

Alternative Splicing of a Multi-Drug Transporter from *Pseudoperonospora cubensis* Generates an RXLR Effector Protein That Elicits a Rapid Cell Death

Elizabeth A. Savory¹, Cheng Zou², Bishwo N. Adhikari², John P. Hamilton², C. Robin Buell², Shin-Han Shiu², Brad Day^{1*}

¹ Department of Plant Pathology, Michigan State University, East Lansing, Michigan, United States of America, ² Department of Plant Biology, Michigan State University, East Lansing, Michigan, United States of America

Abstract

Pseudoperonospora cubensis, an obligate oomycete pathogen, is the causal agent of cucurbit downy mildew, a foliar disease of global economic importance. Similar to other oomycete plant pathogens, *Ps. cubensis* has a suite of RXLR and RXLR-like effector proteins, which likely function as virulence or avirulence determinants during the course of host infection. Using *in silico* analyses, we identified 271 candidate effector proteins within the *Ps. cubensis* genome with variable RXLR motifs. In extending this analysis, we present the functional characterization of one *Ps. cubensis* effector protein, RXLR protein 1 (*PscRXLR1*), and its closest *Phytophthora infestans* ortholog, PITG_17484, a member of the Drug/Metabolite Transporter (DMT) superfamily. To assess if such effector-non-effector pairs are common among oomycete plant pathogens, we examined the relationship(s) among putative ortholog pairs in *Ps. cubensis* and *P. infestans*. Of 271 predicted *Ps. cubensis* effector proteins, only 109 (41%) had a putative ortholog in *P. infestans* and evolutionary rate analysis of these orthologs shows that they are evolving significantly faster than most other genes. We found that *PscRXLR1* was up-regulated during the early stages of infection of plants, and, moreover, that heterologous expression of *PscRXLR1* in *Nicotiana benthamiana* elicits a rapid necrosis. More interestingly, we also demonstrate that *PscRXLR1* arises as a product of alternative splicing, making this the first example of an alternative splicing event in plant pathogenic oomycetes transforming a non-effector gene to a functional effector protein. Taken together, these data suggest a role for *PscRXLR1* in pathogenicity, and, in total, our data provide a basis for comparative analysis of candidate effector proteins and their non-effector orthologs as a means of understanding function and evolutionary history of pathogen effectors.

Citation: Savory EA, Zou C, Adhikari BN, Hamilton JP, Buell CR, et al. (2012) Alternative Splicing of a Multi-Drug Transporter from *Pseudoperonospora cubensis* Generates an RXLR Effector Protein That Elicits a Rapid Cell Death. PLoS ONE 7(4): e34701. doi:10.1371/journal.pone.0034701

Editor: Ching-Hong Yang, University of Wisconsin-Milwaukee, United States of America

Received: January 16, 2012; **Accepted:** March 5, 2012; **Published:** April 5, 2012

Copyright: © 2012 Savory et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Work in the laboratory of BD is supported by a grant from the National Science Foundation (IOS-0641319). EAS is supported by a United States Department of Agriculture Agriculture and Food Research Initiative (AFRI) pre-doctoral fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: bday@msu.edu

Introduction

The identification and characterization of secreted effector proteins from plant pathogens has anchored the recent evolution of molecular plant pathology [1,2,3]. As components of many pathogenic microorganisms' secretomes, effector proteins represent a key component of phytopathogenicity, contributing to both the virulence and avirulence capacity of the invading pathogen. Numerous studies have identified and characterized the activities of secreted effector proteins from a broad range of phytopathogens [1,4,5]. Collectively, these works have revealed two primary functions for pathogen effector molecules. First, as virulence molecules, effector proteins can enhance a pathogen's ability to cause disease, likely through abrogating host processes that would otherwise block pathogen infection, growth, and proliferation within the host [5,6]. Secondly, as avirulence determinants, effector proteins function to modulate the activation of host defense responses by perturbing the activity of host resistance (R) proteins [5,6].

