Title page

Prediction and Characterization of RXLR Effectors in *Pythium* Species

Short title: **RXLR effectors in *Pythium* species**

Gan Ai[1¶], Kun Yang[1¶], Yuee Tian[1,2], Wenwu Ye[1], Hai Zhu[1], Tianli Li[1], Yaxin Du[1], Qingyue Xia[1], Danyu Shen[1], Maofeng Jing[1], Ai Xia[1], Daolong Dou[1*]

[1] College of Plant Protection, Nanjing Agricultural University, Nanjing 210095, China

[2] Department of Plant Protection, Henan University of Science and Technology, Luoyang 471000, China

[*] Corresponding author: Daolong Dou, Email: ddou@njau.edu.cn

Gan Ai, 2016102010@njau.edu.cn

Kun Yang, 2019202016@njau.edu.cn

Yuee Tian, yueetian@haust.edu.cn

Wenwu Ye, yeww@njau.edu.cn

Hai Zhu, 2018102009@njau.edu.cn

Tianli Li, 2018102007@njau.edu.cn

Yaxin Du, duyaxin1997@163.com

Qingyue Xia, 2018802217@njau.edu.cn

Danyu Shen, shendanyu@njau.edu.cn

Maofeng Jing, jingmf@njau.edu.cn

Ai Xia, xiaai@njau.edu.cn

¶These authors contributed equally to this work.

Abstract

Being widely existed in oomycetes, the RXLR effector features conserved RXLR-dEER motifs in its N terminal. Every known *Phytophthora* or *Hyaloperonospora* pathogen harbors hundreds of RXLRs. In *Pythium* species, however, none of the RXLR effectors has been characterized yet. Here, we developed a stringent method for *de novo* identification of RXLRs and characterized 359 putative RXLR effectors from nine tested *Pythium* species. Phylogenetic analysis revealed a single superfamily formed by all oomycetous RXLRs, suggesting they descent from a common ancestor. RXLR effectors from *Pythium* and *Phytophthora* species exhibited similar sequence features, protein structures and genome locations. In particular, the mosquito biological agent *P. guiyangense* contains a significantly larger RXLR repertoire than the other eight *Pythium* species examined, which may result from gene duplication and genome rearrangement events as indicated by synteny analysis. Expression pattern analysis of RXLR-encoding genes in the plant pathogen *P. ultimum* detected transcripts from the vast majority of predicted *RXLRs* with some of them being induced at infection stages. One such RXLRs showed necrosis-inducing activity. Furthermore, all predicted *RXLRs* were cloned from two biocontrol agents *P. oligandrum* and *P. periplocum*. Three of them were found to encode effectors inducing defense response in *Nicotiana benthamiana*. Taken together, our findings represent the first complete synopsis of *Pythium* RXLR effectors, which provides critical clues on their evolutionary patterns as well as the mechanisms of their interactions with diverse hosts.

Keywords: Oomycete; RXLR effector; gene evolution; *Pythium*; pathogenicity

Author summary

Pathogens from the *Pythium* genus are widespread across multiple ecological niches. Most of them are soilborne plant pathogens whereas others cause infectious diseases in mammals. Some *Pythium* species can be used as biocontrol agents for plant diseases or mosquito management. Despite that phylogenetically close oomycete pathogens secrete RXLR effectors to enable infection, no RXLR protein was previously characterized in any *Pythium* species. Here we developed a stringent method to predict *Pythium* RXLR effectors and compared them with known RXLRs from other species. All oomycetous RXLRs form a huge superfamily, which indicates they may share a common ancestor. Our sequence analysis results suggest that the expansion of RXLR repertoire results from gene duplication and genome recombination events. We further demonstrated that most predicted *Pythium RXLRs* can be transcribed and some of them encode effectors exhibiting pathogenic or defense-inducing activities. This work expands our understanding of RXLR evolution in oomycetes in general, and provides novel insights into the molecular interactions between *Pythium* pathogens and their diverse hosts.

Introduction

Coevolution of microbial pathogens and host plants is driven by their endless arms race [1]. Plants respond to the conserved pathogen-associated molecular patterns (PAMPs) of pathogens via a diverse group of cell surface receptors, and thereby induce PAMP-triggered immunity (PTI) [2]. Numerous plant pathogens, including bacteria, nematodes, fungi and oomycetes, can counteract PTI by delivering effector proteins into host cells to suppress plant defense and facilitate infection [1,3].

RXLR proteins are a group of effectors initially identified from *Phytophthora* species. Their N terminals feature conserved RXLR-dEER motifs, which is assumed to function in the host cell translocation process. In contrast, the C-terminals of RXLR proteins are relatively divergent due to their distinct effector activities in modulating host immunity [4,5,6]. The interacting targets of some *Phytophthora* and *Hyaloperonospora* RXLRs have been well-studied. For example, the RXLR effector PexRD2 from *Phytophthora infestans* can perturb host resistance response by interacting with a positive plant immunity regulator MAPKKK [7]. PsAvh262, which is essential for the pathogenicity of *Phytophthora sojae*, suppresses ER stress-dependent immunity via stabilizing ER-luminal binding immunoglobulin proteins (BiPs) [8]. A *Phytophthora capsici* RXLR effector, PcRXLR207, triggers the degradation of BPA1 (Binding partner of ACD11) family proteins to facilitate the biotroph to necrotroph transition [9]. A conserved RXLR effector HaRxL23 from *Hyaloperonospora arabidopsidis* can suppress PTI in tobacco as well as effector-triggered immunity (ETI) in soybean [10]. Furthermore, some RXLRs are

avirulence proteins exhibiting gene-for-gene interactions with specific host resistance proteins [11].

Quantities of RXLR candidates have been identified from *Hyaloperonospora* and *Phytophthora* species via bioinformatic searching approaches. A pilot method was designed to search for protein with a signal peptide (SP) in its first 30 residues and an RXLR motif within the following 30 residues. This method was used for RXLR identification in *Hyaloperonospora parasitica*, *Phytophthora ramorum* and *P. sojae* [12]. Later on, RXLRs were identified in three *Phytophthora* species using the hidden Markov model (HMM) together with regular expression (regex) model [5]. A homology searching method was also developed and used together with HMM to predict 370 and 392 RXLRs in *P. ramorum* and *P. sojae*, respectively [13]. Moreover, the three approaches above were combined to characterize RXLRs in *Plasmopara viticola,* an oomycete pathogen infecting grapevine [14]. RXLR genes and proteins predicted from different species show similar features of diverse variation, disorder region enrichment and predominant localization in gene-sparse regions [13,15,16]. Interestingly, there are also putative RXLRs found in *Albugo* and *Saprolegnia* species. For example, 26 Ac-RxL effectors were identified in *Albugo candida* by searching for the RXL string within proteins with SP. These Ac-RxLs exhibit effector activity despite that they only contain a degenerate RXLR motif [17]. In *Albugo laibachii,* 25 RXLR and 24 RXLQ effector candidates were found using similar methods with two of them being functionally validated [18]. The fish-infecting oomycete *Saprolegnia parasitica* harbors an RXLR effector that enters the host cells to promote virulence

[19]. *Pythium* belongs to *Pythiales*, a sister order of *Peronosporales, Albuginales* and *Saprolegniales* [20]. All these RXL-containing oomycetes are phylogenetically close to *Pythium* species. Nevertheless, previously reported *Pythium* genome analyses did not reveal any RXLR-encoding genes [21,22,23,24].

Unlike *Phytophthora* species which are mostly plant-infecting pathogens, *Pythium* species occupy multiple ecological niches. Many *Pythium* species, including *P. ultmum, P. iwayamai*, and *P. aphanidermatum*, cause a wide variety of diseases in plants [21,22], whereas *P. insidiosum* is a notorious pathogen infecting human and animals [23]. Interestingly, Some *Pythium* species, such as *P. oligandrum* and *P. periplocum*, are used as mycoparasitic biocontrol agents since they can infect fungal hosts and induce plant defense response as well [25,26,27]. Likewise, *P. guiyangense* is a parasite used for mosquito control [24].

It is still unknown whether *Pythium* genomes encode RXLR effectors, or which kinds of effectors participate in the interactions between *Pythium* species and their broad hosts. To answer the questions above, we developed a stringent method for *de novo* identification of RXLRs in the draft genome sequences of 9 *Pythium* species and performed comprehensive genomic analysis on the RXLRs predicted. *Pythium* RXLRs form a single superfamily with all other known oomycetous RXLRs. All RXLR genes and proteins exhibit similar features in sequence divergence, disorder content and genome location. Four novel RXLRs from *P. ultimum, P. oligandrum* and *P. periplocum* were functionally verified as effectors. Our findings reveal the wide occurrence of RXLRs in *Pythium* species at the whole-genome level, demonstrate the

effector activities of selected *Pythium* RXLRs, and provide a solid platform for investigating the diversified roles of *Pythium* RXLRs in pathogenicity.

Results

De novo identification of RXLR effectors

No *Pythium* RXLR effector has been previously identified despite the numerous reports of RXLRs in their phylogenetically close species [21,22,23,24]. Here we developed a *de novo* identification method with stringent threshold to screen RXLR candidates in 9 *Pythium* species (Fig 1A). Meanwhile, well-studied *P. sojae, P. ramorum* and *H. arabidopsidis* genomes were used as positive controls. Some other species including two diatoms were also parallelly examined (S1 Table, Fig 1B). For every gene, all 6 possible open reading frames (ORFs) were retrieved from each genome because almost all *RXLRs* lack intron [13]. Sequences were considered as sORFs (open reading frames encoding putative secretory protein) if their encoding proteins contain an SP but lack transmembrane region (TM).

In total, 311,164 sequences were achieved with 12,490 to 33,431 sORFs being found in each *Pythium* species, which is close to that of *Hyaloperonospora* and *Phytophthora* (12,830 to 25,178 sORFs). Interestingly, non-oomycetes only had 47 to 12,213 sORFs identified (S1 Table). Significantly more sORFs can be found in oomycetes when compared to other species (P<2.2E–16, Fisher's Exact Test), indicating the occurrence of more expansion events in secretory protein-encoding genes in oomycetes.

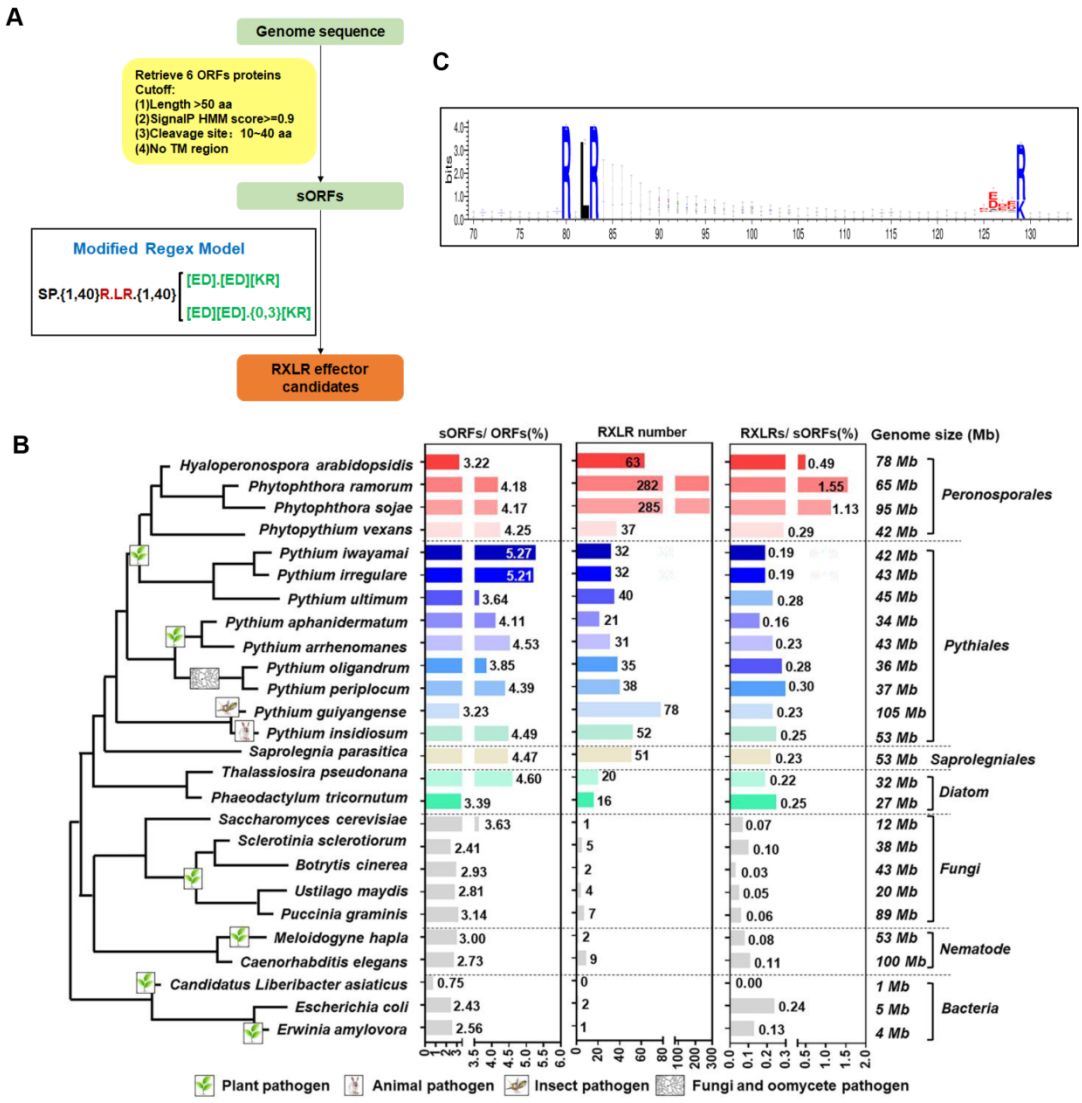


Fig 1. *De novo* identification of RXLR effectors in *Pythium* species. (A) The RXLR effector identification pipeline. sORFs indicate open reading frames encoding secretory proteins with a signal peptide (SP) but without transmembrane region (TM). (B) Summary of predicted RXLRs. Phylogeny of the 26 species is based on data from the Taxonomy Database and previous studies. The ratios of sORFs to whole ORFs, the counts of RXLR candidates and the ratios of RXLRs to sORFs in each species are showed in boxplots following the species names. The genome size of each species is showed after the corresponding boxplots. (C) Weblogo of the RXLR-dEER motifs of *Pythium* RXLR effectors.

