-3	N 39.60323	W 084.87140	Liberty	Union
99 G-3	N 39.60323	W 084.87140	Liberty	Union
101 A-1	N 40.19745	W 084.86183	Union City	Randolph
102 A-1	N 40.20342	W 084.95630	Winchester	Randolph
103 C-1	N 39.24195	W 086.56600	Bloomington	Monroe
104 F-4	N 38.76487	W 086.57853	Mitchell	Lawrence
105 A-1	N 38.67768	W 086.89850	Loogootee	Martin
105 K-4	N 38.67768	W 086.89850	Loogootee	Martin
106 A-1	N 38.66230	W 086.95095	Cannelburg	Daviess
107 C-1	N 38.49293	W 087.30573	Petersburg	Pike
108 C-1	N 38.45258	W 087.29508	Willisville	Pike
109 B-1	N 38.25170	W 087.53410	Fort Branch	Gibson
110 A-1	N 38.09208	W 087.59508	Darmstadt	Vanderburgh
111 A-1	N 38.19442	W 087.78283	Poseyville	Posey
113 B-2	N 38.20110	W 087.77780	Poseyville	Posey
113 E-2	N 38.20110	W 087.77780	Poseyville	Posey
114 F-3	N 40.19990	W 087.07383	Wingate	Montgomery
116 A-1	N 40.08292	W 087.24822	Veedersburg	Fountain
117 N-7	N 39.98230	W 087.24302	Kingman	Fountain
119 C-1	N 39.74553	W 087.25920	Rockville	Parke
120 D-1	N 39.44080	W 087.12565	Prairie City	Clay
121 B-1	N 39.41475	W 087.15493	Prairie City	Clay
122 A-1	N 39.38698	W 087.27763	Riley	Vigo
123 A-1	N 39.36475	W 087.30502	Riley	Vigo
124 C-1	N 39.20808	W 087.39408	Farmersburg	Sullivan

Table 3.2 continued

126 B-1	N 38.86130	W 087.44358	Oaktown	Knox
128 C-2	N 40.09372	W 086.90425	Crawfordsville	Montgomery
129 D-2	N 39.75167	W 086.87467	Bainbridge	Putnam
130 C-2	N 39.56182	W 086.81330	Cloverdale	Putnam
131 D-1	N 39.46290	W 086.77755	Cloverdale	Owen
132 C-1	N 39.20592	W 086.85548	Freedom	Owen
134 B-1	N 38.90387	W 086.92720	Scotland	Green
135 F-2	N 38.36815	W 086.94225	Jasper	Dubois
136 F-2	N 38.32128	W 086.94358	Huntingburg	Dubois
137 B-1	N 38.18507	W 086.99177	Dale	Spencer
137 B-1	N 38.18507	W 086.99177	Dale	Spencer
138 C-2	N 38.14598	W 086.81605	St Meinrad	Spencer
139 D-2	N 38.21225	W 086.73608	Siberia	Perry
140 B-1	N 40.74955	W 085.80962	Wabash	Wabash
141 B-2	N 40.73915	W 085.74463	Wabash	Wabash
142 E-2	N 40.73718	W 085.53127	Mt. Edna	Huntington
143 C-1	N 40.74217	W 085.25655	Bluffton	Wells
144 D-2	N 40.83105	W 085.22252	Uniondale	Wells
145 A-1	N 41.06098	W 085.61470	South Whitley	Whitley
146 B-1	N 41.08325	W 085.41317	Columbia City	Whitley
147 C-1	N 41.67445	W 084.82850	York	Steuben
148 D-2	N 41.66560	W 084.97020	Angola	Steuben
149 B-1	N 41.60707	W 084.95347	Angola	Steuben
150 B-3	N 41.64147	W 085.26250	Brushy Prairie	Lagrange
152 A-1	N 40.48882	W 086.33325	Burlington	Howard
154 B-2	N 40.36760	W 086.16502	Sharpsville	Tipton
155 A-1	N 40.30328	W 086.08532	Tipton	Tipton
156 A-1	N 40.30565	W 085.81825	Elwood	Madison
157 A-1	N 40.28483	W 085.63820	Alexandria	Madison
158 A-1	N 40.13432	W 085.51670	Daleville	Delaware
159 A-1	N 40.06892	W 085.40708	Springport	Henry
160 A-1	N 39.77547	W 085.44742	Mays	Rush
161 A-1	N 39.60178	W 085.49995	Rushville	Rush
163 B-1	N 39.50458	W 085.80985	Shelbyville	Shelby
164 B-2	N 39.49183	W 085.97295	Franklin	Johnson
165 C-2	N 39.54817	W 086.14408	Whiteland	Johnson
166 A-1	N 39.66050	W 086.36810	Mooresville	Hendricks
167 A-1	N 39.98503	W 086.39375	Whitestown	Boone
168 B-1	N 40.04177	W 086.34562	Gadsden	Boone
R2T21 A-1	NA	NA	NA	NA

3.3 Genotyping the recombinants.

SSR markers and insertion/deletion markers (InDel) were identified and designed based on re-sequencing data of the two parental lines. Marker_176kb was a dominant marker that could only be amplified from the donor line. Only the markers with a unique hit at the *Rps11* region were used for fine mapping. Kompetitive allele specific PCR (KASP) makers were also used to identify and genotype the recombinants from the 2017 mapping population. All markers used in this study are listed in Table 3.3.

Table 3.3 List of primers and sequences used for mapping and expression analysis.

Primer Name	Primer Sequences
SSR-07-286F	AAAAATCAGCACCCATCGAC
SSR-07-286R	AGCCCTGGCCTTATTTTGTT
SSR-07-295F	CTCTCCTTTCATTCCCCACA
SSR-07-295R	TTCTTGGAGCTTCGGAGGTA
InDel-626F	GAACTCCACTTAATCATCTCAC
InDel-626R	TTCACTCCGTCCTCGGCGGCG
InDel-43F	ATTTCCTAATTAAGTGAAAGTTTGAAATGTTATATTA
InDel-43R	GATTTATCACACTATCAAAGTGTATGAC
SSR-300F	TCGCAATATTGGCTACGATG
SSR-300R	CTGAAAACAAAATAAAAGAGAACAAA
Marker176F	CTCTGTCCCCACCTCTCC
Marker176R	CATGGTCAGTTTGATAGC
InDel-327F	TAAGTGATTCGTTTGAGTCCT
InDel-327R	TATGGTGTGGCTATGGAGATTG
InDel-5.92F	GCATCAACACTTGGCGCAAGC
InDel-5.92R	GGATAATGCGATAATTGTTCTAGC
InDel-6.04F	AAATATAGCACCCTTTAGAG
InDel-6.04R	AGCCTCACTCTCCACAT
SSR-320F	TTTAACTGAAAATACTCCGGCA
SSR-320R	TCATAATTTAAGAGACCAAACCGA
qRT-PCR-F	TGTGAACATTCGTAGTTGTC
qRT-PCR-R	TTCCACTGACTCACAAAAAG
GmActin11F	CGGTGGTTCTATCTTGGCATC
GmActin11R	GTCTTTCGCTTCAATAACCCTA

3.4 Long and short read genome sequencing

Long-read data was generated using the Pacific BioSciences (Menlo Park, CA, USA) Sequel platform at Corteva Agriscience™ (Johnston, IA, USA). Eight SMRT cells were performed with 10-hr movies and v6 chemistry. Raw subreads were filtered to a minimum of 12 kb generating 77× genome coverage. The raw subread N50 length was 28.9 kb. Linked short-read data were generated by sequencing 10X Genomics (Pleasanton, California) Chromium libraries at Corteva Agriscience™ on the Illumina (San Diego, California) HiSeq2500 platform in a PE151 configuration. The coverage depth and mean molecule length for the Chromium library were 45.2× and 93.8 kb, respectively.

3.5 Genome assembly and sequence polishing

Canu (Koren et al., 2017) v1.8 (https://github.com/marbl/canu) was used to self-correct the raw subreads and to assemble the corrected reads into contigs. The following changes were made to the default parameters: correctedErrorRate=0.065, corMhapSensitivity=normal, and ovlMerDistinct=0.99. A minimum contig length of 30 kb was applied. Additional sequence polishing was performed by aligning raw PacBio subreads to the contig assembly using pbmm2 v0.12.0 (https://github.com/PacificBiosciences/pbmm2) and applying the Arrow algorithm from the Genomic Consensus package v2.3.2 (https://github.com/PacificBiosciences/GenomicConsensus) to identify and correct remaining consensus errors in the contigs. These tools were acquired from pbbioconda (https://github.com/PacificBiosciences/pbbioconda). The consensus sequence accuracy was further enhanced by complementing the long read contig assembly with Chromium linked short-reads. Chromium datasets were aligned to contigs using Long Ranger v2.2.2. The sequence assembly polishing tool Pilon (Walker et al., 2014) v1.22 (https://github.com/broadinstitute/pilon) was used to correct individual base errors and small indels from the Chromium data aligned to the contigs using the "--fix bases –minmq 30" parameters.

3.6 Creating genome maps

Genome maps were generated in the Bionano Saphyr platform at Corteva Agriscience™ using the Direct Label and Stain (DLS) approach (Ou et al., 2020). Nuclear DNA was isolated from leaf

tissue using a modified version of the Bionano Prep™ Plant Tissue DNA Isolation protocol (https://Bionanogenomics.com/wp-content/uploads/2017/01/30068-Bionano-Prep-Plant-Tissue-DNA-Isolation-Protocol.pdf) that did not include a gradient centrifugation step. DLE-1–labeled molecule data were filtered to create a dataset with a molecule N50 of 441 kb and 267× coverage. This dataset was assembled via the Bionano Genomics Access software platform (Solve3.2.2_08222018) with the configuration file optArguments_nonhaplotype_noES_noCut_DLE1_saphyr.xml. The resulting genome maps were filtered to remove coverage and length outliers. The final genome map assembly consisted of 45 maps with a genome map N50 of 26.7 Mb and a total map length of 985 Mb.

3.7 Hybrid scaffolding of genome maps with sequence contigs.

Hybrid scaffolds were generated from the polished contigs and the Bionano genome maps using the Bionano Genomics Access software (Solve3.3_10252018) and the DLE-1 configuration file hybridScaffold_DLE1_config.xml. In addition to auto-conflict resolution performed by the software, manual curation was performed to resolve overlapping and embedded contigs by providing additional "Conflict resolutions" and re-running the hybrid scaffolding. In the final product, the assembly had 43 hybrid scaffolds (Scaffold N50= 26.4 Mb, Total scaffold length= 978.1 Mb) with 229 leftover contigs that were not scaffolded with a combined length of 21.3 Mb.

3.8 Building chromosome-scale pseudomolecules

A reference-based approach was feasible to create chromosome-scale pseudomolecules using the Glycine max Wm82.a2.v1 reference assembly (https://phytozome.jgi.doe.gov/), because only an average of 2.15 scaffolds per chromosome needing to be placed. To map hybrid scaffolds to the reference, each scaffold was chunked into 100-bp fragments and then aligned to the reference genome using minimap2 (Li, 2018) v2.10 (https://github.com/lh3/minimap2). Then, a custom script was used to determine the chromosome position and orientation for each scaffold based on the alignment of each "chunked scaffold cloud". All scaffolds were able to be placed using this method. Leftover unscaffolded contigs were concatenated with 100-bp N-gaps and assigned to Chr00.

3.9 NLR gene annotation and expression analysis

NLR genes were annotated using NLR-Annotator(Steuernagel et al., 2020). RNA samples were extracted from mixed stem tissues from seedlings of each key recombinants using RNeasy Plant Mini Kit (Cat No. 74904, Qiagen) and were treated with RNase-Free DNase Set (Cat No. 79254, Qiagen) to remove DNA. RNA-seq was performed by Purdue Genomic Core Facility (https://www.purdue.edu/hla/sites/genomics/). RNA-seq data were mapped to the genome of the donor line using STAR (Dobin et al., 2013) and expression was calculated based on the number of reads mapped to each NLR gene.

3.10 Plasmid Construction and Transformation

To make the over-expression construct for the *Rps11* candidate gene R6, the CDS of R6 was synthesized by Genscript as 3 fragments and assembled with AtUbi3 promoter and Gateway ATT sites by homologous recombination in yeast to make a Gateway entry vector (Muller et al., 2012). This was recombined into a Gateway destination vector by LR reaction using Gateway Technology with Clonase II (25-0749, Invitrogen) for transformation into *Ochrobactrum*. *Ochrobactrum*-mediated soybean embryonic axis transformation was done as previously described (US20180216123A1; WO2020/005933A1; WO2020/092494A1). Mature dry seeds of soybean cultivar 93Y21 were disinfected using chlorine gas and imbibed on semi-solid medium containing 5g/l sucrose and 6 g/l agar at room temperature in the dark. After an overnight incubation, the seed was soaked in distilled water for an additional 3-4 hrs at room temperature in the dark. Intact embryonic axis explants were isolated and transferred to the deep plate with 15 mL of *Ochrobactrum haywardense* H1-8 suspension (OD 0.5 at 600 nm) in infection medium composed of 1/10X Gamborg B5 basal medium, 30 g/L sucrose, 20 mM MES, 0.25 mg/L GA3, 1.67 mg/L BAP, 200 µM acetosyringone and 1 mM dithiothreitol in pH 5.4. The plates were sealed with parafilm (“Parafilm M” VWR Cat#52858), then sonicated (Sonicator-VWR model 50T) for 30 seconds. After sonication, embryonic axis explants were transferred to a single layer of autoclaved sterile filter paper (VWR#415/Catalog # 28320-020). The plates were sealed with Micropore tape (Catalog # 1530-0, 3M, St. Paul, MN)) and incubated under dim light (5-10 µE/m2/s, cool white fluorescent lamps) for 16 hrs at 21°C for 3 days.

After co-cultivation, the base of each embryonic axis was embedded in shoot induction medium (R7100, PhytoTech Labs) containing 30 g/L sucrose, 6 g/L agar and 25 mg/L spectinomycin (S742, PhytoTech Labs) as a selectable agent and 500 mg/L cefotaxime (GoldBio, ST Louis, MO, USA) in pH5.7. Shoot induction was carried out at 26°C with a photoperiod of 18hrs and a light intensity of 40-70 µE/m2/s. After 4-6 weeks in selection medium, the spectinomycin-resistant shoots were cut and transferred to ½ strength MS rooting medium (M404, PhytoTech Labs) containing 15 g/L sucrose, agar 6 g/L, 10 mg/L spectinomycin and 250 mg/L cefotaxime for further shoot and root elongations.

Marker-free transgenic soybean plants were generated by the Cre-lox site-specific recombination system using heat shock treatment. For heat shock treatment of soybean, 2-4 cm T0 plantlets with roots in 100 x 25 mm Petri dishes on spectinomycin free-rooting medium were transferred into a Percival incubator (Percival Scientific, Perry, IA, USA) at 45°C, 70 % humidity for 2 hrs in the dark. After the heat shock treatment, T0 plantlets were transferred to moistened Berger BM2 soil (Berger, Saint-Modeste, QC, Canada), and kept enclosed in clear plastic tray boxes in a Percival incubator at 26°C with a 16 hr photoperiod at 250-350 µE/m2/s. Leaf punch samples were collected for molecular analyses from newer growth 2 weeks after acclimatization of T0 events. Hardened plantlets were potted in 2-gallon pots containing moistened SunGro 702 and grown to maturity for harvest in a greenhouse. The presence of the construct in the transgenic plants was confirmed by PCR with primers specific to the cloning vector and expression analysis of R6 in the transgenic plants by qPCR.

3.11 Data access

All the Pacbio raw sequence data, the genome assembly, Illumina short-reads sequencing data and RNA-seq data from this article have been deposited in National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) database under BioProject PRJNA718574.

3.12 Material availability

Novel biological materials described in this publication may be available to the academic community and other not-for-profit institutions solely for non-commercial research purposes upon

acceptance and signing of a material transfer agreement between the author's institution and the requestor. In some cases, such materials may originally contain genetic elements described in the manuscript that were obtained from a third party(s), and the authors may not be able to provide materials including third party genetic elements to the requestor because of certain third-party contractual restrictions placed on the author's institution. In such cases, the requester will be required to obtain such materials directly from the third party. The authors and authors' institution do not make any express or implied permission(s) to the requester to make, use, sell, offer for sale, or import third party proprietary materials. Obtaining any such permission(s) will be the sole responsibility of the requestor. In order to protect Corteva Agriscience™ proprietary germplasm, such germplasm will not be made available except at the discretion of Corteva Agriscience™ and then only in accordance with all applicable governmental regulations. Plant germplasm and transgenic material will not be made available except at the discretion of the owner and then only in accordance with all applicable governmental regulations.

CHAPTER 4. RESULTS

4.1 *Rps11* shows broad-spectrum resistance.

In an effort to identify novel sources of resistance to *P. sojae*, we identified a soybean landrace, PI 594527, possessing broad-spectrum resistance (Figure 4.1). A single Rps locus, *Rps11*, had been mapped onto chromosome 7 using a population derived from a cross between PI 594527 and Williams (rps) in our previous study (Ping *et al.*, 2016).

Figure 4.1 Resistance spectrum of PI 594527.

Table 4.1 Resistance spectrum of *Rps11* locus to 14 *P. sojae* isolate based on inoculation of a progeny population.

Genotype	Race1	Race4	Race7	Race25	Race3	OH001	OHC2S1	OH003	MIN1 2004.03.01	MIN1 2004.01.01	MIN1 2002.01.05	MIN1 2002.05.01	MIN1 2005.07.02	MIN1 2002.05.05
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
rps11	S	S	S	S	S	S	S	S	S	S	S	S	S	S
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Rps11	R	R	R	R	R	R	R	R	R	R	R	R	R	R

To confirm that the *Rps11* locus is solely responsible for the broad-spectrum resistance, we inoculated a subpopulation, including 14 progeny lines with *Rps11/Rps11* genotype (PI 594527) and 14 lines with *rps11/rps11* genotype (Williams), using 14 isolates of *P. sojae*. We observed perfect co-segregations between the presence/absence of *Rps11* and the phenotype after inoculation, suggesting that *Rps11* alone is responsible for the resistance in the donor line (Table 4.1). To further explore the resistance spectrum of *Rps11*, we inoculated a progeny line with *Rps11* using 158 additional isolates collected from all Indiana counties. We found that *Rps11* was resistant to 127 out of the 158 of these isolates (Figure 4.2 and Table 4.2). We also compared the resistance spectrum of *Rps11* with two novel Rps loci, *RpsUN1* and *RpsUN2*, previous identified in our lab. It shows that the spectrum of *RpsUN1*, *RpsUN2* and *Rps11* is significantly higher than the control line Williams (*rps*) ($P = 8.6\text{x}10^{-46}$, $3.8\text{x}10^{-51}$ and $5.4\text{x}10^{-63}$, respectively) (Figure 4.3). The spectrum of Rps11 is also significantly higher than that of RpsUN1 and RpsUN2 ($P = 3.6\text{x}10^{-6}$, 0.006, respectively) (Figure 4.3). Taken together, our results demonstrate that *Rps11* possesses a broad resistance spectrum to *P. sojae* and shows excellent potential in managing PRSR. Therefore, cloning of the resistant gene underlying *Rps11* will facilitate its deployment in soybean cultivars.

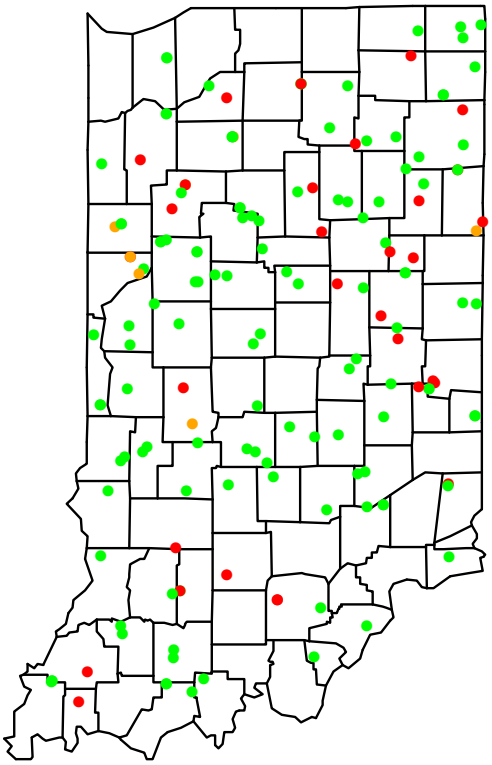


Figure 4.2 Distrabution of the 158 isolates resistant or susceptible to *Rps11*.

Green dots represent isolates that *Rps11* is resistant to, orange dots represents isolates that *Rps11* partially resistant to, red dots represent isolates that *Rps11* are susceptible to.

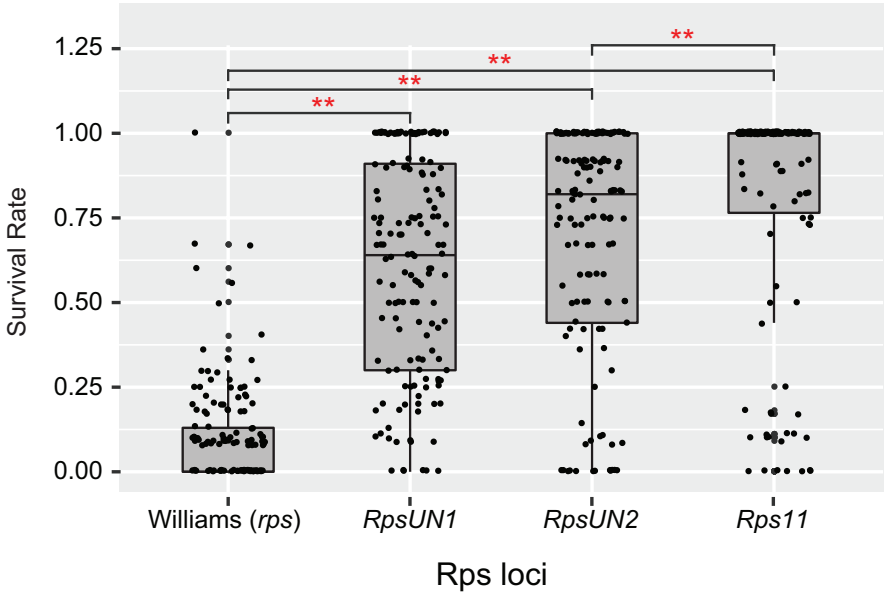


Figure 4.3 Resistance spectrum of *RpsUN1*, *RpsUN2* and *Rps11*.

Red asterisks indicate significance level (P<0.01).

Table 4.2 Resistance spectrum of *Rps11* to 158 *P. sojae* isolates collected from Indiana.

