Whole Genome Re-sequencing Reveals Natural Variation and Adaptive Evolution of *Phytophthora sojae*

Xiong Zhang[1,2†], Bo Liu[3†], Fen Zou[1], Danyu Shen[1], Zhiyuan Yin[2], Rongbo Wang[1], Feng He[1], Yuanchao Wang[1], Brett M. Tyler[4], Wei Fan[3*], Wanqiang Qian[3*] and Daolong Dou[1,2*]

[1] Department of Plant Pathology, Nanjing Agricultural University, Nanjing, China, [2] Department of Plant Pathology, China Agricultural University, Beijing, China, [3] Agricultural Genomic Institute, Chinese Academy of Agricultural Sciences, Shenzhen, China, [4] Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR, United States

Due to the monocultural basis of agricultural crops, mutated plant microbes with increased pathogenicity can easily spread in the field and lead to serious yield losses. As a major threat to a wide range of crop plants, oomycete pathogens continuously undergo adaptive evolution to overcome plant defense barriers. However, the genetic basis of their evolution at the molecular level remains largely unknown. Here, we investigated the nature variation and the population genomics of the soybean pathogen *Phytophthora sojae* by high-throughput genome re-sequencing. Genomic variation analysis revealed uneven "two-speed" evolutionary pattern with genes in gene-sparse regions (GSRs) showing higher rates of structural polymorphisms and positive selection. GSRs are enriched in effector genes and transposase-related genes. Our results also suggested that the NADH oxidase and MIP transporter gene families undergo rapid and diversifying selection. Furthermore, we demonstrated that *P. sojae* isolates possess varying numbers of RxLR effectors with diverse sequences, totaling 471 members. Among them, 42 core RxLR effectors are assumed to be important for infection. Finally, we observed that *Avr* genes exhibit abundant sequence variation in *P. sojae* isolates. Several novel variants lead to the evading of host resistance, including a complete deletion in *Avr3c* and amino acid mutations in *Avr1a*. Taken together, our results provide an adaptive landscape of *P. sojae* at single-nucleotide resolution, as well as resources for further resistance breeding and disease prevention against this important plant pathogen.

Keywords: *Phytophthora sojae*, genome re-sequencing, natural variation, adaptive evolution, RxLR effectors, avirulent genes

INTRODUCTION

Oomycetes are fungal-like eukaryotic organisms classified into the Stramenopila kingdom, which includes multiple plant- and animal-infecting pathogens (Fawke et al., 2015). Oomycetes cause devastating diseases on a wide range of crop plants and result in severe yield loss (Jiang and Tyler, 2012). The most notorious plant-infecting oomycete pathogens belong to the genus *Phytophthora*, which includes over 100 species (Hansen et al., 2012). Being responsible for root rot in soybean

(*Glycine max*), *Phytophthora sojae* is one of the most economically important *Phytophthora* pathogens (Tyler et al., 2006). As the second most destructive pathogen and a major yield-limiting factor for soybean, *P. sojae* can cause up to 100% yield loss in individual fields, resulting in around $2 billion of annual loss worldwide (Tyler, 2007). Despite the importance of *P. sojae* in soybean disease control, the molecular details of its pathogenesis and population genomics remain elusive.

Currently, the most effective way to manage disease caused by *P. sojae* is to develop cultivars harboring major resistance (*Rps*, *Resistance P. sojae*) genes (Dorrance et al., 2003). To date, over 20 *Rps* genes have been identified and mapped to nine chromosomes in soybean, including *Rps1a*, *Rps1b*, *Rps1c*, *Rps1k*, *Rps3a*, and *Rps6*. Some *Rps* genes have been incorporated into commercial soybean cultivars (Chawla et al., 2013; Huang et al., 2016). Meanwhile, 11 *Avr* genes have been identified and cloned from *P. sojae*, including *Avr1a*, *Avr1b*, *Avr1c*, *Avr1d*, *Avr1k*, *Avr3a*, *Avr3b*, Avr3c, *Avr4*, *Avr5*, and *Avr6* (Shan et al., 2004; Dong et al., 2009; Dou et al., 2010; Dong S. et al., 2011; Dong S.M. et al., 2011; Na et al., 2013; Song et al., 2013; Na et al., 2014). Among them, the *Avr3a/Avr5* and *Avr4/Avr6* pairs are different alleles of the same gene, respectively. All known *Avr* genes are predicted to encode effector proteins containing a secretion signal peptide followed by the RxLR (arginine-any amino acid–leucine–arginine) motif (Anderson et al., 2015). To cope with the selection pressure from soybean *Rps* genes, *P. sojae* can overcome *Rps*-mediated immune surveillance through mutations of the *Avr* genes, including amino acid changes, complete or partial deletions, frameshift, insertions and loss of transcripts (Huang et al., 2019). Continuous emergence of new *P. sojae* virulent pathotypes limits the effectiveness of *Rps* genes in the field (Schmitthenner, 1985). To prevent disease outbreaks, it is important to understand the population genomics of *P. sojae* in the field and the associated genetic variation in *Avr* genes.

Apart from their avirulent roles, *Phytophthora* RxLR effectors can play virulent roles to promote infection (Anderson et al., 2015). Every known *Phytophthora* genome harbors 100s of putative RxLR effector genes (Anderson et al., 2015). Structural studies demonstrate that the WY domain common to many RxLR type proteins possesses the core fold, which may involve in avirulence and virulence adaptation (Anderson et al., 2015). Functional studies reveal that RxLR effectors can suppress host immunity by either hijacking plant resistance pathways or utilizing plant susceptibility factors (Dou and Zhou, 2012). Comparative genomic studies indicate that most RxLR effectors are highly divergent across different pathogens (Jiang et al., 2008). Recent research has been focused on conserved effector repertoire shared by different pathogens and core effectors that are common in different strains of the same pathogen. For example, the conserved effectors between *P. sojae* and *Hyaloperonospora arabidopsidis* suppress PAMP- and effector-triggered immunity across diverse plant species (Deb et al., 2018). Four *P. infestans* core effectors contribute to virulence through defense suppression (Yin et al., 2017). Core effectors can be effective for breeding cultivars with durable resistance (Dangl et al., 2013). The re-sequencing strategy has been used to identify core effectors from 65 strains of the bacterial pathogen

Xanthomonas axonopodis pv. *manihotis* (Bart et al., 2013). Currently we have little knowledge on core RxLR effectors of oomycete species. Identification of the core RxLR effector repertoire in *P. sojae* will be critical for screening soybean *Rps* genes conferring broad-spectrum and durable resistance.

Genome plasticity can help organisms adapt to the new environment, which is particularly the case for pathogens as they coevolve with their hosts via "arms races" (Raffaele and Kamoun, 2012; Sanchez-Vallet et al., 2018). Many filamentous microorganisms fit a typical "two-speed" genome model with distinct gene dense regions (GDRs) and gene-sparse regions (GSRs). The dynamic GSRs are enriched in rapidly evolving genes and transposons to drive adaptive evolution (Dong et al., 2015; Schrader and Schmitz, 2018). Comparative analysis of 15 *Clavicipitaceae* genomes has revealed large repeat blocks in their alkaloid biosynthesis gene loci, which can facilitate frequent gene losses, mutations, and duplications for symbiotic adaptation (Schardl et al., 2013). Comparative genomic analysis of four *Phytophthora* species has showed that their "two-speed" genome evolution is driven by host jumps. Genes in GSRs, including effector- and epigenetic-related genes, exhibit higher evolutionary rates of structural polymorphisms and positive selection (Raffaele et al., 2010). Comparative genomic analysis of four *P. sojae* isolates has revealed a relatively covert "two-speed" genome evolution, with effectors overrepresented in positive selection (Ye et al., 2016b). Despite the progress mentioned above, it remains largely unknown how the "two-speed" genome evolution model contributes to *Phytophthora* pathogenic adaptation. Considering the strong selection pressure in both nature and agricultural ecosystems (Thrall et al., 2011), population genomic analysis of *P. sojae* will enable integrated investigation of its genome plasticity and yet unknown pathogenicity determinants. In this study, we dissected the genetic basis of *P. sojae* adaptive evolution with emphasis on pathogen adaptation. We analyzed the whole genomes of 29 field *P. sojae* isolates from China and the North America. Population genomic variation analysis confirmed uneven "two-speed" evolution with genes in GSRs showing higher rates of structural polymorphisms and positive selection. In particular, both NADH oxidase and MIP transporter families are under rapid diversifying selection in GSRs. Pan-genome analysis presumed 42 putative core RxLR effectors. Moreover, *Avr* genes exhibit abundant sequence variation with several novel variants identified. Taken together, our results suggest that *P. sojae* adopts novel mechanisms for virulence, which can be a valuable resource for resistance breeding in soybean.

MATERIALS AND METHODS

Pathogenicity Assays

All *P. sojae* isolates used in this study were routinely maintained on V8 juice agar slant at 12°C in the dark. For pathogenicity assays, all isolates were transferred to V8 juice agar plate medium at 25°C in the dark for 5 to 7 days. Their virulence was evaluated by hypocotyl inoculation (Dou et al., 2010; Song et al., 2013). Pathogenicity determination was repeated twice for each isolate.

Isolates were considered as avirulent if more than 70% of the inoculated seedlings survived. Otherwise they were determined as virulent (Na et al., 2013).