For infection, colonization, and subsequent propagation within their hosts to occur, pathogens must dampen multiple layers of plant defense responses. Often described as basal resistance, the initial perception and elicitation of defenses requires the recognition of pathogen associated molecular patterns (PAMPs; e.g., chitin, flagellin, LPS) [5,6], highly conserved molecules essential for the lifestyle and survival of the microorganism. The recognition of PAMPs, which are highly specific elicitors, occurs through receptors on the host membrane surface, and following initiation of this receptor-ligand interaction, a rapid first response known as PAMP-triggered immunity (PTI) is elicited [5,6]. Overall, the PTI response provides a basal level of resistance against a wide range of microorganisms, often utilizing conserved signaling pathways such as the up-regulation of the mitogen-activated protein kinase (MAPK) pathway, the generation of reactive oxygen species, and the induction of defense-related genes [5,6]. To overcome PTI, phytopathogens, including bacteria and oomycetes, rely on the delivery and activity of secreted effector proteins to abrogate this initial basal level of defense, as well as to

further promote virulence [3,5,6]. In response, pathogen effectors can be recognized by R (resistance) proteins, leading to the activation of effector-triggered immunity (ETI) [5,6] best illustrated as an amplified and sustained layer of defense. ETI is a robust response that is often associated with the activation of a specific type of programmed cell death referred to as the hypersensitive response (HR) [5,6]. Over time, as this cycle of subversion and recognition evolves, host specificity and subsequent interactions between pathogen and host are modulated by the interplay between the activity and recognition of secreted pathogen effector molecules and their host counterparts.

Oomycetes are a phylogenetically distinct eukaryotic lineage within the Stramenopiles, which as a group, are among the best-studied and most economically important plant pathogens. In recent years, the genomes of several agriculturally important oomycete pathogens have been sequenced, including *Phytophthora infestans*, *Phytophthora ramorum*, and *Phytophthora sojae*, the causal agents of late blight of potato and tomato, sudden oak death, and soybean root rot, respectively [7,8]. The genomes of two other oomycete pathogens, *Pythium ultimum*, which causes damping off and root rot on a wide range of hosts, and *Hyaloperonospora arabidopsidis*, a pathogen of *Arabidopsis thaliana*, have also been sequenced [9,10]. These investigations, through the analysis of genome content and structure, have provided a wealth of information, both towards understanding the nature of the host-pathogen interaction (e.g., host specificity, virulence strategies), as well as insight into the evolution of the interaction itself. Central to the analysis of phytopathogen genomes, the identification and characterization of oomycete effector proteins has moved swiftly into the forefront in the field of plant-pathogen interactions, due in large part to the aforementioned available genomic resources. At a primary level, the identification of a highly conserved N-terminal translocation motif (i.e., RXLR; *Arg-X-Leu-Arg*, where “X” is any amino acid) demonstrated to be necessary for effector delivery into host cells, has been a seminal discovery in the field of plant-oomycete interactions [11,12]. Similar in function to phytopathogenic bacterial effector proteins, oomycete RXLR-containing effectors have been demonstrated to suppress PTI [13], as well as to activate ETI [11,14,15,16,17,18]. Structurally, oomycete effector proteins display a modular organization, consisting of a N-terminal signal peptide, a conserved RXLR translocation motif, followed by a variable C-terminal effector domain [3]. It is the function and activity of the variable C-terminal effector domain that drives the activity of these molecules [3,4].

Alternative splicing (AS) of pre-mRNA drives the generation of multiple protein isoforms through assembly of different combinations of splice sites within a single gene. In total, this process represents a conserved mechanism found in eukaryotes which drives proteome complexity within organisms with a finite number of genes [19]. In oomycetes, there are few reports of intron processing [20,21,22], and to date, these analyses has been strictly *in silico* [20,22], with little functional validation [21]. Costanzo et al. [21] characterized alternative processing in *P. sojae* family 5 endoglucanases revealing the generation of both coding and non-coding RNA isoforms. Additionally, based on their large-scale analysis of intronic structure and alternative splicing in *P. sojae*, Shen and colleagues [22] validated splice variants leading to premature translation termination.