Since some functionally verified RXLR effectors contain degenerate dEER motifs (S2 Table), the original regex model (SP.{1,96}R.LR.{1,40}[ED][ED][KR]) [5] was modified (SP.{1,40}R.LR.{1,40}([ED].[ED][KR]|[ED][ED].{0,3}[KR]) (Fig 1A) to allow the match of these atypical RXLRs. In total, 1,146 RXLR candidates were identified using this modified model (Fig 1B).

We found 63, 282 and 285 RXLRs in *H. arabidopsidis*, *P. ramorum* and *P. sojae*, respectively. They are all subsets of previously identified RXLRs and account for 42%, 53% and 42% of the total sets, respectively [5,12,13,28]. In contrast, only 0 to 9 RXLRs were predicted in fungi, nematodes or bacteria. These control results demonstrate the high reliability as well as the low false-positive rate of our method. 21 to 78 putative RXLRs were identified in each *Pythium* species. All predicted *Pythium* RXLR proteins contain a conserved SP, an RXLR motif and a dEER motif, in which two acidic amino acids (D or E) are enriched and followed by a conserved basic amino acid (R or K) (Fig 1C, S1 Fig).

To test the reliability of RXLR prediction in *Pythium*, an enrichment assay was conducted by comparing the number of RXLRs in each oomycete species with that of the fungal pathogen *Botrytis cinerea* (randomly selected as a control). All oomycetes including the *Pythium* species showed significant enrichment of RXLR sequences ($P<0.01$, Fisher's Exact Test) (S3 Table). The enrichment was not observed in any of the non-*Stramenopiles* species ($P>0.05$, Fisher's Exact Test). Next, a permutation test was performed to evaluate whether the predicted RXLR effectors are randomly occurred. We permuted the 90 after-SP residues of all sORFs and re-performed RXLR searching with the modified model for 100 times to estimate the false-positive rates. In general, very few RXLR predictions were generated from the permuted sequences, indicating that the vast majority of oomycete RXLRs were not predicted by chance (S4 Table). For example, 35 RXLRs were predicted in *P. oligandrum* whereas only 1 RXLR on average detected in its permuted sequences.

Besides their presence in all examined oomycetes including *Pythium* species, RXLR sequences can also be identified in diatoms (20 in *Thalassiosira pseudonana* and 16 in *Phaeodactylum tricornutum*) with significant enrichment (P<0.01, Fisher's Exact Test). They may be the precursors of RXLR effectors that help diatoms compete with other organisms in the complicated environment.

The evolutionary patterns of oomycete RXLRs

Despite the highly divergence of RXLR proteins resulted from host-pathogen arms race, RXLR effectors from *P. sojae* and *P. ramorum* exhibit significant relatedness, suggesting that the hundreds of fast-evolving RXLRs in these two species may derive from a single common ancestor [13]. Therefore, we used the BLASTP software to analyze whether the predicted *Pythium* RXLRs belong to the same evolutionary theme. The highly similar SP regions were removed to avoid their interference. Using BLASTP hit *E*-value<1 as the cutoff, 840 RXLR proteins had significant hits to one or more other RXLRs. In total, 15,071 RXLR pairs were identified, which form a huge relatedness network (Fig 2A). *Phytophthora* RXLRs have a higher within-genus relatedness level when compared with those from *Pythium* or *Hyaloperonospora*, suggesting a larger contribution of recent gene expansion events in *Phytophthora*. We found three clusters sharing little relatedness with *Phytophthora* RXLRs, which indicates their relative independence in evolution. Almost all RXLRs in these clusters are from *Pythium*. The three clusters are named as Cluster 1 (C1), Cluster 2 (C2) and Cluster 3 (C3), respectively (Fig 2A). For example,

the average sequence identity of the 22 RXLRs in C1 is 54%. A Peptidase_S8 domain (PF00082.22), which is absent in *Phytophthora*, can be found in the C-terminal of 21 C1 RXLRs.

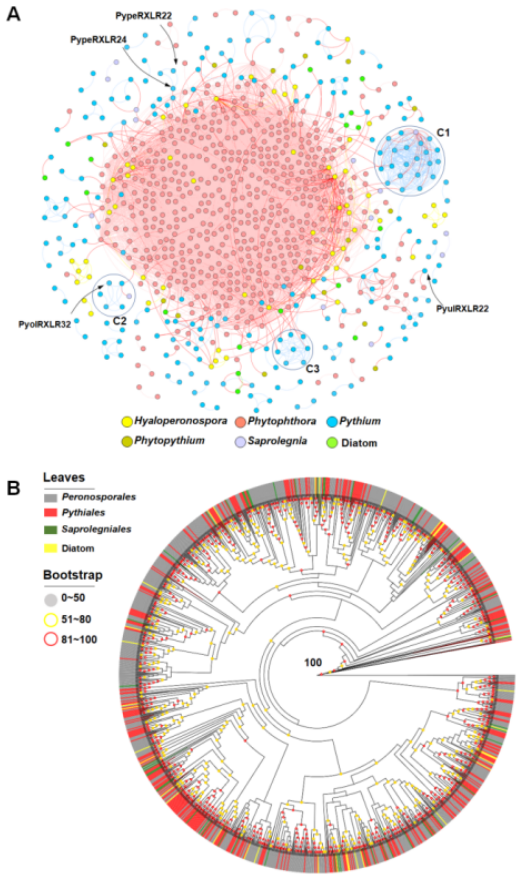


Fig 2. The evolutionary patterns of oomycetes RXLRs. (A). The relatedness network of putative RXLR effectors. Each spot represents an RXLR. Similar RXLRs (*E*-value<1) are joined by edges. Spot colors indicate their genera. Three relatively independent clusters were marked as C1, C2 and C3, respectively. Almost all effectors in these clusters are from *Pythium*. Four RXLRs with effector activities verified in this study are labeled using arrows. (B) The maximum-likelihood phylogenetic tree of the RXLR-dEER motifs in RXLR effectors. Grey, yellow and red circles in the nodes indicate bootstrap values of 0–50, 51–80 and 80–100, respectively. Leaves of the tree are colored base on the families of their indicated species.

Conserved RXLRs may play critical effector roles in pathogen virulence [10]. Analysis of shared sequence identity (peptides with BLASTP hit *E*-value<1e-5 and identity>30%) revealed 67, 93 and 28 *Pythium* RXLRs being intraspecies, intragenus and intergenus conserved, respectively (S5 Table), suggesting that *Pythium* RXLRs share a common ancestor with those from other genera and were only slightly expanded after species divergence. These conserved RXLRs may be essential effectors for pathogens and deserve more attention in future study.

Phylogenetic analysis was performed to further elucidate the evolutionary relationships among oomycete RXLRs as well as RXLR-like proteins from diatoms. Due to the host-pathogen arms race, RXLR C-terminals (after the dEER motif) are

243 much more divergent [12] than their relatively conserved N-terminals [13,29].
244 Therefore, when building the phylogenetic tree, C-terminal and SP regions were
245 discarded to reflect RXLR evolution relatedness with minimal interferences from host
246 coevolution and secretory signal sequences. Except for PyolRXLR33 and
247 HyarRXLR1, all other 1,111 RXLRs formed a huge clade (bootstrap value=100),
248 indicating that they belong to a single superfamily. Using bootstrap value>80 as the
249 criteria, 972 RXLRs can be divided into 105 subfamilies with member sizes ranging
250 from 2 to 112 (S6 Table). No significant divergence was detected among
251 *Peronosparales*, *Pythiales*, *Saprolegniales* or diatom RXLRs (Fig 2B). 39 and 6
252 subfamilies contain 119 and 28 RXLRs exclusively from *Peronosparales* and
253 *Pythiales* species, respectively. All other subfamilies harbor RXLRs from at least two
254 different genera. Taken together, the phylogeny results suggest a common ancestor
255 shared by almost all RXLR and RXLR-like proteins with species-specific expansion
256 much less detected in *Pythium* than in *Phytophthora*.
257

258 ## Rapid evolution of *Pythium* RXLRs

259 Highly divergent *Phytophthora* and *Hyaloperonospor*a RXLRs are one of the
260 most rapidly evolving portions in their proteomes [12,13]. Similarly, *Pythium* RXLRs
261 also share low identities in our study. On average, only 37% identity was detected for
262 all closest RXLR ortholog pairs within the *Pythium* genus. In contrast, we randomly
263 selected 300 *Pythium* proteins and found a much higher average sequence identity
264 (67%) associated with their closest ortholog pairs (Fig 3A). In our study, 233 (65%)

Pythium RXLRs failed to form any ortholog pairs whereas only 10 out of 300 randomly selected *Pythium* proteins lack orthologs (Fig 3A). Furthermore, 236 (66%) *Pythium* RXLRs showed higher than 50% of sequence divergence (Fig 3A).

Disorder content is abundant in *Pythium* RXLRs

Disorder content is abundant in *Phytophthora* RXLRs which contributes to their effector activities [16]. Our analysis demonstrated the abundance of disorder content in *Pythium* RXLRs as well. The mean percentages of disorder region in randomly selected *Pythium* proteins, *Pythium* RXLRs and *P. sojae* RXLRs are 27.17%, 45.48% and 36.89%, respectively (Fig 3B). *Pythium* RXLRs shows obviously abundant disorder content which is even higher than that of *P. sojae*. Over 64% *Pythium* RXLRs have 30% or more disordered residues, which can be found in only 43.67% randomly selected *Pythium* proteins (Fig 3B).

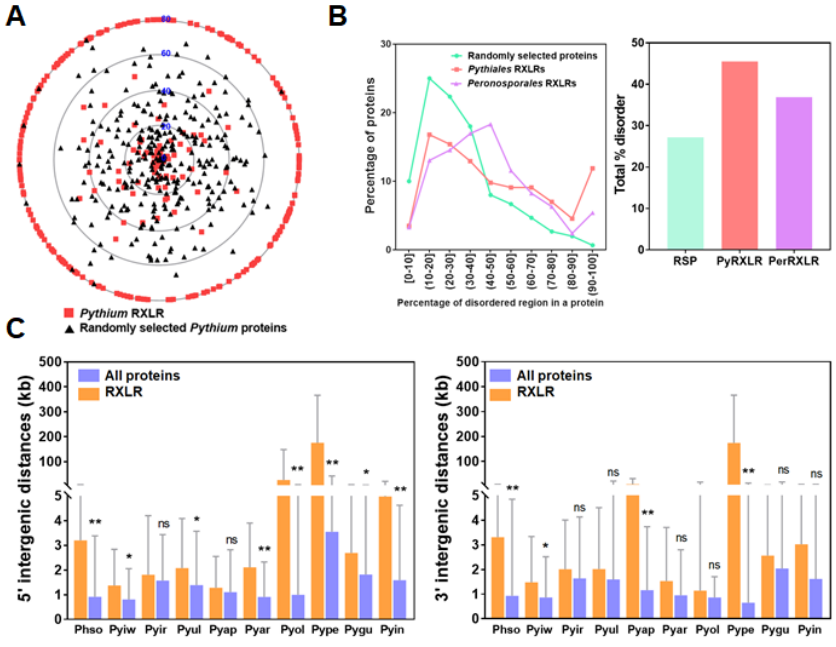


Fig 3. Evolutionary features of *Pythium* RXLRs. (A) Sequence divergence of *Pythium* RXLRs. *Pythium* RXLRs were compared against 300 randomly selected *Pythium* proteins

regarding the sequence identities of their closest ortholog pairs within the *Pythium* genus. Radiuses range from 0 (center) to 80 (outer circle) represent 100% to 20% (or less) sequence identities sequentially. Proteins at the same identity level are randomly distributed along their corresponding circle. **(B) Disorder content in *Peronosporales* and *Pythiales* RXLRs.** The line plot shows the percentages of proteins containing different percentages of disordered regions. The bar plot shows total disorder percentages in *Peronosporales*, *Pythiales* and randomly selected proteins. RSP, Randomly selected proteins; PyRXLR, *Pythiales* RXLRs; PerRXLR, *Peronosporales* RXLRs. **(C) 5' and 3' end intergenic distances of *P. sojae* and *Pythium RXLRs*.** The two bar plots show 5' (left) and 3' (right) end intergenic distances of indicated *RXLR* genes, respectively. (*, $P<0.05$; **, $P<0.001$, Student's *t* test). Phso, *P. sojae*; Pyiw, *P. iwayamai*; Pyir, *P. irregulare*; Pyul, *P. ultimum*; Pyap, *P. aphanidermatum*; Pyar, *P. arrhenomanes*; Pyol, *P. oligandrum*; Pype, *P. periplocum*; Pygu, *P. guiyangense*; Pyin, *P. insidiosum*.

Pythium RXLRs are predominantly located in gene-sparse regions

Phytophthora RXLR genes are preferably located in gene-sparse regions [15]. In *P. sojae*, both the 5' and the 3' end average intergenic distances of *RXLRs* are significantly longer than those of the complete *P. sojae* genes ($P<0.05$, Student's *t* test). The mean 5' end intergenic distance of *RXLRs* is 3.20 kb with the whole-genome mean being 0.91 kb. *RXLRs*' average 3' end intergenic distance of 3.31 kb is also much longer than that of the complete *P. sojae* genes (0.92 kb) (Fig 3C). Both observations are similar to the previous report [15].

Similar gene location preference was detected for *Pythium RXLRs* in our study. Compared with the complete gene sets of their corresponding species, *RXLRs* of all 9 examined *Pythium* species have longer average intergenic distances at both 5' and 3' ends (Fig 3C). For the 5' and the 3' end average intergenic distances, the differences are statistically significant ($P<0.05$, Student's *t* test) in 7 and 3 *Pythium* species, respectively (Fig 3C). These results indicate that *Pythium RXLRs* are predominantly located in gene-sparse regions, which is similar to that of the *Phytophthora RXLRs*.