DNA Extraction, PCR Amplification, and Sequencing

Genomic DNA of *P. sojae* isolates was isolated from mycelia using the CTAB (hexadecyl trimethyl ammonium bromide) method (Murray and Thompson, 1980). Extracted DNA was quantified using the Qubit® 2.0 Fluorometer kit. PCR amplifications were performed as previously described (Song et al., 2013). All primers are listed in **Supplementary Table S6**.

Sequencing libraries were constructed using the standard Illumina protocol. Library quality was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies). All samples were sequenced using Illumina HiSeq X-ten. The paired-end reads have an average insert size of 300 bp. All read sequences were deposited to the NCBI Sequence Read Archive (SRA), under the BioProject: PRJNA578597.

Reads Mapping and Variants Calling

Raw reads were filtered using the clean_adapter and clean_lowqual software[1]. High-quality reads account for 97.03% of the raw reads with average error rate below 0.001. They were then aligned to *P. sojae* reference genome v3.0 (Tyler et al., 2006) using the Burrows–Wheeler Transform Alignment (BWA) software package v0.7.5a (Li and Durbin, 2009). For each sample, duplicate reads were removed from alignments using samtools software package v1.8 (Li et al., 2009). The mpileup function in samtools was used to generate mpileup files for each sample. Bcftools-vcftools (Li, 2011) was used to identify SNPs and Indels in individual samples. Several criteria were applied during SNP filtering: (i) The phred quality score of base sequencing and score of read mapping higher than 60; (ii) Minimum coverage higher than 10; (iii) Allele frequency $\geq$ 80% for homozygous SNPs; (iv) Allele frequency between 20 and 80% for heterozygous SNPs. CNVs were estimated from read depth by the published methodology (Alkan et al., 2009). Rates of synonymous substitution (dS), non-synonymous substitution (dN) and dN/dS were calculated using KaKs_Calculator software v2.0 (Wang et al., 2010). The *P*-value for dN/dS ratio was calculated by the Fisher's exact test. Genes with dN/dS > 1.00 and *P*-value < 0.05 were defined as under positive selection. Genes with over 1.5 kb intergenic regions were considered as GSRs (Dong et al., 2015; Chen et al., 2018). All variants were visualized by the Circos software v0.696 (Krzywinski et al., 2009).

Population Structure Analysis

Phylogenetic analysis and principal component analysis were based on all detected SNPs. Phylogenetic tree was constructed using PhyML v3.0 (Guindon et al., 2010). Principal component analysis was conducted by SNPRelate (Zheng et al., 2012).

---

[1]https://github.com/fanagislab/common_use

RxLR Effectors Identification

Two distinct pipelines were adopted to identify RxLR effectors. The first pipeline is a reads mapping strategy with three steps: (i) Use samtools (Li et al., 2009) to extract mapped reads of each RxLR gene from re-sequenced genomes; (ii) Use SeqMan from the Lasergene software package to assemble these mapped reads; (iii) Modify the assembled sequence manually to get ORF (from start codon to stop codon) from predicted gene model. Three isolates (P7064, P7074, and P7076) were excluded from the analysis due to their low sequencing depth. All RxLR effectors identified by the first pipeline were further used to study their conservation and divergence across all isolates. Additionally, the *de novo* assembly approach (Ye et al., 2016a) was also employed to supplement the repository of RxLR effector candidates. Genomes were assembled via the SOAPdenovo assembly process (Li et al., 2010). Multiple sequence alignments were performed using MUSCLE v3.8.31 (Edgar, 2004) with default setting.

Gene Annotation and Enrichment Analysis

Genes were annotated with PFAM terms against the Pfam database (El-Gebali et al., 2019). Hmmer software package v3.1b2 (Finn et al., 2011) was used for annotation with $1e^{-5}$ as the cut-off E-value. Databases for all species used in this study were obtained from their original sources: *P. capcisi* (Lamour et al., 2012), *P. infestans* (Haas et al., 2009), *P. ramorum* (Tyler et al., 2006), *H. parasitica* (Baxter et al., 2010), *P. ultimum* (Levesque et al., 2010), *S. parasitica* (Jiang et al., 2013), *A. laibachii* (Kemen et al., 2011), and *T. pseudonana* (Armbrust et al., 2004). After curation of the alignment with the G-blocks tool (Castresana, 2000), phylogenetic tree was constructed using PhyML v3.0 (Guindon et al., 2010) with 1000 bootstrap replicates.

Genes were classified as fast-evolving between any two *P. soaje* isolates with the criteria of (i) duplication or deletion event and (ii) dN/dS > 1. Gene enrichment was calculated by the formula: (Genes(g)∩Genes(c)/Genes(c))/(Genes(g)/Genes(a)), where Genes(g) is the number of genes in a family, Genes(c) is the number of fast-evolving genes or genes in GSRs, and Genes(a) is the total number of genes. With calculated value greater than 1, the family can be defined as enriched in fast-evolving genes or GSRs. The enrichment was considered to be significant with value above 2.

Transcriptional Analysis

Transcriptome data was obtained from published literature (Wang et al., 2018). Briefly, RNA samples from mycelia and soybean leaves collected 3 h after infection were obtained as references. RNA samples from soybean leaves collected 0.5, 6, 12, 24, 36, and 48 h after infection were added. The *P. sojae* isolate P6497 and the susceptible soybean cultivar Williams were used in the stages of infection. Three biological replicates were performed for each treatment. Heatmap was generated using the Heml software (Deng et al., 2014). *Z* score normalization was calculated as (Expression$_A$ – Mean expression$_{A1-An}$)/SD$_{A1-An}$, where Expression$_A$ represents the expression level of a gene at one stage, Mean expression$_{A1-An}$ represents the mean expression

level of a gene at all stages, and SD_{A1-An} represents the standard error of the expression level of a gene at all stages.

RESULTS

Virulence Assay and Genome Sequencing

Twenty-nine *P. sojae* isolates were collected from different geographical locations and investigated in this study. Using the hypocotyl split inoculation method (Song et al., 2013), the pathotypes of these isolates were determined on a set of 13 soybean lines carrying a single *Rps* gene, respectively. Twenty-eight distinct pathotypes were identified and shown in **Supplementary Figure S1**. Eleven and 16 isolates were able to overcome 2 to 7 and 8 to 12 *Rps* genes, respectively. The remaining two isolates were able to overcome all 13 *Rps* genes tested. These results suggest that virulence diversity is abundant in these *P. sojae* isolates.

To investigate the genetic basis of virulence diversity across *P. sojae* isolates, genome re-sequencing was performed using Illumina Hiseq X-ten PE150. A total of 93.65 Gb pair-end (PE) data were generated from 25 *P. sojae* isolates with sequencing depth of 27.9 to 51.4X coverage (**Supplementary Table S1**). Three published *P. sojae* genomes were obtained from eumicrobedb.org (formerly VBI Microbial Database) with sequencing depth ranging from 6.8 to 13.2X coverage (Wang et al., 2011). After removing low-quality reads, approximate 83.98 Gb high-quality PE reads were mapped to the reference genome V3.0 of *P. sojae* (Tyler et al., 2006). Including the reference, a total of 29 isolates were used for further studies below.

Genomic Variation and Population Structure in *P. sojae*

Genomic sequence variations, including single-nucleotide polymorphism (SNP), insertion or deletion (Indel), and copy-number variation (CNV), were cataloged using a read-mapping strategy against the *P. sojae* reference genome (Tyler et al., 2006). We detected 207,740 SNPs, 35,324 Indels, and 6,573 CNVs (**Figure 1A**). The proportions of SNPs identified in intergenic, intron, promoter, and coding regions were 30.45, 8.76, 30.16, and 30.63%, respectively (**Figure 1B**). 24,311 synonymous, 39,317 missenses, and 1,168 nonsense SNPs were detected from the coding sequences. The 35,324 Indels identified ranged from 1 to 106 bp in length (**Supplementary Figure S2**), with the majority (86.15%) shorter than 12 bp and low proportion (4.60%) of long (>20 bp) sequences. The proportions of Indels in intergenic, intron, promoter, and coding regions were 31.78, 11.86, 39.62, and 16.74%, respectively (**Figure 1C**). Most (4,445) coding-region Indels caused frame-shift mutations, reflecting the influence of positive selection. The 6,573 CNVs characterized in coding genes included 5,423 duplications, 776 deletions and 374 duplications/deletions (**Figure 1D**).