Ps. cubensis is an obligate biotrophic oomycete pathogen of cucurbits (i.e., cucumber, melon, squash, watermelon, etc.), and is the causal agent of cucurbit downy mildew, an economically important foliar disease [23]. Capable of rapid defoliation of fields in short periods of time (i.e., <10 days), *Ps. cubensis* is the primary factor limiting cucurbit production in the United States.

Despite obvious economic importance, very little is known about the genetic determinants of virulence and pathogenicity of *Ps. cubensis*, as well as the molecular-genetic basis of resistance in the cucurbits.

Similar to related oomycete pathogens of plants, *Ps. cubensis* possesses a suite of effector proteins that likely function to promote virulence and suppress host defense responses [3,24]. Recent work by Tian et al. [24] identified and characterized a preliminary set of effector proteins from a draft genome sequence of *Ps. cubensis* obtained using 454 pyrosequencing. In brief, this set of 61 candidate effectors included a large class of variants with sequence similarity to the canonical RXLR motif found in other oomycete plant pathogens [24]. Specifically, this work characterized the function of a QXLR-containing effector, designated *PcQNE*, which was shown to be a member of a large family of *Ps. cubensis* QXLR nuclear-localized effectors, up-regulated during infection of cucumber. Additionally, internalization of *PcQNE* was shown to require the QXLR-EER motif, thereby establishing a basic homology to the well-characterized *Phytophthora* spp. effector proteins.

In the current study, we describe the identification and evolutionary potential of the *Ps. cubensis* effector repertoire. First, through characterization of a RXLR effector protein, *PscRXLR1*, we investigated the localization and *in planta* activity, and similarly to some oomycete effector proteins described to date, *PscRXLR1* induces a rapid cell death response when delivered into plant cells. Additionally, using whole transcriptome sequencing analyses, as well as RT-PCR, we show that *PscRXLR1* is a product of alternative splicing of the *Psc_781.4* gene which encodes a putative multi-drug transporter. Coupled with the induction and expression of *PscRXLR1* mRNA during *Ps. cubensis* infection of cucumber, as well as a complement of bioinformatic, cell biology and *in vivo* analyses, we provide evidence suggestive of a virulence role for *PscRXLR1*. Finally, we used *PscRXLR1* as template for assessing the conservation and evolutionary potential of oomycete effector proteins from *Ps. cubensis*, identifying and analyzing orthologous pairs of *Ps. cubensis* effector proteins and *P. infestans* non-effector proteins. Using more robust methods, we identified additional candidate effectors from *Ps. cubensis* for these analyses and showed that, like other oomycete effectors, they tend to be influenced by positive selection. Assessment of evolutionary rate and conservation of secretion signals between orthologous pairs revealed that *Ps. cubensis* effectors are undergoing adaptive evolution and conservation of signal peptides are similar among effector and non-effector proteins in *Ps. cubensis*. Overall, our study provides support for the investigation of relationships among oomycete effectors and their non-effector orthologs, and in total, the analysis presented herein establishes a foundation for understanding the evolution of effector repertoires and host-pathogen specificity.

Results

Genome sequencing of *Ps. cubensis*

Next generation sequencing with the Illumina Genome Analyzer II platform was used to generate an assembly of the *Ps. cubensis* MSU-1 genome. A total of 4.5 Gb of cleaned paired end reads from two libraries were used to generate the assembly using Velvet, a *de novo* short read assembler [25]. The final assembly contains 35,546 contigs with an N50 contig size of 4.0 Kbp representing 64.4 Mbp. Protein coding genes in the draft assembly were annotated using MAKER [26] which incorporated *ab initio* gene predictions, protein evidence, and transcript evidence from other sequenced oomycete genomes. In total, 23,519 loci and 23,522 gene models were predicted.