Isolate ID	Williams (Control)		RIL *(Rps11)*	
	Dead/Total	Dead rate	D/T	Dead Rate
1 A-1	10/11	91%	9/10	90%
2 A-1	12/12	100%	1/9	11%
3 B-1	9/11	82%	3/11	27%
4 E-2	9/9	100%	8/9	89%
5 A-1	11/12	92%	10/12	83%
6 E-2	9/9	100%	1/11	9%
7 E-2	10/10	100%	0/6	0%
8 A-1	9/11	82%	1/12	8%
9 A-1	8/11	73%	4/8	50%
9 A-1	7/10	70%	6/8	75%
10 B-1	7/8	88%	0/11	0%
11 B-1	11/12	92%	11/11	100%
12 E-2	9/9	100%	0/9	0%
14 A-1	9/10	90%	0/11	0%
15 A-1	12/12	100%	10/12	83%
16 C-3	11/12	92%	5/11	45%
16b C-2	4/10	40%	0/10	0%
16b C-2	6/10	60%	0/9	0%
16b H-5	11/11	100%	0/12	0%
17 A-1	10/10	100%	0/10	0%
18 H-3	11/11	100%	0/11	0%
19 A-1	12/12	100%	1/11	9%
20 B-2	10/11	91%	0/12	0%
21 A-1	10/10	100%	0/12	0%
22 A-1	12/12	100%	0/12	0%
25 C-2	10/13	77%	0/12	0%
29 A-2	11/12	92%	0/9	0%
29 E-2	5/10	50%	0/12	0%
30 C-1	9/10	90%	0/12	0%
32 B-1	11/12	92%	0/12	0%
33 B-1	8/10	80%	0/9	0%
35 A-1	11/11	100%	0/10	0%
35 F-3	11/11	100%	0/11	0%
36 A-1	0/12	0%	0/10	0%
36 A-1	4/9	44%	0/11	0%
36 B-2	10/10	100%	12/12	100%
37 E-3	11/11	100%	0/11	0%
38 G-4	9/9	100%	9/10	90%
39 B-2	12/12	100%	0/10	0%

Table 4.2 Continued

40 G-4	10/10	100%	9/9	100%
42 B-1	9/9	100%	0/11	0%
43 A-1	11/11	100%	0/10	0%
43 B-1	9/9	100%	0/7	0%
44 B-1	8/8	100%	9/9	100%
45 B-1	9/10	90%	8/8	100%
46 C-2	11/11	100%	0/11	0%
47 A-1	11/11	100%	0/12	0%
48 A-1	11/11	100%	12/12	100%
48 B-1	8/9	89%	5/9	56%
49 A-1	11/12	92%	0/9	0%
50 B-1	11/11	100%	3/10	30%
51 M-4	9/11	82%	0/10	0%
52 B-1	7/11	64%	0/9	0%
52 B-1	11/11	100%	0/11	0%
54 F-4	10/10	100%	2/11	18%
55 A-3	10/10	100%	0/9	0%
56 C-1	12/12	100%	0/9	0%
57 E-1	12/12	100%	0/9	0%
58 J-4	8/11	73%	0/10	0%
58 L-5	9/10	90%	0/9	0%
59 A-1	11/11	100%	9/10	90%
61 A-1	7/7	100%	9/9	100%
62 B-2	9/9	100%	0/8	0%
65 B-1	9/9	100%	0/10	0%
67 F-3	10/10	100%	0/10	0%
69 E-2	11/11	100%	0/10	0%
70 A-1	8/11	73%	10/10	100%
70 B-1	12/12	100%	12/12	100%
71 D-1	9/9	100%	0/11	0%
72 B-1	12/12	100%	3/11	27%
73 E-2	3/9	33%	0/7	0%
73 E-2	11/11	100%	0/12	0%
75 E-3	7/9	78%	0/8	0%
77 F-4	9/9	100%	0/10	0%
79 A-1	9/10	90%	0/12	0%
82 A-1	10/11	91%	0/11	0%
85 A-1	12/12	100%	1/9	11%
85 B-2	9/10	90%	0/11	0%
86 A-1	10/11	91%	0/11	0%
87 A-1	8/10	80%	0/12	0%
89 G-2	4/12	33%	2/10	20%

Table 4.2 continued

89 I-3	10/11	91%	0/11	0%
90 B-2	12/12	100%	0/11	0%
91 A-1	8/12	67%	0/11	0%
91 B-1	11/11	100%	2/11	18%
92 C-2	10/10	100%	9/11	82%
93 E-3	9/12	75%	9/10	90%
94 E-3	10/10	100%	9/9	100%
95 E-2	12/12	100%	8/8	100%
96 A-1	11/11	100%	0/11	0%
96 C-2	9/12	75%	0/11	0%
97 A-1	10/11	91%	8/8	100%
98 E-4	9/11	82%	10/10	100%
99 E-3	5/7	71%	0/10	0%
99 G-3	11/11	100%	0/11	0%
101 A-1	8/10	80%	0/10	0%
102 A-1	10/12	83%	0/10	0%
103 C-1	7/9	78%	0/11	0%
104 F-4	12/12	100%	8/9	89%
105 A-1	7/8	88%	0/6	0%
105 K-4	12/12	100%	12/12	100%
106 A-1	10/10	100%	0/11	0%
107 C-1	8/9	89%	0/12	0%
108 C-1	9/11	82%	0/6	0%
109 B-1	8/10	80%	10/10	100%
110 A-1	9/12	75%	13/13	100%
111 A-1	10/11	91%	0/10	0%
113 B-2	9/12	75%	0/12	0%
113 E-2	8/12	67%	0/14	0%
114 F-3	11/11	100%	0/14	0%
116 A-1	9/10	90%	0/11	0%
117 N-7	12/13	92%	0/12	0%
119 C-1	9/9	100%	0/11	0%
120 D-1	6/8	75%	0/9	0%
121 B-1	8/9	89%	0/10	0%
122 A-1	9/9	100%	2/8	25%
123 A-1	11/12	92%	0/11	0%
124 C-1	9/10	90%	0/9	0%
126 B-1	12/12	100%	0/9	0%
128 C-2	7/9	78%	0/6	0%
129 D-2	8/10	80%	5/6	83%
130 C-2	9/12	75%	5/10	50%
131 D-1	10/11	91%	0/8	0%

Table 4.2 continued

132 C-1	8/8	100%	0/10	0%
134 B-1	10/11	91%	9/9	100%
135 F-2	7/8	88%	0/11	0%
136 F-2	7/8	88%	1/11	9%
137 B-1	7/10	70%	0/11	0%
137 B-1	11/12	92%	2/12	17%
138 C-2	12/12	100%	0/9	0%
139 D-2	10/11	91%	0/11	0%
140 B-1	10/10	100%	0/10	0%
141 B-2	10/11	91%	2/11	18%
142 E-2	10/10	100%	0/8	0%
143 C-1	9/11	82%	7/7	100%
144 D-2	10/10	100%	0/12	0%
145 A-1	12/12	100%	0/12	0%
146 B-1	10/11	91%	0/7	0%
147 C-1	10/10	100%	2/11	18%
148 D-2	10/10	100%	0/9	0%
149 B-1	11/11	100%	0/11	0%
150 B-3	11/11	100%	0/12	0%
152 A-1	10/10	100%	0/12	0%
154 B-2	9/9	100%	0/11	0%
155 A-1	10/10	100%	1/11	9%
156 A-1	9/9	100%	10/11	91%
157 A-1	11/12	92%	0/9	0%
158 A-1	10/11	91%	8/9	89%
159 A-1	10/10	100%	1/8	13%
160 A-1	11/11	100%	0/11	0%
161 A-1	8/8	100%	0/12	0%
163 B-1	9/9	100%	0/9	0%
164 B-2	11/11	100%	0/11	0%
165 C-2	10/10	100%	0/9	0%
166 A-1	11/11	100%	0/9	0%
167 A-1	9/9	100%	0/7	0%
168 B-1	10/10	100%	0/5	0%
R2T21 A-1	10/10	100%	2/8	25%

4.2 Genome sequencing of PI 594527 and NLR gene annotation.

Based on the responses of 209 F2:3 families derived from a cross between PI 594527 and a susceptible variety Williams to *P. sojae* Race 1, *Rps11* was initially mapped to a 348-kb genomic region on chromosome 7, with 12 sequencing gaps of unknown sizes according to the Williams 82 reference genome assembly v2.0 (www.soybase.org). In order to determine the correct genome structure and the copy number of NLR genes in the mapping region in the two parental lines, we performed whole genome sequencing and *de novo* assembly for PI 594527. The genome was built with 34 kb PacBio reads which were assembled into 424 contigs with a contig N50 of 13.8 Mb and further polished with Chromium 10X data. The contigs were scaffolded with 45 BioNano maps into 43 hybrid scaffolds with a scaffold N50 of 26.4 Mb, essentially 1-2 scaffolds per chromosome. Then, we performed NLR gene annotation across the entire genome and identified 512 NLR genes in total (Figure 4.4, Table 4.3, and Table 4.4).

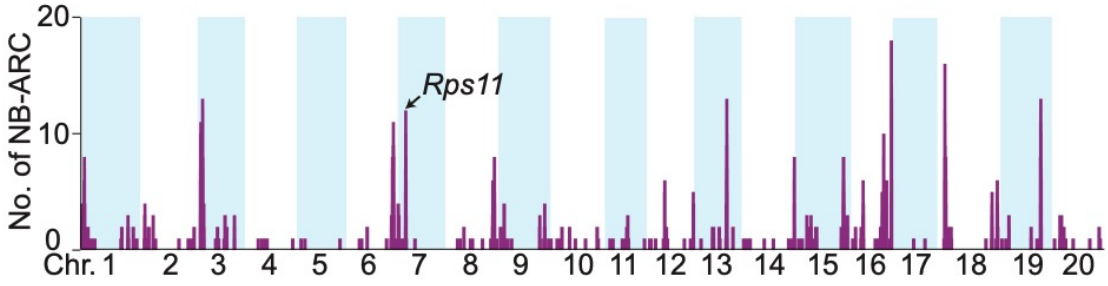


Figure 4.4 Physical distributuions of NLR genes across the PI 594527 genome.

Number of NLR genes was counted in 1-Mb sliding windows with 100-kb steps. y axis is the number of NLR genes and x axis is the soybean chromosomes. The arrow points the position of the NLR gene cluster in the *Rps11* region.

Table 4.3 Number of NLR genes on each chromosome.

Chromosome ID	Length (bp)	No. of NLR
Chr01	58479646	32
Chr02	51823354	18
Chr03	46994370	58
Chr04	52535076	6
Chr05	44312942	3
Chr06	50373303	45
Chr07	45788982	16
Chr08	48806881	30
Chr09	50310076	22
Chr10	53640677	11
Chr11	39637742	12
Chr12	42723935	22
Chr13	46929662	28
Chr14	51027704	17
Chr15	53394395	36
Chr16	37902597	58
Chr17	42197827	2
Chr18	59935911	49
Chr19	50705441	28
Chr20	49315660	19
Total	976836181	512

Table 4.4 List of NLR genes across entire genome in the *Rps11* donor line.

Chromosome	Start	End	NLR ID
Chr01	1018243	1020772	Chr01_NLR_1
Chr01	1028735	1032332	Chr01_NLR_2
Chr01	1155718	1158247	Chr01_NLR_3
Chr01	1205536	1208569	Chr01_NLR_32
Chr01	3108328	3109752	Chr01_NLR_4
Chr01	3346470	3350153	Chr01_NLR_5
Chr01	3402501	3406211	Chr01_NLR_6
Chr01	3429226	3433313	Chr01_NLR_7
Chr01	3457895	3461060	Chr01_NLR_8
Chr01	3475540	3479598	Chr01_NLR_9
Chr01	3667231	3669484	Chr01_NLR_10
Chr01	3721305	3728439	Chr01_NLR_11
Chr01	3753256	3756121	Chr01_NLR_31
Chr01	4101324	4105346	Chr01_NLR_12
Chr01	4170801	4178689	Chr01_NLR_13
Chr01	5428123	5431064	Chr01_NLR_30
Chr01	6938035	6943111	Chr01_NLR_14
Chr01	7105590	7106438	Chr01_NLR_15
Chr01	10181969	10184537	Chr01_NLR_16
Chr01	11671567	11673783	Chr01_NLR_17
Chr01	13489864	13499459	Chr01_NLR_29
Chr01	38589414	38594185	Chr01_NLR_28
Chr01	39249408	39252531	Chr01_NLR_18
Chr01	39269138	39273150	Chr01_NLR_19
Chr01	45032137	45035928	Chr01_NLR_27
Chr01	45111318	45119410	Chr01_NLR_26
Chr01	45318880	45323686	Chr01_NLR_20
Chr01	45628458	45631884	Chr01_NLR_25
Chr01	50211917	50213723	Chr01_NLR_21
Chr01	50271665	50275918	Chr01_NLR_22
Chr01	52519285	52521826	Chr01_NLR_23
Chr01	53523755	53526335	Chr01_NLR_24
Chr02	2314819	2317423	Chr02_NLR_18
Chr02	2681939	2686763	Chr02_NLR_17
Chr02	2735020	2737510	Chr02_NLR_16
Chr02	2956476	2959676	Chr02_NLR_1
Chr02	3068461	3069006	Chr02_NLR_15
Chr02	3837003	3838822	Chr02_NLR_2
Chr02	6583035	6589230	Chr02_NLR_3
Chr02	6591196	6593970	Chr02_NLR_14

Table 4.4 continued

Chr02	10602852	10604713	Chr02_NLR_13
Chr02	10607147	10609574	Chr02_NLR_4
Chr02	10832025	10833422	Chr02_NLR_5
Chr02	12936214	12939798	Chr02_NLR_6
Chr02	35383398	35385830	Chr02_NLR_12
Chr02	44251441	44258749	Chr02_NLR_11
Chr02	45820920	45832428	Chr02_NLR_7
Chr02	48551783	48555285	Chr02_NLR_8
Chr02	49799087	49803294	Chr02_NLR_9
Chr02	49818472	49822579	Chr02_NLR_10
Chr03	3930708	3934092	Chr03_NLR_58
Chr03	3956375	3959744	Chr03_NLR_57
Chr03	4090213	4093576	Chr03_NLR_56
Chr03	4113203	4114943	Chr03_NLR_55
Chr03	4140643	4144005	Chr03_NLR_54
Chr03	4182585	4185870	Chr03_NLR_53
Chr03	4284140	4286212	Chr03_NLR_52
Chr03	4337323	4340740	Chr03_NLR_51
Chr03	4356115	4369096	Chr03_NLR_50
Chr03	4393941	4397316	Chr03_NLR_49
Chr03	4472889	4474961	Chr03_NLR_48
Chr03	4501961	4505342	Chr03_NLR_47
Chr03	4513721	4517102	Chr03_NLR_46
Chr03	4531439	4534760	Chr03_NLR_45
Chr03	4797133	4800508	Chr03_NLR_1
Chr03	4823707	4827085	Chr03_NLR_2
Chr03	4887729	4892596	Chr03_NLR_3
Chr03	4943437	4946812	Chr03_NLR_4
Chr03	5115680	5118146	Chr03_NLR_5
Chr03	5141852	5145263	Chr03_NLR_6
Chr03	5189490	5192859	Chr03_NLR_7
Chr03	5206807	5208003	Chr03_NLR_8
Chr03	5688001	5690673	Chr03_NLR_44
Chr03	5792617	5795980	Chr03_NLR_9
Chr03	5809948	5811232	Chr03_NLR_10
Chr03	5846170	5849416	Chr03_NLR_11
Chr03	5860616	5863933	Chr03_NLR_12
Chr03	5877802	5881161	Chr03_NLR_13
Chr03	5899792	5902182	Chr03_NLR_14
Chr03	5921218	5924647	Chr03_NLR_15
Chr03	5936482	5939783	Chr03_NLR_16
Chr03	5974597	5977840	Chr03_NLR_17

Table 4.4 continued

Chr03	6128247	6133997	Chr03_NLR_43
Chr03	6238667	6242327	Chr03_NLR_18
Chr03	6264949	6271338	Chr03_NLR_42
Chr03	6278264	6280484	Chr03_NLR_41
Chr03	6294071	6300428	Chr03_NLR_40
Chr03	6311699	6313272	Chr03_NLR_39
Chr03	6354860	6363248	Chr03_NLR_38
Chr03	6383992	6385955	Chr03_NLR_37
Chr03	7041854	7042733	Chr03_NLR_19
Chr03	7078849	7079728	Chr03_NLR_36
Chr03	7227958	7228837	Chr03_NLR_20
Chr03	7303058	7303619	Chr03_NLR_35
Chr03	18585749	18588665	Chr03_NLR_34
Chr03	20228389	20231818	Chr03_NLR_33
Chr03	20298719	20303653	Chr03_NLR_21
Chr03	20812492	20816980	Chr03_NLR_32
Chr03	21309219	21309653	Chr03_NLR_22
Chr03	26634900	26635179	Chr03_NLR_23
Chr03	27408074	27408640	Chr03_NLR_31
Chr03	27472235	27476310	Chr03_NLR_24
Chr03	27480237	27482624	Chr03_NLR_25
Chr03	29139630	29141430	Chr03_NLR_30
Chr03	29149303	29150734	Chr03_NLR_29
Chr03	36362249	36365017	Chr03_NLR_28
Chr03	36473066	36475271	Chr03_NLR_27
Chr03	36577956	36580353	Chr03_NLR_26
Chr04	12326557	12326962	Chr04_NLR_1
Chr04	15980406	15982526	Chr04_NLR_2
Chr04	16531386	16532809	Chr04_NLR_6
Chr04	18782329	18783164	Chr04_NLR_5
Chr04	20598925	20601307	Chr04_NLR_3
Chr04	45221421	45221743	Chr04_NLR_4
Chr05	655197	657864	Chr05_NLR_1
Chr05	4246763	4250008	Chr05_NLR_2
Chr05	37736806	37739602	Chr05_NLR_3
Chr06	11988495	11991066	Chr06_NLR_1
Chr06	14029540	14031946	Chr06_NLR_45
Chr06	19374375	19377248	Chr06_NLR_2
Chr06	19383583	19384912	Chr06_NLR_3
Chr06	37896220	37899094	Chr06_NLR_44
Chr06	42089645	42092978	Chr06_NLR_4
Chr06	42577878	42581025	Chr06_NLR_43

Table 4.4 continued

Chr06	42599307	42601070	Chr06_NLR_42
Chr06	42610088	42613773	Chr06_NLR_41
Chr06	43380806	43383929	Chr06_NLR_40
Chr06	43391983	43399958	Chr06_NLR_39
Chr06	43461638	43464744	Chr06_NLR_38
Chr06	43492892	43496074	Chr06_NLR_37
Chr06	43503552	43504828	Chr06_NLR_36
Chr06	43557763	43560898	Chr06_NLR_35
Chr06	43593291	43595720	Chr06_NLR_5
Chr06	43609826	43610506	Chr06_NLR_6
Chr06	43822768	43825912	Chr06_NLR_34
Chr06	43863310	43866454	Chr06_NLR_33
Chr06	44121485	44124597	Chr06_NLR_7
Chr06	44136215	44144088	Chr06_NLR_8
Chr06	44168499	44172160	Chr06_NLR_9
Chr06	44194575	44197591	Chr06_NLR_10
Chr06	44203838	44207073	Chr06_NLR_11
Chr06	44258853	44262803	Chr06_NLR_12
Chr06	44282926	44285919	Chr06_NLR_13
Chr06	44314947	44318545	Chr06_NLR_14
Chr06	44580513	44582148	Chr06_NLR_15
Chr06	44587025	44589525	Chr06_NLR_32
Chr06	44590414	44591771	Chr06_NLR_31
Chr06	44650664	44651340	Chr06_NLR_16
Chr06	44687452	44688999	Chr06_NLR_17
Chr06	44752046	44756131	Chr06_NLR_18
Chr06	44756632	44759135	Chr06_NLR_30
Chr06	44760565	44762104	Chr06_NLR_29
Chr06	45616008	45620432	Chr06_NLR_28
Chr06	46465764	46469223	Chr06_NLR_27
Chr06	48822781	48826136	Chr06_NLR_19
Chr06	48924327	48926754	Chr06_NLR_26
Chr06	48931594	48935711	Chr06_NLR_25
Chr06	48939995	48941924	Chr06_NLR_20
Chr06	49392722	49395009	Chr06_NLR_24
Chr06	49559933	49561939	Chr06_NLR_23
Chr06	49593109	49596220	Chr06_NLR_21
Chr06	49846999	49848968	Chr06_NLR_22
Chr07	575635	580634	Chr07_NLR_16
Chr07	3037671	3041520	Chr07_NLR_1
Chr07	5532547	5537035	Chr07_NLR_15(R1)
Chr07	5568238	5572804	Chr07_NLR_14(R2)

Table 4.4 continued

Chr07	5605573	5610003	Chr07_NLR_13(R3)
Chr07	5648442	5652964	Chr07_NLR_12(R4)
Chr07	5698247	5702803	Chr07_NLR_2(R5)
Chr07	5737137	5741668	Chr07_NLR_3(R6)
Chr07	5782086	5786645	Chr07_NLR_4(R7)
Chr07	5816684	5821388	Chr07_NLR_5(R8)
Chr07	5846404	5850961	Chr07_NLR_6(R9)
Chr07	5874676	5886104	Chr07_NLR_7(R10)
Chr07	5932950	5937508	Chr07_NLR_8(R11)
Chr07	5974753	5986619	Chr07_NLR_9(R12)
Chr07	6200569	6206829	Chr07_NLR_11
Chr07	14511580	14514560	Chr07_NLR_10
Chr08	9499111	9501898	Chr08_NLR_30
Chr08	11718607	11720233	Chr08_NLR_29
Chr08	15408755	15410819	Chr08_NLR_1
Chr08	15587862	15590980	Chr08_NLR_2
Chr08	21721011	21724281	Chr08_NLR_3
Chr08	23086954	23088019	Chr08_NLR_4
Chr08	23624236	23626717	Chr08_NLR_28
Chr08	33232220	33232832	Chr08_NLR_27
Chr08	41262599	41263486	Chr08_NLR_26
Chr08	41787851	41792602	Chr08_NLR_25
Chr08	42808320	42811413	Chr08_NLR_24
Chr08	42862126	42865264	Chr08_NLR_23
Chr08	43030810	43035774	Chr08_NLR_22
Chr08	43218113	43219004	Chr08_NLR_21
Chr08	43229056	43231690	Chr08_NLR_20
Chr08	43270256	43270964	Chr08_NLR_5
Chr08	43489752	43490193	Chr08_NLR_19
Chr08	43851062	43851852	Chr08_NLR_18
Chr08	44073574	44076921	Chr08_NLR_6
Chr08	44538408	44540731	Chr08_NLR_17
Chr08	44560840	44562834	Chr08_NLR_16
Chr08	44574248	44576795	Chr08_NLR_15
Chr08	44596979	44598973	Chr08_NLR_14
Chr08	44611899	44614455	Chr08_NLR_13
Chr08	44729505	44732079	Chr08_NLR_12
Chr08	44741100	44743659	Chr08_NLR_11
Chr08	44761318	44763856	Chr08_NLR_10
Chr08	45069363	45071937	Chr08_NLR_9
Chr08	45557169	45564646	Chr08_NLR_8
Chr08	46826848	46829182	Chr08_NLR_7