To investigate the phylogenetic relationship of *P. sojae* populations, a phylogenetic tree was first constructed with detected 20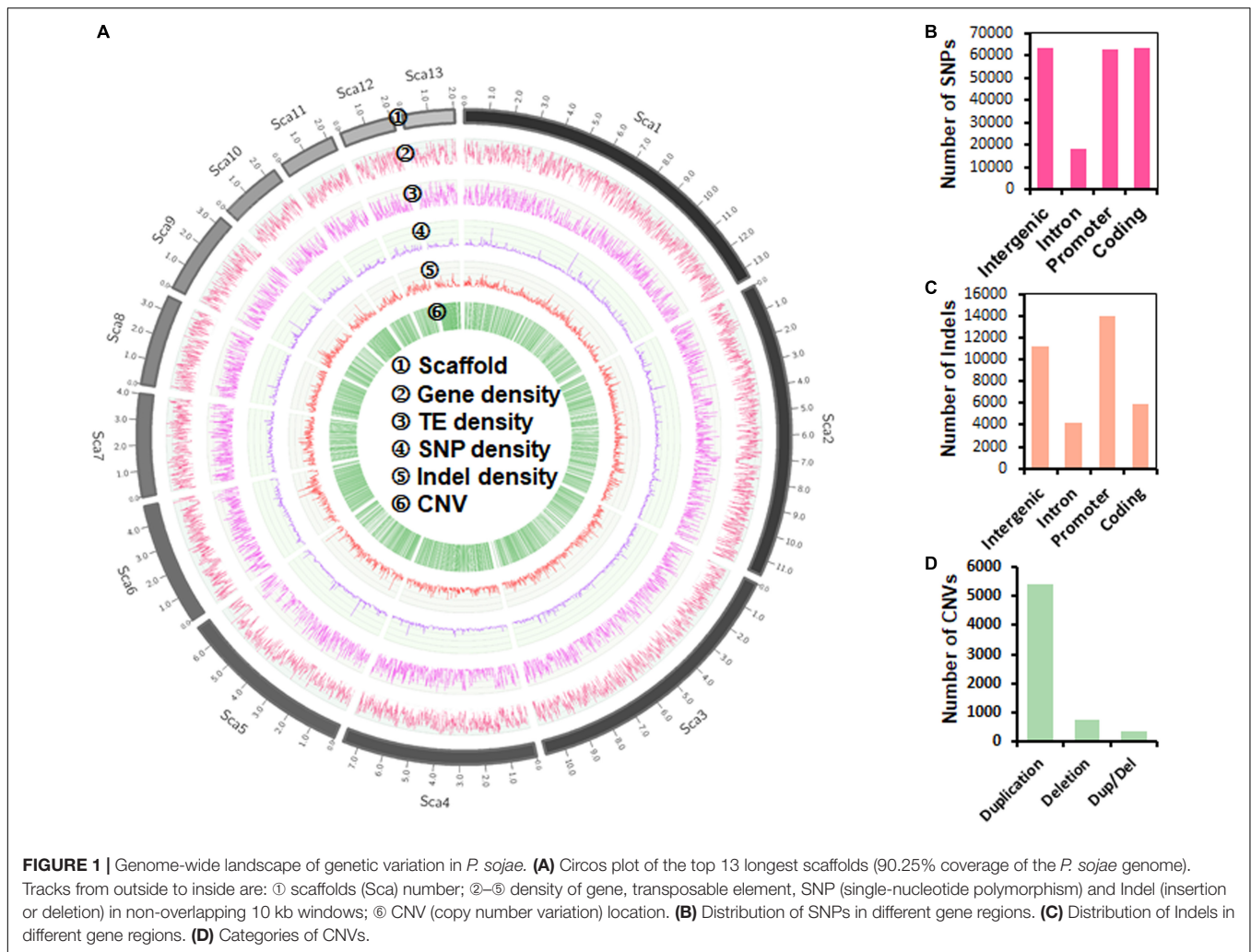7,740 SNPs among the 29 isolates. Results showed no clustering of isolates by geography. We found that isolates from China could be roughly divided into two groups (**Figure 2A**). Group I was clustered with all four isolates from North America, and group II contained isolates only form China, indicating that isolates from China might not be part of the same clonal lineage. Principal component analysis also got the similar results (**Figure 2B**).

Genomic Variation Uncovers Adaptive Evolution in *P. sojae* Populations

To explore whether distinct genomic regions evolve at significantly different rates, we calculated and compared genomic variations in GSRs and GDRs, including SNP and Indel densities per kb, CNVs and dN/dS ratios in genes (ORFs) (**Supplementary Figure S3A**). Compared to GDRs, GSRs showed significantly higher densities of both SNP and Indel. Averages of CNVs and dN/dS ratios were also significantly higher in GSRs than in GDRs. Moreover, we compared the transcription induction levels at seven stages versus the mycelia stage (Wang et al., 2018). Genes in GSRs showed significantly higher induction levels than those in GDRs within 24 h of soybean leave inoculation (**Supplementary Figure S3B**), implicating the adaptation of *P. sojae* genome evolution to plant infection.

Using previously reported method (Raffaele et al., 2010), we performed domain enrichment analysis on predicted proteomes encoded by rapidly evolving genes and genes in GSRs, respectively. A total of 274 domains met the criteria of appearing in 10 or more proteins. We demonstrated that 27 and 33 domains were statistically enriched in proteomes encoded by rapidly evolving genes and GSRs, respectively. There were 23 domains shared by both groups (**Figure 3A**). As expected, some RxLR effectors, pectate lyase, NLP and protease were enriched in both groups whereas CRN effectors and elicitins were only enriched in the proteome encoded by rapidly evolving genes. Some domains, such as transposase, retrotransposon gag, and ribonuclease H, were also enriched in both groups. Interestingly, we found that histone methylation and DNA-binding domains, which were thought to be key regulators of gene expression (Brennan and Matthews, 1989; Raffaele et al., 2010), were enriched in the proteome encoded by rapidly evolving genes. We then calculated the pairwise sequence similarity of DNA-binding proteins between two related species *P. sojae* and *P. ramorum*. The highest sequence identity within any given DNA-binding protein paralogs was only 38.18%, whilst the average sequence identity of 1,000 randomly selected protein paralogs was 69.27%. The result indicates that DNA-binding proteins have above-average divergence levels (**Figure 3B**). Transcriptome analysis further revealed that most of their encoding genes exhibited stage-specific expression patterns (**Figure 3C**). Among them, 20 and 141 genes exhibited developmental- and infection-specific expression patterns, respectively.

NADH oxidases play key roles in the development and virulence of filamentous fungi (Segmuller et al., 2008; Heller and Tudzynski, 2011; Liu et al., 2014). MIP transporters are potential targets of antifungal drugs against pathogenic fungi (Verma et al., 2014). Surprisingly, domains annotated as NADH

FIGURE 1 | Genome-wide landscape of genetic variation in *P. sojae*. **(A)** Circos plot of the top 13 longest scaffolds (90.25% coverage of the *P. sojae* genome). Tracks from outside to inside are: ① scaffolds (Sca) number; ②–⑤ density of gene, transposable element, SNP (single-nucleotide polymorphism) and Indel (insertion or deletion) in non-overlapping 10 kb windows; ⑥ CNV (copy number variation) location. **(B)** Distribution of SNPs in different gene regions. **(C)** Distribution of Indels in different gene regions. **(D)** Categories of CNVs.
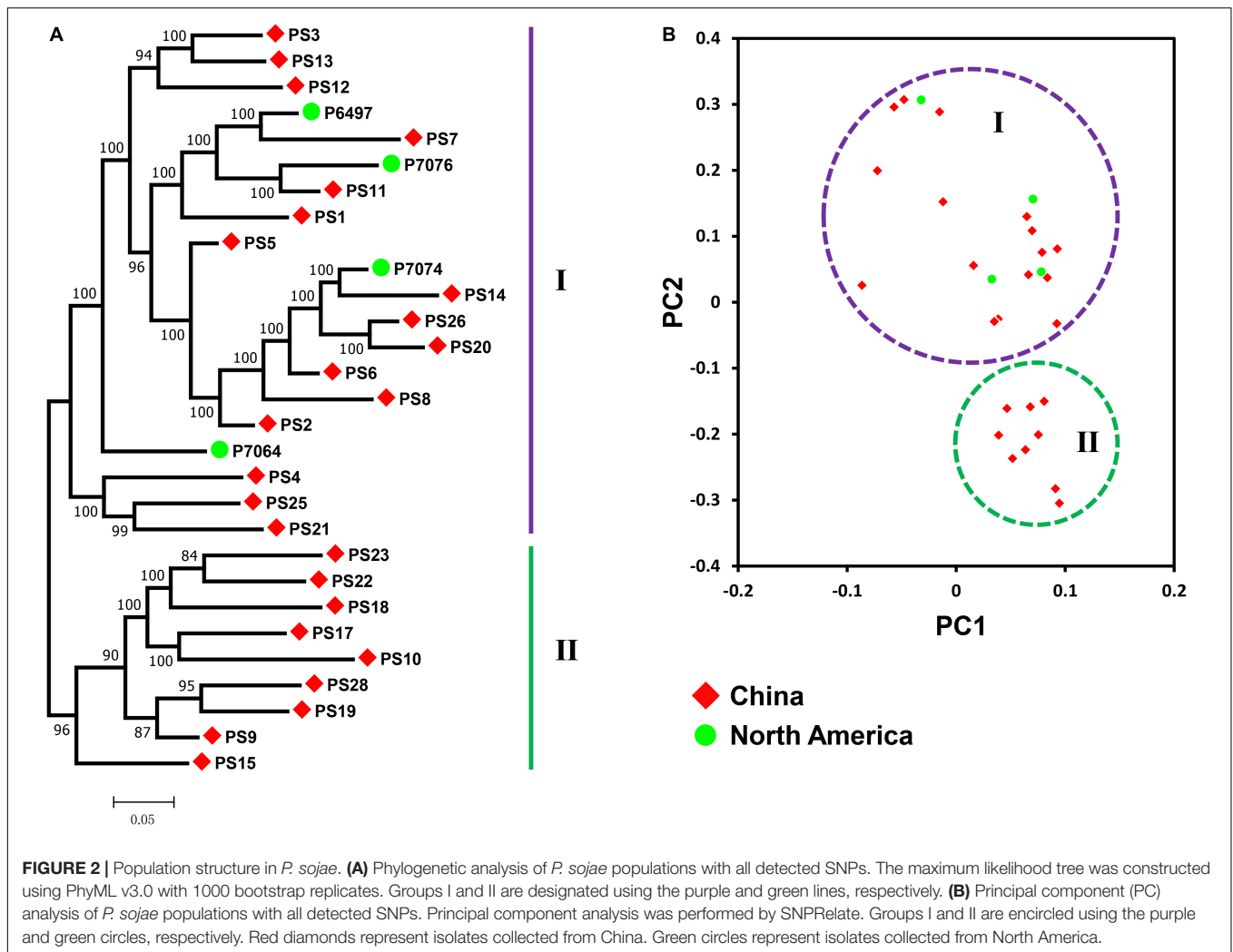
oxidase and MIP transporter families were significantly enriched in the proteome encoded by rapidly evolving genes. Further analysis revealed that the NADH oxidase gene family expanded in *P. sojae* when compared to other oomycetes (**Figure 4A** and **Supplementary Table S2**). We identified 65 NADH oxidase genes in *P. sojae* and 26, 20, 27, 4, 29, 9, 1, and 1 homologous genes in *P. capcisi*, *P. infestans*, *P. ramorum*, *H. parasitica*, *Pythium ultimum*, *Saprolegnia parasitica*, *Albugo laibachii*, and *Thalassiosira pseudonana*, respectively. Gene expansion in *P. sojae* was mainly found in the NADH-III clade. Transcriptome analysis revealed that 38 NADH oxidase genes showed highest expression at the mycelia stage and 58 genes of this family were downregulated during the early biotrophic phase (**Figure 4B**). Additionally, we identified 35, 25, 22, 35, 4, 11, 5, 2, and 2 MIP genes in *P. sojae*, *P. capcisi*, *P. infestans*, *P. ramorum*, *H. parasitica*, *P. ultimum*, *S. parasitica*, *A. laibachii,* and *T. pseudonana*, respectively (**Supplementary Table S3**). Unlike NADH oxidase genes, MIPs had no significant expansion in *P. sojae*.