Genome rearrangement contribute to the emergence of novel *RXLR* genes in *P. guiyangense*

P. guiyangense has a larger RXLR repertoire than all other *Pythium* species examined (Fig 1B). Our previous synteny analysis showed that *P. guiyangense* has a hybrid genome derived from two distinct parental species, which leads to a nearly two-fold expansion of its predicted genes when compared with other *Pythium* species. Consequently, most homology gene pairs in *P. guiyangense* show high degrees of collinearity [24]. However, we only found 4 *RXLRs* from different scaffolds showing collinearity, indicating that additional mechanisms may also contribute to the expansion of *RXLR* genes in *P. guiyangense*. We noticed that 29 *PyguRXLRs* form several clusters (<40 kb in size) with each cluster located in the rearrangement region of a same contig. Similar phenomenon was not observed in other *Pythium* species (S2 Fig). For example, *PyguRXLR6/7/8* (Fig 4A) and *PyguRXLR32/33/34* (Fig 4B) are clustered in the genome rearrangement regions in contigs 2 and 9, respectively. Furthermore, 22 clustered RXLRs show high similarities in N-terminals but have less conserved or even highly divergent C-terminals within each cluster. For example, PyguRXLR6/7/8 share a conserved N-terminal but their C-terminals are highly divergent (Fig 4C). Similar pattern was observed in PyguRXLR32/33 versus PyguRXLR34 (Fig 4D). Taken together, these results suggest that both gene duplications and genome rearrangements contribute to the expansion and the emergence of novel *RXLRs* in *P. guiyangense*.

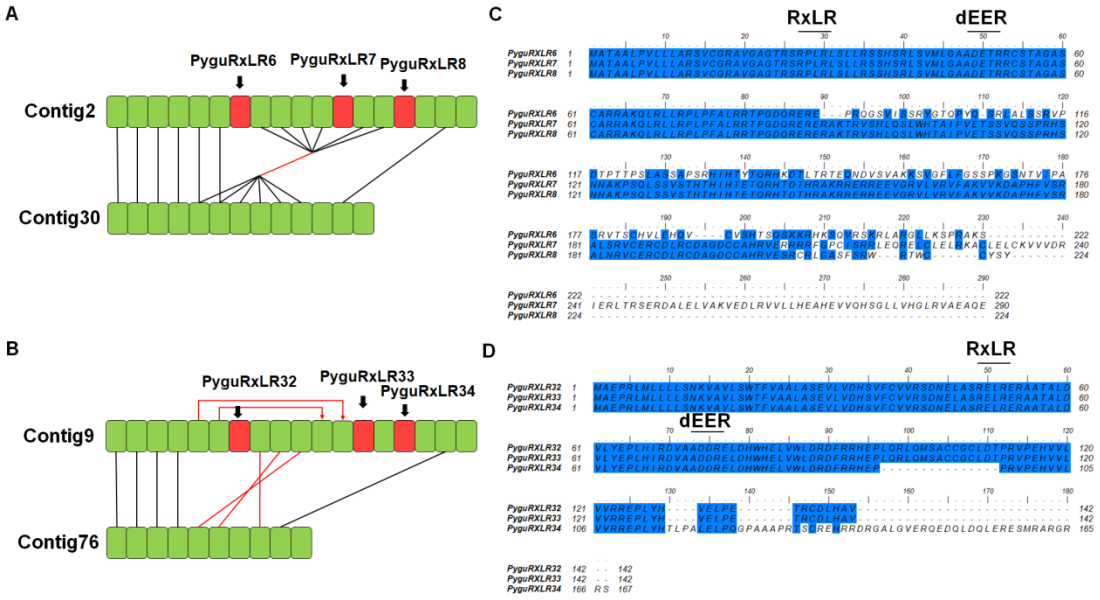


Fig 4. Synteny analysis and protein sequence alignments of clustered RXLRs in *P. guiyangense*. (A and B) Synteny analysis of clustered *P. guiyangense RXLRs*. Red blocks indicate clustered *PyguRXLRs* described in the text. Green blocks indicate genes flanking to these *PyguRXLRs*. Homologous genes with similar and opposite genome orientations are joined with black and edges, respectively. (C and D) Protein sequence alignments of two RXLR clusters. Identical or similar residues are colored in blue. The RXLR-dEER motifs are labeled in the alignments.

Most predicted *P. ultimum RXLRs* are truly transcribed genes

To further elucidate whether these predicted *Pythium RXLRs* are truly transcribed genes, an RNA-seq analysis was carried out to examine the transcript accumulations of 40 *RXLR* candidates in *P. ultimum* during mycelium and 3, 6, 12, 24, 36 h post host-infection stages. As one of the most important oomycetous plant pathogens with a high-quality assembled genome sequence available [21,30], *P. ultimum* was selected as the investigation target. Except for 2 candidates (PyulRXLR25/26) whose RPKM (reads per kilobase transcript length per million reads mapped) values were consistently below 1 at all stages, transcripts of the remaining 38 predicted *RXLRs* can be detected at one or more stages (Fig 5A), suggesting that most *P. ultimum RXLRs*

predicted in our study are truly transcribed genes.

P. ultimum RXLR genes showed stage-specific expression patterns. They were assigned to four gene expression clusters (Fig 5A) via hierarchical clustering (HCL). Cluster 1 contains 2 putative pseudogenes and 5 *RXLRs* exhibiting zero to low expression. Despite their low transcript accumulations, all *RXLRs* in this cluster showed elevated expression during early infection stages. The 14 *RXLRs* in Cluster 2 exhibited elevated transcript accumulations in both mycelium and late infection stages. The expression of 11 Cluster 3 *RXLRs* were largely constant across all the stages with slightly more transcripts detected at 3/6 hours post infection. The 8 Cluster 4 *RXLRs* showed a similar expression pattern as that of the *RXLRs* in Cluster 2 but with much higher RPKM values. These results revealed major transcription shifts of *PyulRXLRs* at mycelium and different host infection stages.

PyulRXLR22 induces cell death in *Nicotiana benthamiana*

Phytophthora RXLRs suppress PAMP-induced plant cell death. Their ectopic expression can trigger cell death in plants [31]. To evaluate the effector activity of predicted PyulRXLRs, we selected 4 candidates (PyulRXLR3/4/10/22, 2 from Cluster 1 and 2 from Cluster 4) for functional characterizations (Fig 5A). When co-expressed with *INF1* [32] in *N. benthamiana* leaves, none of these four RXLRs can suppress cell death triggered by the *P. infestans* PAMP INF1 (S3 Fig). When evaluating their ability to trigger host cell death by ectopic expression, only PyulRXLR22 induced dramatic cell death in *N. benthamiana* leaves (Fig 5B). *PyulRXLR22* expression is

highly upregulated at late infection stages (Fig 5A). Fused with a GFP tag at N-terminal, the four PyulRXLRs were transiently expressed in *N. benthamiana* epidermal cells to reveal their subcellular localizations. PyulRXLR22 is located in endoplasmic reticulum (ER), plasma membrane and the nucleus (Fig 5C). In contrast, PyulRXLR3/4/10 were found in plasma membrane and the nucleus but not in ER (Fig 5C). Bioinformatic analysis showed that PyulRXLR22 has two homologs in *P. ramorum* (BLASTP *E*-value<1) (Fig 2A). The demonstrated effector activity makes PyulRXLR22 and its homologs promising targets for further investigation.

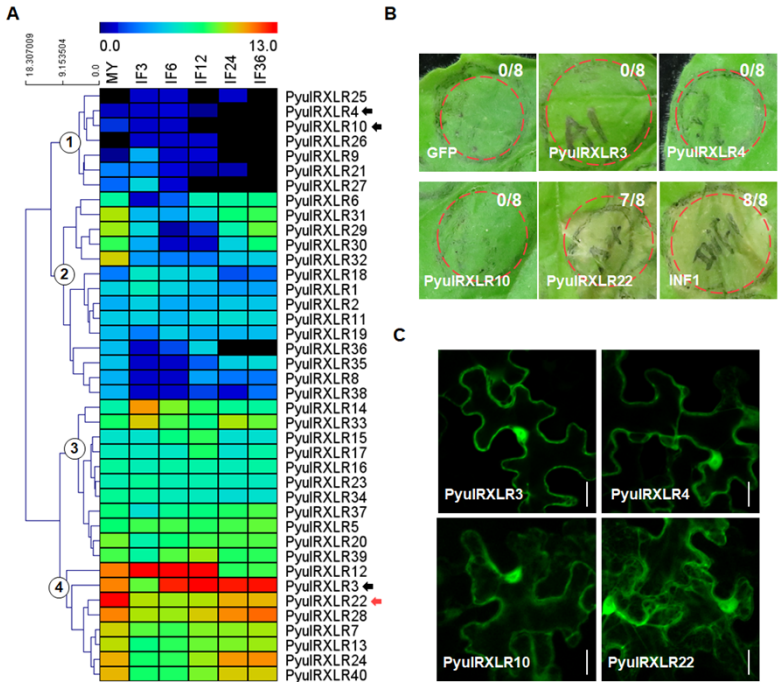
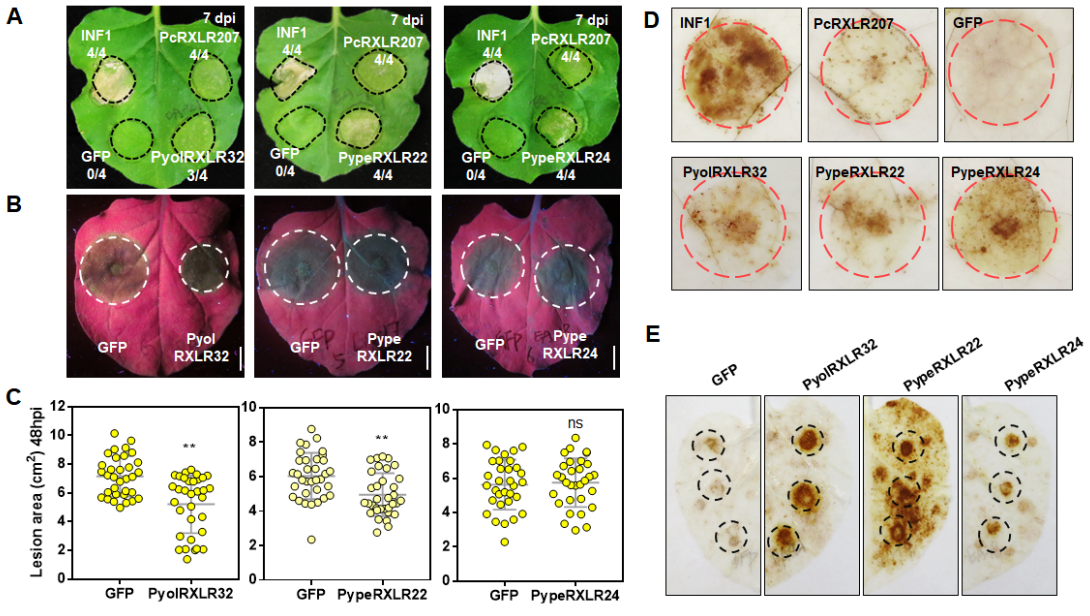


Fig 5. Transcription patterns of *P. ultimum RXLRs* and the effector activity of PyulRXLR22. (A) Transcription patterns of *P. ultimum RXLR* candidates. The heat map shows gene transcription patterns at the mycelia (MY) stage and stages of 3, 6, 12, 24 and 36 h post-inoculation of soybean leaves. Color bars represent log2 of gene RPKM values, ranging from dark blue (0) to red (13). Black bars indicate no expression detected. RXLRs selected for the effector activity assay are marked with arrows with red arrow indicating ability of inducing cell death in *N. benthamiana*. Based on transcript accumulation levels and patterns, *P. ultimum RXLRs* are divided into 4 clusters using HCL methods and the MeV software. Clusters are marked with circled numbers (1 to 4). (B) PyulRXLR22 induced cell death in *N. benthamiana*. *N. benthamiana* leaves were transformed with the indicated constructs by agro-infiltration. Dashed circles indicate infiltration sites and the ratios are sites with necrotic lesion versus total infiltration sites. PyulRXLR3/4/10 did not induce

392 cell death. GFP and INF1 were used as negative and positive controls, respectively. (C)
393 Subcellular localization of PyulRXLRs. Green fluorescence indicates the localizations of
394 PyulRXLRs in *N. benthamiana* epidermal cells. Photographs were taken at 48 hpi. Bar =20 μm.
395

396 **RXLRs from biocontrol *Pythium* species induce plant defense in *N.***
397 ***benthamiana***

398 *P. oligandrum* and *P. periplocum* are two unique *Pythium* species that are widely
399 used for controlling soilborne plant diseases. They can infect pathogens and induce
400 plant defense reactions as well [25,26,27]. To elucidate whether their RXLRs
401 participate in this biocontrol process, all 73 RXLRs of these two species were cloned
402 and expressed in *N. benthamiana* leaves individually to test their ability of triggering
403 cell death. PyolRXLR32, PypeRXLR22 and PypeRXLR24 triggered cell death at 7
404 days post infiltration (dpi) (Fig 6A). All other 70 RXLRs lack this ability.
405 PyolRXLR32 is located in plasma membrane and the nucleus. Both PypeRXLR22
406 and PypeRXLR24 are located in plasma and nuclear membranes (S4 Fig).
407



408

Fig 6. RXLRs from *P. oligandrum* and *P. periplocum* induce plant defense response. (A) Expression of PyolRXLR32, PypeRXLR22 or PypeRXLR24 induce cell death in *N. benthamiana* leaves. *N. benthamiana* leaves were transformed with the indicated constructs by agro-infiltration. Dashed circles indicate infiltration sites and the ratios are sites with necrotic lesion versus total infiltration sites. INF1 and the RXLR effector RXLR207 from *P. capsici* were used as positive controls. Empty vector expressing GFP was used as a negative control. Same controls were also used in the H_2O_2 accumulation assay. (B and C) *P. capsici* infection can be reduced by PyolRXLR32 and PypeRXLR22, but not PypeRXLR24. 24 h after infiltration, equal amount of mycelium was inoculated onto each infiltrated area expressing *PyolRXLR32*, PypeRXLR22, or PypeRXLR24. Photographs were taken under UV light at 48 hpi and lesion areas were measured at the same time. Lesion area data shown in the dot plots were collected from three biological replicates. At least five leaves were inoculated for each replicate (**, $P<0.001$, Student's *t* test). Bar=1cm. (D) H_2O_2 accumulation in *N. benthamiana* leaves expressing PyolRXLR32, PypeRXLR22 or PypeRXLR24. DAB staining was performed 48 hpi. Red dashed circles indicate infiltration sites. (E) H_2O_2 accumulation in *P. capsica*-inoculated *N. benthamiana* leaves expressing PyolRXLR32, PypeRXLR22, or PypeRXLR24. *P. capsica* was inoculated 24 h after infiltration. DAB staining was performed at 12 hpi. Black dashed circles indicate inoculation sites.