Table 4.4 continued

Chr09	1664357	1667201	Chr09_NLR_22
Chr09	1685051	1687889	Chr09_NLR_21
Chr09	3488082	3490219	Chr09_NLR_20
Chr09	5121063	5126222	Chr09_NLR_19
Chr09	5132852	5135160	Chr09_NLR_18
Chr09	5178947	5183424	Chr09_NLR_1
Chr09	5187662	5191629	Chr09_NLR_2
Chr09	5932668	5936323	Chr09_NLR_17
Chr09	7746329	7748278	Chr09_NLR_3
Chr09	8246465	8249461	Chr09_NLR_4
Chr09	8732309	8734557	Chr09_NLR_5
Chr09	12227037	12229572	Chr09_NLR_16
Chr09	39321069	39324284	Chr09_NLR_6
Chr09	39381914	39382573	Chr09_NLR_7
Chr09	39645058	39647509	Chr09_NLR_15
Chr09	43350673	43356393	Chr09_NLR_14
Chr09	44087511	44090841	Chr09_NLR_13
Chr09	44095685	44098223	Chr09_NLR_12
Chr09	44248184	44251257	Chr09_NLR_8
Chr09	44297213	44298897	Chr09_NLR_9
Chr09	47818943	47821304	Chr09_NLR_11
Chr09	50151815	50154466	Chr09_NLR_10
Chr10	5653748	5655800	Chr10_NLR_1
Chr10	8449827	8450405	Chr10_NLR_11
Chr10	10224206	10230814	Chr10_NLR_2
Chr10	10254152	10255595	Chr10_NLR_10
Chr10	17483998	17491238	Chr10_NLR_3
Chr10	17507105	17508071	Chr10_NLR_4
Chr10	23026533	23027326	Chr10_NLR_9
Chr10	32722976	32725983	Chr10_NLR_5
Chr10	43827450	43830396	Chr10_NLR_6
Chr10	43836750	43839705	Chr10_NLR_7
Chr10	44876514	44880495	Chr10_NLR_8
Chr11	2542567	2546354	Chr11_NLR_12
Chr11	5358965	5361512	Chr11_NLR_11
Chr11	14324247	14328618	Chr11_NLR_10
Chr11	15579667	15584507	Chr11_NLR_1
Chr11	18327280	18329679	Chr11_NLR_9
Chr11	18558704	18562605	Chr11_NLR_8
Chr11	18930094	18930304	Chr11_NLR_7
Chr11	19437285	19443806	Chr11_NLR_2
Chr11	19442243	19443806	Chr11_NLR_3

Table 4.4 continued

Chr11	19582824	19584643	Chr11_NLR_4
Chr11	20180848	20184987	Chr11_NLR_6
Chr11	35429448	35431792	Chr11_NLR_5
Chr12	863768	866390	Chr12_NLR_22
Chr12	1971747	1975444	Chr12_NLR_1
Chr12	6352712	6355154	Chr12_NLR_2
Chr12	13851350	13854204	Chr12_NLR_3
Chr12	15043240	15046062	Chr12_NLR_21
Chr12	15057697	15060656	Chr12_NLR_20
Chr12	15078837	15081682	Chr12_NLR_4
Chr12	15088192	15088746	Chr12_NLR_19
Chr12	15096325	15099214	Chr12_NLR_18
Chr12	15220789	15223665	Chr12_NLR_17
Chr12	16280331	16283178	Chr12_NLR_16
Chr12	16629657	16632781	Chr12_NLR_15
Chr12	16785201	16788384	Chr12_NLR_14
Chr12	17692710	17694367	Chr12_NLR_5
Chr12	33866793	33869580	Chr12_NLR_13
Chr12	39792405	39796698	Chr12_NLR_6
Chr12	40437016	40439383	Chr12_NLR_12
Chr12	42159629	42161873	Chr12_NLR_11
Chr12	42440942	42444124	Chr12_NLR_7
Chr12	42482624	42486279	Chr12_NLR_10
Chr12	42491409	42495643	Chr12_NLR_8
Chr12	42518426	42521135	Chr12_NLR_9
Chr13	6984488	6986586	Chr13_NLR_1
Chr13	17881337	17885328	Chr13_NLR_2
Chr13	17912334	17915115	Chr13_NLR_3
Chr13	18581083	18584299	Chr13_NLR_4
Chr13	18959267	18962364	Chr13_NLR_5
Chr13	19065359	19067005	Chr13_NLR_28
Chr13	24311336	24312967	Chr13_NLR_6
Chr13	24319731	24320587	Chr13_NLR_27
Chr13	24903170	24903609	Chr13_NLR_26
Chr13	30854565	30857796	Chr13_NLR_25
Chr13	31152218	31156250	Chr13_NLR_7
Chr13	31181253	31191119	Chr13_NLR_8
Chr13	31198396	31202453	Chr13_NLR_9
Chr13	31364517	31367793	Chr13_NLR_10
Chr13	31369960	31373751	Chr13_NLR_11
Chr13	31387380	31390884	Chr13_NLR_12
Chr13	31488757	31493851	Chr13_NLR_24

Table 4.4 continued

Chr13	31587868	31589176	Chr13_NLR_13
Chr13	31590980	31594184	Chr13_NLR_14
Chr13	31610516	31613780	Chr13_NLR_23
Chr13	31674131	31677434	Chr13_NLR_15
Chr13	31700584	31703741	Chr13_NLR_22
Chr13	31705054	31706511	Chr13_NLR_21
Chr13	31711039	31714348	Chr13_NLR_16
Chr13	31743906	31747284	Chr13_NLR_20
Chr13	31796683	31800060	Chr13_NLR_17
Chr13	37417686	37421079	Chr13_NLR_18
Chr13	37432745	37436902	Chr13_NLR_19
Chr14	673609	677931	Chr14_NLR_17
Chr14	2284057	2284979	Chr14_NLR_1
Chr14	3890663	3895736	Chr14_NLR_16
Chr14	6810943	6815569	Chr14_NLR_2
Chr14	20354037	20359468	Chr14_NLR_15
Chr14	29038674	29044844	Chr14_NLR_14
Chr14	43679196	43679856	Chr14_NLR_13
Chr14	47095870	47102831	Chr14_NLR_3
Chr14	48378315	48380727	Chr14_NLR_12
Chr14	48914189	48918247	Chr14_NLR_4
Chr14	48931023	48934698	Chr14_NLR_5
Chr14	48951932	48957368	Chr14_NLR_6
Chr14	48974366	48980101	Chr14_NLR_7
Chr14	49010712	49015977	Chr14_NLR_11
Chr14	49028923	49034411	Chr14_NLR_10
Chr14	49180248	49185319	Chr14_NLR_8
Chr14	49202538	49206464	Chr14_NLR_9
Chr15	2021636	2024853	Chr15_NLR_36
Chr15	5478551	5480957	Chr15_NLR_35
Chr15	9936867	9939200	Chr15_NLR_1
Chr15	10049226	10052067	Chr15_NLR_34
Chr15	10059234	10062081	Chr15_NLR_33
Chr15	12652234	12656021	Chr15_NLR_2
Chr15	12661227	12665112	Chr15_NLR_3
Chr15	13704450	13709762	Chr15_NLR_4
Chr15	13973230	13976501	Chr15_NLR_32
Chr15	14180667	14182340	Chr15_NLR_31
Chr15	14979058	14982479	Chr15_NLR_30
Chr15	17235990	17236643	Chr15_NLR_5
Chr15	18285541	18287598	Chr15_NLR_6
Chr15	18590564	18592354	Chr15_NLR_29

Table 4.4 continued

Chr15	19294391	19296734	Chr15_NLR_7
Chr15	19354355	19357205	Chr15_NLR_28
Chr15	42737399	42742409	Chr15_NLR_8
Chr15	42899470	42902803	Chr15_NLR_9
Chr15	44590127	44591904	Chr15_NLR_27
Chr15	44656207	44659033	Chr15_NLR_10
Chr15	44665191	44668230	Chr15_NLR_11
Chr15	44674389	44677215	Chr15_NLR_12
Chr15	44766231	44768119	Chr15_NLR_13
Chr15	44848528	44851408	Chr15_NLR_14
Chr15	44953099	44955904	Chr15_NLR_15
Chr15	44989601	44992429	Chr15_NLR_16
Chr15	45092856	45095814	Chr15_NLR_17
Chr15	45176628	45179463	Chr15_NLR_26
Chr15	45270344	45271893	Chr15_NLR_25
Chr15	45291210	45293268	Chr15_NLR_18
Chr15	45297225	45300088	Chr15_NLR_19
Chr15	45316132	45319036	Chr15_NLR_24
Chr15	45904754	45909003	Chr15_NLR_23
Chr15	48664646	48671589	Chr15_NLR_22
Chr15	48720027	48722970	Chr15_NLR_21
Chr15	48762810	48765463	Chr15_NLR_20
Chr16	493765	496916	Chr16_NLR_58
Chr16	2906589	2913285	Chr16_NLR_1
Chr16	3089227	3094852	Chr16_NLR_57
Chr16	8134130	8137343	Chr16_NLR_56
Chr16	8466583	8468739	Chr16_NLR_55
Chr16	10000338	10002893	Chr16_NLR_54
Chr16	10015641	10018050	Chr16_NLR_53
Chr16	10164533	10167124	Chr16_NLR_2
Chr16	10311063	10314508	Chr16_NLR_52
Chr16	10460023	10463249	Chr16_NLR_51
Chr16	10494043	10497093	Chr16_NLR_50
Chr16	10600576	10603954	Chr16_NLR_49
Chr16	21839575	21839779	Chr16_NLR_3
Chr16	24147712	24148058	Chr16_NLR_48
Chr16	26914284	26915226	Chr16_NLR_47
Chr16	26957011	26958873	Chr16_NLR_4
Chr16	28443093	28446102	Chr16_NLR_5
Chr16	28469764	28473534	Chr16_NLR_46
Chr16	28488672	28491819	Chr16_NLR_6
Chr16	28526208	28530128	Chr16_NLR_45

Table 4.4 continued

Chr16	28535488	28540255	Chr16_NLR_44
Chr16	29823950	29832367	Chr16_NLR_43
Chr16	29866590	29869589	Chr16_NLR_42
Chr16	29873108	29877738	Chr16_NLR_41
Chr16	29892052	29893623	Chr16_NLR_40
Chr16	29913207	29919677	Chr16_NLR_39
Chr16	29942008	29947274	Chr16_NLR_38
Chr16	29985870	29986526	Chr16_NLR_7
Chr16	29995984	30000154	Chr16_NLR_37
Chr16	30013191	30016592	Chr16_NLR_36
Chr16	30030099	30033126	Chr16_NLR_35
Chr16	31356970	31360468	Chr16_NLR_34
Chr16	31401021	31404078	Chr16_NLR_8
Chr16	32532682	32535821	Chr16_NLR_9
Chr16	32537558	32550451	Chr16_NLR_10
Chr16	32559448	32570210	Chr16_NLR_11
Chr16	32575467	32580611	Chr16_NLR_12
Chr16	32582667	32585512	Chr16_NLR_13
Chr16	32586997	32588880	Chr16_NLR_14
Chr16	36064092	36067149	Chr16_NLR_15
Chr16	37020266	37023300	Chr16_NLR_16
Chr16	37033924	37036952	Chr16_NLR_17
Chr16	37049094	37050958	Chr16_NLR_18
Chr16	37078112	37081371	Chr16_NLR_19
Chr16	37137817	37142148	Chr16_NLR_33
Chr16	37235288	37238304	Chr16_NLR_32
Chr16	37241524	37244598	Chr16_NLR_31
Chr16	37248795	37251800	Chr16_NLR_30
Chr16	37254543	37257612	Chr16_NLR_29
Chr16	37261902	37264973	Chr16_NLR_28
Chr16	37279909	37281277	Chr16_NLR_27
Chr16	37292504	37295512	Chr16_NLR_26
Chr16	37299247	37305637	Chr16_NLR_25
Chr16	37318417	37321876	Chr16_NLR_24
Chr16	37323990	37327765	Chr16_NLR_23
Chr16	37334887	37337967	Chr16_NLR_22
Chr16	37340600	37341779	Chr16_NLR_21
Chr16	37348654	37354069	Chr16_NLR_20
Chr17	20522337	20527359	Chr17_NLR_2
Chr17	31347748	31349596	Chr17_NLR_1
Chr18	7342824	7344921	Chr18_NLR_1
Chr18	7889361	7891554	Chr18_NLR_49

Table 4.4 continued

Chr18	7910752	7913227	Chr18_NLR_48
Chr18	7921859	7922883	Chr18_NLR_47
Chr18	7928246	7930799	Chr18_NLR_46
Chr18	7957312	7959874	Chr18_NLR_45
Chr18	7977415	7979821	Chr18_NLR_44
Chr18	8022920	8025120	Chr18_NLR_43
Chr18	8033902	8036459	Chr18_NLR_42
Chr18	8233968	8236170	Chr18_NLR_2
Chr18	8260277	8262737	Chr18_NLR_3
Chr18	8269885	8274572	Chr18_NLR_4
Chr18	8280681	8281026	Chr18_NLR_5
Chr18	8305564	8306827	Chr18_NLR_6
Chr18	8310849	8313333	Chr18_NLR_7
Chr18	8333328	8334814	Chr18_NLR_8
Chr18	8333328	8338334	Chr18_NLR_9
Chr18	8342661	8345144	Chr18_NLR_10
Chr18	8877895	8880462	Chr18_NLR_11
Chr18	8890029	8892597	Chr18_NLR_12
Chr18	8957974	8960557	Chr18_NLR_13
Chr18	8964856	8967442	Chr18_NLR_14
Chr18	9001127	9003677	Chr18_NLR_15
Chr18	9008232	9010084	Chr18_NLR_16
Chr18	9053102	9055011	Chr18_NLR_17
Chr18	9073235	9075146	Chr18_NLR_18
Chr18	10360331	10362918	Chr18_NLR_41
Chr18	10399186	10402592	Chr18_NLR_40
Chr18	11195972	11198489	Chr18_NLR_19
Chr18	11209589	11212476	Chr18_NLR_20
Chr18	13808494	13811653	Chr18_NLR_21
Chr18	14094676	14097711	Chr18_NLR_39
Chr18	14310604	14311049	Chr18_NLR_22
Chr18	47343047	47345609	Chr18_NLR_23
Chr18	52795319	52797631	Chr18_NLR_24
Chr18	52996804	53004716	Chr18_NLR_25
Chr18	53082056	53089911	Chr18_NLR_26
Chr18	53157519	53164584	Chr18_NLR_27
Chr18	53223770	53229987	Chr18_NLR_28
Chr18	53331648	53340069	Chr18_NLR_29
Chr18	57293643	57297270	Chr18_NLR_30
Chr18	58062843	58064997	Chr18_NLR_31
Chr18	58065962	58070382	Chr18_NLR_32
Chr18	58179322	58181449	Chr18_NLR_38

Table 4.4 continued

Chr18	58189906	58192009	Chr18_NLR_37
Chr18	58207090	58209259	Chr18_NLR_36
Chr18	58342210	58344649	Chr18_NLR_35
Chr18	58644139	58646674	Chr18_NLR_33
Chr18	58647822	58650539	Chr18_NLR_34
Chr19	2551179	2554564	Chr19_NLR_28
Chr19	6327382	6332201	Chr19_NLR_27
Chr19	9243997	9246091	Chr19_NLR_1
Chr19	9334424	9337524	Chr19_NLR_26
Chr19	9375900	9379354	Chr19_NLR_25
Chr19	30472973	30475388	Chr19_NLR_2
Chr19	30509432	30515052	Chr19_NLR_3
Chr19	36191718	36194493	Chr19_NLR_4
Chr19	36695540	36698342	Chr19_NLR_24
Chr19	39155746	39156670	Chr19_NLR_5
Chr19	39409447	39411937	Chr19_NLR_23
Chr19	39423087	39424992	Chr19_NLR_22
Chr19	39437795	39440285	Chr19_NLR_21
Chr19	39553652	39555956	Chr19_NLR_6
Chr19	39565519	39567655	Chr19_NLR_7
Chr19	39575181	39577599	Chr19_NLR_8
Chr19	39587274	39592035	Chr19_NLR_9
Chr19	39607519	39610024	Chr19_NLR_10
Chr19	39731711	39734129	Chr19_NLR_20
Chr19	39738178	39738964	Chr19_NLR_19
Chr19	39747565	39750058	Chr19_NLR_18
Chr19	39824643	39826107	Chr19_NLR_17
Chr19	39954477	39956892	Chr19_NLR_11
Chr19	39967395	39969810	Chr19_NLR_12
Chr19	39981296	39983711	Chr19_NLR_13
Chr19	39995227	39997714	Chr19_NLR_14
Chr19	40016924	40019342	Chr19_NLR_15
Chr19	40044020	40046408	Chr19_NLR_16
Chr20	2082969	2086781	Chr20_NLR_19
Chr20	7141725	7142697	Chr20_NLR_1
Chr20	7366848	7369665	Chr20_NLR_18
Chr20	7374248	7375104	Chr20_NLR_2
Chr20	7785340	7787890	Chr20_NLR_17
Chr20	7847194	7852073	Chr20_NLR_3
Chr20	8669833	8670974	Chr20_NLR_4
Chr20	8776246	8779717	Chr20_NLR_5
Chr20	8795261	8798732	Chr20_NLR_6

Table 4.4 continued

Chr20	9335103	9336259	Chr20_NLR_16
Chr20	9691034	9694907	Chr20_NLR_7
Chr20	10901516	10903273	Chr20_NLR_8
Chr20	10955542	10958938	Chr20_NLR_9
Chr20	12669762	12681675	Chr20_NLR_10
Chr20	19828433	19831986	Chr20_NLR_11
Chr20	35755519	35757603	Chr20_NLR_12
Chr20	44620356	44624430	Chr20_NLR_13
Chr20	44637360	44639559	Chr20_NLR_14
Chr20	45670091	45672685	Chr20_NLR_15

There are 12 NLR genes annotated in the *Rps11* region in the *Rps11* donor line. Among the twelve NLR genes, R2, R3, R5 and R8 are truncated with 2 to 4 exons lost at the 3' end; there is a 1.4 kb insertion in the first exon of R10, resulting in disruption of the NBS domain; The remaining seven NLR genes (R1, R4, R6, R7, R9, R11, and R12) are intact (Figure 4.5a), but only five (R1, R4, R6, R9, R12) are expressed based on the RNA-seq analysis (Figure 4.5b).

In order to figure out the correct genome structure and copy number of NLR genes in the susceptible parent, Williams, we obtained a sequence contig covering the *Rps11* corresponding region from Williams 82 assembly v3.0 (Chu et al., 2021), which was assumed to be the same as Williams at this region. Sequence comparison between the Williams 82 assembly v2.0 and v3.0 reveals a number of assembly errors in the *rps11* region in the Williams 82 assembly v2.0 (Figure 4.6). We then annotated the NLR genes in the *Rps11* corresponding region and identified only 8 NLR genes, dubbed r1 to r8, which have the similar gene models as R1-R12 in PI 594527 (Figure 4.7).

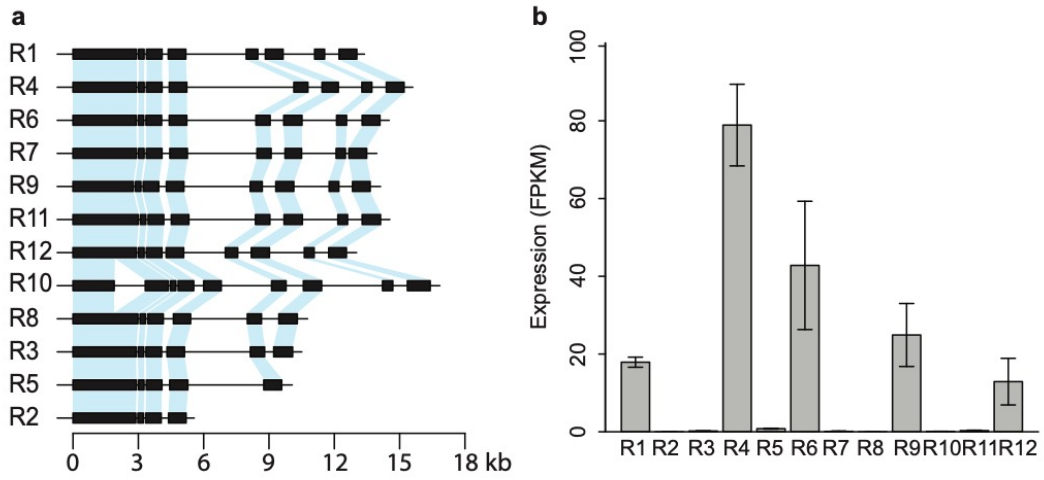


Figure 4.5 Gene models and expression pattern of the NLR genes in the *Rps11* region.

a, Gene models and alignments among twelve NBS-LRR genes in PI 594527. Black boxes represent exons of each NBS-LRR gene. Light-blue shades represent alignments.
b, Expression profile of the twelve NBS-LRR genes in the *Rps11* region. Y axis is the expression level (FPKM) based on RNA-seq data. The error bars represent standard errors.

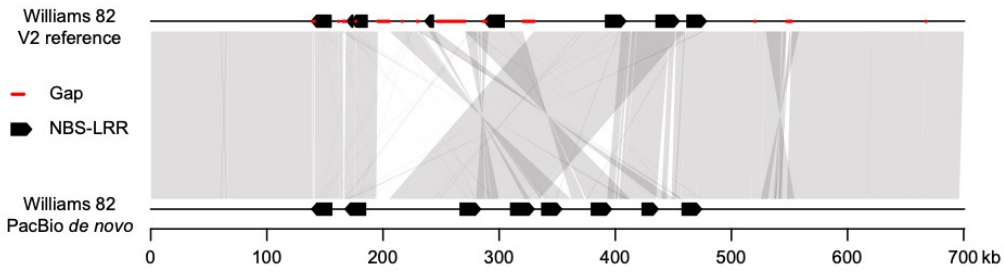


Figure 4.6 Sequence comparison between Williams 82 assembly v2.0 and assembly v3.0.

Black boxes represent NBS-LRR genes and grey shades represent syntenic blocks between two versions. Red lines/dots represent sequencing gaps in the assembly v2.0.

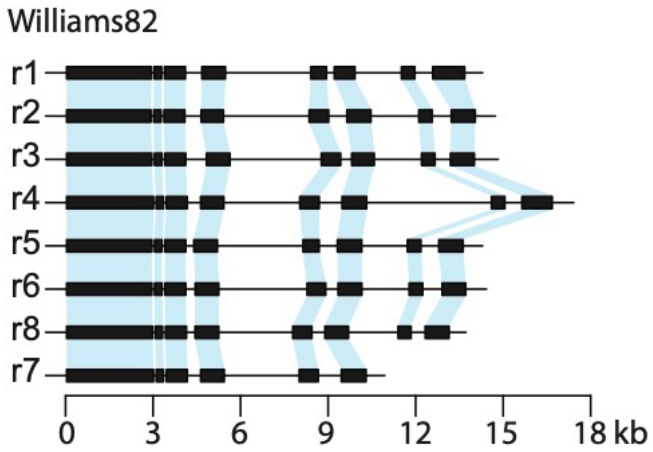


Figure 4.7 Gene models of the 8 NLR genes in the *Rps11* corresponding region in Williams 82. Black boxes represent exons of each NBS-LRR gene. Light-blue shades represent alignments.

4.3 Fine mapping of the *Rps11* locus

Sequence comparison reveals only a small proportion of the *Rps11* region shared between PI 594527 and Williams 82 as syntenic blocks (Figure 4.8), which allowed us to design markers for fine mapping. To determine whether a single copy or a combination of multiple copies of the five expressed NLR genes was responsible for the broad resistance, we screened a total of 17,050 progenies using the boundary markers, SSR_286 and SSR_320, and identified 43 recombinants for finer mapping. We inoculated each recombinant with isolate Race 1 and designed additional markers within the NLR gene cluster at the syntenic blocks shared by the two parental lines. We eventually delimited *Rps11* to a 151-kb genome interval defined by Marker176 and InDel327, harboring 4 NLR genes (R5, R6, R7 and R8, Figure 4.9). Interestingly, R6 was the only expressed NLR gene within the final mapping interval, suggesting it was the best candidate responsible for the resistances (Figure 4.9). Furthermore, we examined the expression of the NLR genes in 9 key recombinants with different combinations of NLR genes. Recombinants with R1 expressed alone or a combination of R1 and R4 expressed together were susceptible, suggesting R1 and R4 do not confer resistance (Figure 4.10). Meanwhile, recombinants with a combination of R9 and R12 expressed together were also susceptible, which ruled out R9 and R12 (Figure 4.10). Only the recombinants with R6 expressed were resistant (Figure 4.10), further supporting R6 alone as responsible for the resistance. Then, we examined the expression pattern of R6 in different tissues and found it is also expressed in other tissues (Figure 4.11). The genomic DNA of R6, excluding

its regulatory sequences, is 14.1-kb, but it possesses a 13.1-kb 5' untranslated region (UTR) as determined by 5' Rapid amplification of cDNA ends (RACE) (Figure 4.12) and a 0.5-kb 3'UTR, totaling 27.7 kb (Figure 4.13). R6 is predicted to encode an NLR protein of 2,463 amino acids (Figure 4.13).