The *P. sojae* RxLR Effector Repertoire

During virulence and colonization, *Phytophthora* pathogens target different subcellular compartments of the plant cell via deploying a large number of RxLR effectors (Jiang and Tyler, 2012; Anderson et al., 2015; Qiao et al., 2015; Kong et al., 2017). Genome sequencing unveiled 374 RxLR effectors in *P. sojae* isolate P6497 previously (Jiang et al., 2008). In the updated of reference genome of *P. sojae* isolate P6497, a total of 423 RxLR effectors were obtained in our re-identification. Considering the low sequencing depth of *P. sojae* isolates P7064, P7074 and P7076, we used the remaining 25 isolates for RxLR effector analysis. First, we identified 390 to 418 homogeneous RxLR effectors in 25 isolates using the reads mapping approach (**Figure 5A**). In addition, we identified 3 to 28 novel RxLR effectors in 25 isolates using the *de novo* assembly approach (**Figure 5A**). Altogether, the numbers of RxLR effectors ranged from 407 to 430 in these isolates (**Figure 5A**). There were totally 471 RxLR effectors across isolates with 97 novel ones being identified for the first time.

Next, we investigated the conservation and divergence of the RxLR effectors that identified by reads mapping across *P. sojae* populations. Generally, the RxLR effectors were highly divergent in *P. sojae* isolates. There were 157 RxLR effectors that could be found in no more than 20 *P. sojae* isolates (**Figure 5B**). In contrast, 132 conserved RxLR effectors existed in all 26 isolates with 100% similarity (**Figure 5B**). In particular, 42 conserved

FIGURE 2 | Population structure in *P. sojae*. **(A)** Phylogenetic analysis of *P. sojae* populations with all detected SNPs. The maximum likelihood tree was constructed using PhyML v3.0 with 1000 bootstrap replicates. Groups I and II are designated using the purple and green lines, respectively. **(B)** Principal component (PC) analysis of *P. sojae* populations with all detected SNPs. Principal component analysis was performed by SNPRelate. Groups I and II are encircled using the purple and green circles, respectively. Red diamonds represent isolates collected from China. Green circles represent isolates collected from North America.

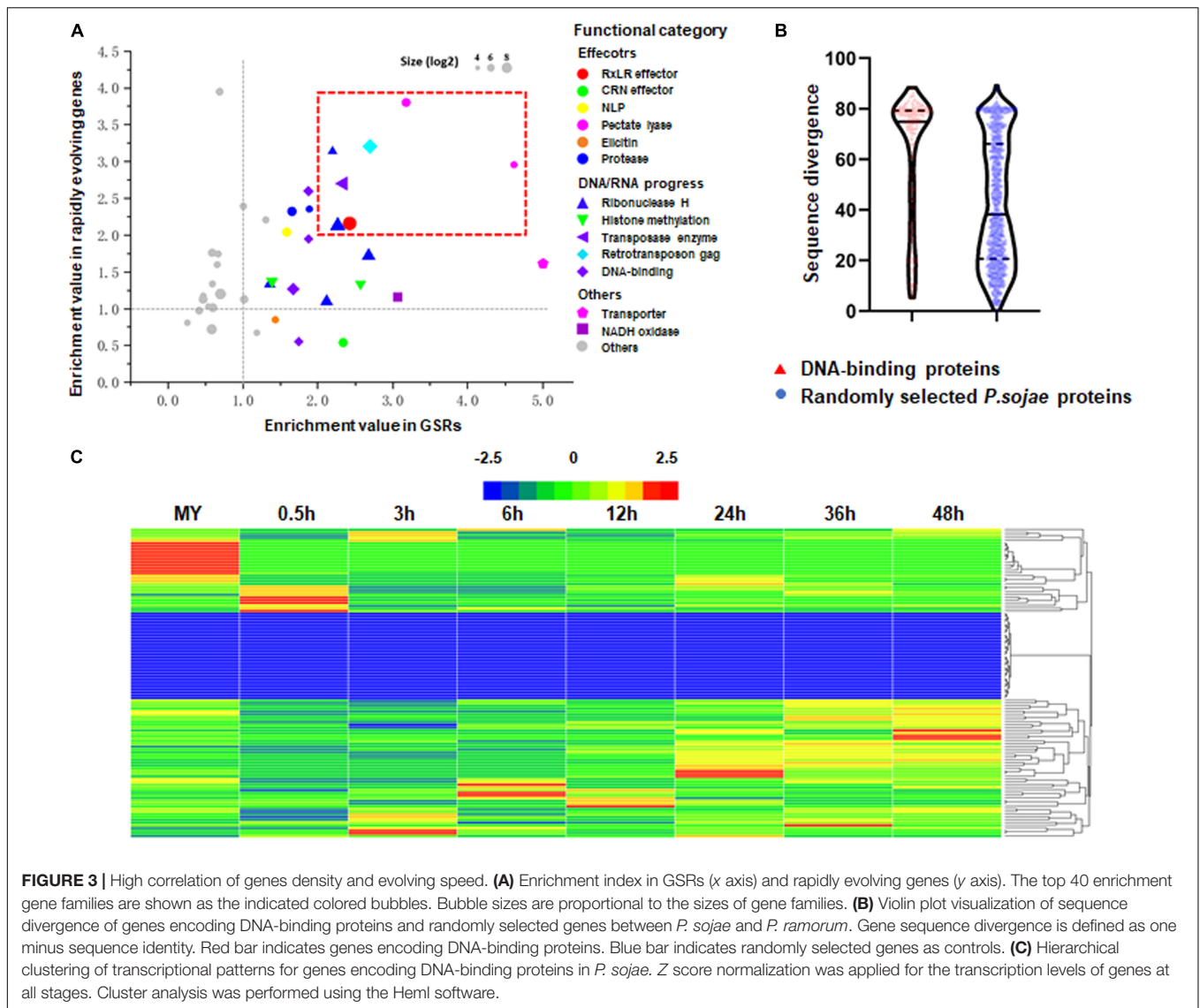RxLR effectors had detectable expression (FPKM > = 2) at one or more stages of infection (**Figures 5C,D** and **Supplementary Table S4**), indicating their roles in virulence. We considered these 42 RxLR effectors as core effectors based on their sequence and transcriptional patterns. Among them, PsAvh241 and PsAvh23 have been demonstrated to be essential for the full virulence of *P. sojae* (Yu et al., 2012; Kong et al., 2017). The other 40 core effectors may also play critical roles in the infection process.