To investigate whether the three RXLR effectors can induce plant defense, we infiltrated *N. benthamiana* leaves with *Agrobacterium tumefaciens* harboring a *GFP* or indicated *RXLR* fusion construct and then inoculated the leaves with *P. capsici*. Compared with the GFP control, both PyolRXLR32 and PypeRXLR22 significantly reduced the lesion area caused by *P. capsici* whereas PypeRXLR24 did not have noticeable effect on lesion control (Fig 6B and 6C).

Reactive oxygen species (ROS) play a central role in plant immunity with ROS burst being a defense response indicator [33]. Our DAB staining assay showed that, compared with the GFP control, ectopic expression of any of the three RXLRs induced ROS burst in *N. benthamiana* leaves (Fig 6D). Furthermore, all three RXLRs can enhance ROS burst triggered by *P. capsici* inoculation (Fig 6E).

Interestingly, PyolRXLR32 is both intragenus and intergenus conserved (S5 Fig). Its *P. infestans* homolog (BLASTP *E*-value=2e-115), previously reported as SFI4

(Suppressor of early Flg22-induced Immune response 4), can inhibit host immunity and promote infection [34], which is opposite to PyolRXLR32's positive role in triggering cell death and defense response. Likewise, although PypeRXLR21/22/24 share high sequence similarity (S6 Fig), PyolRXLR21 cannot trigger host cell death as PyolRXLR22/24 did. These results indicate that slight sequence difference may lead to diverse or even contrary functions of homologous RXLR effectors.

Discussion

In this study, we performed a *de novo* identification of RXLR effector genes from the genomes of nine *Pythium* species and analyzed their evolutionary features. Selected *Pythium* RXLR effectors were functionally verified. Despite the stringent prediction method we adopted, RXLRs can be found in a wide range of oomycetes including the *Pythium* species, and are almost absent outside *Stramenopiles*. Our phylogenetic analysis indicates the existence of a common ancestor for RXLRs with all oomycetous RXLRs clustered into a single superfamily. *Pythium* and *Phytophthora RXLR* genes share several similar features including rapid evolution, abundant disorder content and the preference of locating in gene-sparse regions. In particular, *P. guiyangense* harbors a much larger RXLR reservoir than other *Pythium* species due to frequent gene duplication and genome rearrangement events. Using *P. ultimum* as a model, most of its predicted *RXLRs* were demonstrated to be truly transcribed genes that cluster into four distinct expression patterns. The effector activity of PyulRXLRs22 was functionally verified in this study. Finally, we

identified three functional RXLR effectors from two biocontrol *Pythium* species.

Although no *Pythium* RXLRs has been identified before [21,22], there are clues indicating their existence. Several RXLR effectors have been experimentally validated in species phylogenetically close to *Pythium*, including *A. candida*, *A. laibachii* and *S. parasitica* [17][18] [19]. The necrotrophic infection type of *Pythium* pathogens also match the observations that many RXLR effectors can induce necrosis. Largely due to the stringent definition of the dEER motif, existing genome annotation models may be inadequate for *RXLR* gene prediction in *Pythium*. Hence, we developed a modified regex model to allow the search of degenerate dEER motifs. Previously reported effectors were adopted as references and all 6 possible ORFs were retrieved from the whole genomes. To reduce false-positive rates, we omitted the BLASTP step following RXLR searching. By adopting our modified regex model, we predicted 359 new RXLR proteins in nine *Pythium* species as well as re-identification of RXLRs from *H. arabidopsidis*, *P. ramorum* and *P. sojae*. The high accuracy of our new model is demonstrated by the observations that very few RXLRs was predicted outside *Stramenopiles* and all re-identified RXLRs are subsets of previously known RXLRs. Notably, our detection of RXLR-like proteins in diatoms suggests the origin of oomycetous RXLRs from an early ancestor of the *Stramenopiles* lineage. *Pythium* and other oomycetous RXLRs form a superfamily in our phylogenetic analysis, which also support the common-ancestor hypothesis of RXLR evolution.

Despite the high divergence of RXLR sequences, oomycetous RXLRs identified

484 in our study are close enough to form a huge relatedness network with *Phytophthora*
485 RXLRs exhibiting higher overall sequence similarities. There are also RXLR clusters
486 predominantly associated with *Pythium* or *Hyaloperonospora* species, suggesting the
487 occurrence of independent evolution in these effectors.

488 *Phytophthora* and *Pythium* RXLRs share several common features. *Phytophthora*
489 RXLR effectors promote virulence by manipulating host defense processes. Thus,
490 they undergo rapid sequence evolution to escape the surveillance of plant immune
491 system. In our study, this fast evolution pattern was also observed in *Pythium* RXLRs,
492 which indicate their participation in host pathogenesis. As a consequence of effector
493 evolution, abundant disorder regions exist in *Phytophthora* RXLRs [16]. These
494 disorder regions are even more prevalent in *Pythium* RXLRs. *Phytophthora* genomes
495 feature a so-called 'two-speed' structure of gene-dense and gene-sparse regions [35],
496 with *RXLRs* enriched in the more dynamic gene-sparse regions. Gene-sparse regions
497 are the heat sites of genome rearrangement and tend to accelerate pathogen evolution
498 [36]. In *Pythium* species, we also found that both 5' and 3' end average intergenic
499 distances of *RXLRs* are significantly longer than the whole genome averages, which
500 demonstrates the similar location preference of *Pythium RXLRs* in the gene-sparse
501 regions.

502 Due to their highly divergent C-terminals, few RXLR paralogs can be found in
503 *Pythium* species except for *P. guiyangense*. We characterized the *P. guiyangense*
504 genome as hybrid and identified a large set of *RXLR* genes. *P. guiyangense RXLRs*
505 share low degrees of synteny with *PyguRXLR* paralogs often clustered in the same

506 contig and located in genome rearrangement regions. Furthermore, physically close
507 paralog pairs generally share a conserved N-terminal but have less conserved or even
508 highly divergent C-terminals. Our results indicate that recently happend gene
509 duplication and genome rearrangement events shape the evolution theme of *P.*
510 *guiyangense RXLR*s. Previous research suggests that the duplications of genes
511 encoding virulence-associated effectors can facilitate pathogen adjustment towards a
512 host jump [37]. In the case of *P. guiyangense*, the duplication and rearrangement
513 events detected may generate novel *RXLR* genes which can help its adaption against
514 insects. *RXLR* genes also clustered closely (<100kb) in *P. ramorum* [12], suggesting
515 that *RXLR* gene cluster may be a common host adaptation strategy shared by
516 oomycete species.

517 Our RNA-seq analysis on *P. ultimum* demonstrated that most predicted *RXLRs*
518 are transcribed genes, which further verified the reliability of our redesigned RXLR
519 annotation method. Meanwhile, these *RXLRs* exhibit distinct expression patterns
520 during infection. RXLR expression stages are associated with their specific functions.
521 In *P. sojae*, early-induced RXLRs inhibit host cell death triggered by PAMPs, while
522 necrosis is caused by late-induced RXLR effectors [31]. Another *Phytophthora*
523 pathogen, *P. capsici*, secretes a RXLR effector to induce host defense and cell death,
524 and trigger the biotroph-to-necrotroph transition [9]. Since most *Pythium* pathogens
525 are necrotrophic, their RXLRs may trigger host cell death instead of acting as cell
526 death repressors. Consistent with this expectation, none of our tested PyulRXLRs can
527 suppress host cell death induced by INF1. Instead, ectopic expression of

528 PyulRXLR22 triggers necrosis. The function diversity of RXLRs in *Phytophthora* and
529 *Pythium* species may result from their different nutrient acquisition strategies during
530 infection.

531 *P. oligandrum* and *P. periplocum* are two biocontrol agents with their
532 defense-inducing mechanisms largely unknown. Previous reports suggest that plant
533 defense may be induced by Microbe-Associated Molecular Patterns (MAMPs) from
534 these two *Pythium* species [38]. In this study, we identified three RXLR effectors
535 from them and demonstrated the plant defense induction ability of these RXLRs. This
536 finding unveils the participation of cytoplasmic effectors in triggering host defense
537 reactions. Since the *P. oligandrum* hyphae can penetrate into the root tissues, its
538 RXLR effectors may be secreted to host cells via a similar mechanism detected in
539 *Phytophthora* pathogens.

540 In conclusion, we for the first time identified RXLR effectors in *Pythium* species
541 at the genome-wide scale. Their evolutionary features, sequence characteristics and
542 expression patterns were analyzed in detail. The effector activities of selected *Pythium*
543 RXLRs were experimentally verified, including their diversified functions in inducing
544 cell death and host defense. This study expands our current understanding of
545 oomycetous RXLR effectors to the *Pythium* genus. Resources generated in this work
546 facilitate future in-depth investigations on the interactions between these novel
547 *Pythium* RXLRs and their diversified hosts.

548

549

Materials and methods

Data sets

The genome sequences of *P. ramorum* and *P. sojae* [4] were obtained from the Joint Genome Institute website (http://genome.jgi-psf.org). The genome sequences of *H. arabidopsidis* [28], *Ph. vexans* [22], *P. aphanidermatum* [22], *P. arrhenomanes* [22], *P. irregulare* [22], *P. iwayamai* [22], *P. oligandrum* [39], *P. periplocum* [39], *P. ultimum* [21], *P insidiosum* [23], *P. guiyangense* [24], *S. parasitica* [40], *P. tricornutum* [41], *T. pseudonana* [42], *S. cerevisiae* [43], *S. sclerotiorum* [44], *U. maydis* [45], *P. graminis* [46], *B. cinerea* [47], *C. elegans* [48], *M. hapla* [49], *E. coli* [50], *E. amylovora* [51] and *Ca* L. asiaticus [52] were obtained from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/). Detailed genomic information of all species is listed in S7 Table.

Identification of RXLR effectors

For RXLR effector identification, all open reading frames (ORFs) encoding proteins longer than 50 amino acids were retrieved. Signal peptides (SP) were predicted using SignalP v3 with the criteria that cleavage sites are located between residues 10 and 40 and HMM prob>=0.9 [53]. Transmembrane regions (TMs) were predicted using TMHMM v2.0 [54]. Sequences containing SP but lacking TM were retained and screened for RXLRs with a modified regex model described as SP.{1,40}R.LR.{1,40}([ED].[ED][KR]|[ED][ED].{0,3}KR]). All matched protein sequences were predicted as RXLR effectors. For redundant RXLRs caused by a set of alternative overlapping ORFs within the same reading frame, only the single ORF

572 encoding the longest RXLR was selected.

573 ## Phylogenetic analysis

574 For phylogenetic analysis, SPs and C-terminals after the dEER motif were
575 removed from RXLR sequences. The truncated sequences were aligned using
576 MUSCLE [55] and then manually edited. Maximum-likelihood (ML) trees were
577 constructed using the IQ-TREE software with the parameter of -bb 1000 [56].
578 VT+F+R6 was selected as the best-fit model for this assay by using the ModelFinder
579 tool in IQ-TREE. The phylogenetic tree was showed using EvolView v3 [57].

580 ## Disorder region prediction and synteny analysis

581 Disorder region prediction was performed by using ESpritz
582 (http://protein.bio.unipd.it/espritz/) with the X-ray prediction type and the best-sw
583 threshold [58]. Synteny analysis was conducted using MCScanX [59].

584 ## Vector construction

585 For transient expression in *N. benthamiana*, *PyulRXLRs, PyolRXLRs* and
586 *PypeRXLRs* (lacking the SP-encoding region) were amplified from their
587 corresponding genomic DNA and inserted into the pBinGFP2 vector [60]. Primers
588 used for PCR amplification are listed in S8 Table.

589 ## Plant materials and agroinfiltration

590 *N. benthamiana* plants were maintained in a greenhouse with an environmental
591 temperature of 25℃ under a 16-h/8-h light/dark photoperiod. The *A. tumefaciens*
592 strain GV3101 was previously stored in our lab. Before infiltration, *A. tumefaciens*
593 clones harboring indicated constructs were cultured at 28℃, 200 rpm for 48 h to reach

the appropriate concentration. The *Agrobacterium* cultures were then resuspended in 10 mM $MgCl_2$ to an appropriate optical density (OD). For the necrosis induction assay, *A. tumefaciens* cells were infiltrated into the 6-week-old leaves of *N. benthamiana* at a final OD of 0.2 at 600 nm for each construct. For the cell death suppression assay, *A. tumefaciens* cells expressing the indicated RXLR effectors (OD_{600}=0.3) were infiltrated into *N. benthamiana* leaves one day before the infiltration of *A. tumefaciens* expressing INF1 (OD_{600}=0.2). Symptom development was examined 7 days after infiltration.

Oomycete culture conditions and the inoculation assay

The *P. oligandrum* strain CBS 530.74 and the *P. periplocum* strain CBS 532.74 were kindly provided by Dr. Laura J. Grenville-Briggs at the Swedish University of Agricultural Sciences and were stored at 25 ℃ on 10% (v/v) V8 juice medium in the dark. *P. ultimum* and the *P. capsici* strain LT263 were previously stored in our lab and grown at 25 ℃ on the V8 medium in the dark. For *P. capsici* mycelium inoculation, 7-mm disks of 4-day growth medium were inoculated on *N. benthamiana* leaves 24 h after infiltration. Inoculated leaves were photographed at 48 hpi under UV light. Lesion areas were measured at the indicated time points. This assay was independently repeated for at least three times. For the inoculation of *P. ultimum*, 4-mm plugs from the growing edge of a 2-day V8 agar culture were placed in the *N. benthamiana* leaves, and the samples were harvested at the indicated time points.

Transcriptome sequencing and bioinformatic analysis

For mycelium samples, *P. ultimum* mycelium growing in V8 agar culture for 2

616 days were harvested. *P. ultimum* samples at different infection stages (3, 6, 12, 24, 36
617 hpi) were harvested using a 5 cm punch. Three biological replicates were performed
618 and used to produce an independent pool. The quality of total RNA was evaluated by
619 Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Samples
620 with $OD_{260/230} \geq 1.8$ and $OD_{260/280} \geq 1.8$ were used for RNA-seq. Raw reads were
621 filtered using the FastQC software
622 (https://www.bioinformatics.babraham.ac.uk/index.html) to remove poor-quality
623 reads and adapters. Sequences with a Phred score above 20 were considered as
624 clean reads. All clean reads were mapped to the *P. ultimum* reference genome using
625 Hisat2 with the parameters of --min-intronlen 30 --max-intronlen 5000 –dta [61].
626 SAMtools was used to convert and sort SAM files [62]. Based on the lengths of each
627 gene and the counts of uniquely mapped reads, gene expression levels were
628 normalized as RPKM (reads per kilobase transcript length per million reads mapped)
629 using the Stringtie tool [61]. Heat maps and clustering (using hierarchical clustering
630 methods) of gene expression patterns were performed using the MeV software [63].