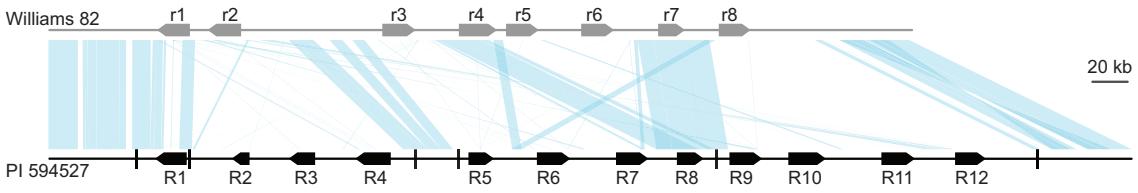


Figure 4.8 Comparison of the NLR gene clusters between Williams 82 and PI 594527.

Black boxes represent NLR genes and light-blue shades represent syntenic blocks between two genomes.

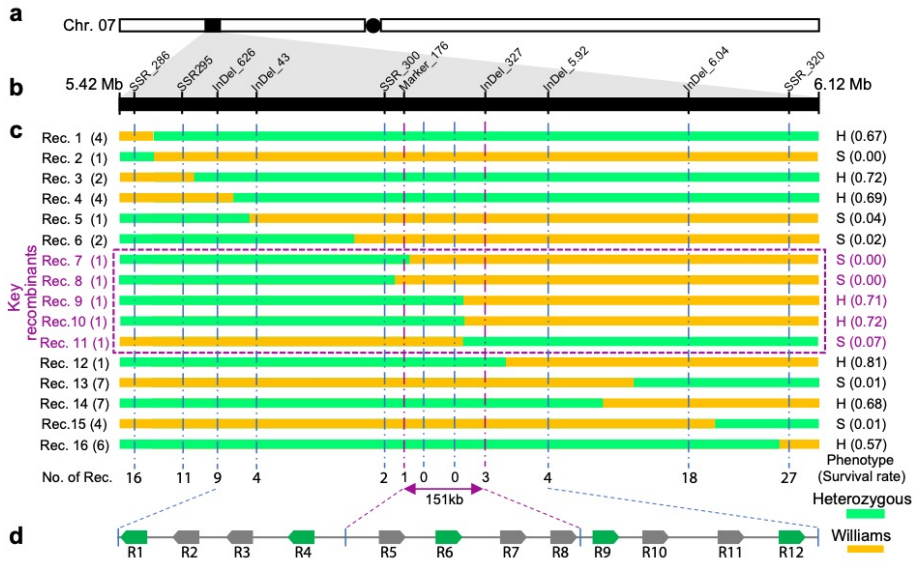


Figure 4.9 Fine mapping of the *Rps11* locus.

a, Physical position of the *Rps11* locus on chromosome 7.
b, Physical position of the markers used for fine mapping.
c, Genotype and phenotype of the recombinants. The numbers in parentheses at left side are the number of recombinants with same haplotype. Green bars represent heterozygous genotype. Orange bars represent Williams genotype. Phenotype (survival rate after inoculation) of each recombinant type are shown at right side. H means heterozygous, S means susceptible. Numbers under each dashed line are the number of recombinants at each marker.
d, NLR gene cluster in the *Rps11* region. Green bars represent expressed NLR genes. Grey bars represent non-expressed NLR genes.

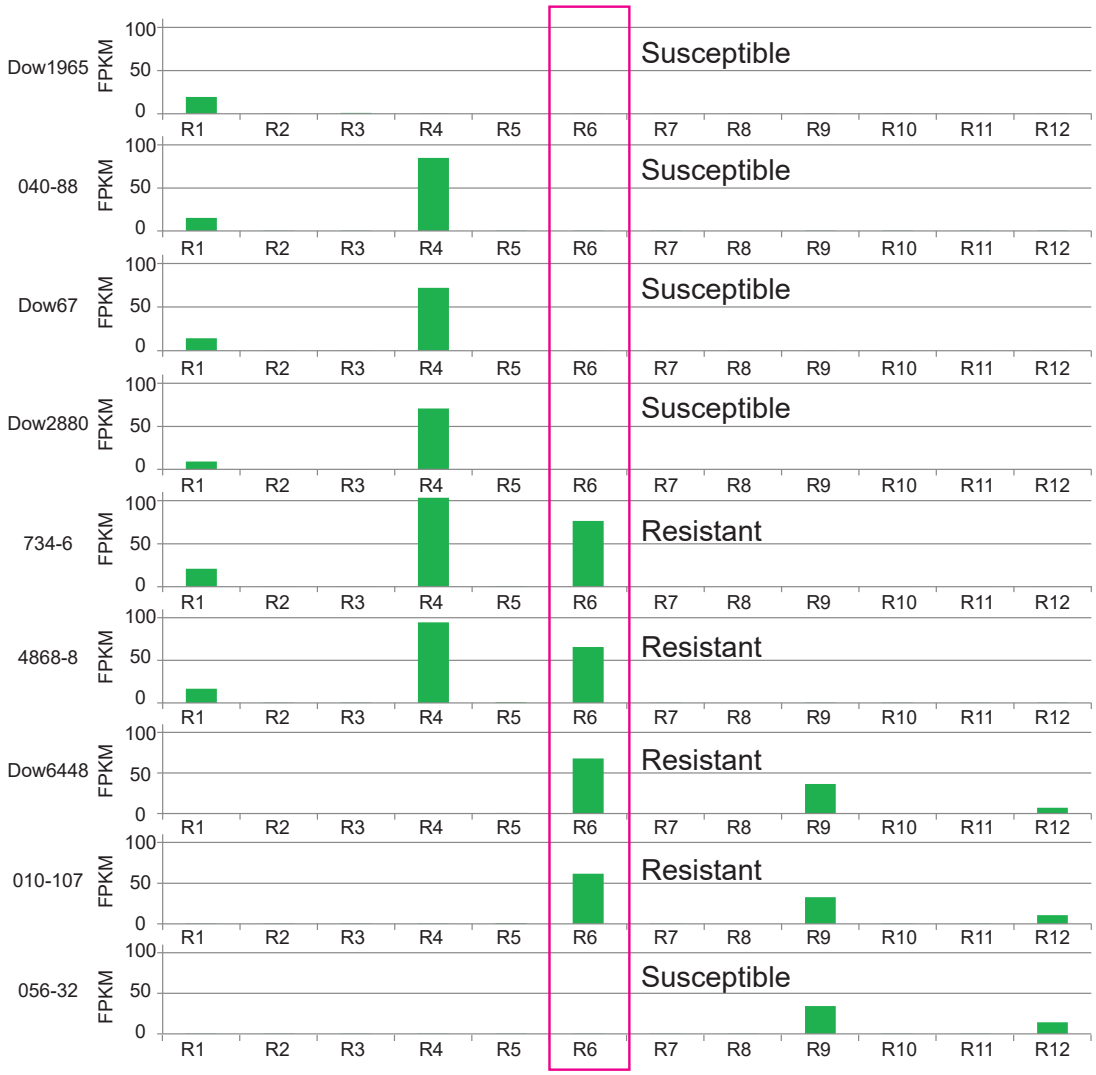


Figure 4.10 Detection of the expression of the twelve NLR genes in each key recombinant.

The phenotype of each recombinant is labeled. Y axis is expression level (FPKM) based on RNA-seq data.

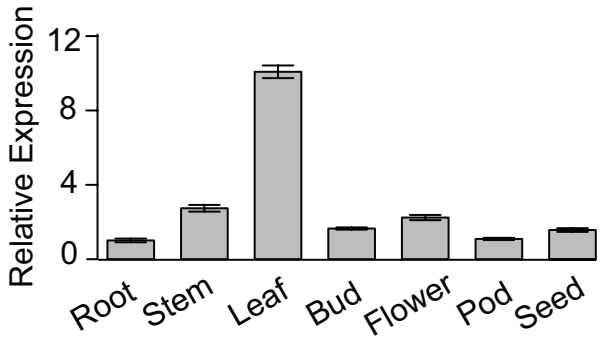


Figure 4.11 Expression profile of Rps11 in different tissues.

The error bars represent standard errors.

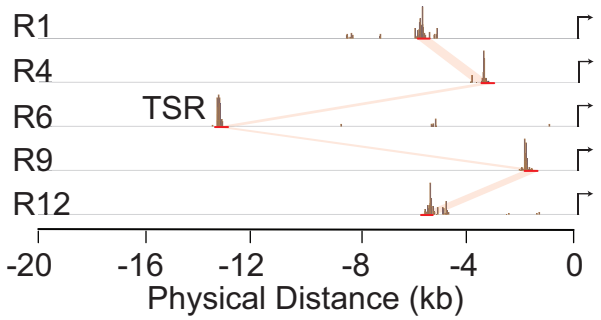


Figure 4.12 5' Rapid amplification of cDNA ends (RACE) performed for the 5 expressed NLR genes in PI594527.

x axis represents the distance upstream of the first exon. Brown bars represents the 5' RACE reads mapped to each NLR gene. Red lines/shades show the promoter regions sharing sequence similarity. Arrows at the right side indicate the direction of the NLR genes.

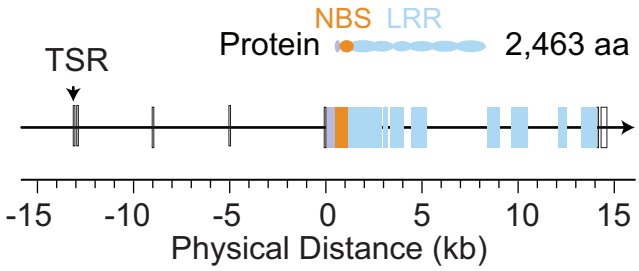


Figure 4.13 Gene model of *Rps11* (R6).

The vertical Arrow indicates the transcription start region (TSR). Orange color indicates the region encoding the NBS domain; Light-blue color indicates the regions encoding the LRR domains; Grey color indicates the region without a domain detected. Open boxes represent UTRs.

4.4 Functional validation of the *Rps11* candidate gene R6

To validate the function of R6, its CDS was synthesized and assembled with the ubiquitin promoter. We introduced the construct (pUbi:CDS-*Rps11*) into the soybean cultivar 93Y21, which is susceptible to *P. sojae* Races 25, 31 and OH12108-06-03. T_0 plants with a single copy of the insertion were obtained and advanced to the T_1 generation. The expression level of R6 in the homozygous T_2 transgenic lines was confirmed to be high and was not detected in non-transgenic lines (Figure 4.14). Segregating T_2 seedling families were inoculated with Race 25, 31 and OH12108-06-03. As exemplified in Figure 4.15, resistance to these three isolates in homozygous transgenic families were significantly increased compared to non-transgenic families (Figure 4.16), indicating R6 confers resistance to multiple isolates. Furthermore, we observed strong correlations between the expression levels of R6 and the resistance to all three isolates in T_2 families (Figure 4.17). These results demonstrate that R6 is responsible for the broad-spectrum resistance to *P. sojae* underlying *Rps11*.

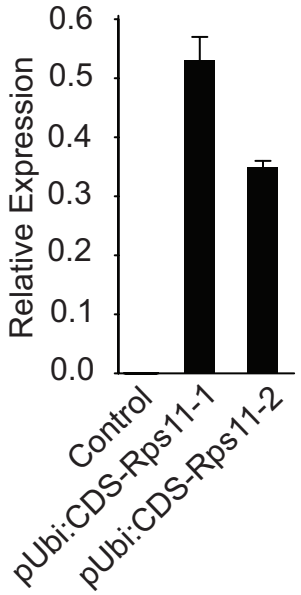


Figure 4.14 Relative expression of the transgene (R6) in T2 families
The error bars represent standard errors.

Figure 4.15 Photographic illustration of the resistance in two independent transgenic events. In each transgenic event, three homozygous T2 families, a non-transgenic line (Control), and a F_5 RIL (*Rps11/Rps11*) were inoculated. Scale bars = 2.5 cm.

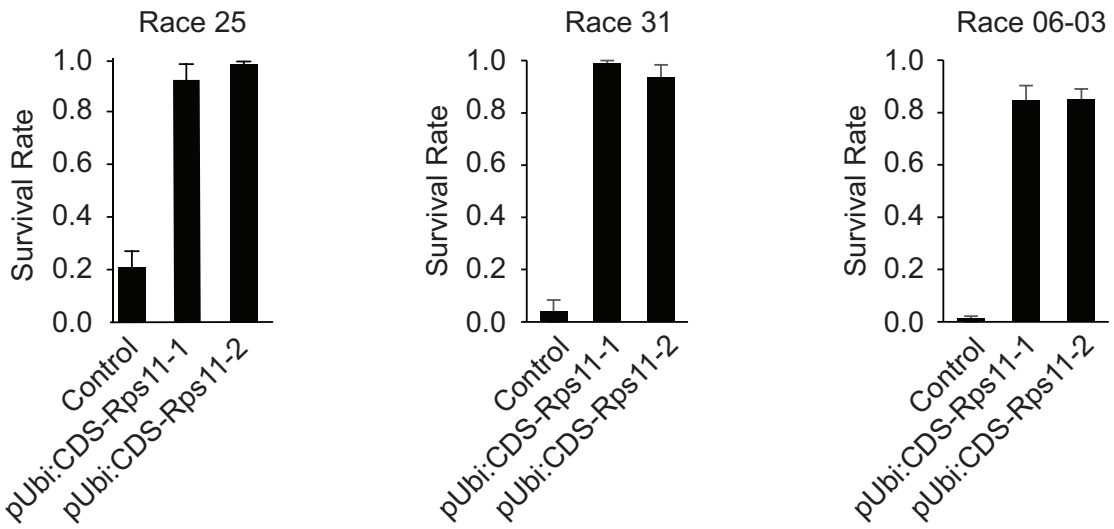


Figure 4.16 Statistics of the resistance test of homozygous T2 families

compared with non-transgenic lines (Control). The Y-axis is the survival rate after inoculation. The error bars represent standard errors.

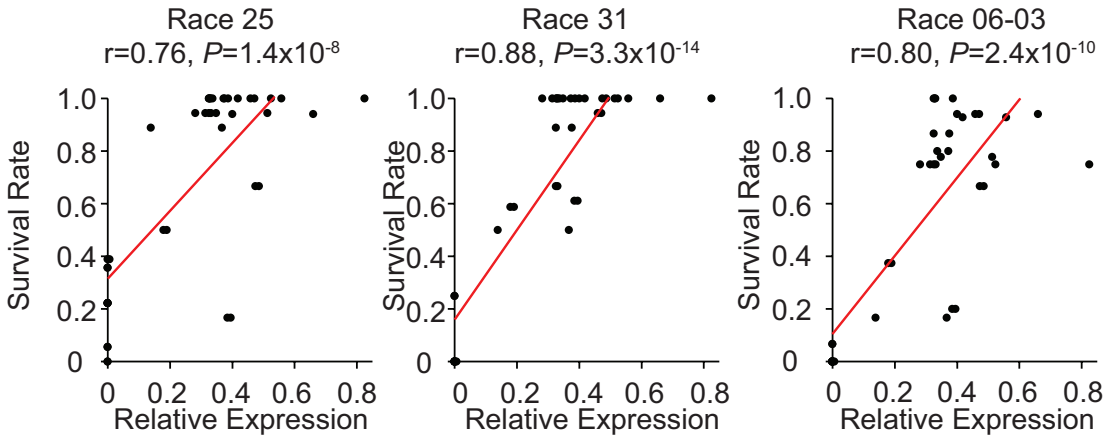


Figure 4.17 Correlation between the expression of R6 (Rps11) and the survival rate after inoculation in T2 population.

x axes are relative expression of R6, and y axes are survival rate after inoculation with Race 25 (left), Race 31 (middle), Race 06-03 (right).

4.5 Evolutionary history of *Rps11* and the NLR genes cluster in PI 594527

One of the most noticeable and unique features of the NLR genes within the *Rps11* region is their giant size. All intact NLR genes in this region have a CDS longer than 7 kb and a protein containing more than 2,000 amino acids (Figure 4.5a and Figure 4.7), larger than 97% of the soybean NLR genes and 99% of the NLR genes from 10 other plant species examined (Figure 4.18). Interestingly, all of these giant NLR genes clustered into the same clade on the phylogenetic tree (Figure 4.19), suggesting a single origin, which was echoed by the conserved gene models shared among all of the giant NLR genes (Figure 4.5a and Figure 4.7). To understand how the giant NLR genes reached this size, we compared the protein sequence of R6 with a typically sized NLR gene from the closest neighboring branch and found that the NBS domain was conserved between the giant NLR gene and the typically sized NLR gene (Figure 4.20). However, several rounds of tandem duplication were identified within the LRR region (Figure 4.20), suggesting the giant size of these NLR genes mainly resulted from tandem duplication of the LRR domains. Given that the NLR genes from both the *Rps11* region and its whole genome duplication (WGD) region (Chr16.R1) have similar sizes and gene models (Figure 4.21), the LRR tandem duplications must have occurred before the whole genome duplication ~13 million years ago (Doyle and Egan, 2010; Gill et al., 2009; Innes et al., 2008; Schmutz *et al.*, 2010).

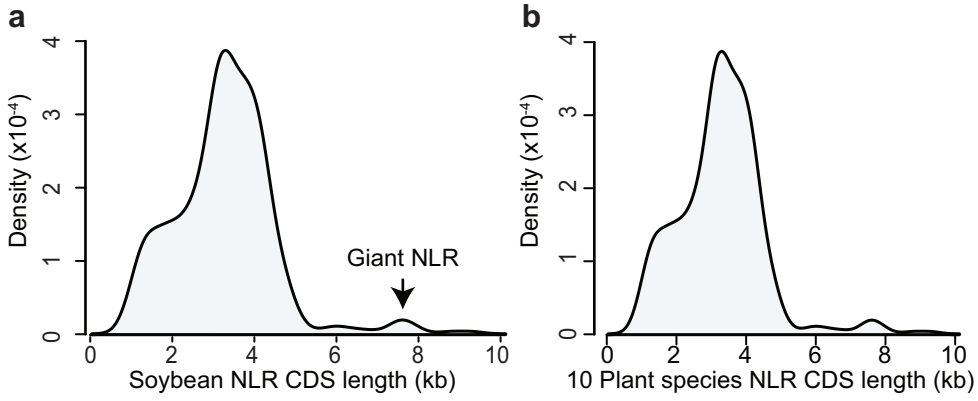


Figure 4.18 Distribution of CDS length in soybean genome and 10 representative plant species

10 plant species are Rice, Maize, Cotton, Common bean, Medicago, Apple, Peach, Banana, Tomato, Cucumber.

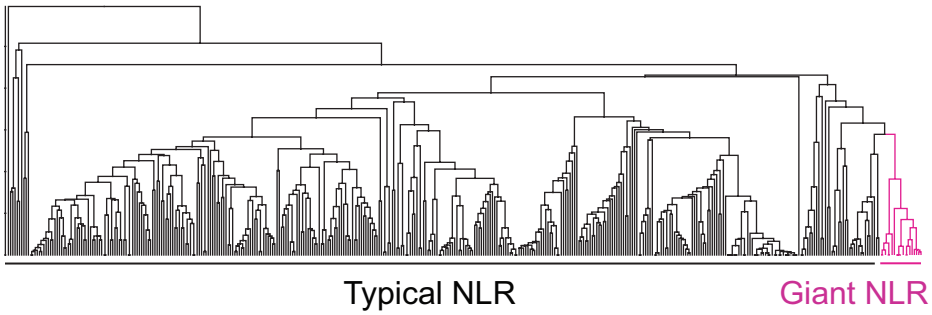


Figure 4.19 Phylogenetic tree of all the NLR genes in PI 594527 built using the conserved NB-ARC domain region.

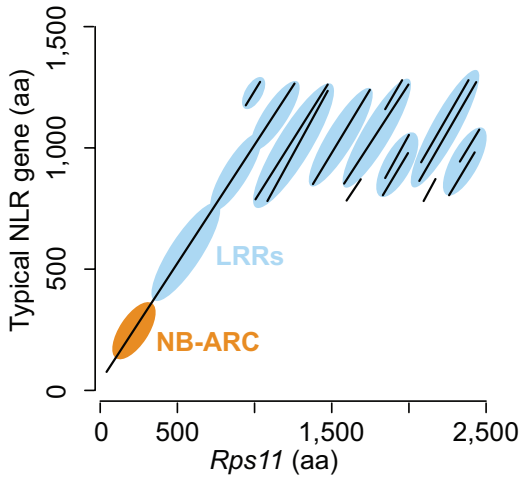


Figure 4.20 Comparison of protein sequences between a giant NLR gene and a typical NLR gene.

Lines represent the alignments between two protein sequences. Orange color highlights the NB-ARC domain region. Light-blue color highlights the LRR domain regions.

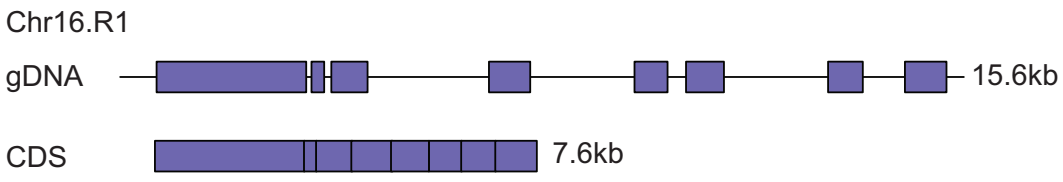


Figure 4.21Gene model of Chr16.R1, the NLR in the WGD region of *Rps11*

The similar sizes and conserved gene models of the giant NLR genes suggested they were originally derived from a single gene. Given that only one NLR gene was found in its WGD region and the homologous region in common bean (Schmutz et al., 2014), this NLR gene cluster must have been formed after the WGD. To understand how this copy number variation evolved, we blasted the genome sequence from the *Rps11* region against itself, which revealed several segmental duplication events. First, R1-R2 were formed from a duplication of R3-R4, and then an inverted duplication event occurred to produce R11-R12 (Figure 4.22, D1), which was further supported by their relationship on the phylogenetic tree, where R1, R3 and R12 were grouped together, and R2, R4 and R11 were grouped together (Figure 4.23, D1). Second, R8 and R10 were formed by duplications of R7 (Figure 4.22, D2); Therefore, R7, R8 and R10 were grouped into the same clade on the phylogenetic tree (Figure 4.23, D2). All the NLR genes are likely derived from R5, because R5 is the closest to Chr16.R1 from the WGD region (Figure 4.13). In addition, the phylogenetic tree also revealed that R9 and R6 were derived from R11 and R3 respectively (Figure 4.23, D3 and D4). To summarize, the giant size of the NLR genes in the *Rps11* region stems from tandem duplications that occurred in the LRR domain of the original NLR gene prior to the WGD; The NLR gene cluster is mainly a result of segmental inversions and tandem duplication events after the WGD (Figure 4.24).