Avr Genes Exhibit Abundant Sequence Variation

The gene-for-gene resistance to *Phytophthora* is often overcome by the rapid evolution of *Avr* genes under field conditions. To date, nine *Avr* genes recognized by 11 cognate soybean *R* genes have been cloned in *P. sojae* (Anderson et al., 2015). To better understand the gene-for-gene interactions between *P. sojae* and soybean, we examined 261 (29*9) interaction combinations between 29 *P. sojae* isolates and 9 *Rps* genes. Among them, 182 interactions confirmed previous studies on *Avr* haplotypes and associated phenotypes. Taking *Avr1a* as

an example, 12 isolates harbored the *Avr* haplotype and the other 17 isolates exhibited three different virulent haplotypes. Five virulent isolates contained *Avr1a-I* haplotype, which may undergo gene silencing to escape the recognition of *Rps1a*. Nine virulent isolates lost *Avr1a*. Both haplotypes were consistent to previous report (Na et al., 2014). Interestingly, we found the three isolates left constituted a novel *Avr1a* haplotype (named as *Avr1a-II*). In this haplotype, substitution of 14 nucleotides led to the mutation of 9 amino acid residues (**Figure 6A**). Taken together, 17 new interactions were identified in 4 *Avr* genes and their corresponding haplotypes (**Figure 6B** and **Supplementary Table S5**). Among them, 15 newly discovered homozygous *Avr* haplotypes were verified by PCR amplification and sequencing (**Supplementary Table S5**).

We summarized five strategies for overcoming gene-for-gene resistance in *P. sojae*, including amino acid mutation, gene absence, pseudogenization, gene silencing, and CNV (**Figure 6B**). Among them, amino acid mutation was the most common escape mechanism. *Avr1b*, *Avr1c*, and *Avr3c* all harbored new virulent alleles (**Figure 6B** and **Supplementary Figures S4–S6**). In contrast, pseudogenization of *Avr* genes is rare and only in

FIGURE 3 | High correlation of genes density and evolving speed. **(A)** Enrichment index in GSRs (x axis) and rapidly evolving genes (y axis). The top 40 enrichment gene families are shown as the indicated colored bubbles. Bubble sizes are proportional to the sizes of gene families. **(B)** Violin plot visualization of sequence divergence of genes encoding DNA-binding proteins and randomly selected genes between *P. sojae* and *P. ramorum*. Gene sequence divergence is defined as one minus sequence identity. Red bar indicates genes encoding DNA-binding proteins. Blue bar indicates randomly selected genes as controls. **(C)** Hierarchical clustering of transcriptional patterns for genes encoding DNA-binding proteins in *P. sojae*. Z score normalization was applied for the transcription levels of genes at all stages. Cluster analysis was performed using the HemI software.

response to *Avr1k* (Song et al., 2013). Seven haplotypes were detected in *Avr1b*. *Avr1b-II* showed a nucleotide substitution at -2nd base upstream of the start codon in three virulent isolates. Avr1b-VI exhibited two-amino-acid polymorphisms in two virulent isolates (**Supplementary Figure S4**). There were also seven haplotypes detected in *Avr1c* (**Supplementary Figure S5**). In four virulent isolates, Avr1c-IV showed five-amino-acid polymorphisms resulted from substitution of 20 nucleotides. Interestingly, *Avr1c-III* was heterozygous with one avirulent copy of *Avr1c-I*. For *Avr3c*, six haplotypes were identified (**Supplementary Figure S6**). *Avr3c-VI* showed a complete deletion in one virulent isolate. *Avr3c-V* displayed 14th-amino-acid polymorphisms caused by substitution of 19 nucleotides in one virulent isolate. Avr3c-I showed 10-amino-acid polymorphisms mediated by substitution of 12 nucleotides in one avirulent isolate.
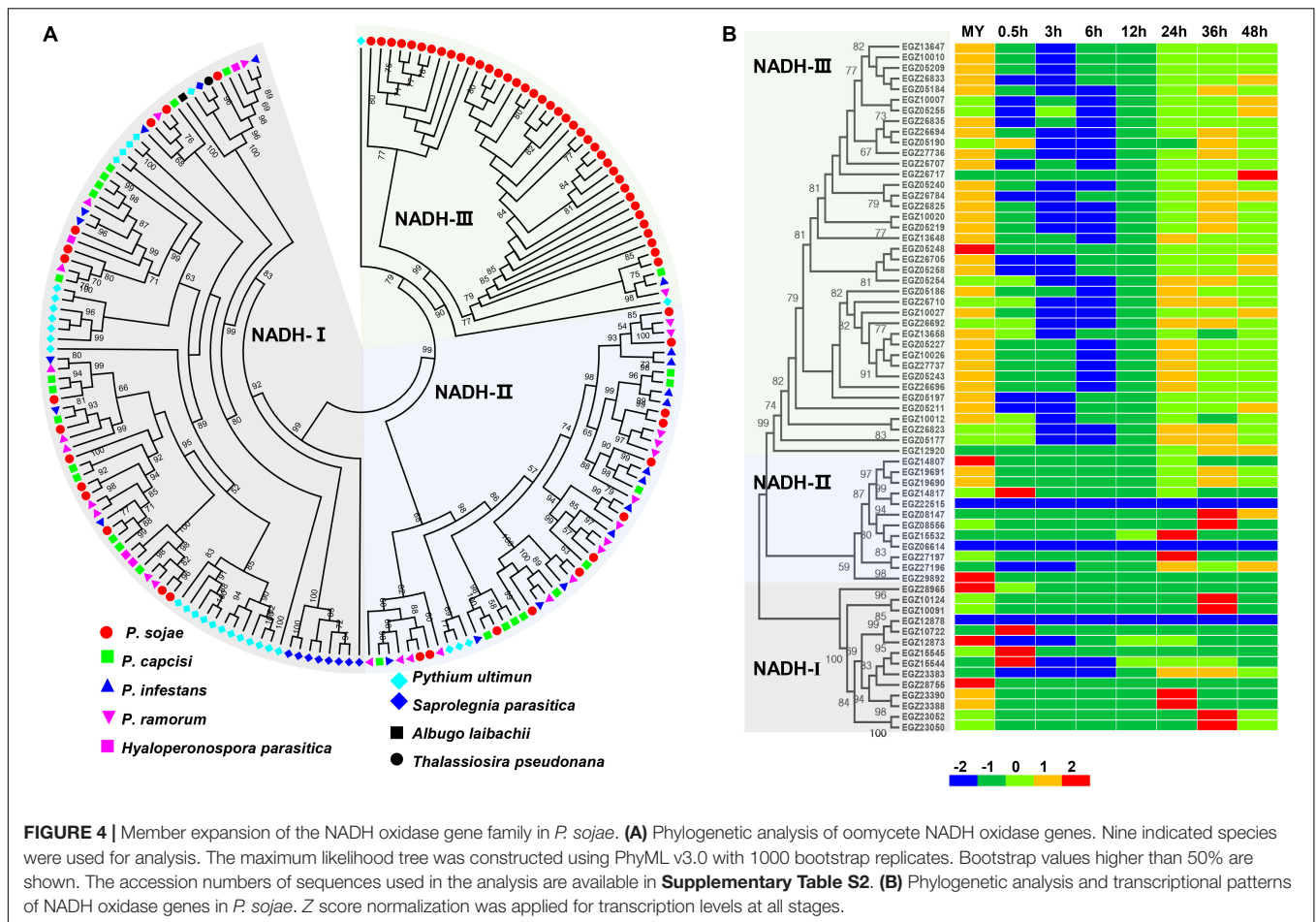
Phytophthora sojae isolates exhibited distinct evolution themes. For example, isolates PS9 and PS15 could overcome all 11 *Rps* genes. Four evolutionary mechanisms, including amino acid mutation, gene absence, pseudogenization and gene silencing, existed in these two isolates with the first two being dominant (**Figure 6C**). Isolates PS1 and P6497, which overcame only one *Rps*, showed either pseudogenization or gene silencing (**Figure 6C**).

DISCUSSION

Plant diseases caused by *Phytophthora* are a major threat to crop production. However, little is known about the genetic basis of adaptive evolution, especially pathogen adaptation, in *Phytophthora*. In this study, we sequence the genomes of 28 field *P. sojae* isolates and compare them to the reference genome.

Previous studies indicate that several *Phytophthora* genomes contain a typical "two-speed" architecture, with GSRs being more plastic than GDRs and serving as a cradle for adaptive

FIGURE 4 | Member expansion of the NADH oxidase gene family in *P. sojae*. **(A)** Phylogenetic analysis of oomycete NADH oxidase genes. Nine indicated species were used for analysis. The maximum likelihood tree was constructed using PhyML v3.0 with 1000 bootstrap replicates. Bootstrap values higher than 50% are shown. The accession numbers of sequences used in the analysis are available in **Supplementary Table S2**. **(B)** Phylogenetic analysis and transcriptional patterns of NADH oxidase genes in *P. sojae*. *Z* score normalization was applied for transcription levels at all stages.

evolution (Raffaele and Kamoun, 2012; Dong et al., 2015). We confirm that genomic variation in *P. sojae* shows uneven "two-speed" evolutionary rates with GSRs experiencing accelerated evolution. This observation is consistent with previously reported comparative genome analysis results of four *Phytophthora* clade 1c species (Raffaele et al., 2010). The "two-speed" strategy for driving adaptive evolution may be conserved across *Phytophthora* species. Under strong selection pressures from plant immune systems and the environment (Croll and McDonald, 2017), many pathogen genes tend to evolve rapidly for various adaptation. Consistent with previous analysis on four *Phytophthora* clade 1c species (Raffaele et al., 2010), *P. sojae* GSRs are enriched in transposons and rapidly evolving genes encoding virulence effectors, cell wall hydrolases, transposon-related proteins, and proteins related to epigenetic maintenance, suggesting the existence of a conserved *Phytophthora* adaptation and pathogenicity strategy. Interestingly, *P. sojae* GSRs are also enriched in genes encoding NADH oxidases and MIP transporters. NADH oxidases are involved in sexual differentiation (Malagnac et al., 2004), virulence (Segmuller et al., 2008), cell degeneration and host defense (Haedens et al., 2005). MIP transporters are involved in a wide range of cellular processes in human (Preston et al., 1992), plant (Wong et al., 2018) and protozoan

parasites (Dean et al., 2014). They are attractive targets for antifungal drugs in pathogenic fungi (Verma et al., 2014). No NADH oxidase or MIP transporter has been functionally investigated in *Phytophthora*. They can be promising targets for future research.