631 **DAB staining**

632 *N. benthamiana* leaves with ectopic expression of indicated RXLR effectors were
633 detached at 48 hpi and then stained with 1 mg/ml DAB solution for 8 h in the dark.
634 After destaining with ethanol, leaves were examined under light microscopy. *P.*
635 *capsica*-inoculated *N. benthamiana* leaves were stained with 1 mg/ml DAB solution
636 for 8 h in the dark at 12 hpi, and then destained with ethanol before photographing.
637
638

References

1. Jones JD, Dangl JL. 2006. The plant immune system. Nature. 444(7117): 323-9.
2. Monaghan J, Zipfel C. 2012. Plant pattern recognition receptor complexes at the plasma membrane. Curr Opin Plant Biol. 15(4): 349-57.
3. Latijnhouwers M, de Wit PJ, Govers F. 2003. Oomycetes and fungi: similar weaponry to attack plants. Trends Microbiol. 11(10): 462-9.
4. Tyler BM, Tripathy S, Zhang X, et al. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science. 313(5791): 1261-6.
5. Whisson SC, Boevink PC, Moleleki L, et al. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature. 450(7166): 115-8.
6. Dou D, Zhou JM. 2012. Phytopathogen effectors subverting host immunity: different foes, similar battleground. Cell Host Microbe. 12(4): 484-95.
7. King SR, McLellan H, Boevink PC, et al. 2014. *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKK ε to suppress plant immune signaling. Plant Cell. 26(3): 1345-59.
8. Jing M, Guo B, Li H, et al. 2016. A *Phytophthora sojae* effector suppresses endoplasmic reticulum stress-mediated immunity by stabilizing plant Binding immunoglobulin Proteins. Nat Commun. 7: 11685.
9. Li Q, Ai G, Shen D, et al. 2019. A *Phytophthora capsici* Effector Targets ACD11 Binding Partners that Regulate ROS-Mediated Defense Response in Arabidopsis. Mol Plant. 12(4): 565-581.
10. Deb D, Anderson RG, How-Yew-Kin T, Tyler BM, McDowell JM. 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(3): 374-385.
11. Anderson RG, Deb D, Fedkenheuer K, McDowell JM. 2015. Recent Progress in RXLR Effector Research. Mol Plant Microbe Interact. 28(10): 1063-72.
12. Win J, Morgan W, Bos J, et al. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. Plant Cell. 19(8): 2349-69.
13. Jiang RH, Tripathy S, Govers F, Tyler BM. 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. 2008. 105(12): 4874-9.
14. Yin L, An Y, Qu J, Li X, Zhang Y, et al. (2017) Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism. Sci Rep 7: 46553.
15. Haas BJ, Kamoun S, Zody MC, et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature. 2009. 461(7262): 393-8.
16. Shen D, Li Q, Sun P, Zhang M, Dou D. 2017. Intrinsic disorder is a common structural characteristic of RxLR effectors in oomycete pathogens. Fungal Biol. 121(11): 911-919.
17. Links MG, Holub E, Jiang RH, et al. 2011. De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. BMC Genomics. 12: 503.
18. Kemen E, Gardiner A, Schultz-Larsen T, et al. 2011. Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. PLoS Biol. 9(7):

e1001094.

19. Wawra S, Bain J, Durward E, et al. 2012. Host-targeting protein 1 (SpHtp1) from the oomycete *Saprolegnia parasitica* translocates specifically into fish cells in a tyrosine-O-sulphate-dependent manner. Proc Natl Acad Sci U S A. 109(6): 2096-101.

20. McCarthy C, Fitzpatrick DA. 2017. Phylogenomic Reconstruction of the Oomycete Phylogeny Derived from 37 Genomes. mSphere. 2(2).

21. Lévesque CA, Brouwer H, Cano L, et al. 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. Genome Biol. 11(7): R73.

22. Adhikari BN, Hamilton JP, Zerillo MM, Tisserat N, Lévesque CA, Buell CR. 2013. Comparative genomics reveals insight into virulence strategies of plant pathogenic oomycetes. PLoS One. 8(10): e75072.

23. Rujirawat T, Patumcharoenpol P, Lohnoo T, Yingyong W, Kumsang Y, et al. (2018) Probing the Phylogenomics and Putative Pathogenicity Genes of *Pythium insidiosum* by Oomycete Genome Analyses. Sci Rep 8: 4135.

24. Shen D, Tang Z, Wang C, et al. 2019. Infection mechanisms and putative effector repertoire of the mosquito pathogenic oomycete *Pythium guiyangense* uncovered by genomic analysis. PLoS Genet. 15(4): e1008116.

25. Benhamou N, Rey P, Chérif M, Hockenhull J, Tirilly Y. 1997. Treatment with the Mycoparasite *Pythium oligandrum* Triggers Induction of Defense-Related Reactions in Tomato Roots When Challenged with *Fusarium oxysporum* f. sp. radicis-lycopersici. Phytopathology. 87(1): 108-22.

26. Paul B. 1999. *Pythium periplocum*, an aggressive mycoparasite of Botrytis cinerea causing the gray mould disease of grape-vine. FEMS Microbiol Lett. 181(2): 277-80.

27. Takenaka S, Nishio Z, Nakamura Y. 2003. Induction of Defense Reactions in Sugar Beet and Wheat by Treatment with Cell Wall Protein Fractions from the Mycoparasite *Pythium oligandrum*. Phytopathology. 93(10): 1228-32.

28. Baxter L, Tripathy S, Ishaque N, et al. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science. 330(6010): 1549-1551.

29. Kale SD, Gu B, Capelluto DG, et al. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell. 142(2): 284-95.

30. Kamoun S, Furzer O, Jones JD, et al. 2015. The Top 10 oomycete pathogens in molecular plant pathology. Mol Plant Pathol. 16(4): 413-434.

31. Wang Q, Han C, Ferreira AO, et al. 2011. Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. Plant Cell. 23(6): 2064-86.

32. Kanneganti TD, Huitema E, Cakir C, Kamoun S. 2006. Synergistic interactions of the plant cell death pathways induced by *Phytophthora infestans* Nepl-like protein PiNPP1.1 and INF1 elicitin. Mol Plant Microbe Interact. 19(8): 854-63.

33. Qi J, Wang J, Gong Z, Zhou JM. 2017. Apoplastic ROS signaling in plant immunity. Curr Opin Plant Biol. 38: 92-100.

34. Zheng X, McLellan H, Fraiture M, et al. 2014. Functionally redundant RXLR effectors from *Phytophthora infestans* act at different steps to suppress early flg22-triggered immunity. PLoS Pathog. 10(4): e1004057.

35. Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol. 10(6): 417-30.
36. Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Genet Dev. 35: 57-65.
37. Raffaele S, Farrer RA, Cano LM, et al. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science. 330(6010): 1540-3.
38. Benhamou N, le Floch G, Vallance J, Gerbore J, Grizard D, Rey P (2012) *Pythium oligandrum*: an example of opportunistic success. Microbiology 158: 2679-2694.
39. Kushwaha SK, Vetukuri RR, Grenville-Briggs LJ (2017) Draft Genome Sequence of the Mycoparasitic Oomycete *Pythium oligandrum* Strain CBS 530.74. Genome Announc 5 .
40. Jiang RH, de Bruijn I, Haas BJ, Belmonte R, Löbach L, et al. (2013) Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. PLoS Genet 9: e1003272.
41. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, et al. (2008) The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature 456: 239-244.
42. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, et al. (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306: 79-86.
43. Jacq C, Alt-Mörbe J, Andre B, Arnold W, Bahr A, et al. (1997) The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV. Nature 387: 75-78.
44. Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, et al. (2011) Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet 7: e1002230.
45. Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, et al. (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444: 97-101.
46. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, et al. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A 108: 9166-9171.
47. Van Kan JA, Stassen JH, Mosbach A, Van Der Lee TA, Faino L, et al. (2017) A gapless genome sequence of the fungus *Botrytis cinerea*. Mol Plant Pathol 18: 75-89.
48. C. elegans Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282: 2012-2018.
49. Opperman CH, Bird DM, Williamson VM, Rokhsar DS, Burke M, et al. (2008) Sequence and genetic map of *Meloidogyne hapla*: A compact nematode genome for plant parasitism. Proc Natl Acad Sci U S A 105: 14802-14807.
50. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, et al. (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277: 1453-1462.
51. Smits TH, Rezzonico F, Kamber T, Blom J, Goesmann A, et al. (2010) Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other Erwinia spp. Mol Plant Microbe Interact 23: 384-393.
52. Duan Y, Zhou L, Hall DG, Li W, Doddapaneni H, et al. (2009) Complete genome sequence of citrus huanglongbing bacterium, 'Candidatus Liberibacter asiaticus' obtained through metagenomics. Mol Plant Microbe Interact 22: 1011-1020.
53. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol. 340(4): 783-95.

770 54. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein
771 topology with a hidden Markov model: application to complete genomes. J Mol Biol. 305(3):
772 567-80.
773 55. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and
774 space complexity. BMC Bioinformatics. 5: 113.
775 56. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective
776 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):
777 268-74.
778 57. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH (2019) Evolview v3: a webserver for
779 visualization, annotation, and management of phylogenetic trees. Nucleic Acids Res 47:
780 W270-270W275.
781 58. Walsh I, Martin AJ, Di Domenico T, Tosatto SC (2012) ESpritz: accurate and fast prediction of
782 protein disorder. Bioinformatics 28: 503-509.
783 59. Wang Y, Tang H, Debarry JD, et al. 2012. MCScanX: a toolkit for detection and evolutionary
784 analysis of gene synteny and collinearity. Nucleic Acids Res. 40(7): e49.
785 60. Song T, Ma Z, Shen D, Li Q, Li W, et al. (2015) An Oomycete CRN Effector Reprograms
786 Expression of Plant *HSP* Genes by Targeting their Promoters. PLoS Pathog 11: e1005348.
787 61. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis
788 of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 11: 1650-1667.
789 62. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map
790 format and SAMtools. Bioinformatics 25: 2078-2079.
791 63. Saeed AI, Sharov V, White J, et al. 2003. TM4: a free, open-source system for microarray data
792 management and analysis. Biotechniques. 34(2): 374-8.
793
794

795 # Figure Legends

796 **Fig 1. *De novo* identification of RXLR effectors in *Pythium* species.** (A) The
797 RXLR effector identification pipeline. sORFs indicate open reading frames encoding
798 secretory proteins with a signal peptide (SP) but without transmembrane region (TM).
799 (B) Summary of predicted RXLRs. Phylogeny of the 26 species is based on data from
800 the Taxonomy Database and previous studies. The ratios of sORFs to whole ORFs,
801 the counts of RXLR candidates and the ratios of RXLRs to sORFs in each species are
802 showed in boxplots following the species names. The genome size of each species is
803 showed after the corresponding boxplots. (C) Weblogo of the RXLR-dEER motifs of
804 *Pythium* RXLR effectors.

805

806 **Fig 2. The evolutionary patterns of oomycetes RXLRs.** (A). The relatedness
807 network of putative RXLR effectors. Each spot represents an RXLR. Similar RXLRs

(*E*-value<1) are joined by edges. Spot colors indicate their genera. Three relatively independent clusters were marked as C1, C2 and C3, respectively. Almost all effectors in these clusters are from *Pythium*. Four RXLRs with effector activities verified in this study are labeled using arrows. (B) The maximum-likelihood phylogenetic tree of the RXLR-dEER motifs in RXLR effectors. Grey, yellow and red circles in the nodes indicate bootstrap values of 0–50, 51–80 and 80–100, respectively. Leaves of the tree are colored base on the families of their indicated species.

Fig 3. Evolutionary features of *Pythium* RXLRs. (A) Sequence divergence of *Pythium* RXLRs. *Pythium* RXLRs were compared against 300 randomly selected *Pythium* proteins regarding the sequence identities of their closest ortholog pairs within the *Pythium* genus. Radiuses range from 0 (center) to 80 (outer circle) represent 100% to 20% (or less) sequence identities sequentially. Proteins at the same identity level are randomly distributed along their corresponding circle. **(B) Disorder content in *Peronosporales* and *Pythiales* RXLRs.** The line plot shows the percentages of proteins containing different percentages of disordered regions. The bar plot shows total disorder percentages in *Peronosporales*, *Pythiales* and randomly selected proteins. RSP, Randomly selected proteins; PyRXLR, *Pythiales* RXLRs; PerRXLR, *Peronosporales* RXLRs. **(C) 5' and 3' end intergenic distances of *P. sojae* and *Pythium RXLRs*.** The two bar plots show 5' (left) and 3' (right) end intergenic distances of indicated *RXLR* genes, respectively. (*, $P<0.05$; **, $P<0.001$, Student's *t* test). Phso, *P. sojae*; Pyiw, *P. iwayamai*; Pyir, *P. irregulare*; Pyul, *P. ultimum*; Pyap, *P. aphanidermatum*; Pyar, *P. arrhenomanes*; Pyol, *P. oligandrum*; Pype, *P. periplocum*; Pygu, *P. guiyangense*; Pyin, *P. insidiosum*.