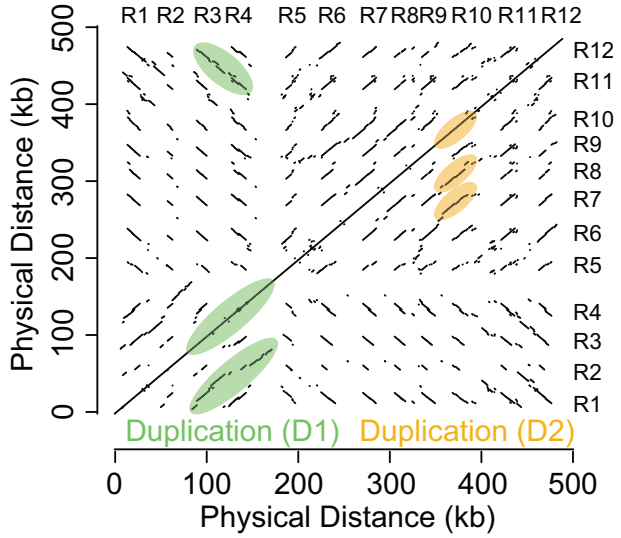


Figure 4.22 Dot plot of sequence comparison within the *Rps11* region in PI 594527.

Light-green color highlights the segmental duplication of R1-R2, R3-R4 and R11-R12 (Duplication 1, D1). Light-orange color highlights the segmental duplication of R7, R8 and R10 (Duplication 2, D2).

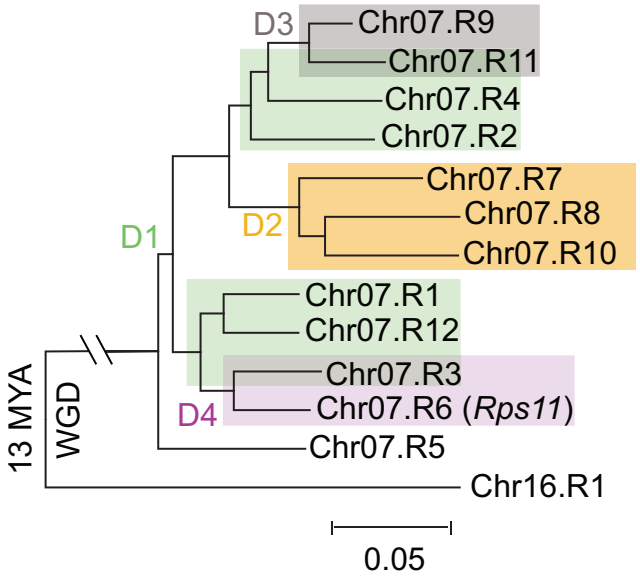


Figure 4.23 Phylogenetic relationship of all the NLR genes underlying Rps11 and its WGD region.

Chr16.R1 is the NLR gene from *Rps11*'s whole-genome duplication region, constructed using transcript sequence. Background colors highlight the groups produced by each duplication event.

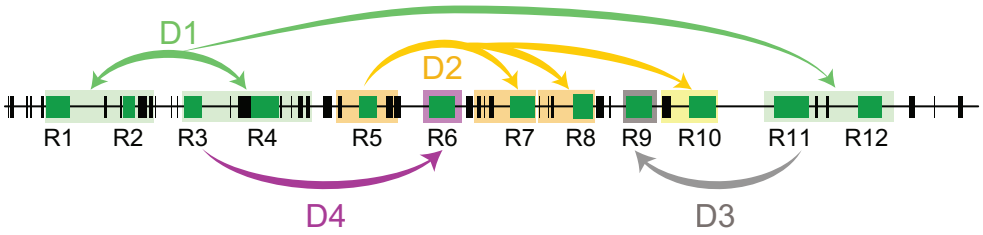


Figure 4.24 Illustration of the evolutionary history of the giant NLR gene cluster in the Rps11 region.

Green boxes represent NLR genes and black boxes represent predicted non-NLR genes.

4.6 Complex diversification of the NLR gene cluster across 30 soybean genomes

To explore the diversification of the NLR gene cluster in the *Rps11*/ "*rps11*" regions, we annotated the NLR genes for 28 additional soybean genomes (Liu *et al.*, 2020; Shen et al., 2019; Shen *et al.*, 2018; Xie *et al.*, 2019). We identified a total of 316 NLR genes in the regions corresponding to *Rps11* across the 30 genomes (Table 4.5); Surprisingly, we observed drastic structural and copy number variations, ranging from 5 copies in SoyL09 to 23 copies in SoyC10. To examine if any of these genomes also carry *Rps11*(R6), we blasted the genome sequence from the *Rps11* donor line against all other genomes. Interestingly, none of them carried R6 in its entirety, with only SoyC03.R6 sharing high sequence similarity with R6 in the first ~10kb from the 5'end while being highly divergent from R6 in the remaining ~4kb at the 3' end (Figure 4.25).

Table 4.5 Information of NLR genes in the Rps11 corresponding region across 30 soybean genomes.

Chromosome	Start	End	NLR ID	Strand	Expression	TSR
Chr07	5524572	5537604	PI594527.R1	-	9.58	yes
Chr07	5568113	5573308	PI594527.R2	-	0.03	no
Chr07	5600488	5610572	PI594527.R3	-	0.14	yes
Chr07	5638345	5653547	PI594527.R4	-	37.89	yes
Chr07	5697682	5707265	PI594527.R5	+	0.42	no
Chr07	5736565	5750658	PI594527.R6	+	15.39	yes
Chr07	5781515	5795000	PI594527.R7	+	0.01	no
Chr07	5816113	5826412	PI594527.R8	+	0.02	no
Chr07	5845991	5859637	PI594527.R9	+	9.69	yes
Chr07	5879422	5895822	PI594527.R10	+	0.11	no
Chr07	5932310	5946431	PI594527.R11	+	0.16	yes
Chr07	5974185	5986739	PI594527.R12	+	4.04	yes
GWHACEK00000007	5544398	5556747	SoyC01.R1	-	13.83	yes
GWHACEK00000007	5580474	5594087	SoyC01.R2	-	3.80	yes
GWHACEK00000007	5632196	5645868	SoyC01.R3	-	17.61	yes
GWHACEK00000007	5696751	5710922	SoyC01.R4	+	1.28	no
GWHACEK00000007	5741132	5746369	SoyC01.R5	+	0.22	no
GWHACEK00000007	5774498	5787019	SoyC01.R6	+	11.94	yes
GWHACEK00000007	5807317	5824380	SoyC01.R7	+	16.84	yes
GWHACEK00000007	5841733	5852009	SoyC01.R8	+	0.19	no
GWHACEK00000007	5869634	5884394	SoyC01.R9	+	3.59	yes
GWHACEK00000007	5908884	5922083	SoyC01.R10	+	5.77	yes
GWHACEK00000007	5946523	5959565	SoyC01.R11	+	0.86	yes

Table 4.5 continued

GWHACEK00000007	5983746	5995843	SoyC01.R12	+	1.55	yes
GWHACEK00000007	6023644	6040696	SoyC01.R13	+	17.32	yes
GWHACEK00000007	6053718	6068702	SoyC01.R14	+	0.00	no
GWHACEK00000007	6112665	6125929	SoyC01.R15	+	1.09	yes
GWHACEK00000007	6141814	6154111	SoyC01.R16	+	0.28	yes
GWHACEK00000007	6177299	6190296	SoyC01.R17	+	4.96	yes
GWHACEL00000007	5540642	5552939	SoyC02.R1	-	4.31	yes
GWHACEL00000007	5578572	5583810	SoyC02.R2	-	0.08	no
GWHACEL00000007	5621396	5635102	SoyC02.R3	-	27.95	yes
GWHACEL00000007	5685595	5699099	SoyC02.R4	+	0.75	no
GWHACEL00000007	5737718	5747999	SoyC02.R5	+	0.18	no
GWHACEL00000007	5767540	5786535	SoyC02.R6	+	8.33	yes
GWHACEL00000007	5802306	5815661	SoyC02.R7	+	0.71	yes
GWHACEL00000007	5835363	5850378	SoyC02.R8	+	0.41	no
GWHACEL00000007	5884946	5898365	SoyC02.R9	+	2.07	yes
GWHACEL00000007	5923537	5933269	SoyC02.R10	+	1.38	no
GWHACEL00000007	5992102	6005104	SoyC02.R11	+	21.71	yes
GWHACEL00000007	6027465	6040960	SoyC02.R12	+	5.51	yes
GWHACEM00000007	5631789	5644891	SoyC03.R1	-	5.81	yes
GWHACEM00000007	5672790	5677985	SoyC03.R2	-	0.09	no
GWHACEM00000007	5709965	5723277	SoyC03.R3	-	12.20	yes
GWHACEM00000007	5755537	5770590	SoyC03.R4	-	19.01	yes
GWHACEM00000007	5824887	5830161	SoyC03.R5	+	0.34	no
GWHACEM00000007	5859389	5873334	SoyC03.R6	+	14.95	yes
GWHACEM00000007	5893427	5903711	SoyC03.R7	+	0.07	no
GWHACEM00000007	5923109	5936937	SoyC03.R8	+	10.04	no
GWHACEM00000007	5956282	5971323	SoyC03.R9	+	0.16	no
GWHACEM00000007	6004594	6020960	SoyC03.R10	+	2.40	yes
GWHACEM00000007	6042567	6055123	SoyC03.R11	+	1.59	yes
GWHACEM00000007	6066036	6079773	SoyC03.R12	+	9.22	yes
GWHACEM00000007	6108338	6125249	SoyC03.R13	+	2.19	yes
GWHACEM00000007	6149355	6159844	SoyC03.R14	+	0.61	no
GWHACEN00000007	5575068	5588755	SoyC04.R1	-	3.91	yes
GWHACEN00000007	5603799	5617851	SoyC04.R2	-	5.36	yes
GWHACEN00000007	5698035	5712042	SoyC04.R3	+	3.43	no
GWHACEN00000007	5741526	5758215	SoyC04.R4	+	0.64	no
GWHACEN00000007	5768316	5781953	SoyC04.R5	+	5.96	yes
GWHACEN00000007	5811011	5824734	SoyC04.R6	+	3.67	yes
GWHACEN00000007	5854684	5864979	SoyC04.R7	+	0.59	no
GWHACEN00000007	5889129	5902277	SoyC04.R8	+	3.96	no
GWHACEO00000007	5584980	5598387	SoyC05.R1	-	38.54	yes
GWHACEO00000007	5652691	5662453	SoyC05.R2	+	0.67	no

Table 4.5 continued

GWHACEO00000007	5668792	5674035	SoyC05.R3	+	0.45	no
GWHACEO00000007	5698787	5711090	SoyC05.R4	+	2.13	yes
GWHACEO00000007	5734500	5749552	SoyC05.R5	+	0.64	no
GWHACEO00000007	5793605	5807611	SoyC05.R6	+	4.61	yes
GWHACEO00000007	5826686	5843168	SoyC05.R7	+	0.22	no
GWHACEO00000007	5850293	5864562	SoyC05.R8	+	1.67	yes
GWHACEO00000007	5890586	5903252	SoyC05.R9	+	1.11	yes
GWHACEO00000007	5911795	5925543	SoyC05.R10	+	10.94	yes
GWHACEO00000007	5945281	5961703	SoyC05.R11	+	0.54	no
GWHACEO00000007	6018188	6035253	SoyC05.R12	+	2.36	yes
GWHACEO00000007	6059152	6070506	SoyC05.R13	+	0.56	no
GWHACEP00000007	5526536	5539004	SoyC06.R1	-	22.49	no
GWHACEP00000007	5562805	5576343	SoyC06.R2	-	8.40	yes
GWHACEP00000007	5605594	5617604	SoyC06.R3	-	4.15	no
GWHACEP00000007	5625508	5637841	SoyC06.R4	-	4.89	yes
GWHACEP00000007	5661936	5667181	SoyC06.R5	-	0.32	no
GWHACEP00000007	5673529	5683245	SoyC06.R6	-	3.99	no
GWHACEP00000007	5736662	5746687	SoyC06.R7	+	2.88	yes
GWHACEP00000007	5766158	5781173	SoyC06.R8	+	0.66	no
GWHACEP00000007	5813616	5827218	SoyC06.R9	+	1.28	yes
GWHACEP00000007	5851943	5864245	SoyC06.R10	+	2.73	yes
GWHACEP00000007	5916061	5926613	SoyC06.R11	+	1.75	no
GWHACEQ00000007	5517496	5531215	SoyC07.R1	-	11.24	yes
GWHACEQ00000007	5572209	5586580	SoyC07.R2	-	1.64	yes
GWHACEQ00000007	5625367	5634239	SoyC07.R3	-	3.71	yes
GWHACEQ00000007	5666364	5679509	SoyC07.R4	-	20.45	yes
GWHACEQ00000007	5730484	5744658	SoyC07.R5	+	1.06	no
GWHACEQ00000007	5774843	5780081	SoyC07.R6	+	0.12	no
GWHACEQ00000007	5808248	5820801	SoyC07.R7	+	8.33	yes
GWHACEQ00000007	5841097	5858045	SoyC07.R8	+	1.43	yes
GWHACEQ00000007	5875346	5885617	SoyC07.R9	+	0.06	no
GWHACEQ00000007	5903219	5916755	SoyC07.R10	+	8.72	yes
GWHACEQ00000007	5941215	5953421	SoyC07.R11	+	0.33	yes
GWHACEQ00000007	5981351	5998298	SoyC07.R12	+	2.15	yes
GWHACEQ00000007	6011311	6026287	SoyC07.R13	+	0.01	no
GWHACEQ00000007	6070250	6083489	SoyC07.R14	+	1.81	yes
GWHACEQ00000007	6099374	6111673	SoyC07.R15	+	1.53	yes
GWHACEQ00000007	6136910	6145537	SoyC07.R16	+	1.09	yes
GWHACER00000007	5514579	5528260	SoyC08.R1	-	1.64	yes
GWHACER00000007	5543334	5557370	SoyC08.R2	-	13.82	yes
GWHACER00000007	5637497	5651511	SoyC08.R3	+	1.13	no
GWHACER00000007	5680963	5697557	SoyC08.R4	+	0.01	no

Table 4.5 continued

GWHACER00000007	5707594	5721155	SoyC08.R5	+	12.49	yes
GWHACER00000007	5750207	5763918	SoyC08.R6	+	6.00	yes
GWHACER00000007	5793878	5804170	SoyC08.R7	+	0.01	no
GWHACER00000007	5828314	5841465	SoyC08.R8	+	11.12	no
GWHACES00000007	5497707	5511463	SoyC09.R1	-	0.66	yes
GWHACES00000007	5526287	5540378	SoyC09.R2	-	8.34	yes
GWHACES00000007	5569188	5582479	SoyC09.R3	-	11.66	yes
GWHACES00000007	5662579	5676624	SoyC09.R4	+	0.72	no
GWHACES00000007	5706371	5723060	SoyC09.R5	+	0.29	no
GWHACES00000007	5733184	5746782	SoyC09.R6	+	10.49	yes
GWHACES00000007	5775863	5789577	SoyC09.R7	+	2.66	yes
GWHACES00000007	5825451	5839186	SoyC09.R8	-	0.46	no
GWHACES00000007	5877849	5888182	SoyC09.R9	+	0.24	no
GWHACES00000007	5912297	5925459	SoyC09.R10	+	5.91	no
GWHACET00000007	5559384	5571684	SoyC10.R1	-	2.35	yes
GWHACET00000007	5597313	5602517	SoyC10.R2	-	0.19	no
GWHACET00000007	5639942	5653344	SoyC10.R3	-	6.87	yes
GWHACET00000007	5749481	5763123	SoyC10.R4	+	13.12	yes
GWHACET00000007	5800396	5805634	SoyC10.R5	+	0.40	no
GWHACET00000007	5833797	5846333	SoyC10.R6	+	6.43	yes
GWHACET00000007	5873549	5890543	SoyC10.R7	+	1.61	yes
GWHACET00000007	5907825	5918111	SoyC10.R8	+	0.07	no
GWHACET00000007	5938223	5951650	SoyC10.R9	+	1.04	yes
GWHACET00000007	5975841	5988947	SoyC10.R10	+	0.78	yes
GWHACET00000007	6013340	6025208	SoyC10.R11	+	0.14	yes
GWHACET00000007	6103823	6116199	SoyC10.R12	-	0.12	yes
GWHACET00000007	6140540	6153745	SoyC10.R13	-	0.66	yes
GWHACET00000007	6232788	6249840	SoyC10.R14	+	2.51	no
GWHACET00000007	6267219	6277518	SoyC10.R15	+	0.13	no
GWHACET00000007	6295150	6308687	SoyC10.R16	+	7.68	yes
GWHACET00000007	6333205	6346405	SoyC10.R17	+	3.18	yes
GWHACET00000007	6370829	6383260	SoyC10.R18	+	0.36	yes
GWHACET00000007	6410847	6427896	SoyC10.R19	+	2.61	yes
GWHACET00000007	6440933	6455910	SoyC10.R20	+	0.00	no
GWHACET00000007	6499886	6513124	SoyC10.R21	+	0.81	yes
GWHACET00000007	6529013	6541312	SoyC10.R22	+	0.72	yes
GWHACET00000007	6564498	6577494	SoyC10.R23	+	1.41	yes
GWHACEU00000007	5728714	5741014	SoyC11.R1	-	1.54	yes
GWHACEU00000007	5766645	5771883	SoyC11.R2	-	0.13	no
GWHACEU00000007	5809483	5823190	SoyC11.R3	-	20.10	yes
GWHACEU00000007	5873609	5887784	SoyC11.R4	+	0.74	no
GWHACEU00000007	5917957	5923195	SoyC11.R5	+	0.10	no

Table 4.5 continued

GWHACEU00000007	5951357	5963894	SoyC11.R6	+	6.64	yes
GWHACEU00000007	5991109	6008079	SoyC11.R7	+	0.50	yes
GWHACEU00000007	6025319	6035572	SoyC11.R8	+	0.01	no
GWHACEU00000007	6053129	6066610	SoyC11.R9	+	4.70	yes
GWHACEU00000007	6091135	6104276	SoyC11.R10	+	1.15	yes
GWHACEU00000007	6128738	6137765	SoyC11.R11	+	0.06	yes
GWHACEU00000007	6168417	6185419	SoyC11.R12	+	0.52	yes
GWHACEU00000007	6202699	6212984	SoyC11.R13	+	0.01	no
GWHACEU00000007	6230618	6244155	SoyC11.R14	+	7.67	yes
GWHACEU00000007	6268673	6281873	SoyC11.R15	+	5.19	yes
GWHACEU00000007	6306305	6318617	SoyC11.R16	+	0.32	yes
GWHACEU00000007	6346325	6363371	SoyC11.R17	+	0.87	yes
GWHACEU00000007	6376407	6391382	SoyC11.R18	+	0.00	no
GWHACEU00000007	6435357	6448595	SoyC11.R19	+	0.47	yes
GWHACEU00000007	6464482	6476781	SoyC11.R20	+	0.28	no
GWHACEU00000007	6499965	6512959	SoyC11.R21	+	0.50	yes
GWHACEV00000007	2887648	2900925	SoyC12.R1	-	4.44	yes
GWHACEV00000007	2925340	2938919	SoyC12.R2	-	10.01	yes
GWHACEV00000007	2967666	2977313	SoyC12.R3	-	0.00	no
GWHACEV00000007	2988250	3000572	SoyC12.R4	-	6.46	yes
GWHACEV00000007	3024676	3029919	SoyC12.R5	-	1.71	no
GWHACEV00000007	3036279	3046017	SoyC12.R6	-	0.90	no
GWHACEV00000007	3099556	3109608	SoyC12.R7	+	2.13	yes
GWHACEV00000007	3129088	3144107	SoyC12.R8	+	0.66	no
GWHACEV00000007	3179301	3196195	SoyC12.R9	+	0.45	yes
GWHACEV00000007	3220314	3230672	SoyC12.R10	+	0.40	no
GWHACEW00000007	5527016	5540702	SoyC13.R1	-	2.02	yes
GWHACEW00000007	5555815	5569856	SoyC13.R2	-	16.26	yes
GWHACEW00000007	5650034	5664072	SoyC13.R3	+	1.22	no
GWHACEW00000007	5693574	5710248	SoyC13.R4	+	0.26	no
GWHACEW00000007	5720350	5733976	SoyC13.R5	+	16.55	yes
GWHACEW00000007	5763027	5776738	SoyC13.R6	+	5.00	yes
GWHACEW00000007	5806706	5817011	SoyC13.R7	+	0.17	no
GWHACEW00000007	5841158	5854314	SoyC13.8	+	11.44	no
GWHACEX00000007	5518259	5531522	SoyC14.R1	-	2.33	yes
GWHACEX00000007	5555948	5569367	SoyC14.R2	-	2.22	yes
GWHACEX00000007	5598846	5610847	SoyC14.R3	-	0.23	no
GWHACEX00000007	5618803	5631121	SoyC14.R4	-	6.01	yes
GWHACEX00000007	5643183	5653527	SoyC14.R5	-	7.62	no
GWHACEX00000007	5677648	5694547	SoyC14.R6	-	1.52	yes
GWHACEX00000007	5729726	5744745	SoyC14.R7	-	0.02	no
GWHACEX00000007	5764214	5774257	SoyC14.R8	-	14.81	yes

Table 4.5 continued

GWHACEX00000007	5827769	5837498	SoyC14.R9	+	0.22	no
GWHACEX00000007	5843852	5861213	SoyC14.R10	+	0.15	no
GWHACEB00000007	5516144	5529824	SoyL01.R1	-	1.73	yes
GWHACEB00000007	5544854	5558906	SoyL01.R2	-	16.79	yes
GWHACEB00000007	5638983	5652997	SoyL01.R3	+	2.17	no
GWHACEB00000007	5682411	5698991	SoyL01.R4	+	0.15	no
GWHACEB00000007	5709035	5722565	SoyL01.R5	+	6.73	no
GWHACEB00000007	5751354	5764930	SoyL01.R6	+	3.68	yes
GWHACEB00000007	5794683	5811315	SoyL01.R7	+	0.17	no
GWHACEB00000007	5821324	5834848	SoyL01.R8	+	10.29	yes
GWHACEB00000007	5863621	5877210	SoyL01.R9	+	4.14	yes
GWHACEB00000007	5906904	5917167	SoyL01.R10	+	0.19	no
GWHACEB00000007	5941311	5954453	SoyL01.R11	+	10.78	no
GWHACEC00000007	5586419	5600586	SoyL02.R1	-	20.63	yes
GWHACEC00000007	5655134	5668543	SoyL02.R2	+	1.73	no
GWHACEC00000007	5689544	5703511	SoyL02.R3	+	1.22	yes
GWHACEC00000007	5722901	5737879	SoyL02.R4	+	0.30	no
GWHACEC00000007	5771302	5787654	SoyL02.R5	+	1.85	yes
GWHACEC00000007	5809368	5822857	SoyL02.R6	+	0.33	no
GWHACEC00000007	5856079	5869462	SoyL02.R7	+	0.96	yes
GWHACED00000007	5583480	5597039	SoyL03.R1	-	6.02	yes
GWHACED00000007	5629556	5644661	SoyL03.R2	-	21.54	yes
GWHACED00000007	5698812	5708695	SoyL03.R3	+	0.84	no
GWHACED00000007	5720307	5737592	SoyL03.R4	+	0.65	no
GWHACED00000007	5749468	5763180	SoyL03.R5	+	16.69	yes
GWHACED00000007	5787462	5800590	SoyL03.R6	+	19.32	yes
GWHACED00000007	5828269	5841006	SoyL03.R7	+	6.17	yes
GWHACEE00000007	5498063	5510524	SoyL04.R1	-	2.78	yes
GWHACEE00000007	5534243	5547774	SoyL04.R2	-	0.06	yes
GWHACEE00000007	5591870	5604023	SoyL04.R3	+	1.63	no
GWHACEE00000007	5650168	5660475	SoyL04.R4	+	0.31	no
GWHACEE00000007	5684866	5697007	SoyL04.R5	+	1.12	yes
GWHACEE00000007	5722933	5736802	SoyL04.R6	+	0.26	yes
GWHACEE00000007	5763063	5776025	SoyL04.R7	+	0.91	yes
GWHACEE00000007	5797669	5812695	SoyL04.R8	+	0.29	no
GWHACEE00000007	5852037	5865363	SoyL04.R9	+	2.11	yes
GWHACEE00000007	5895125	5912024	SoyL04.R10	+	1.05	yes
GWHACEF00000007	5612701	5624996	SoyL05.R1	-	4.10	yes
GWHACEF00000007	5650652	5656129	SoyL05.R2	-	0.21	yes
GWHACEF00000007	5693473	5707181	SoyL05.R3	-	39.66	yes
GWHACEF00000007	5757623	5767368	SoyL05.R4	+	0.90	no
GWHACEF00000007	5773634	5778865	SoyL05.R5	+	0.14	no