Effectors, especially highly-conserved core effectors, have become powerful tools to develop disease resistance cultivars against various pathogens (Dangl et al., 2013; Vleeshouwers and Oliver, 2014). For example, by referring to an effector repertoire predicted from the *P. infestans* genome, two cognate *Rps* genes (*Rpi-sto1* and *Rpi-pta1*) were rapidly cloned from incompatible *Solanum* species (Vleeshouwers et al., 2008). Unfortunately, little is known about the core effectors in *P. sojae* to date. In this study, we identify a set of 471 RxLR effectors across 26 *P. sojae* genomes. We demonstrate that *P. sojae* harbors a significantly larger RxLR repertoire than other species. Being conserved in protein coding sequences and expressed at one or more infection stages, 42 members are defined as core RxLR effectors. Core effectors are generally considered important to pathogen virulence (Dangl et al., 2013). For example, PsAvh241 and PsAvh23 are two core effectors that are essential for full virulence of *P. sojae* (Yu et al., 2012; Kong et al., 2017). Transcriptome comparison uncovers 18 core RxLR effectors from five *P. infestans* isolates. Nine of them contribute
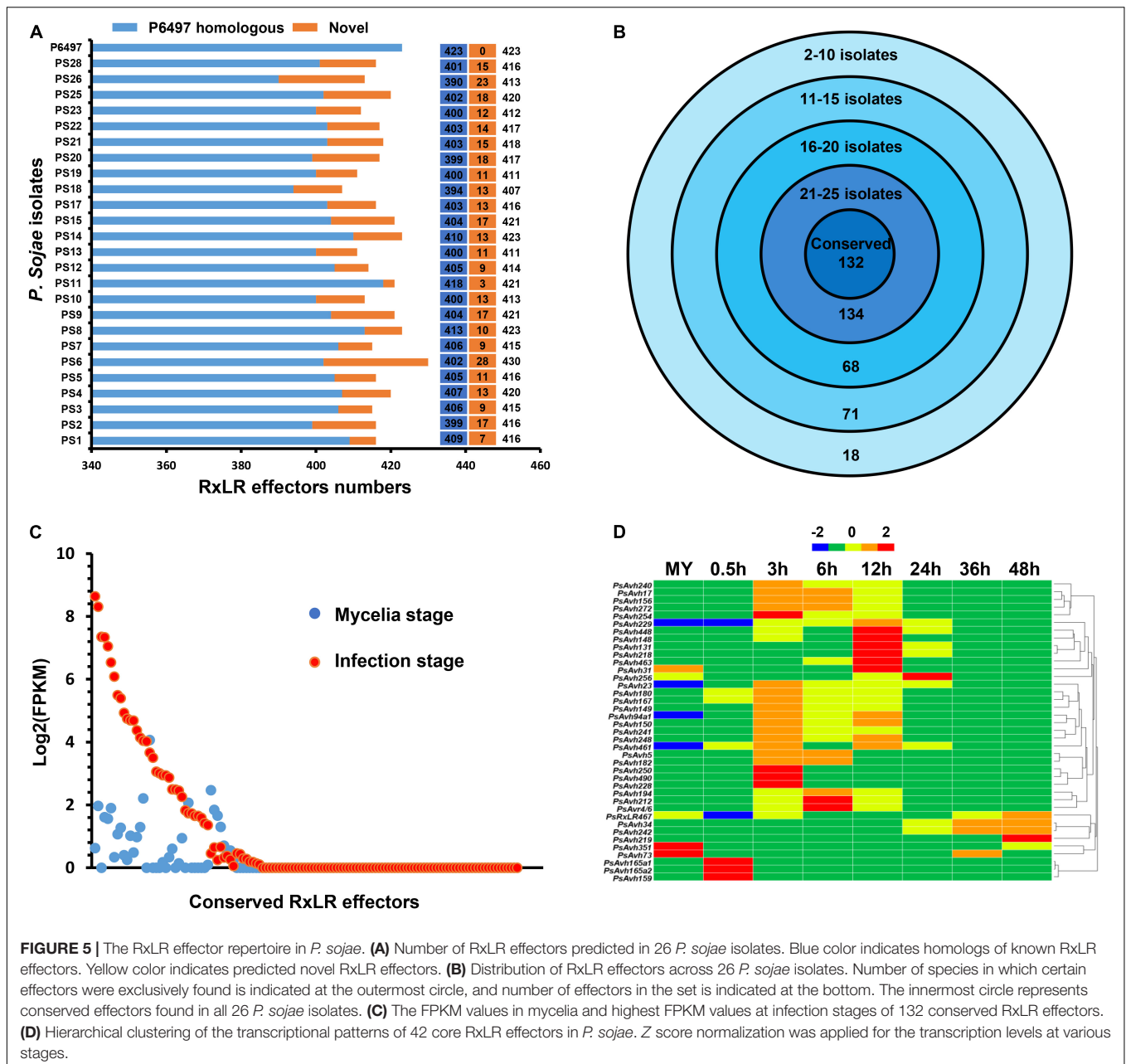
FIGURE 5 | The RxLR effector repertoire in *P. sojae*. **(A)** Number of RxLR effectors predicted in 26 *P. sojae* isolates. Blue color indicates homologs of known RxLR effectors. Yellow color indicates predicted novel RxLR effectors. **(B)** Distribution of RxLR effectors across 26 *P. sojae* isolates. Number of species in which certain effectors were exclusively found is indicated at the outermost circle, and number of effectors in the set is indicated at the bottom. The innermost circle represents conserved effectors found in all 26 *P. sojae* isolates. **(C)** The FPKM values in mycelia and highest FPKM values at infection stages of 132 conserved RxLR effectors. **(D)** Hierarchical clustering of the transcriptional patterns of 42 core RxLR effectors in *P. sojae*. Z score normalization was applied for the transcription levels at various stages.

to virulence through defense suppression (Yin et al., 2017). Soybean Rps proteins may recognize certain core effectors to activate defense against *P. sojae* isolates. Our RxLR effector repertoire provides a valuable resource for the search of novel *Rps* genes.

Nine *Avr* genes have been cloned and their gain-of-virulence alleles have been analyzed in *P. sojae* (Anderson et al., 2015). However, little is known about their polymorphism across *P. sojae* populations and interactions with *Rps* genes. In this study, we examine *Avr* polymorphism as well as haplotype-phenotype matches in 261 interactions between 9 *Rps* genes and 29 *P. sojae* isolates. Recently, re-sequencing of 31 *P. sojae* isolates has revealed stable predictive markers and variants for five *Avr*

genes (Arsenault-Labrecque et al., 2018). Our results not only confirm these existing markers but also identify new markers for a broader range of *Avr* genes. Inexplicably, unexpected phenotypic outcomes were often observed when taking into account *Avr* haplotypes and corresponding phenotype. Previous study has revealed that gene silencing and epistatic effects might confuse the phenotypic interactions (Qutob et al., 2013; Arsenault-Labrecque et al., 2018). Nevertheless, these inexplicable interactions require further investigation.

In our study, Avr effectors containing new amino acid polymorphisms will offer new resources for functional characterizations of *Avr* genes in plant–pathogen interactions. Meanwhile, CNV and gene silencing raise new challenges to

FIGURE 6 | *Avr* genes exhibit abundant sequence variation among 29 *P. sojae* isolates. **(A)** Structural and nucleotide diversity at the *Avr1a* locus among 29 *P. sojae* isolates. The green arrow indicates gene-coding region. Red vertical lines indicate variant regions of the gene. Variant positions and types are connected by dotted lines. Schematic graphs of the variant positions in all isolates are grouped by haplotypes. Black circles indicate previously reported haplotypes and corresponding phenotypes. Red circles indicate novel haplotypes and corresponding phenotypes. Disease outcomes on soybean plants harboring *Rps1a* are recorded as avirulent (A) or virulent (V). Variation positions are from *P. sojae* reference genome V3.0. **(B)** Summary of different variant types of the 9 *Avr* genes. *X* axis indicates variant types of corresponding *Avr* genes. "Novel variants" indicates new variant types identified in this study. "Reported variants" indicates variant types previously described. "No variants" indicates no variant being found. *Y* axis denotes cluster analysis generated by the Heml software using the variant types of *Avr* genes. **(C)** Summary of *Avr* variant types and virulence formulas across 29 *P. sojae* isolates. Bubble sizes are proportional to the numbers of corresponding variant types. Cluster analysis was generated by Heml using the virulence formulas of *P. sojae*.

traditional PCR detection of *Avr* genes. Other methods, such as RT-PCR, are needed to measure the transcript levels and CNVs of *Avr* genes. 26 out of 29 collected isolates can overcome *Rps5*. In contrast, only 13 isolates can overcome *Rps1d*, indicating that *Rps1d* is overall more effective for controlling *Phytophthora* diseases. This result is consistent with a previous report by Moriwaki (2010). No *Rps* gene can render resistance to all 29 isolates in our test. Therefore, efforts should be made to identify novel *Rps* genes conferring durable and broad-spectrum resistance. The core RxLR effector repertoire we report can play a key role in the *Rps* search.