Fig 4. Synteny analysis and protein sequence alignments of clustered RXLRs in *P. guiyangense*. (A and B) Synteny analysis of clustered *P. guiyangense RXLRs*. Red blocks indicate clustered *PyguRXLRs* described in the text. Green blocks indicate genes flanking to these *PyguRXLRs*. Homologous genes with similar and opposite genome orientations are joined with black and edges, respectively. (C and D) Protein

838 sequence alignments of two RXLR clusters. Identical or similar residues are colored
839 in blue. The RXLR-dEER motifs are labeled in the alignments.
840
841 **Fig 5. Transcription patterns of *P. ultimum RXLRs* and the effector activity of**
842 **PyulRXLR22.** (A) Transcription patterns of *P. ultimum RXLR* candidates. The heat
843 map shows gene transcription patterns at the mycelia (MY) stage and stages of 3, 6,
844 12, 24 and 36 h post-inoculation of soybean leaves. Color bars represent log2 of gene
845 RPKM values, ranging from dark blue (0) to red (13). Black bars indicate no
846 expression detected. RXLRs selected for the effector activity assay are marked with
847 arrows with red arrow indicating ability of inducing cell death in *N. benthamiana*.
848 Based on transcript accumulation levels and patterns, *P. ultimum RXLRs* are divided
849 into 4 clusters using HCL methods and the MeV software. Clusters are marked with
850 circled numbers (1 to 4). (B) PyulRXLR22 induced cell death in *N. benthamiana*. *N.*
851 *benthamiana* leaves were transformed with the indicated constructs by
852 agro-infiltration. Dashed circles indicate infiltration sites and the ratios are sites with
853 necrotic lesion versus total infiltration sites. PyulRXLR3/4/10 did not induce cell
854 death. GFP and INF1 were used as negative and positive controls, respectively. (C)
855 Subcellular localization of PyulRXLRs. Green fluorescence indicates the localizations
856 of PyulRXLRs in *N. benthamiana* epidermal cells. Photographs were taken at 48 hpi.
857 Bar =20 μm.
858
859 **Fig 6. RXLRs from *P. oligandrum* and *P. periplocum* induce plant defense**
860 **response.** (A) Expression of PyolRXLR32, PypeRXLR22 or PypeRXLR24 induce
861 cell death in *N. benthamiana* leaves. *N. benthamiana* leaves were transformed with
862 the indicated constructs by agro-infiltration. Dashed circles indicate infiltration sites
863 and the ratios are sites with necrotic lesion versus total infiltration sites. INF1 and the
864 RXLR effector RXLR207 from *P. capsici* were used as positive controls. Empty
865 vector expressing GFP was used as a negative control. Same controls were also used
866 in the H_2O_2 accumulation assay. (B and C) *P. capsici* infection can be reduced by
867 PyolRXLR32 and PypeRXLR22, but not PypeRXLR24. 24 h after infiltration, equal

amount of mycelium was inoculated onto each infiltrated area expressing *PyolRXLR32*, PypeRXLR22, or PypeRXLR24. Photographs were taken under UV light at 48 hpi and lesion areas were measured at the same time. Lesion area data shown in the dot plots were collected from three biological replicates. At least five leaves were inoculated for each replicate (**, $P<0.001$, Student's *t* test). Bar=1cm. (D) H_2O_2 accumulation in *N. benthamiana* leaves expressing PyolRXLR32, PypeRXLR22 or PypeRXLR24. DAB staining was performed 48 hpi. Red dashed circles indicate infiltration sites. (E) H_2O_2 accumulation in *P. capsica*-inoculated *N. benthamiana* leaves expressing PyolRXLR32, PypeRXLR22, or PypeRXLR24. *P. capsica* was inoculated 24 h after infiltration. DAB staining was performed at 12 hpi. Black dashed circles indicate inoculation sites.

Supporting information

S1 Fig. Sequence alignment of *Pythium* RXLRs. Signal peptide and the RXLR-dEER motifs are labeled above their corresponding regions. Arg and Lys residues in the RXLR-dEER motifs are colored in blue. Glu and Asp residues are colored in red.

S2 Fig. Numbers of clustered *RXLR* effectors in each species. *RXLR* pairs with distances less than 40 kb in the same scaffold are showed. Hyar, *H. arabidopsidis*; Phra, *P. ramorum*; Phso, *P. sojae*; Phve, *P. vexans*; Pyiw, *P. iwayamai*; Pyir, *P. irregulare*; Pyul, *P. ultimum*; Pyap, *P. aphanidermatum*; Pyar, *P. arrhenomanes*; Pyol, *P. oligandrum*; Pype, *P. periplocum*; Pygu, *P. guiyangense*; Pyin, *P. insidiosum*; Sapa, *S parasitica*; Thps, *T. pseudonana*; Phtr, *P. tricornutum*.

S3 Fig. PyulRXLRs do not inhibit cell death triggered by INF1. *A. tumefaciens* harboring indicated constructs were infiltrated into *N. benthamiana* leaves 24 h before

893 the infiltrations of INF1. Photos were taken 4 days after infiltration.

894 **S4 Fig. Subcellular localization of PyolRXLR32, PypeRXLR22 and PyolRXLR24.**

895 Indicated constructs were expressed in *N. benthamiana* epidermal cells to show

896 protein localizations. Images were taken at 48 hpi. Bar=20μm.

897 **S5 Fig. Sequence alignment of PyolRXLR32 and its homologs.** The RXLR-dEER

898 motifs were labeled above their corresponding regions.

899 **S6 Fig. Sequence alignment of PyolRXLR21, 22 and 24**. The RXLR-dEER motifs

900 were labeled above their corresponding sites.

901

902 **S1 Table.** Overall counts of predicted RXLR effectors per species.

903 **S2 Table.** Known functional RXLR with degernerate dEER motif

904 **S3 Table.** Enrichment assay result comparing *B. cinerea* to the other species

905 **S4 Table.** Permutation test of RXLR motif in each species

906 **S5 Table.** Conserved RXLR effectors

907 **S6 Table.** RXLRs summary

908 **S7 Table.** Genomic information used in this study

909 **S8 Table.** Primers used in this study

910 **S9 Table.** RPKM value of *PyRXLRs* during different stages

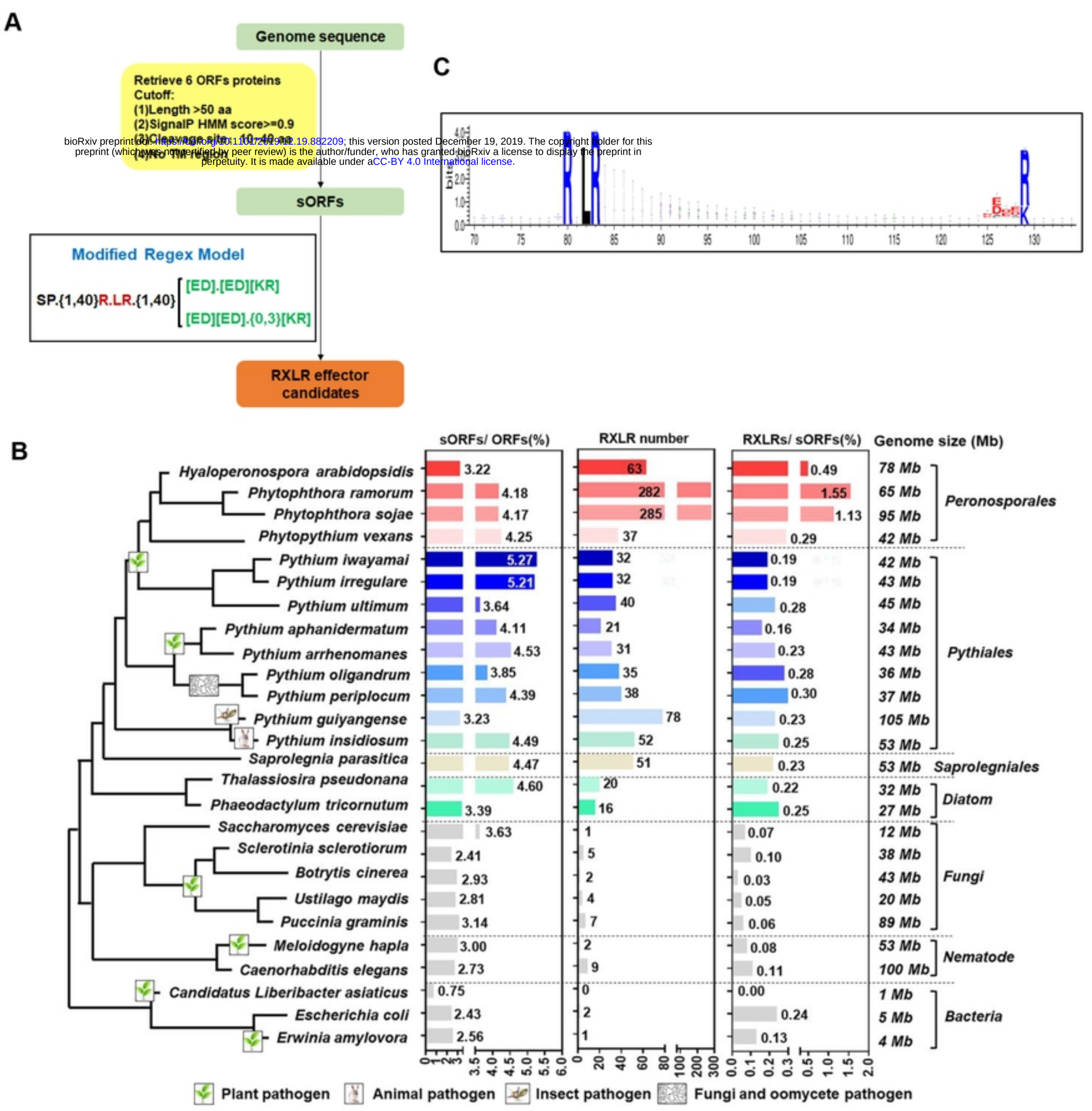


Figure1

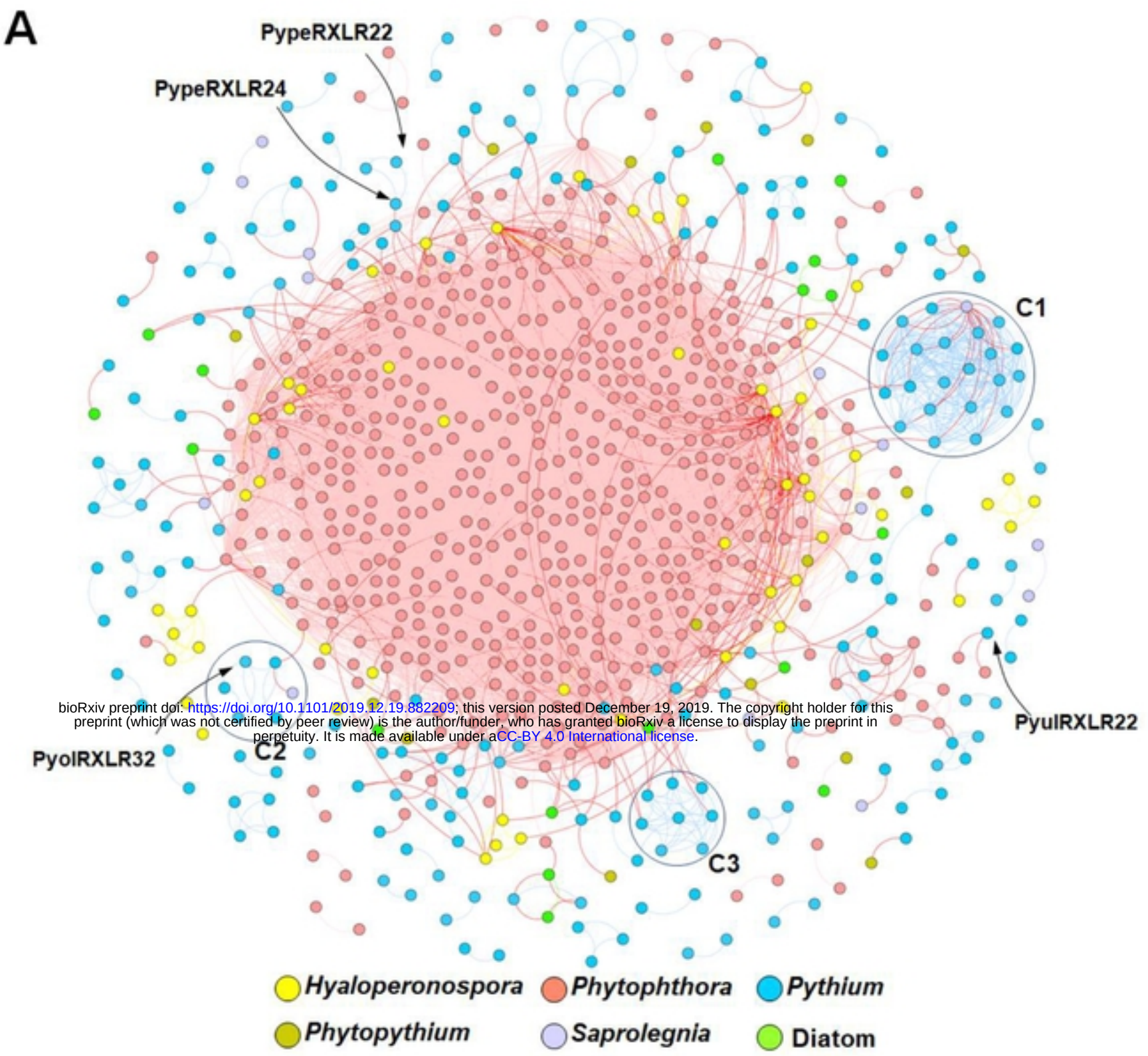
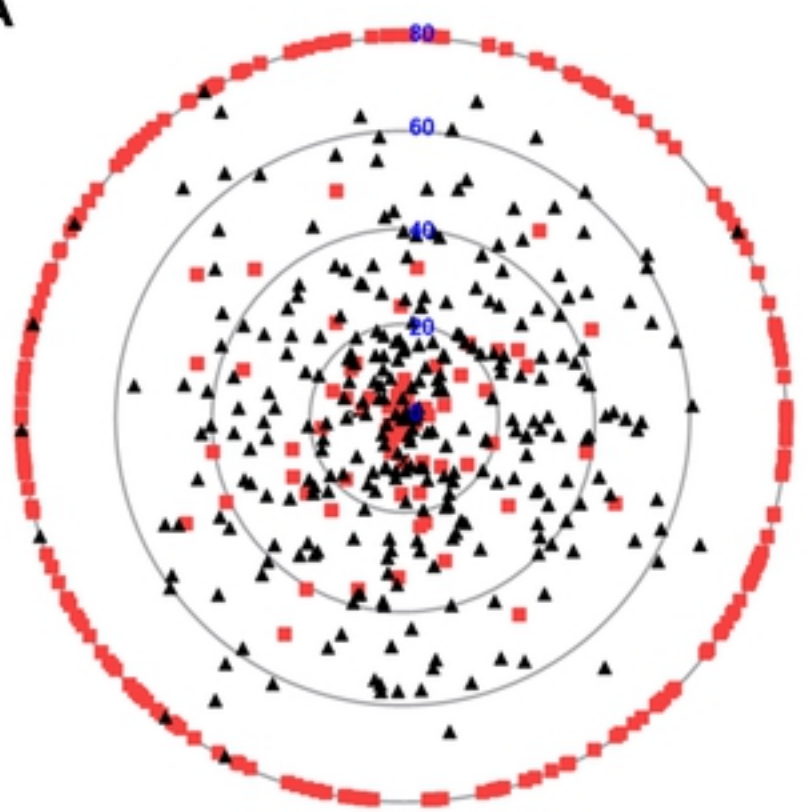
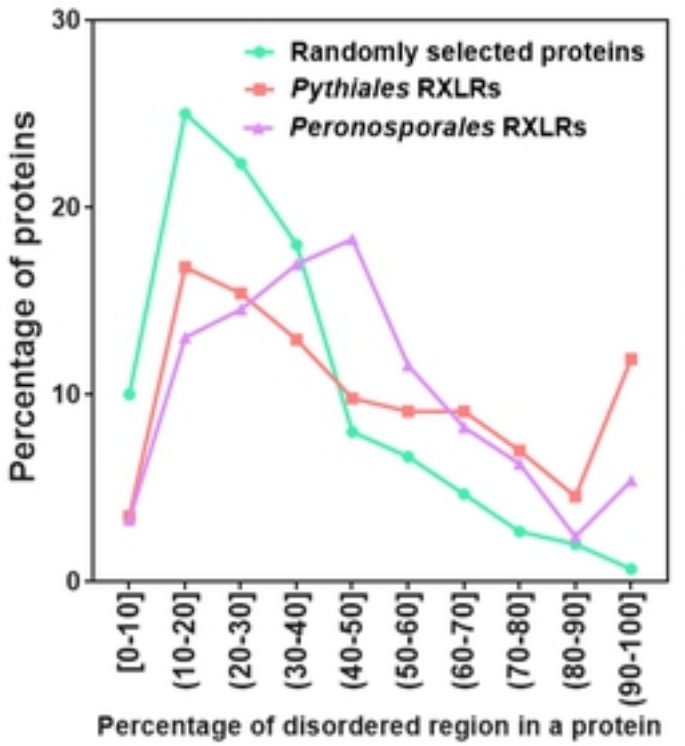