Table 4.5 continued

GWHACEF00000007	5830160	5843140	SoyL05.R6	+	16.94	yes
GWHACEF00000007	5865493	5878992	SoyL05.R7	+	7.17	yes
GWHACEG00000007	5552819	5566115	SoyL06.R1	-	7.55	yes
GWHACEG00000007	5590551	5604132	SoyL06.R2	-	3.70	yes
GWHACEG00000007	5633499	5645522	SoyL06.R3	-	0.92	no
GWHACEG00000007	5653490	5665823	SoyL06.R4	-	14.39	no
GWHACEG00000007	5689948	5695191	SoyL06.R5	-	0.08	no
GWHACEG00000007	5701553	5711300	SoyL06.R6	-	0.59	no
GWHACEG00000007	5764879	5774932	SoyL06.R7	+	26.63	yes
GWHACEG00000007	5794421	5809451	SoyL06.R8	+	0.64	no
GWHACEG00000007	5844669	5861583	SoyL06.R9	+	4.11	yes
GWHACEG00000007	5885731	5896094	SoyL06.R10	+	1.01	no
GWHACEH00000007	5581385	5594677	SoyL07.R1	-	3.37	yes
GWHACEH00000007	5619111	5632691	SoyL07.R2	-	2.68	yes
GWHACEH00000007	5662058	5674081	SoyL07.R3	-	0.57	no
GWHACEH00000007	5682049	5694384	SoyL07.R4	-	12.88	yes
GWHACEH00000007	5740965	5757880	SoyL07.R5	-	2.08	yes
GWHACEH00000007	5793104	5808130	SoyL07.R6	-	0.39	no
GWHACEH00000007	5827618	5837672	SoyL07.R7	-	19.76	no
GWHACEH00000007	5891244	5900989	SoyL07.R8	+	0.29	no
GWHACEH00000007	5907350	5912593	SoyL07.R9	+	0.05	no
GWHACEI00000007	5539358	5552658	SoyL08.R1	-	6.57	yes
GWHACEI00000007	5577099	5590681	SoyL08.R2	-	3.19	no
GWHACEI00000007	5620059	5632082	SoyL08.R3	-	0.83	no
GWHACEI00000007	5640050	5652387	SoyL08.R4	-	13.25	yes
GWHACEI00000007	5698972	5715887	SoyL08.R5	-	5.35	yes
GWHACEI00000007	5751118	5766144	SoyL08.R6	-	0.28	no
GWHACEI00000007	5785632	5795687	SoyL08.R7	-	23.29	yes
GWHACEI00000007	5849276	5859024	SoyL08.R8	+	0.53	no
GWHACEI00000007	5865392	5870635	SoyL08.R9	+	0.04	no
GWHACEJ00000007	5546928	5560620	SoyL09.R1	-	2.55	yes
GWHACEJ00000007	5575798	5590135	SoyL09.R2	-	17.95	yes
GWHACEJ00000007	5670055	5684092	SoyL09.R3	+	0.91	no
GWHACEJ00000007	5713590	5723904	SoyL09.R4	+	0.18	no
GWHACEJ00000007	5748058	5761218	SoyL09.R5	+	8.54	no
GWHACDY00000007	5506580	5519846	SoyW01.R1	-	7.76	no
GWHACDY00000007	5552390	5567052	SoyW01.R2	-	3.09	yes
GWHACDY00000007	5605711	5620748	SoyW01.R3	-	0.23	no
GWHACDY00000007	5652798	5663993	SoyW01.R4	-	0.19	no
GWHACDY00000007	5711183	5724983	SoyW01.R5	+	13.22	yes
GWHACDY00000007	5749791	5762391	SoyW01.R6	+	2.66	yes
GWHACDY00000007	5790528	5808931	SoyW01.R7	+	5.60	yes

Table 4.5 continued

GWHACDZ00000007	6745129	6757588	SoyW02.R1	-	1.91	yes
GWHACDZ00000007	6777241	6790598	SoyW02.R2	-	1.08	yes
GWHACDZ00000007	6814978	6828697	SoyW02.R3	-	14.13	yes
GWHACDZ00000007	6875651	6889696	SoyW02.R4	+	0.61	no
GWHACDZ00000007	6932488	6948419	SoyW02.R5	+	3.83	yes
GWHACDZ00000007	6977089	6990574	SoyW02.R6	+	3.89	yes
GWHACEA00000007	2949730	2962040	SoyW03.R1	-	5.22	yes
GWHACEA00000007	2992734	3006262	SoyW03.R2	-	0.23	yes
GWHACEA00000007	3046348	3059866	SoyW03.R3	-	0.83	yes
GWHACEA00000007	3085433	3095885	SoyW03.R4	-	6.12	yes
GWHACEA00000007	3107656	3124240	SoyW03.R5	-	2.12	no
GWHACEA00000007	3142093	3159019	SoyW03.R6	-	0.44	yes
GWHACEA00000007	3184924	3190159	SoyW03.R7	-	2.75	no
GWHACEA00000007	3222552	3235890	SoyW03.R8	-	0.69	no
GWHACEA00000007	3302232	3314531	SoyW03.R9	+	79.78	yes
GWHACEA00000007	3365030	3376806	SoyW03.R10	+	2.74	no
Chr07	5627901	5640373	W05.R1	-	16.49	yes
Chr07	5664158	5677694	W05.R2	-	9.89	yes
Chr07	5721832	5733992	W05.R3	+	2.81	no
Chr07	5780227	5790544	W05.R4	+	0.26	no
Chr07	5814984	5827135	W05.R5	+	14.39	yes
Chr07	5853198	5867107	W05.R6	+	32.52	yes
Chr07	5893478	5906473	W05.R7	+	21.49	yes
Chr07	5928147	5943192	W05.R8	+	0.78	no
Chr07	5982577	5995909	W05.R9	+	1.53	no
Chr07	6025675	6042586	W05.R10	+	5.69	yes
chr7	5530236	5543928	Wms82.R1	-	5.08	yes
chr7	5558994	5573049	Wms82.R2	-	22.37	yes
chr7	5653256	5667279	Wms82.R3	+	0.87	no
chr7	5696765	5713468	Wms82.R4	+	0.06	no
chr7	5723574	5737210	Wms82.R5	+	3.93	yes
chr7	5766276	5780002	Wms82.R6	+	11.29	yes
chr7	5809978	5820276	Wms82.R7	+	0.05	no
chr7	5844431	5857585	Wms82.R8	+	5.24	no
GWHAAEV00000007.1	5569122	5583289	ZH13.v2.R1	-	13.83	yes
GWHAAEV00000007.1	5637865	5651291	ZH13.v2.R2	+	0.92	no
GWHAAEV00000007.1	5672587	5686283	ZH13.v2.R3	+	1.66	yes
GWHAAEV00000007.1	5705779	5720774	ZH13.v2.R4	+	0.52	no
GWHAAEV00000007.1	5754298	5770669	ZH13.v2.R5	+	0.93	yes
GWHAAEV00000007.1	5792405	5805898	ZH13.v2.R6	+	0.38	yes
GWHAAEV00000007.1	5839204	5852602	ZH13.v2.R7	+	0.86	yes

[a] Transcriptional Start Region.

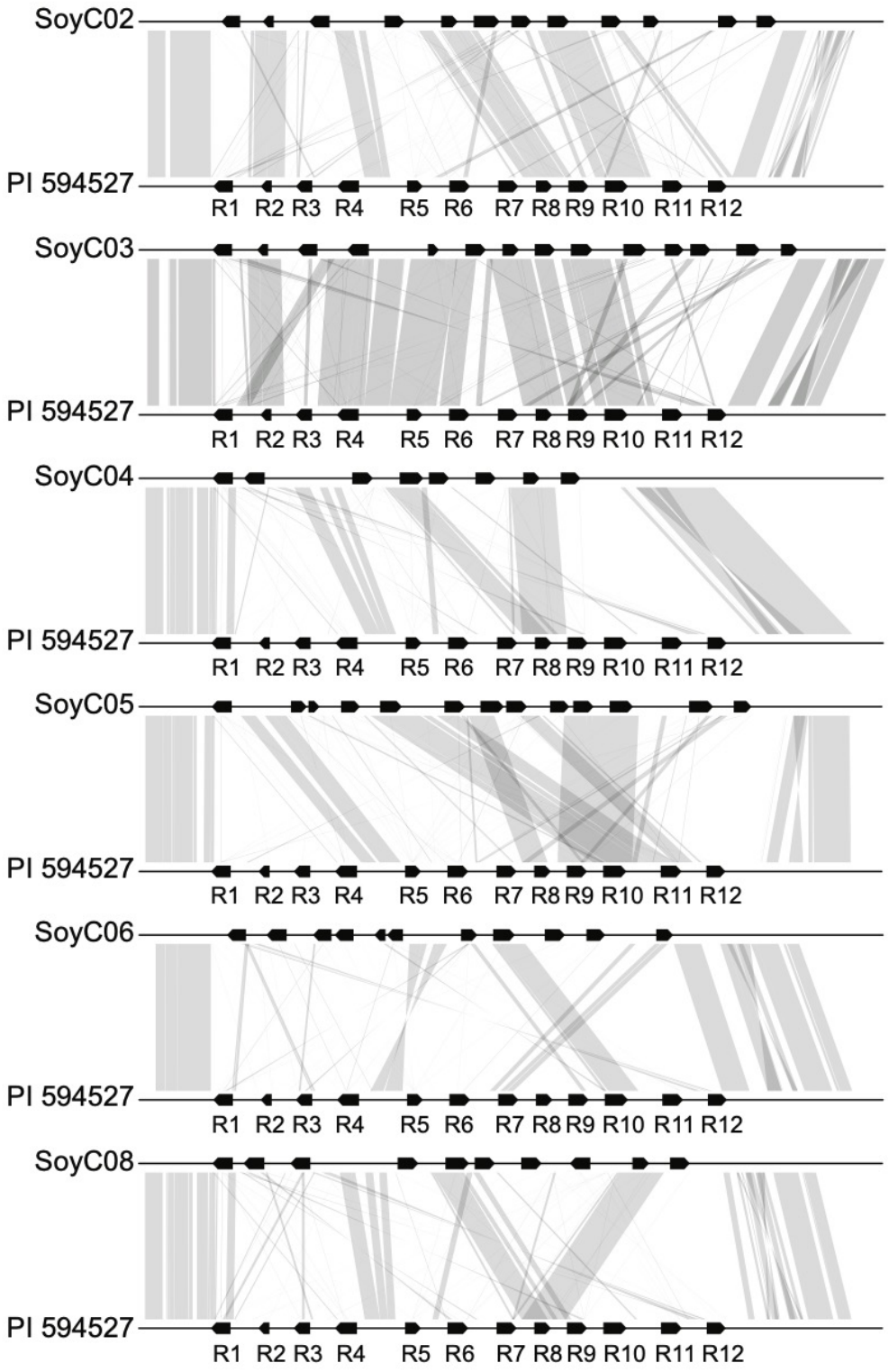


Figure 4.25 Sequence comparisons between PI 594527 and 29 additional soybean genome in the *Rps11*/ "*rps11*" region

Figure 4.25 continued

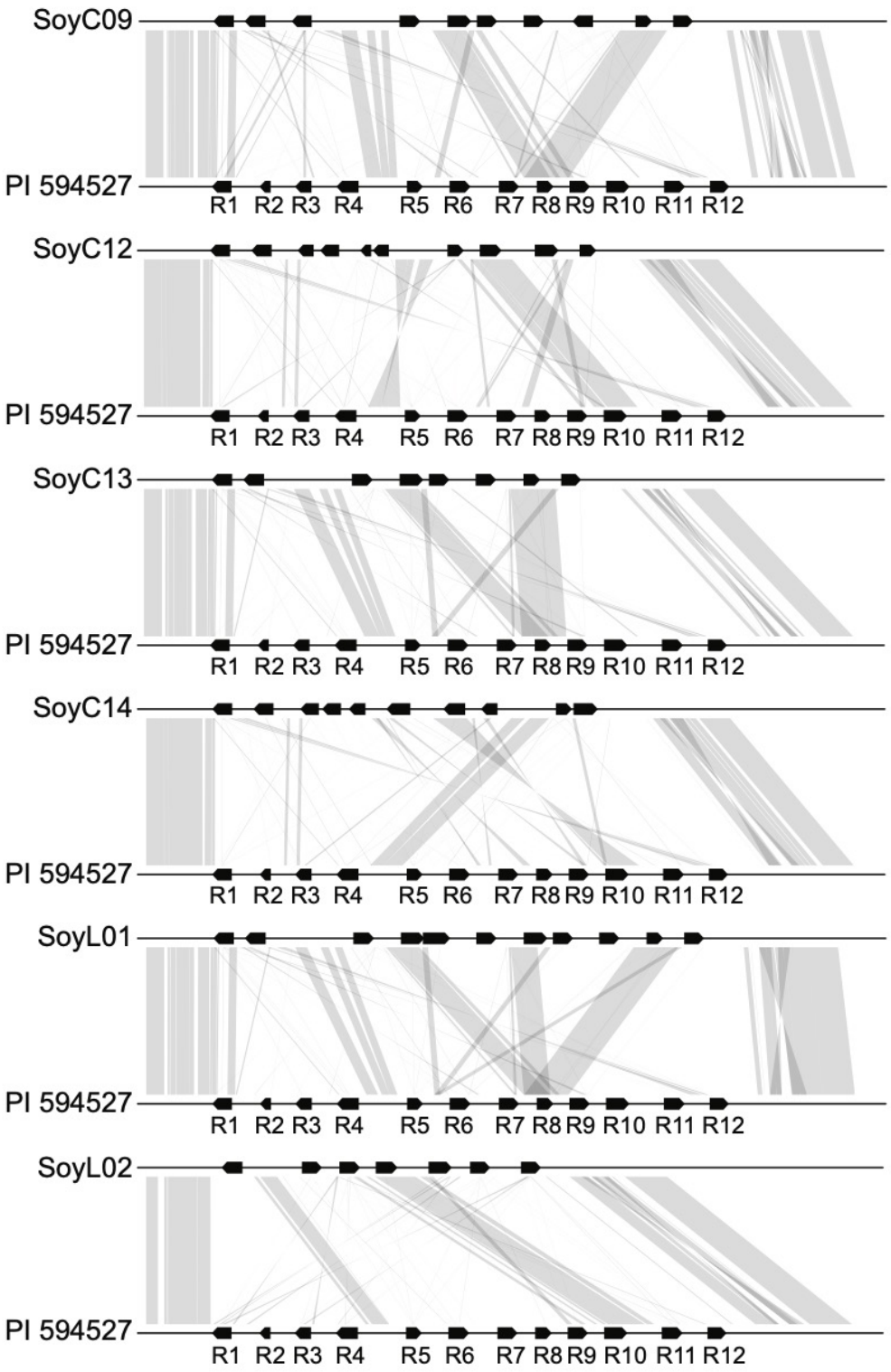


Figure 4.25 continued

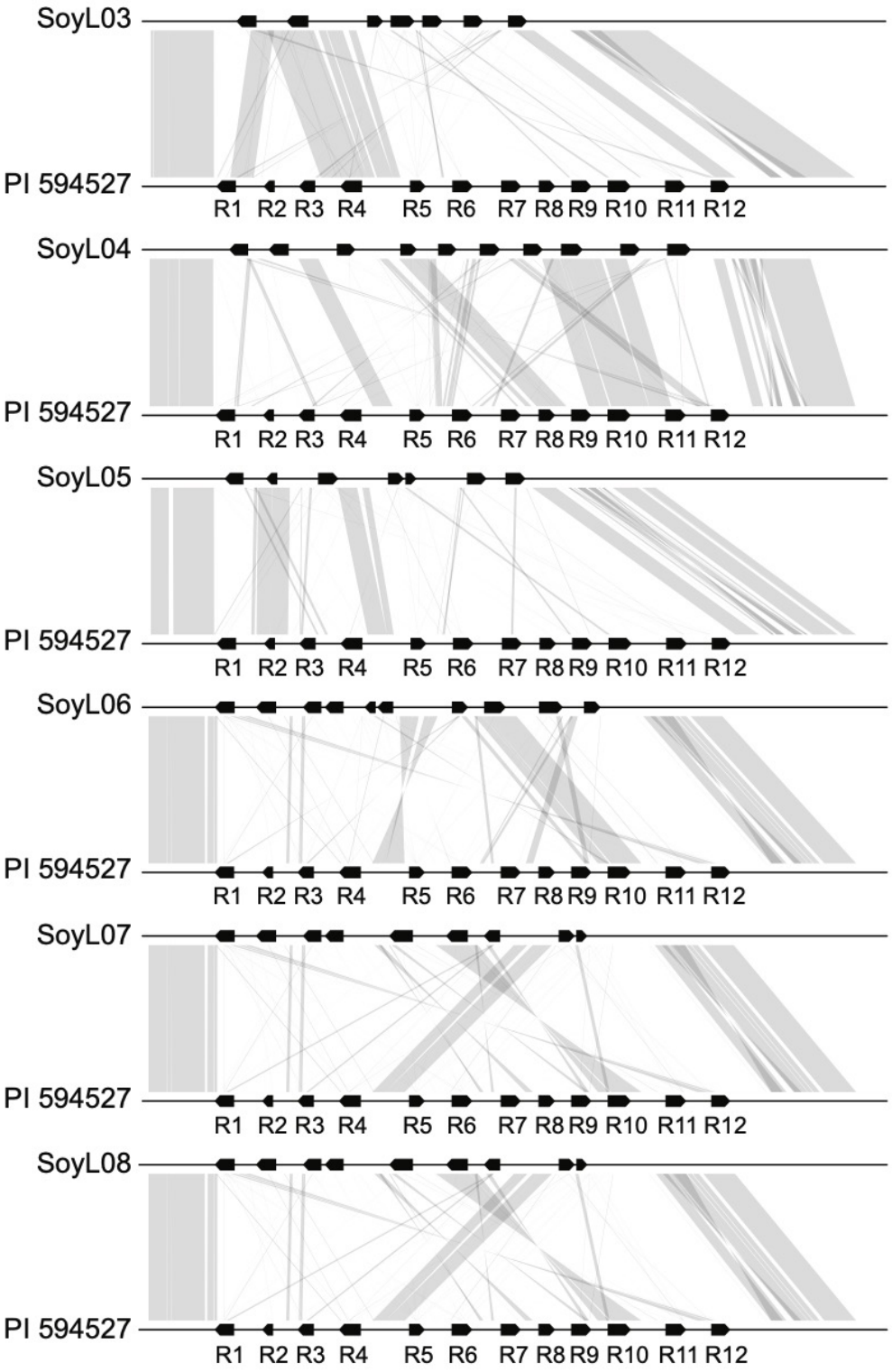


Figure 4.25 continued

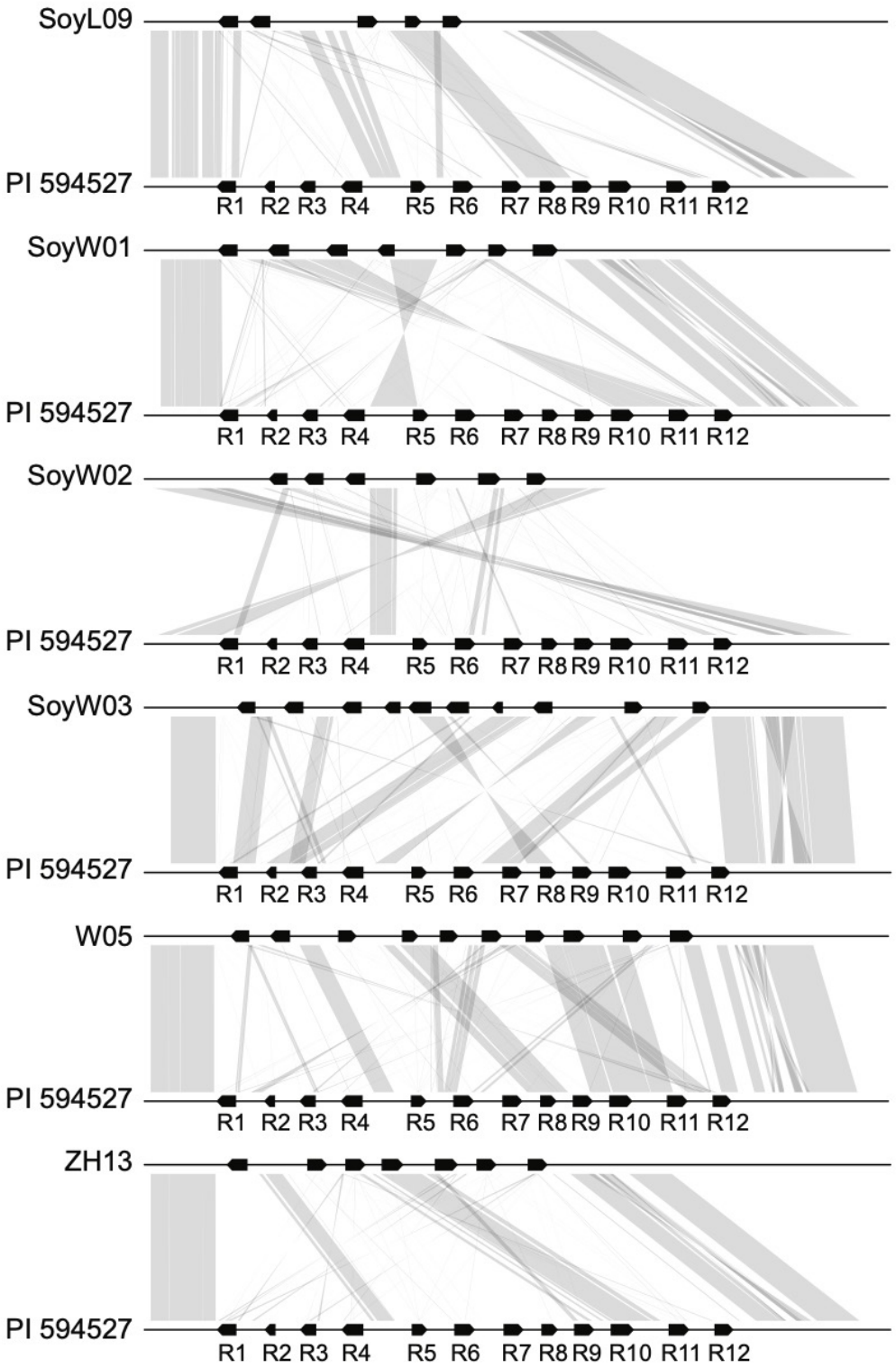
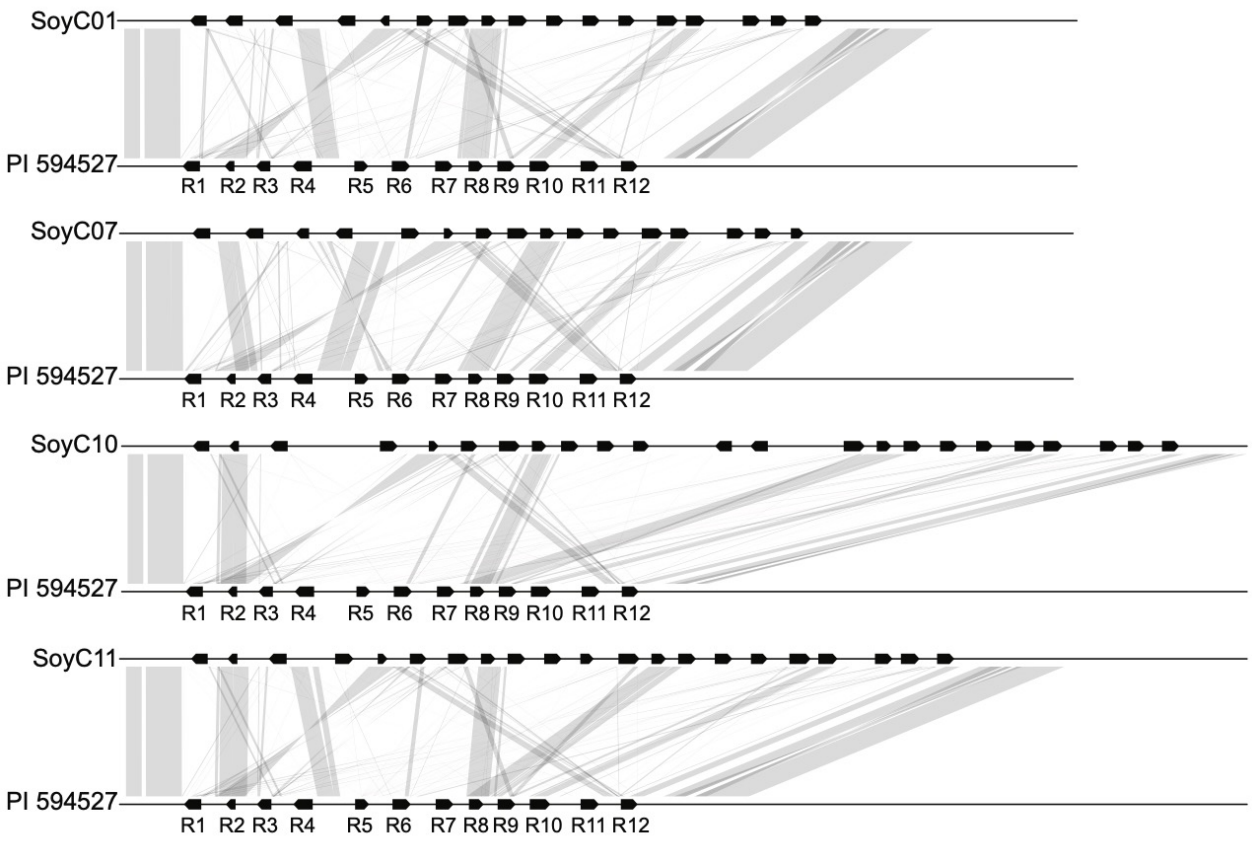