In recent years, genome re-sequencing has emerged as a powerful tool for dissecting the genetic basis of trait adaptation in fungal pathogens, such as virulence, lifestyles,

fungicide resistance, host jump, and host specialization (Grunwald et al., 2016). However, few cases have been reported in oomycete pathogens. In this study, we re-sequence the genomes of 26 field *P. sojae* isolates. Our comparative genomic analysis results provide new clues for adaptive evolution, especially pathogen adaptation, in *Phytophthora*. We demonstrate that *P. sojae* genome undergoes uneven "two-speed" evolution with genes in GSRs exhibiting higher rates of sequence polymorphism and positive selection. GSRs are enriched by effector genes, transposase-related genes and gene families encoding NADH oxidases and MIP transporters. In addition, we report an updated RxLR effector repertoire with 42 core effectors. Finally, we show that *Avr* genes exhibit abundant sequence variation and present several novel variants. Our work provides a genomic platform

for studying soybean-*Phytophthora* interaction and breeding new resistant cultivars.

DATA AVAILABILITY STATEMENT

AUTHOR CONTRIBUTIONS

DD, WF, WQ, and BT directed the project. XZ, BL, DS, ZY, and WF analyzed the data. FZ, RW, and FH performed the experiments. YW, BT, and WQ provided very critical suggestions. DD and XZ wrote the manuscript with input from all other authors.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb. 2019.02792/full#supplementary-material

REFERENCES

Alkan, C., Kidd, J. M., Marques-Bonet, T., Aksay, G., Antonacci, F., Hormozdiari, F., et al. (2009). Personalized copy number and segmental duplication maps using next-generation sequencing. *Nat. Genet.* 41, 1061–U1029. doi: 10.1038/ng.437

Anderson, R. G., Deb, D., Fedkenheuer, K., and McDowell, J. M. (2015). Recent progress in RXLR effector research. *Mol. Plant Microbe Interact.* 28, 1063–1072. doi: 10.1094/MPMI-01-15-0022-CR

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306, 79–86. doi: 10.1126/science.1101156

Arsenault-Labrecque, G., Sonah, H., Lebreton, A., Labbe, C., Marchand, G., Xue, A., et al. (2018). Stable predictive markers for *Phytophthora sojae* avirulence genes that impair infection of soybean uncovered by whole genome sequencing of 31 isolates. *BMC Biol.* 16:9. doi: 10.1186/s12915-018-0549-9

Bart, R., Cohn, M., Kassen, A., McCallum, E. J., Shybut, M., Petriello, A., et al. (2013). High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19969–19969. doi: 10.1073/pnas.1319913110

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., et al. (2010). Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330, 1549–1551. doi: 10.1126/science.1195203

Brennan, R. G., and Matthews, B. W. (1989). The helix-turn-helix DNA binding motif. *J. Biol. Chem.* 264, 1903–1906.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chawla, S., Bowen, C. R., Slaminko, T. L., Hobbs, H. A., and Hartman, G. L. (2013). A public program to evaluate commercial soybean cultivars for pathogen and pest resistance. *Plant Dis.* 97, 568–578. doi: 10.1094/Pdis-05-12-0479-Fe

Chen, H., Shu, H. D., Wang, L. Y., Zhang, F., Li, X., Ochola, S. O., et al. (2018). *Phytophthora* methylomes are modulated by 6mA methyltransferases and associated with adaptive genome regions. *Genome Biol.* 19:181. doi: 10.1186/s13059-018-1564-4

Croll, D., and McDonald, B. A. (2017). The genetic basis of local adaptation for pathogenic fungi in agricultural ecosystems. *Mol. Ecol.* 26, 2027–2040. doi: 10.1111/mec.13870

Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi: 10.1126/science.1236011

Dean, P., Major, P., Nakjang, S., Hirt, R. P., and Embley, T. M. (2014). Transport proteins of parasitic protists and their role in nutrient salvage. *Front. Plant Sci.* 5:153. doi: 10.3389/fpls.2014.00153

Deb, D., Anderson, R. G., How-Yew-Kin, T., Tyler, B. M., and McDowell, J. M. (2018). Conserved RxLR effectors from oomycetes *Hyaloperonospora arabidopsidis* and *Phytophthora sojae* suppress PAMP- and effector-triggered immunity in diverse plants. *Mol. Plant Microbe Interact.* 31, 374–385. doi: 10.1094/MPMI-07-17-0169-FI

Deng, W. K., Wang, Y. B., Liu, Z. X., Cheng, H., and Xue, Y. (2014). HemI: a toolkit for illustrating heatmaps. *PLoS One* 9:e0111988. doi: 10.1371/journal.pone.0111988

Dong, S., Yu, D., Cui, L., Qutob, D., Tedman-Jones, J., Kale, S. D., et al. (2011). Sequence variants of the *Phytophthora sojae* RXLR effector Avr3a/5 are differentially recognized by Rps3a and Rps5 in soybean. *PLoS One* 6:e20172. doi: 10.1371/journal.pone.0020172

Dong, S. M., Yin, W. X., Kong, G. H., Yang, X. Y., Qutob, D., Chen, Q. H., et al. (2011). *Phytophthora sojae* avirulence effector Avr3b is a secreted NADH and ADP-ribose pyrophosphorylase that modulates plant immunity. *PLoS Pathog.* 7:e1002353. doi: 10.1371/journal.ppat.1002353

Dong, S. M., Qutob, D., Tedman-Jones, J., Kuflu, K., Wang, Y. C., Tyler, B. M., et al. (2009). The *Phytophthora sojae* avirulence locus Avr3c encodes a multi-copy RXLR effector with sequence polymorphisms among pathogen strains. *PLoS One* 4:e0005556. doi: 10.1371/journal.pone.0005556

Dong, S. M., Raffaele, S., and Kamoun, S. (2015). The two-speed genomes of filamentous pathogens: waltz with plants. *Curr. Opin. Genet. Dev.* 35, 57–65. doi: 10.1016/j.gde.2015.09.001

Dorrance, A. E., McClure, S. A., and Martin, S. K. S. (2003). Effect of partial resistance on *Phytophthora* stem rot incidence and yield of soybean in Ohio. *Plant Dis.* 87, 308–312. doi: 10.1094/Pdis.2003.87.3.308

Dou, D. L., Kale, S. D., Liu, T. L., Tang, Q. H., Wang, X., Arredondo, F. D., et al. (2010). Different domains of *Phytophthora sojae* effector Avr4/6 are recognized by soybean resistance genes Rps4 and Rps6. *Mol. Plant Microbe Interact.* 23, 425–435. doi: 10.1094/Mpmi-23-4-0425

Dou, D. L., and Zhou, J. M. (2012). Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host Microbe* 12, 484–495. doi: 10.1016/j.chom.2012.09.003

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995

Fawke, S., Doumane, M., and Schornack, S. (2015). Oomycete interactions with plants: infection strategies and resistance principles. *Microbiol. Mol. Biol. Rev.* 79, 263–280. doi: 10.1128/MMBR.00010-15

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Grunwald, N. J., McDonald, B. A., and Milgroom, M. G. (2016). Population genomics of fungal and oomycete pathogens. *Annu. Rev. Phytopathol.* 54, 323–346. doi: 10.1146/annurev-phyto-080614-115913

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461, 393–398. doi: 10.1038/nature08358

Haedens, V., Malagnac, F., and Silar, P. (2005). Genetic control of an epigenetic cell degeneration syndrome in *Podospora anserina*. *Fungal Genet. Biol.* 42, 564–577. doi: 10.1016/j.fgb.2005.03.011

Hansen, E. M., Reeser, P. W., and Sutton, W. (2012). *Phytophthora* beyond agriculture. *Annu. Rev. Phytopathol.* 50, 359–378. doi: 10.1146/annurev-phyto-081211-172946

Heller, J., and Tudzynski, P. (2011). Reactive oxygen species in phytopathogenic fungi: signaling, development, and disease. *Annu. Rev. Phytopathol.* 49, 369–390. doi: 10.1146/annurev-phyto-072910-095355

Huang, J., Chen, L., Lu, X., Peng, Q., Zhang, Y., Yang, J., et al. (2019). Natural allelic variations provide insights into host adaptation of *Phytophthora* avirulence effector PsAvr3c. *New Phytol.* 221, 1010–1022. doi: 10.1111/nph.15414

Huang, J., Guo, N., Li, Y., Sun, J., Hu, G., Zhang, H., et al. (2016). Phenotypic evaluation and genetic dissection of resistance to *Phytophthora sojae* in the Chinese soybean mini core collection. *BMC Genet.* 17:85. doi: 10.1186/s12863-016-0383-4