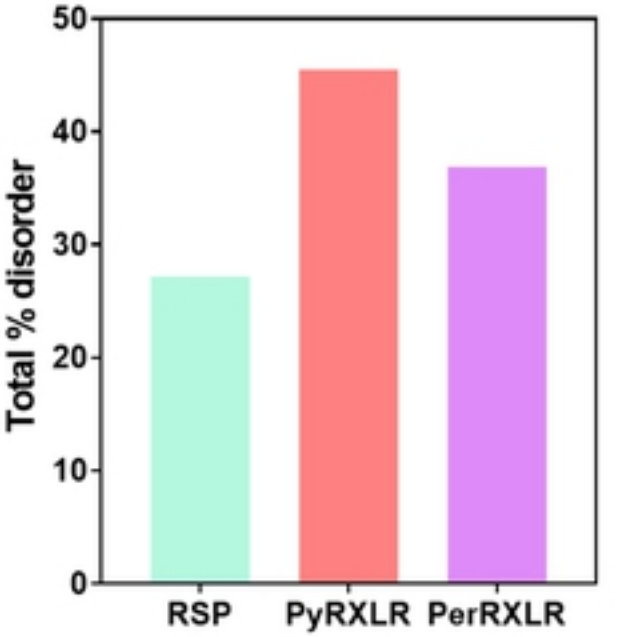
Figure2

A

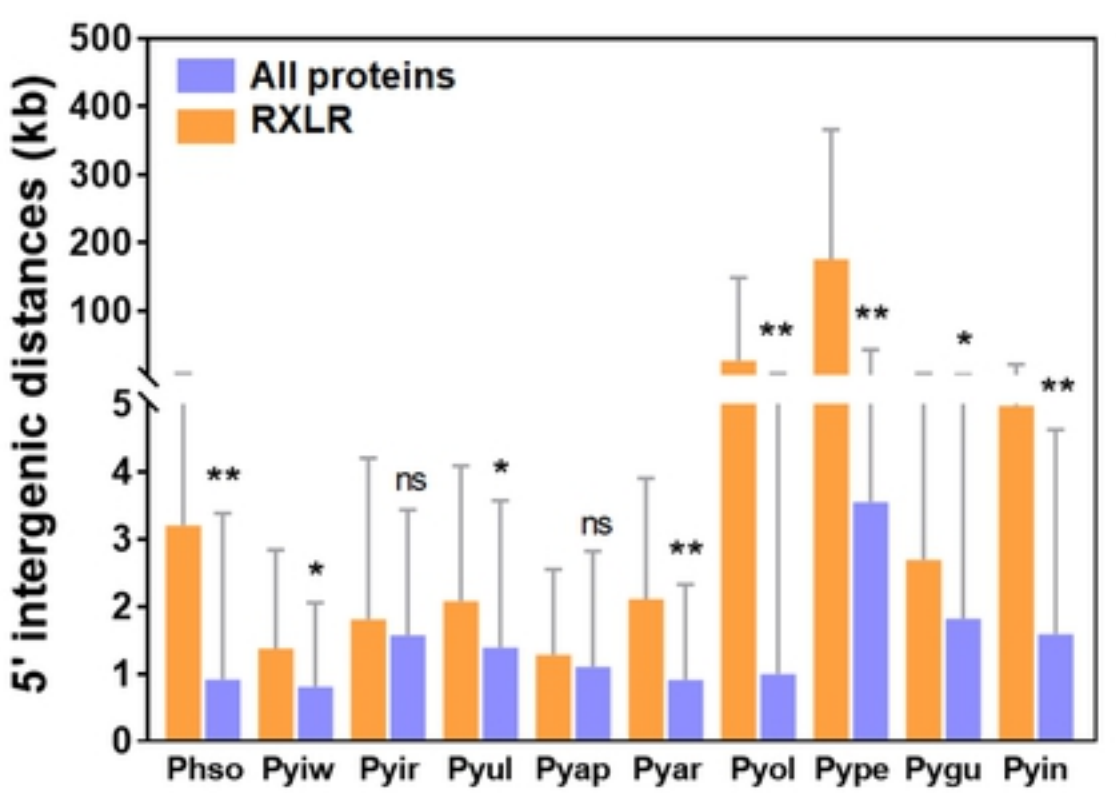

80
60
40
20
0
Pythium RXLR
Randomly selected Pythium proteins

B


Randomly selected proteins
Pythiales RXLRs
Peronosporales RXLRs
Percentage of proteins
30
20
10
0
[0-10]
(10-20]
(20-30]
(30-40]
(40-50]
(50-60]
(60-70]
(70-80]
(80-90]
(90-100]
Percentage of disordered region in a protein


Total % disorder
50
40
30
20
10
0
RSP
PyRXLR
PerRXLR

C


All proteins
RXLR
5' intergenic distances (kb)
500
400
300
200
100
5
4
3
2
1
0
**
*
ns
*
ns
**
**
**
*
**
Phso
Pyiw
Pyir
Pyul
Pyap
Pyar
Pyol
Pype
Pygu
Pyin

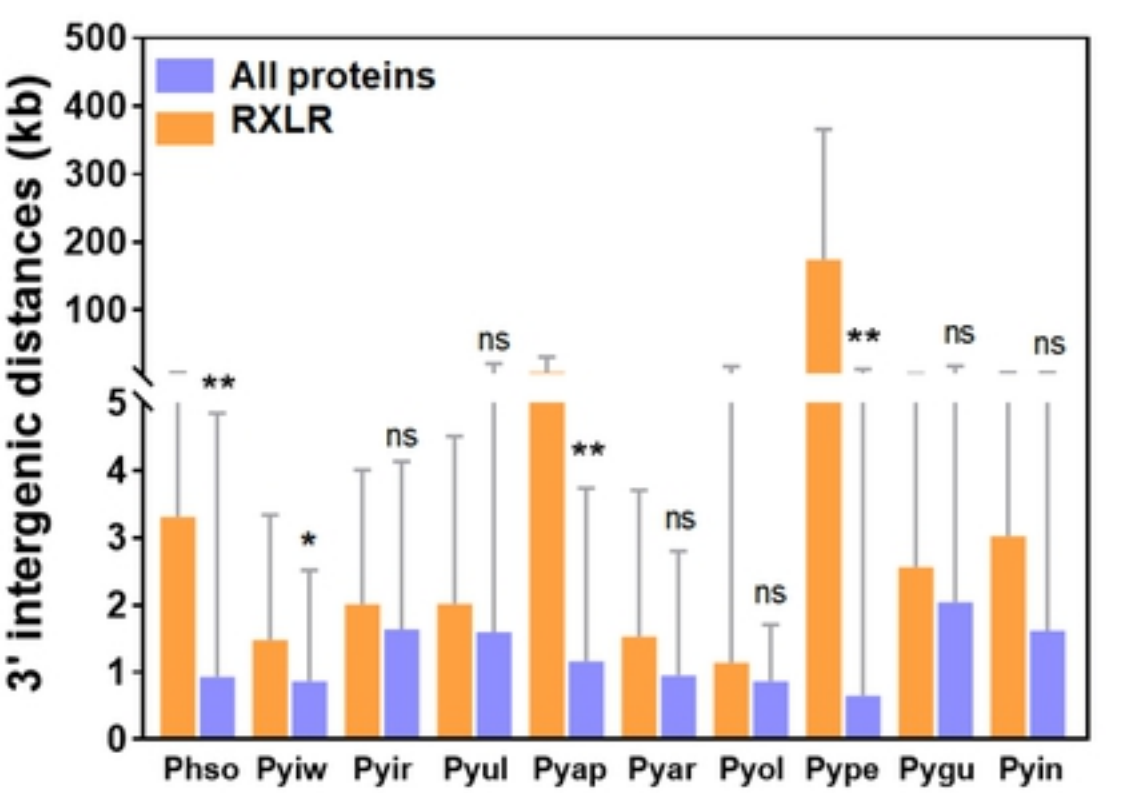

All proteins
RXLR
3' intergenic distances (kb)
500
400
300
200
100
5
4
3
2
1
0
**
*
ns
ns
ns
**
ns
ns
**
ns
ns
Phso
Pyiw
Pyir
Pyul
Pyap
Pyar
Pyol
Pype
Pygu
Pyin