Figure 4.25 continued



Furthermore, we constructed a phylogenetic tree using the SNPs from the *Rps11* region and identified four major haplotype groups with high sequence similarity within but not between groups (Figure 4.26). Group I included SoyC06, SoyC12, SoyL06, SoyL07, SoyL08 and SoyC14; There was an inversion event within this group: R5-R6-R7-R8-R9-R10 in SoyC12 and SoyL06 was inverted from R10-R9-R8-R7-R6-R5 in SoyL07, SoyL08 and SoyC14 (Figure 4.26). In Group II, SoyC03 shared the highest sequence similarity with the *Rps11* donor line; SoyL04 and ZH13 share exactly the same haplotype with W05 and L02, respectively (Figure 4.26). Group III included SoyC07, SoyC01, SoyC11, SoyC10, SoyC02 and SoyL05. Several inversions and segmental duplications involving 2 to 5 NLR genes resulted in this group having the highest copy number (Figure 4.26). A deletion, potentially induced by unequal recombination, was detected between SoyC01 and SoyC07 (Box 1, Figure 4.26). Group IV shared a similar haplotype with the reference genome, Williams 82, with a tandem duplication event in SoyL01 induced by unequal recombination (Box 2, Figure 4.26). Taken together, our analyses revealed the complexity of the

NLR gene cluster across diverse soybean genomes and suggested that *Rp11* (R6) was unique to PI 594527. Cloning of *Rps11* will enable more precise selection of *Rps11* for cultivar improvement.

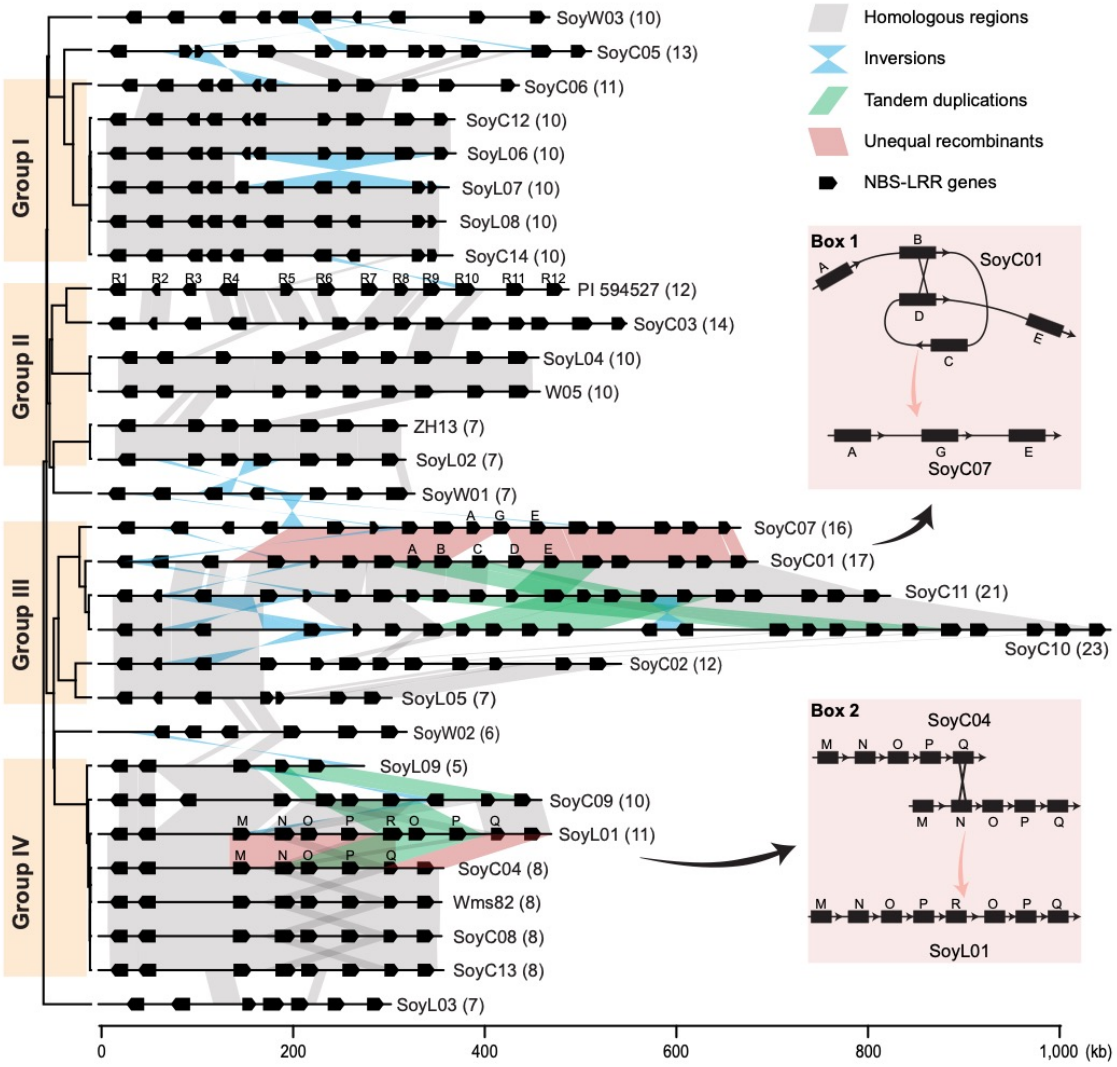


Figure 4.26 Diversification of the NBS-LRR gene cluster across 30 diverse soybean genomes.

The phylogenetic tree on the left side was built using SNPs data with the *Rps11* region. Orange color highlights the four major haplotype groups. Each black box represents an NBS-LRR gene. Grey shades represent syntenic blocks among genomes. Light-blue highlight inversion events. Green highlight tandem duplication events. Light-red highlight potential unequal recombination events. Box 1 and Box 2 exemplify a deletion and a duplication event induced by unequal recombination, respectively. The name of each genome was labeled at right side of each cluster and the numbers in parentheses are the total copy number of NBS-LRR genes in each genome at the *Rps11* corresponding region.

In addition to the structural and copy number variations, we also observed drastic variations in the expression levels of the NLR genes. While 97.8% (45 out of 46) NLR genes in the WGD region are expressed, only 59.8% (189 out of 316) of the NLR genes in the *Rps11* region are expressed (Table 4.5). To explain the expression differences of the NLR genes between the *Rps11* region and its WGD region and the variations within the *Rps11* region, we performed 5' Rapid amplification of cDNA ends (RACE) for the 5 expressed NLR genes in the *Rps11* region and Chr16.R1 in the WGD region in the *Rps11* donor line. We identified two independent transcription start regions (TSR), dubbed TSR1 and TSR2, at ~4kb and ~2.5kb upstream of Chr16.R1 (Figure 4.27a), both of which were found in every NLR gene in the WGD regions in all 29 genomes. Interestingly, only TSR2 was found in each of the 5 NLR genes in the *Rps11* region while TSR1 was completely absent, which might explain the significantly higher expression level of the NLR genes in the WGD region compared to those from the *Rps11* region (Figure 4.27b). Furthermore, we observed that the expression levels of NLR genes in the *Rps11* region were significantly associated with the presence/absence of TSR2 (Figure 4.27c), indicating that loss of TSR2 might be responsible for the expression variations of NLR genes within the *Rps11* region. By comparing the sequence of the *Rps11* donor line to the other genomes, we found that the *Rps11* promoter region shares high similarity with segments of the promoter regions of W05.R5 and W05.R6, indicating that an ancient unequal recombination event may have brought portions of R5 and R6 together to form the extremely long promoter region of *Rps11* (Figure 4.27d).

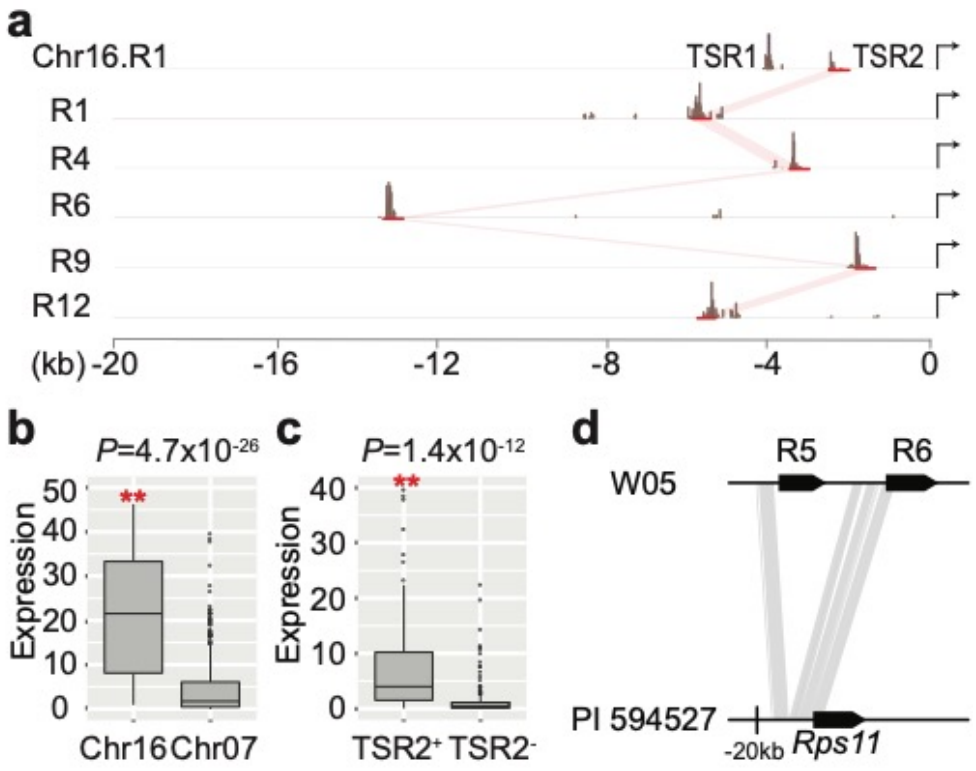


Figure 4.27 Analysis of the transcription start region of the NLR genes.

a, 5' Rapid amplification of cDNA ends (RACE) for the 5 expressed NBS-LRR genes (R1, R4, R6, R9 and R12) as well as the NBS-LRR gene at the WGD region (Chr16.R1). x axis represents the distance to the first exon. Brow bars represents the 5' RACE reads mapped to each gene.
b, Comparison of the expression levels of NBS-LRR genes from Rps11 region (Chr07) and its WGD region (Chr16). Red asterisks indicate the significance.
c, Comparison of the expression levels of NBS-LRR with TSR2 (TSR^+) and that without TSR2 ($TSR2^-$) at the *Rps11* region. Red asterisks indicate the significance.
d, Sequence comparison of the promoter regions of *Rps11* and that of W05.r5 and W05.r6. Light-blue shades represent syntenic blocks.

CHAPTER 5. DISCUSSION

5.1 Significance of cloning the *Rps11* locus

In this study, we demonstrate that *Rps11* (R6), a ~27.7-kb NLR gene, confers broad spectrum resistance to *P. sojae*. Although Gao *et al.* claimed to have cloned *Rps1-k* (Gao and Bhattacharyya, 2008; Gao et al., 2005), neither of the two NLR genes they reported were found in any versions of the genome assembly of Williams 82, which carries *Rps1-k*. Furthermore, the sequence contig used in their study was found to be an incorrect assembly of sequences from different chromosomes (Figure 5.1). The similarity between the two NLR genes identified by Gao *et al*. and the two closest NLR genes found in Williams 82 v3.0 is only around 95% (Figure 5.1). Therefore, the gene(s) underlying *Rps1-k* still remain unclear and *Rps11* is most likely the first gene cloned to confer resistance to *P. sojae*. PacBio sequencing was also shown to be much more powerful for assembling NLR gene clusters compared to short-read sequencing.

In addition, we found the NLR gene cluster was mainly formed by inversions and segmental tandem duplications after the WGD, some of which were induced by unequal recombination among different NLR genes. Unequal recombination has been proposed as a mechanism for creating novel resistance specificities (Hammond-Kosack and Jones, 1997; Hulbert, 1997; McHale et al., 2006; Parniske et al., 1997; Pryor and Ellis, 1993; Richter et al., 1995), but high frequencies of unequal recombination also put the resistant gene at high risk of disruption. Precise isolation of the NLR gene underlying resistance will significantly accelerate its application in breeding programs towards effective management of PRSR.

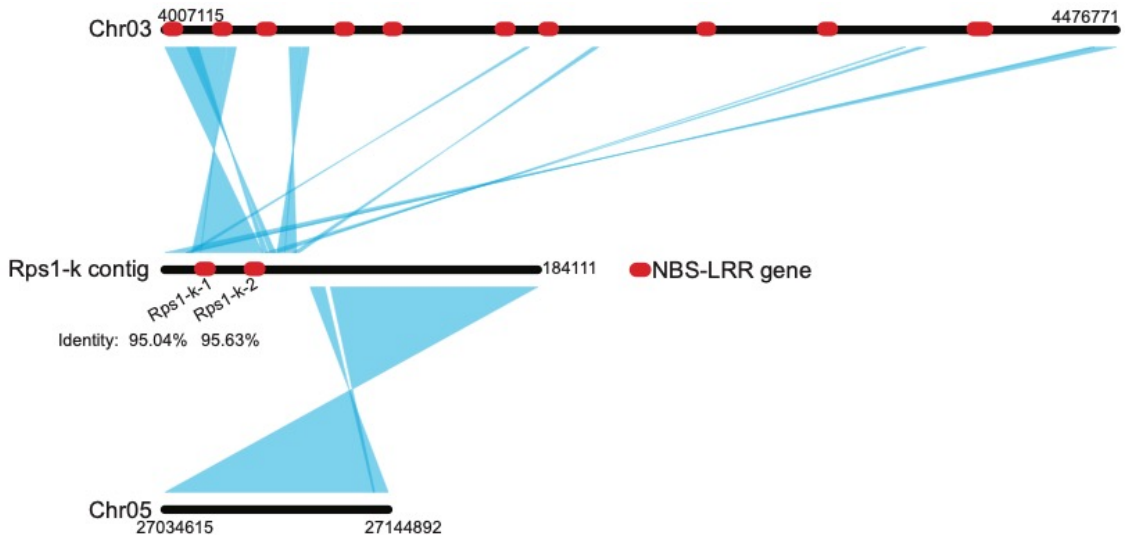


Figure 5.1 Sequence comparison between the Rps1-k contig and the Williams 82 genome. Black bars represent genome sequence of chromosome 3 (top), *Rps1-k* contig (middle) used for cloning, and chromosome 5 (bottom). Red bars represent NBS-LRR genes Light-blue shades represent alignments among different sequences.

5.2 Possible mechanisms underlying the broad resistance spectrum of *Rps11*

Sequence analyses suggested the giant size of *Rps11* is mainly a result of LRR tandem duplications which occurred before the whole genome duplication event in ~ 13 million years ago. The LRR domains in NLR genes are involved in determining the plant's ability to recognize specific pathogen effectors (McHale *et al.*, 2006). Whether the tandem duplications of the LRR domains found in *Rps11* are responsible for its broad resistance spectrum remains to be tested. Two possible mechanisms might explain the broad resistance spectrum of *Rps11*. First, the effector recognized by *Rps11* could be a conserved effector shared by a high proportion of the *P. sojae* population. Second, it is possible that *Rps11* recognize multiple effectors from different *P. sojae* isolates, which combine together responsible for the broad resistance spectrum of *Rps11*.

In this study, we have screened the *Rps11* locus with 158 isolates collected from fields across Indiana and identify 31 isolates that can defeat *Rps11*. Given that these isolates distributed across Indiana evenly, they might show high diversity. It would be interesting to sequence these isolates as well as a set of representative isolates that are not able to defeat *Rps11* to perform a genome-wide association study. If a single major association signal were detected, it is more likely that

Rps11 can recognize a single conserved effector from *P. sojae*. If multiple association signals were detected, it is more likely that *Rps11* can recognize different effectors. Sequence comparisons between the two categories of the isolates can help to isolate the effector gene(s) in *P. sojae* interacting with *Rps11*. One limitation of this genome-wide association study would be the small population size as well as the narrow genetic diversity of the tested isolates in this study since they were all collected from Indiana. Dorrance *et al*. has collected 213 unique isolates across eleven States across US (Dorrance et al., 2016), it would be more powerful if these isolates could be tested on *Rps11* and sequenced for identification of effector gene(s) interacting with *Rps11*.

5.3 Plant disease resistance not involving NLR genes

Besides disease resistance involving NLR genes, plants also have other different kinds of defense strategies when they are attacked by their own pathogens. In general, these resistances depend on either structural characteristics that can avoid the pathogens from entering and spreading in the plant cell and tissues, or biochemical reactions with products that can inhibit or kill the pathogens (Jones and Dangl, 2006). Most of the physical barriers are preexist. For example, the waxes on the surfaces of some plant leaves or stem can prevent pathogens from entering the plants. Some plants have very thick wall of epidermal cells which also important barriers to pathogens.

On contrast, biochemical defenses can be either preexist before infection or induced by pathogens. Some plants can produce one or more inhibitory compounds in young leaves, fruits, or seeds, for example phenolic compounds like tannins, or fatty acid-like compounds like dienes, and these compounds were proposed to play important roles in plant resistance. However, some toxic substances were produced by plants only when they are attacked by pathogens. Those pathogens can induce the expression of genes in plants to produce compounds that reduce the damage caused by the pathogens. All the resistances discussed above are nonspecific to a certain pathogen. They can prevent attacks from different pathogens. Transmembrane pattern recognition receptors of some plants, however, can recognize a specific pathogen by recognizing molecular patterns in the pathogen such as flagellin, which can trigger resistance reaction in plants.

Besides these natural resistance in plants. Engineered resistance is also very important, especially for crops (Salomon and Sessa, 2012). In addition to resistance genes, plants also have susceptibility

genes (S gene) that can facilitate the proliferation of pathogens. Genome editing on these susceptibility genes (for example using CRISPR/CAS9 to knock out these S genes) can also improve resistance of plants. Another good example of engineered resistance is pathogen-derived resistance. Transgenic over-expression of viral RNA can produce double strand RNA and eventually trigger RNAi to prevent the infection of virial pathogens (Voinnet, 2001).

5.4 Factors affecting durability of a R gene

Factors from both R gene side and pathogen side will affect the durability of a R gene. R genes with narrower resistance spectrum might be overcome faster because the pathogen isolates in a field changed over time due to quick evolution or long-distance spread. On contrast, R genes with wide resistance spectrum might interact with core effectors in the pathogen which shared by most isolates of the pathogens. Therefore, the duration of these R genes might be longer. R genes located at R gene enriched region (several copies of R gene at same locus) might be overcome faster because these R genes are easier to loss function due to mutations caused by unequal recombination. On contrast, R genes located at genome region with low mutation and recombination will be more stable. R genes with partners during recognition might be overcome faster because once their partner genes mutated, the recognition will also fail. Mutations on these partners will cause the loss of resistance of these R genes.

In order to maximize the durability of a R gene, we first need to select R genes with broad resistance spectrum, which means the R gene carries resistance to multiple strains of a pathogen. Therefore, it will not be easily defeated by a single strain. In addition, combination of different R genes will also dramatically increase the durability of resistance because it is less likely all the R genes are defeated simultaneously. We also need to regularly check the sequence of the R gene after deployment to make sure no mutations in the R gene occur.

From the pathogen side, it is very important to understand the virulence gene interacting with the R gene. So, we can check the diversity of the pathogen population in the field before deployment of the R gene to see if any pathogen strains carry virulence genes to the R gene. In addition, combining R gene with other means of disease control, such as chemical management, will also increase durability.

REFERENCES

Abney, T., Melgar, J., Richards, T., Scott, D., Grogan, J., and Young, J. (1997). New races of Phytophthora sojae with Rps 1-d virulence. Plant disease **81**:653-655.

Anderson, J.P., Gleason, C.A., Foley, R.C., Thrall, P.H., Burdon, J.B., and Singh, K.B. (2010). Plants versus pathogens: an evolutionary arms race. Functional plant biology **37**:499-512.