Jiang, R. H., and Tyler, B. M. (2012). Mechanisms and evolution of virulence in oomycetes. *Annu. Rev. Phytopathol.* 50, 295–318. doi: 10.1146/annurev-phyto-081211-172912

Jiang, R. H. Y., de Bruijn, I., Haas, B. J., Belmonte, R., Lobach, L., Christie, J., et al. (2013). Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. *PLoS Genet.* 9:e1003272. doi: 10.1371/journal.pgen.1003272

Jiang, R. H. Y., 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. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4874–4879. doi: 10.1073/pnas.0709303105

Kemen, E., Gardiner, A., Schultz-Larsen, T., Kemen, A. C., Balmuth, A. L., Robert-Seilaniantz, A., et al. (2011). Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. *PLoS Biol.* 9:e1001094. doi: 10.1371/journal.pbio.1001094

Kong, L., Qiu, X., Kang, J., Wang, Y., Chen, H., Huang, J., et al. (2017). A *Phytophthora* effector manipulates host histone acetylation and reprograms defense gene expression to promote infection. *Curr. Biol.* 27, 981–991. doi: 10.1016/j.cub.2017.02.044

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., et al. (2012). Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Mol. Plant Microbe Interact.* 25, 1350–1360. doi: 10.1094/Mpmi-02-12-0028-R

Levesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., et al. (2010). Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol.* 11:R73. doi: 10.1186/gb-2010-11-7-r73

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20, 265–272. doi: 10.1101/gr.097261.109

Liu, J., Yin, Y., Song, Z., Li, Y., Jiang, S., Shao, C., et al. (2014). NADH: flavin oxidoreductase/NADH oxidase and ROS regulate microsclerotium development in *Nomuraea rileyi*. *World J. Microbiol. Biotechnol.* 30, 1927–1935. doi: 10.1007/s11274-014-1610-7

Malagnac, F., Lalucque, H., Lepere, G., and Silar, P. (2004). Two NADPH oxidase isoforms are required for sexual reproduction and ascospore germination in the filamentous fungus *Podospora anserina*. *Fungal Genet. Biol.* 41, 982–997. doi: 10.1016/j.fgb.2004.07.008

Moriwaki, J. (2010). Aiming of the construction of the race distinction system of *Phytophthora sojae*. *Plant Prot.* 64, 508–510.

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325.

Na, R., Yu, D., Chapman, B. P., Zhang, Y., Kuflu, K., Austin, R., et al. (2014). Genome re-sequencing and functional analysis places the *Phytophthora sojae* avirulence genes Avr1c and Avr1a in a tandem repeat at a single locus. *PLoS One* 9:e0089738. doi: 10.1371/journal.pone.0089738

Na, R., Yu, D., Qutob, D., Zhao, J., and Gijzen, M. (2013). Deletion of the *Phytophthora sojae* avirulence gene Avr1d causes gain of virulence on Rps1d. *Mol. Plant Microbe Interact.* 26, 969–976. doi: 10.1094/Mpmi-02-13-0036-R

Preston, G. M., Carroll, T. P., Guggino, W. B., and Agre, P. (1992). Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* 256, 385–387. doi: 10.1126/science.256.5055.385

Qiao, Y. L., Shi, J. X., Zhai, Y., Hou, Y. N., and Ma, W. B. (2015). *Phytophthora* effector targets a novel component of small RNA pathway in plants to promote infection. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5850–5855. doi: 10.1073/pnas.1421475112

Qutob, D., Chapman, B. P., and Gijzen, M. (2013). Transgenerational gene silencing causes gain of virulence in a plant pathogen. *Nat. Commun.* 4:1349. doi: 10.1038/ncomms2354

Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., et al. (2010). Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science* 330, 1540–1543. doi: 10.1126/science.1193070

Raffaele, S., and Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* 10, 417–430. doi: 10.1038/nrmicro2790

Sanchez-Vallet, A., Fouche, S., Fudal, I., Hartmann, F. E., Soyer, J. L., Tellier, A., et al. (2018). The genome biology of effector gene evolution in filamentous plant pathogens. *Annu. Rev. Phytopathol.* 56, 21–40. doi: 10.1146/annurev-phyto-080516-035303

Schardl, C. L., Florea, S., Pan, J., Nagabhyru, P., Bec, S., and Calie, P. J. (2013). The epichloae: alkaloid diversity and roles in symbiosis with grasses. *Curr. Opin. Plant Biol.* 16, 480–488. doi: 10.1016/j.pbi.2013.06.012

Schmitthenner, A. F. (1985). Problems and progress in control of *Phytophthora* root-rot of soybean. *Plant Dis.* 69, 362–368. doi: 10.1094/Pd-69-362

Schrader, L., and Schmitz, J. (2018). The impact of transposable elements in adaptive evolution. *Mol. Ecol.* 28, 1537–1549. doi: 10.1111/mec.14794

Segmuller, N., Kokkelink, L., Giesbert, S., Odinius, D., van Kan, J., and Tudzynski, P. (2008). NADPH Oxidases are involved in differentiation and pathogenicity in *Botrytis cinerea*. *Mol. Plant Microbe Interact.* 21, 808–819. doi: 10.1094/Mpmi-21-6-0808

Shan, W. X., Cao, M., Dan, L. U., and Tyler, B. M. (2004). The Avr1b locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps1b. *Mol. Plant Microbe Interact.* 17, 394–403. doi: 10.1094/Mpmi.2004.17.4.394

Song, T. Q., Kale, S. D., Arredondo, F. D., Shen, D. Y., Su, L. M., Liu, L., et al. (2013). Two RxLR avirulence genes in *Phytophthora sojae* determine soybean *Rps1k*-mediated disease resistance. *Mol. Plant Microbe Interact.* 26, 711–720. doi: 10.1094/Mpmi-12-12-0289-R

Thrall, P. H., Oakeshott, J. G., Fitt, G., Southerton, S., Burdon, J. J., Sheppard, A., et al. (2011). Evolution in agriculture: the application of evolutionary approaches to the management of biotic interactions in agro-ecosystems. *Evol. Appl.* 4, 200–215. doi: 10.1111/j.1752-4571.2010.00179.x

Tyler, B. M. (2007). *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Mol. Plant Pathol.* 8, 1–8. doi: 10.1111/j.1364-3703.2006.00373.x

Tyler, B. M., Tripathy, S., Zhang, X. M., Dehal, P., Jiang, R. H. Y., Aerts, A., et al. (2006). *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266. doi: 10.1126/science.1128796

Verma, R. K., Prabh, N. D., and Sankararamakrishnan, R. (2014). New subfamilies of major intrinsic proteins in fungi suggest novel transport properties in fungal channels: implications for the host-fungal interactions. *BMC Evol. Biol.* 14:4. doi: 10.1186/s12862-014-0173-4

Vleeshouwers, V. G., Rietman, H., Krenek, P., Champouret, N., Young, C., Oh, S. K., et al. (2008). Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS One* 3:e2875. doi: 10.1371/journal.pone.0002875

Vleeshouwers, V. G. A. A., and Oliver, R. P. (2014). Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Mol. Plant Microbe Interact.* 27, 196–206. doi: 10.1094/Mpmi-10-13-0313-Ia

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* 8, 77–80. doi: 10.1016/S1672-0229(10)60008-3

Wang, Q., Han, C., Ferreira, A. O., Yu, X., Ye, W., Tripathy, S., et al. (2011). Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* 23, 2064–2086. doi: 10.1105/tpc.111.086082

Wang, Y., Ye, W., and Wang, Y. (2018). Genome-wide identification of long non-coding RNAs suggests a potential association with effector gene transcription in *Phytophthora sojae*. *Mol. Plant Pathol.* 19, 2177–2186. doi: 10.1111/mpp.12692

Wong, D. C. J., Zhang, L., Merlin, I., Castellarin, S. D., and Gambetta, G. A. (2018). Structure and transcriptional regulation of the major intrinsic protein gene family in grapevine. *BMC Genomics* 19:248. doi: 10.1186/s12864-018-4638-5

Ye, W., Wang, Y., Shen, D., Li, D., Pu, T., Jiang, Z., et al. (2016a). Sequencing of the litchi downy blight pathogen reveals it is a *Phytophthora* species with downy mildew-like characteristics. *Mol. Plant Microbe Interact.* 29, 573–583. doi: 10.1094/MPMI-03-16-0056-R

Ye, W., Wang, Y., Tyler, B. M., and Wang, Y. (2016b). Comparative genomic analysis among four representative isolates of *Phytophthora sojae* reveals genes under evolutionary selection. *Front. Microbiol.* 7:1547. doi: 10.3389/fmicb.2016.01547

Yin, J. L., Gu, B., Huang, G. Y., Tian, Y., Quan, J. L., Lindqvist-Kreuze, H., et al. (2017). Conserved RXLR effector genes of *Phytophthora infestans* expressed at the early stage of potato infection are suppressive to host defense. *Front. Plant Sci.* 8:2155. doi: 10.3389/fpls.2017.02155

Yu, X., Tang, J., Wang, Q., Ye, W., Tao, K., Duan, S., et al. (2012). The RxLR effector Avh241 from *Phytophthora sojae* requires plasma membrane localization to induce plant cell death. *New Phytol.* 196, 247–260. doi: 10.1111/j.1469-8137.2012.04241.x

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328. doi: 10.1093/bioinformatics/bts606