Figure3

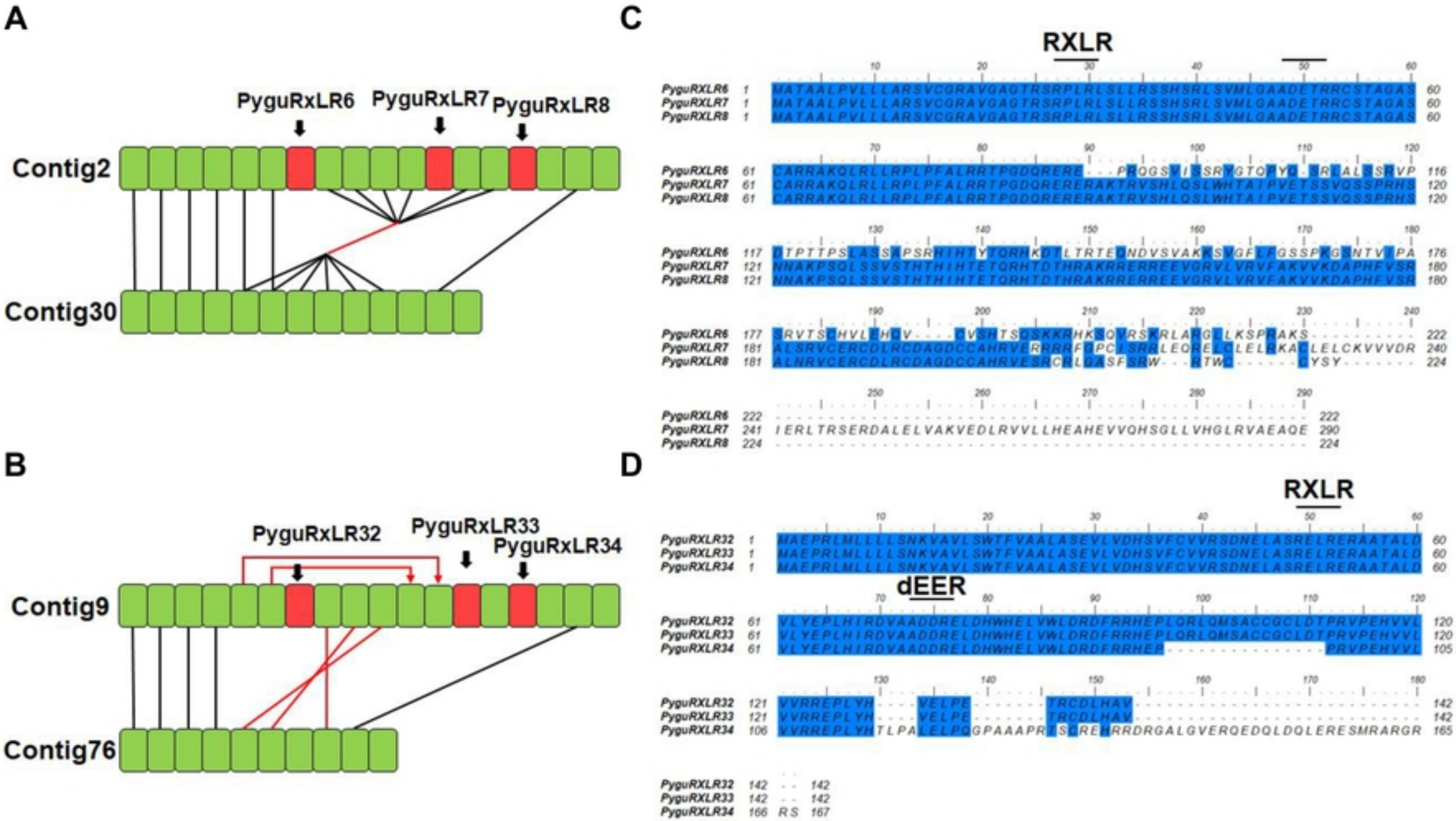
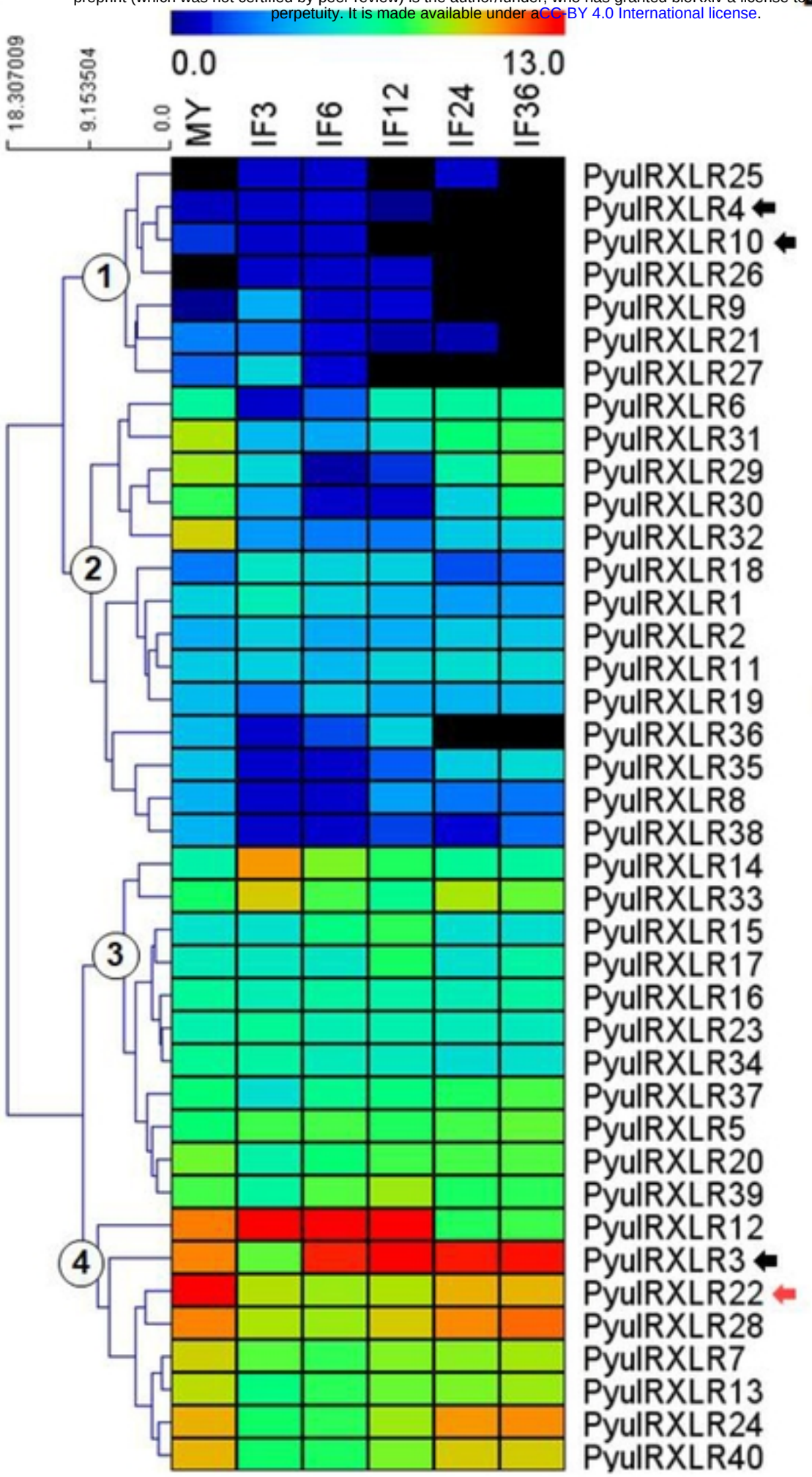
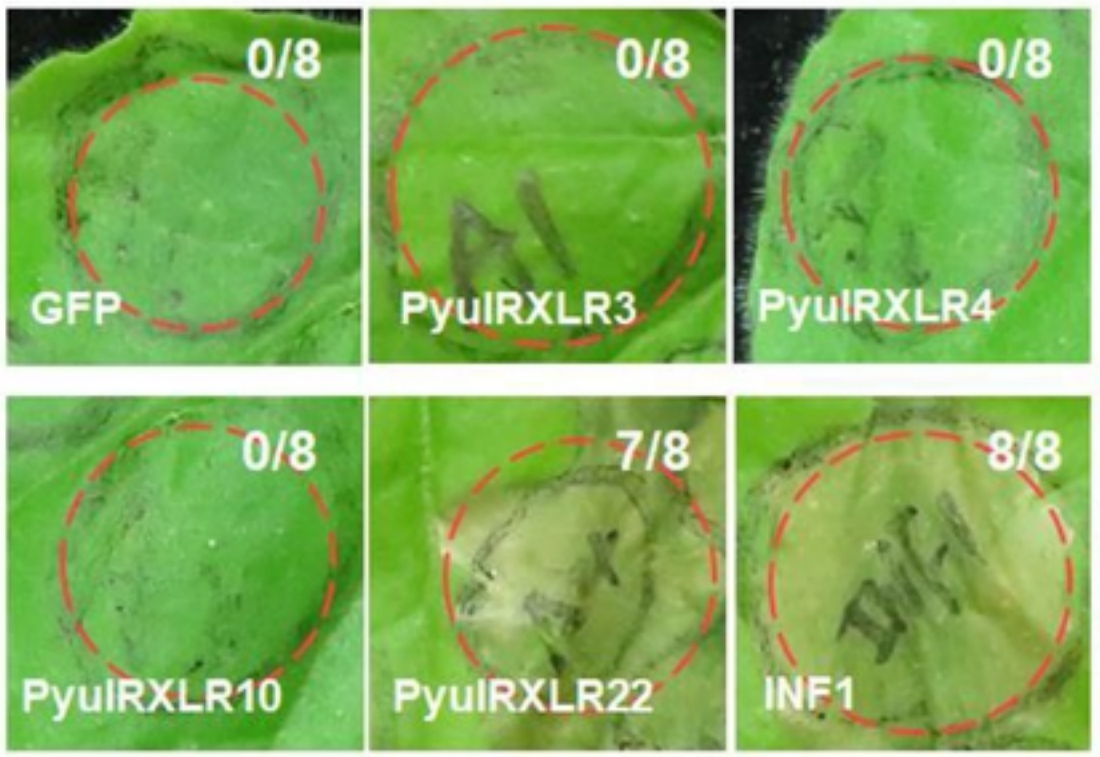

A
PyguRxLR6
PyguRxLR7
PyguRxLR8
Contig2
Contig30
B
PyguRxLR32
PyguRxLR33
PyguRxLR34
Contig9
Contig76
C
RXLR
PyguRXLR6
PyguRXLR7
PyguRXLR8
D
RXLR
dEER
PyguRXLR32
PyguRXLR33
PyguRXLR34

Figure4





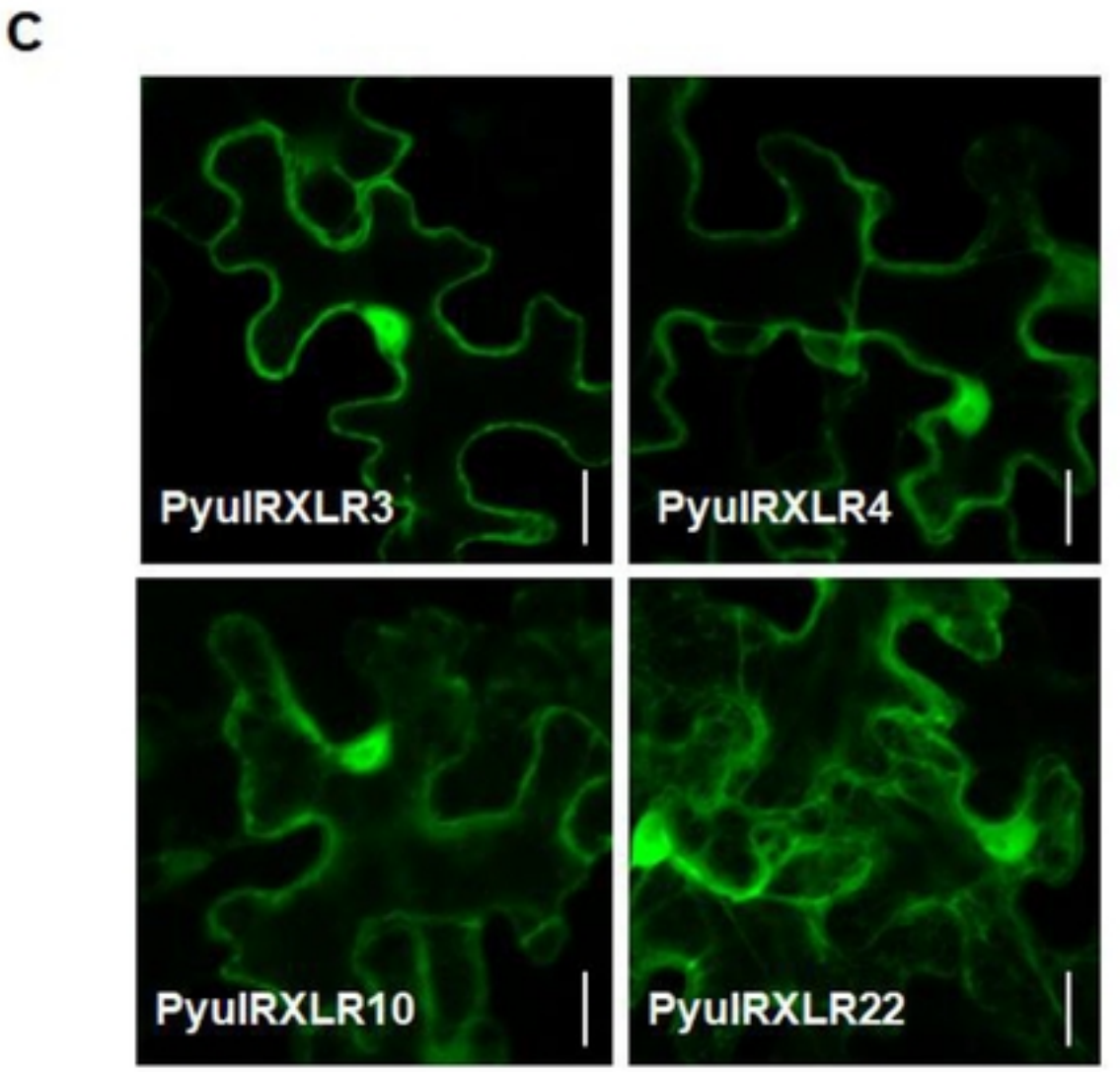


Figure5

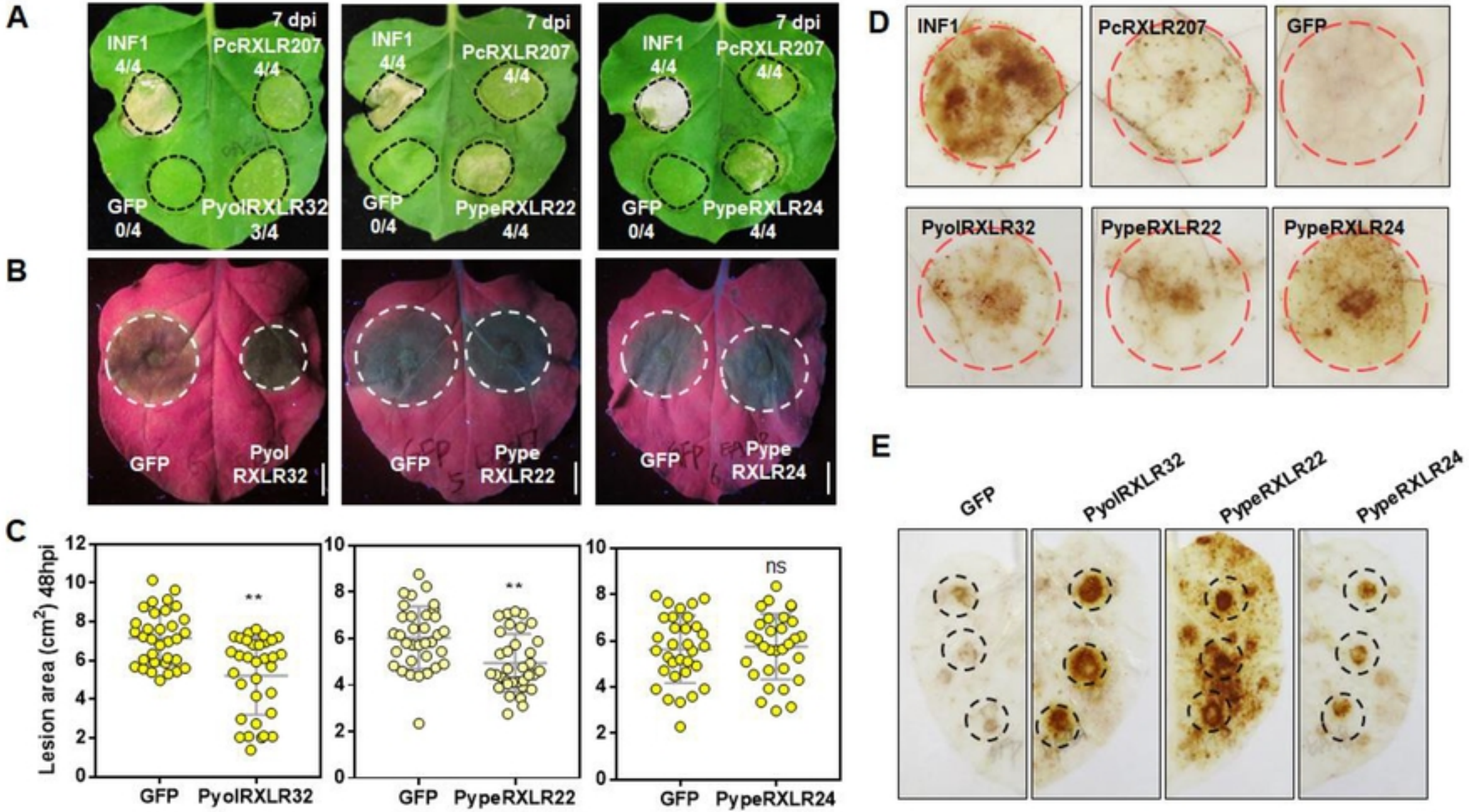

A
7 dpi
INF1
4/4
PcRXLR207
4/4
GFP
0/4
PyolRXLR32
3/4
7 dpi
INF1
4/4
PcRXLR207
4/4
GFP
0/4
PypeRXLR22
4/4
7 dpi
INF1
4/4
PcRXLR207
4/4
GFP
0/4
PypeRXLR24
4/4
B
GFP
Pyol
RXLR32
GFP
Pype
RXLR22
GFP
Pype
RXLR24
C
Lesion area (cm2) 48hpi
**
**
ns
GFP
PyolRXLR32
GFP
PypeRXLR22
GFP
PypeRXLR24
D
INF1
PcRXLR207
GFP
PyolRXLR32
PypeRXLR22
PypeRXLR24
E
GFP
PyolRXLR32
PypeRXLR22
PypeRXLR24

Figure6