Anderson, T., and Buzzell, R. (1992). Inheritance and linkage of the Rps7 gene for resistance to Phytophthora rot of soybean. Plant disease **76**:958-959.

Athow, K., and Laviolette, F. (1982). Rps 6, a major gene for resistance to Phytophthora megasperma f. sp. glycinea in soybean. Phytopathology **72**:1564-1567.

Athow, K., Laviolette, F., Mueller, E., and Wilcox, J. (1980). A new major gene for resistance to Phytophthora megasperma var. sojae in soybean. Phytopathology **70**:977-980.

Bernard, R., and Cremeens, C. (1981). An allele at the rps1 locus from the variety'Kingwa'. Soybean Genetics Newsletter **8**:14.

Bernard, R., Smith, P., Kaufmann, M., and Schmitthenner, A. (1957). Inheritance of Resistance to Phytophthora Root and Stem Rot in the Soybean 1. Agronomy Journal **49**:391-391.

Burnham, K., Dorrance, A., Francis, D., Fioritto, R., and St. Martin, S. (2003). Rps8, a new locus in soybean for resistance to Phytophthora sojae. Crop science **43**:101-105.

Buzzell, R. (1981). Another major gene for resistance to Phytophthora megasperma var. sojae in soybean. Soybean Genet. Newsl. **18**:30-33.

Buzzell, R., and Anderson, T. (1992). Inheritance and race reaction of a new soybean Rps1 allele. Plant disease **76**:600-601.

Carter, T., Hymowitz, T., and Nelson, R. (2004). Biogeography, local adaptation, Vavilov, and genetic diversity in soybean. In Biological resources and migration, (Springer: pp. 47-59.

Cheng, Y., Ma, Q., Ren, H., Xia, Q., Song, E., Tan, Z., Li, S., Zhang, G., and Nian, H. (2017). Fine mapping of a Phytophthora-resistance gene RpsWY in soybean (Glycine max L.) by high-throughput genome-wide sequencing. Theoretical and applied genetics **130**:1041-1051.

Chu, J., Peng, B., Tang, K., Yi, X., Zhou, H., Wang, H., Li, G., Leng, J., Chen, N., and Feng, X. (2021). Construction and analysis of eight soybean reference genomes reveal structural variations important for domestication and genetic breeding. Sientific Data **(In press)**.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**:15-21.

Dorrance, A., and Schmitthenner, A. (2000). New sources of resistance to Phytophthora sojae in the soybean plant introductions. Plant Disease **84**:1303-1308.

Dorrance, A., Kurle, J., Robertson, A., Bradley, C., Giesler, L., Wise, K., and Concibido, V. (2016). Pathotype diversity of Phytophthora sojae in eleven states in the United States. Plant disease **100**:1429-1437.

Dorrance, A.E., Mills, D., Robertson, A.E., Draper, M.A., Giesler, L., and Tenuta, A. (2007). Phytophthora root and stem rot of soybean. The Plant Health Instructor:1.

Dou, D., Kale, S.D., Wang, X., Jiang, R.H., Bruce, N.A., Arredondo, F.D., Zhang, X., and Tyler, B.M. (2008). RXLR-mediated entry of Phytophthora sojae effector Avr1b into soybean cells does not require pathogen-encoded machinery. The Plant Cell **20**:1930-1947.

Doughari, J. (2015). An overview of plant immunity. J. Plant Pathol. Microbiol **6**:10.4172.

Doyle, J.J., and Egan, A.N. (2010). Dating the origins of polyploidy events. New Phytologist **186**:73-85.

Gao, H., and Bhattacharyya, M.K. (2008). The soybean-Phytophthora resistance locus Rps1-k encompasses coiled coil-nucleotide binding-leucine rich repeat-like genes and repetitive sequences. BMC Plant Biology **8**:1-14.

Gao, H., Narayanan, N.N., Ellison, L., and Bhattacharyya, M.K. (2005). Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the Rps1-k locus encode Phytophthora resistance in soybean. Molecular plant-microbe interactions **18**:1035-1045.

Gill, N., Findley, S., Walling, J.G., Hans, C., Ma, J., Doyle, J., Stacey, G., and Jackson, S.A. (2009). Molecular and chromosomal evidence for allopolyploidy in soybean. Plant physiology **151**:1167-1174.

Goheen, E., and Frankel, S. (2009). Phytophthoras in forests and natural ecosystems. Gen. Tech. Rep. PSW-GTR-221:311-314.

Gordon, S.G., St. Martin, S.K., and Dorrance, A.E. (2006). Rps8 maps to a resistance gene rich region on soybean molecular linkage group F. Crop science **46**:168-173.

Hammond-Kosack, K.E., and Jones, J.D. (1997). Plant disease resistance genes. Annual review of plant biology **48**:575-607.

Hartman, G., Sinclair, J., and Rupe, J. (1999). Compendium of soybean diseases (Soybean Disease Compendium).

Hartwig, E., Keeling, B., and Edwards Jr, C. (1968). Inheritance of Reaction to Phytophthora Rot in the Soybean 1. Crop Science **8**:634-636.

Hulbert, S. (1997). Structure and evolution of the rp1 complex conferring rust resistance in maize. Annual review of phytopathology **35**:293-310.

Innes, R.W., Ameline-Torregrosa, C., Ashfield, T., Cannon, E., Cannon, S.B., Chacko, B., Chen, N.W., Couloux, A., Dalwani, A., and Denny, R. (2008). Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. Plant Physiology **148**:1740-1759.

Jiang, B., Cheng, Y., Cai, Z., Li, M., Jiang, Z., Ma, R., Yuan, Y., Xia, Q., and Nian, H. (2020). Fine mapping of a Phytophthora-resistance locus RpsGZ in soybean using genotyping-by-sequencing. BMC genomics **21**:1-11.

Jiang, R.H., Tripathy, S., Govers, F., and Tyler, B.M. (2008). RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. Proceedings of the National Academy of Sciences **105**:4874-4879.

Jones, J.D., and Dangl, J.L. (2006). The plant immune system. nature **444**:323-329.

Kaitany, R., Hart, L., and Safir, G. (2001). Virulence composition of Phytophthora sojae in Michigan. Plant Disease **85**:1103-1106.

Kaufmann, M.J., and Gerdemann, J. (1958). Root and stem rot of soybean caused by Phytophthora sojae n. sp. Phytopathology **48**.

Kilen, T., Hartwig, E., and Keeling, B. (1974). Inheritance of a second major gene for resistance to Phytophthora rot in soybeans 1. Crop science **14**:260-262.

Kim, M.Y., Van, K., Kang, Y.J., Kim, K.H., and Lee, S.-H. (2012). Tracing soybean domestication history: From nucleotide to genome. Breeding Science **61**:445-452.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome research **27**:722-736.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics **34**:3094-3100.

Li, Y., Sun, S., Zhong, C., Wang, X., Wu, X., and Zhu, Z. (2017). Genetic mapping and development of co-segregating markers of RpsQ, which provides resistance to Phytophthora sojae in soybean. Theoretical and Applied Genetics **130**:1223-1233.

Lin, F., Zhao, M., Ping, J., Johnson, A., Zhang, B., Abney, T.S., Hughes, T.J., and Ma, J. (2013). Molecular mapping of two genes conferring resistance to Phytophthora sojae in a soybean landrace PI 567139B. Theoretical and applied genetics **126**:2177-2185.

Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., Zhou, G.-A., Zhang, H., Liu, Z., and Shi, M. (2020). Pan-genome of wild and cultivated soybeans. Cell **182**:162-176. e113.

McHale, L., Tan, X., Koehl, P., and Michelmore, R.W. (2006). Plant NBS-LRR proteins: adaptable guards. Genome biology **7**:1-11.

Mueller, E., Athow, K., and Laviolette, F. (1978). Inheritance of resistence to four physiologic races of Phytophthora megasperma var. sojae. Phytopathology.

Muller, H., Annaluru, N., Schwerzmann, J.W., Richardson, S.M., Dymond, J.S., Cooper, E.M., Bader, J.S., Boeke, J.D., and Chandrasegaran, S. (2012). Assembling large DNA segments in yeast. In Gene Synthesis, (Springer: pp. 133-150.

Niu, J., Guo, N., Sun, J., Li, L., Cao, Y., Li, S., Huang, J., Zhao, J., Zhao, T., and Xing, H. (2017). Fine mapping of a resistance gene RpsHN that controls Phytophthora sojae using recombinant inbred lines and secondary populations. Frontiers in plant science **8**:538.

Nowicki, M., Foolad, M.R., Nowakowska, M., and Kozik, E.U. (2012). Potato and tomato late blight caused by Phytophthora infestans: an overview of pathology and resistance breeding. Plant disease **96**:4-17.

Osbourn, A.E. (1996). Preformed antimicrobial compounds and plant defense against fungal attack. The plant cell **8**:1821.

Ou, S., Liu, J., Chougule, K.M., Fungtammasan, A., Seetharam, A.S., Stein, J.C., Llaca, V., Manchanda, N., Gilbert, A.M., and Wei, S. (2020). Effect of sequence depth and length in long-read assembly of the maize inbred NC358. Nature communications **11**:1-10.

Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B., and Jones, J.D. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell **91**:821-832.

Ping, J., Fitzgerald, J.C., Zhang, C., Lin, F., Bai, Y., Wang, D., Aggarwal, R., Rehman, M., Crasta, O., and Ma, J. (2016). Identification and molecular mapping of Rps11, a novel gene conferring resistance to Phytophthora sojae in soybean. Theoretical and Applied Genetics **129**:445-451.

Pryor, T., and Ellis, J. (1993). The genetic complexity of fungal resistance genes in plants. Advances in Plant Pathology **10**:281-305.

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.

Sahoo, D.K., Abeysekara, N.S., Cianzio, S.R., Robertson, A.E., and Bhattacharyya, M.K. (2017). A novel Phytophthora sojae resistance Rps12 gene mapped to a genomic region that contains several Rps genes. PloS one **12**:e0169950.

Salomon, D., and Sessa, G. (2012). Biotechnological strategies for engineering plants with durable resistance to fungal and bacterial pathogens. Plant Biotechnology and Agriculture:329-342.

Schmitthenner, A. (1985). Problems and progress in control of Phytophthora root rot of soybean. Plant disease **69**:362-368.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., and Cheng, J. (2010). Genome sequence of the palaeopolyploid soybean. nature **463**:178-183.

Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., and Chavarro, C. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. Nature genetics **46**:707-713.

Shen, Y., Du, H., Liu, Y., Ni, L., Wang, Z., Liang, C., and Tian, Z. (2019). Update soybean Zhonghuang 13 genome to a golden reference. Sci China Life Sci **62**:1257-1260.

Shen, Y., Liu, J., Geng, H., Zhang, J., Liu, Y., Zhang, H., Xing, S., Du, J., Ma, S., and Tian, Z. (2018). De novo assembly of a Chinese soybean genome. Science China Life Sciences **61**:871-884.

Song, Q., Hyten, D.L., Jia, G., Quigley, C.V., Fickus, E.W., Nelson, R.L., and Cregan, P.B. (2015). Fingerprinting soybean germplasm and its utility in genomic research. G3: Genes, genomes, genetics **5**:1999-2006.

Steuernagel, B., Witek, K., Krattinger, S.G., Ramirez-Gonzalez, R.H., Schoonbeek, H.-j., Yu, G., Baggs, E., Witek, A.I., Yadav, I., and Krasileva, K.V. (2020). The NLR-Annotator tool enables annotation of the intracellular immune receptor repertoire. Plant physiology **183**:468-482.

Sun, J., Li, L., Zhao, J., Huang, J., Yan, Q., Xing, H., and Guo, N. (2014). Genetic analysis and fine mapping of RpsJS, a novel resistance gene to Phytophthora sojae in soybean [Glycine max (L.) Merr.]. Theoretical and Applied Genetics **127**:913-919.

Sun, S., Wu, X., Zhao, J., Wang, Y., Tang, Q., Yu, D., Gai, J., and Xing, H. (2011). Characterization and mapping of RpsYu25, a novel resistance gene to Phytophthora sojae. Plant breeding **130**:139-143.

Tooley, P.W., and Grau, C.R. (1982). Identification and quantitative characterization of rate-reducing resistance to Phytophthora megasperma f. sp. glycinea in soybean seedlings. Phytopathology **72**:727-733.

Tyler, B.M. (2007). Phytophthora sojae: root rot pathogen of soybean and model oomycete. Molecular plant pathology **8**:1-8.

Tyler, B.M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R.H., Aerts, A., Arredondo, F.D., Baxter, L., Bensasson, D., and Beynon, J.L. (2006). Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science **313**:1261-1266.

Underwood, W. (2012). The plant cell wall: a dynamic barrier against pathogen invasion. Frontiers in plant science **3**:85.

Valliyodan, B., Brown, A.V., Wang, J., Patil, G., Liu, Y., Otyama, P.I., Nelson, R.T., Vuong, T., Song, Q., and Musket, T.A. (2021). Genetic variation among 481 diverse soybean accessions, inferred from genomic re-sequencing. Scientific data **8**:1-9.

Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. TRENDS in Genetics **17**:449-459.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., and Young, S.K. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one **9**:e112963.

Wrather, A., and Koenning, S. (2009). Effects of diseases on soybean yields in the United States 1996 to 2007. Plant Health Progress **10**:24.

WU, X.-l., ZHANG, B.-q., Shi, S., ZHAO, J.-m., Feng, Y., Na, G., GAI, J.-y., and Han, X. (2011a). Identification, genetic analysis and mapping of resistance to Phytophthora sojae of Pm28 in soybean. Agricultural Sciences in China **10**:1506-1511.

Wu, X., Zhou, B., Sun, S., Zhao, J., Chen, S., Gai, J., and Xing, H. (2011b). Genetic analysis and mapping of resistance to Phytophthora sojae of Pm14 in soybean. Scientia Agricultura Sinica **44**:456-460.

Xie, M., Chung, C.Y.-L., Li, M.-W., Wong, F.-L., Wang, X., Liu, A., Wang, Z., Leung, A.K.-Y., Wong, T.-H., and Tong, S.-W. (2019). A reference-grade wild soybean genome. Nature communications **10**:1-12.

Yu, A., Xu, P., Wang, J., Zhang, S., Wu, J., Li, W., Chen, W., Li, N., Fan, S., and Wang, X. (2010). Genetic analysis and SSR mapping of gene resistance to Phytophthora sojae race 1 in soybean cv Suinong 10. Chinese Journal of Oil Crop Sciences **32**:462-466.

Zhang, J., Xia, C., Wang, X., Duan, C., Sun, S., Wu, X., and Zhu, Z. (2013a). Genetic characterization and fine mapping of the novel Phytophthora resistance gene in a Chinese soybean cultivar. Theoretical and Applied Genetics **126**:1555-1561.

Zhang, J., Xia, C., Duan, C., Sun, S., Wang, X., Wu, X., and Zhu, Z. (2013b). Identification and candidate gene analysis of a novel Phytophthora resistance gene Rps10 in a Chinese soybean cultivar. PloS one **8**:e69799.

ZHANG, X.-C., ZHONG, C., DUAN, C.-X., SUN, S.-L., and ZHU, Z.-D. (2020). Fine mapping of Phytophthora resistance gene RpsZheng in soybean cultivar Zheng 97196. Acta Agronomica Sinica **46**:997-1005.

Zhendong, Z., Yunlong, H., Xiaoming, W., Junbin, H., and Xiaofei, W. (2010). Molecular identification of a novel Phytophthora resistance gene in soybean.

Zhong, C., Sun, S., Li, Y., Duan, C., and Zhu, Z. (2018a). Next-generation sequencing to identify candidate genes and develop diagnostic markers for a novel Phytophthora resistance gene, RpsHC18, in soybean. Theoretical and applied genetics **131**:525-538.

Zhong, C., Li, Y., Sun, S., Duan, C., and Zhu, Z. (2019). Genetic mapping and molecular characterization of a broad-spectrum Phytophthora sojae resistance gene in Chinese soybean. International journal of molecular sciences **20**:1809.

Zhong, C., Sun, S., Yao, L., Ding, J., Duan, C., and Zhu, Z. (2018b). Fine mapping and identification of a novel Phytophthora root rot resistance locus RpsZS18 on chromosome 2 in soybean. Frontiers in plant science **9**:44.

Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., Yu, Y., Shu, L., Zhao, Y., and Ma, Y. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nature biotechnology **33**:408-414.

VITA

Weidong Wang

Department of Agronomy, Purdue University

West Lafayette, IN 47906

(a) Professional Preparation

China Agricultural University	Beijing, China	Agronomy	B.S. 2010
China Agricultural University	Beijing, China	Crop Genetics & Breeding	M.S. 2013
Purdue University	West Lafayette, USA	Agronomy	Ph.D. 2021

(b) Appointments

2014-2015 Graduate Research Assistant, Plant Biological Science Department, University of Minnesota-twin cities, MN, USA

2013-2014 Research Intern, Genetics Department, University of Wisconsin-Madison, Madison, WI, USA

(c) Honors and Awards

Graduate Research and Travel Award by Center for Plant Biology, Purdue University	2019
Outstanding Ph.D. Research Award by Agronomy Department, Purdue University	2018
George W. Bailey Travel Scholarship by Agronomy Department, Purdue University	2017
Dow AgroScience Graduate Scholarship, Agronomy Department, Purdue University	2017
George W. Bailey Travel Scholarship by Agronomy Department, Purdue University	2016
Top 10 Excellent M.S. Degree thesis award by China Agricultural University	2013

(d) Publications

1. Mohsen Mohammadi, Alencar Xavier, Travis Beckett, Savannah Beyer, Liyang Chen, Habte Chikssa, Valerie Cross, Fabiana Freitas Moreira, Elizabeth French, Rupesh Gaire, Stefanie Griebel, Miguel Angel Lopez, Samuel Prather, Blake Russell, **Weidong Wang**. "Identification, deployment, and transferability of quantitative trait loci from genome-wide association studies in plants." Current Plant Biology (2020): 100145.

2. Stephen A. Swarm, Lianjun Sun, Xutong Wang, **Weidong Wang**, Patrick J. Brown, Jianxin Ma, and Randall L. Nelson. "Genetic dissection of domestication-related traits in soybean through genotyping-by-sequencing of two interspecific mapping populations." **Theoretical and Applied Genetics** (2019): 1-15.

3. Chin Jian Yang, Luis Fernando Samayoa, Peter J. Bradbury, Bode A. Olukolu, Wei Xue, Alessandra M. York, Michael R. Tuholski, **Weidong Wang**, Lora L. Daskalska, Michael A. Neumeyer, Jose de Jesus Sanchez-Gonzalez, Maria Cinta Romay, Jeffrey C. Glaubitz, Qi Sun, Edward S. Buckler, James B. Holland, John F. Doebley. "The genetic architecture of teosinte catalyzed and constrained maize domestication." **Proceedings of the National Academy of Sciences** (2019): 201820997.

4. Dajian Zhang[#], Lianjun Sun[#], Shuai Li[#], **Weidong Wang[#]**, Yanhua Ding[#], Stephen A Swarm, Linghong Li, Xutong Wang, Xuemin Tang, Zhifang Zhang, Zhixi Tian, Patrick J Brown, Chunmei Cai, Randall L Nelson, Jianxin Ma. "Elevation of soybean seed oil content through selection for seed coat shininess" **Nature Plant** (2018) **(#These authors contribute equally to the paper)**

5. Dajian Zhang, Meixia Zhao, Shuai Li, Lianjun Sun, **Weidong Wang**, Chunmei Cai, Emily C. Dierking, and Jianxin Ma. "Plasticity and innovation of regulatory mechanisms underlying seed oil content mediated by duplicated genes in the palaeopolyploid soybean." **The Plant Journal** (2017).

6. Linghong Li, Feng Lin, **Weidong Wang**, Jieqing Ping, Joshua C. Fitzgerald, Meixia Zhao, Shuai Li, Lianjun Sun, Chunmei Cai, and Jianxin Ma. "Fine mapping and candidate gene analysis of two loci conferring resistance to Phytophthora sojae in soybean." **Theoretical and Applied Genetics** 129, no. 12 (2016): 2379-2386.

7. Yingjie Xiao, Hao Tong, Xiaohong Yang...(many authors) **Weidong Wang**, Debo Zheng, Jianbing Yan. Genome-wide dissection of the maize ear genetic architecture using multiple populations. ***New Phytologist***, 2016, DOI: 10.1111/nph.13814

8. Hui Li[#], Zhiyu Peng[#], Xiaohong Yang[#], **Weidong Wang**[#], Junjie Fu[#], Jianhua Wang[#], Yinjia Han, Yuchao Chai, Tingting Guo, Ning Yang, Jie Liu, Marilyn L. Warburton, Yanbing Cheng, Xiaomin Hao, Pan Zhang, Jinyang Zhao, Yunjun Liu, Guoying Wang, Jiansheng Li, Jianbing Yan. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. ***Nature Genetics***, 2013, 45: 43-50 ([#]**These authors contribute equally to the paper**)

9. Junjie Fu, Yanbing Cheng, Jingjing Linghu, Xiaohong Yang, Lin Kang, Zuxin Zhang, Bo Wang, Zhiyu Peng, Jie Zhang, Lihong Zhai, Changmin Dai, Cheng He, Jiabao Xu, **Weidong Wang**, Xiangru Li, Jun Zheng, Li Chen, Longhai Luo, Junjie Liu, Xiaoju Qian, Jianbing Yan, Jun Wang, Guoying Wang. RNA sequencing reveals the complex regulatory network in maize kernel. ***Nature Communication***, 2013, 4

10. Shutu Xu, Dalong Zhang, Ye Cai, Yi Zhou, Shah Trushar, Qing Li, Zhigang Li, **Weidong Wang**, Jiansheng Li, Xiaohong Yang, Jianbing Yan. Dissecting tocopherols content in maize (Zea mays L.), using two segregating populations and high density SNP markers. ***BMC Plant Biology*** 2012, 12:201

(e) Conference Poster Presentations

Weidong Wang, Ping Jieqing, Fitzgerald Joshua C, *et al.* Identification and molecular mapping of *Rps11*, a novel gene conferring resistance to *Phytophthora sojae* in soybean. PAG XXVI, January 13-17, 2018, San Diego, CA, USA

Weidong Wang, Ping Jieqing, Fitzgerald Joshua C, *et al.* Identification and molecular mapping of *Rps11*, a novel gene conferring resistance to *Phytophthora sojae* in soybean. PAG XXV, January 14-18, 2017, San Diego, CA, USA

Weidong Wang, Ping Jieqing, Fitzgerald Joshua C, *et al.* Identification and molecular mapping of *Rps11*, a novel gene conferring resistance to *Phytophthora sojae* in soybean. Soy2016 16th Biennial Conference of the Molecular and Cellular Biology of the Soybean, poster presentation, Columbus, OH, U.S.

Weidong Wang, Irina Makarevitch, Amanda J Waters, Nathan M Springer. Allele Specific Responses to Salt and UV in Maize. 2015 Maize Genetics Conference, poster presentation, St. Charles, IL, U.S.

Weidong Wang, Yan Li, Yingyu Liu, Jiansheng Li, Xiaohong Yang, Jianbing Yan. Identification of expression regulatory hotspots in developing maize kernels. 2013 Maize Genetics Conference, poster presentation, St. Charles, IL, U.S.

(f) Academic Services

Serve as reviewer for Crop Journal, Theoretical and Applied Genetics, Molecular Breeding and The Plant Journal.

(g) Teaching Experience

Teaching Assistant in AGRY 320 Genetics	2019 Spring
Teaching Assistant in AGRY 320 Genetics	2018 Spring
Teaching Assistant in AGRY320 Genetics	2017 Spring