Identification of the *Ps. cubensis* effector repertoire

Our initial analysis of the effector complement of *Ps. cubensis* in an earlier draft assembly [24] identified 61 sequences containing the conserved RXLR, or novel QXLR, motif found in known oomycete effector proteins. This number is significantly less than the effector count predicted for other plant pathogenic oomycetes (i.e., 563 effectors in *P. infestans*, 396 in *P. sojae*, 374 in *P. ramorum*, and 134 in *H. arabidopsidis*; [7,10,27]), and is likely the result of limited coverage generated from an initial 454 pyrosequencing [24]. Generation of genomic sequences using the Illumina Genome Analyzer platform and their subsequent assembly generated a more comprehensive dataset. Using this resource, 269 additional sequences were identified as putative effector proteins. Interestingly, the putative *Ps. cubensis* effectors showed more variation at the R1 position of the RXLR motif than previously shown [24], with 18 amino acids predicted at the R1 position, in addition to R and Q (Table S1). Moreover, we have evidence for expression for at least one predicted effector with any one of 19 amino acids (except Y, Tyr) at position R1, during the course of infection on a susceptible cucumber cultivar (Table S1; Savory et al., PLoS ONE, in press), supporting the hypothesis of an expanded translocation motif repertoire in *Ps. cubensis*. In total, including the previously characterized *PcQNE*, the current predicted effector complement of *Ps. cubensis* contains 271 members.

Nature of selection on *Ps. cubensis* paralogs

Based on comparative genomic analyses of several oomycete plant pathogens, positive selection has been postulated to act disproportionately on effectors gene[s] compared to other genes in the genome [27,28]. To this end, we examined the strength of selection acting upon the predicted effector complement of *Ps. cubensis* by estimating ω , the ratio of the non-synonymous substitution rate (Ka) to the synonymous substitution rate (Ks). Among all *Ps. cubensis* effector paralogs, the median ω is 0.54, which is significantly higher than that of *Ps. cubensis* paralogous genes in general ($\omega = 0.24$, Wilcoxon Rank Sum Test, $p < 2.2 \times 10^{-16}$). For comparison, we also examined *P. infestans* effectors and arrived at the same conclusion (Wilcoxon Rank Sum Test, $p < 4.0 \times 10^{-14}$). Because more recent duplicates tend to have elevated ω , we examined if the higher ω values among effector paralogs can be attributed to recent duplication. We found that, using Ks as a proxy of time, the ω values for effector pairs are in general significantly higher than other paralogs in *Ps. cubensis* in a Ks bin (Figure 1A). Thus, the elevated ω values among effectors are not exclusively due to relaxation of selection among recent duplicates. The results for *P. infestans* effectors are similar (Figure 1B), although the ω values of *P. infestans* are in general higher than those in *Ps. cubensis*.

Taken together, *Ps. cubensis* effectors either have experienced a significantly lower degree of selective constraints, or tend to be positively selected. Consistent with the latter, 6.3% of *Ps. cubensis* effector paralog pairs have $\omega > 1$, compared to 3.2% of all other paralogous gene pairs. In parallel to our observations in *Ps. cubensis*, 4.6% of *P. infestans* effector paralog pairs have $\omega > 1$, compared to 3.2% for all other paralogs. Although there is no clear evidence suggesting that most effectors are subjected to positive selection, it is interesting that even among relatively ancient effector duplicates, the rate of evolution among effectors is significantly higher than most genes. Given that older duplicates that survive for tens to hundreds of millions of years tend to be subjected to substantially stronger selective constraints than young duplicates [29], this would suggest that, perhaps, effectors function in a way that do not require as strong a constraint on their primary

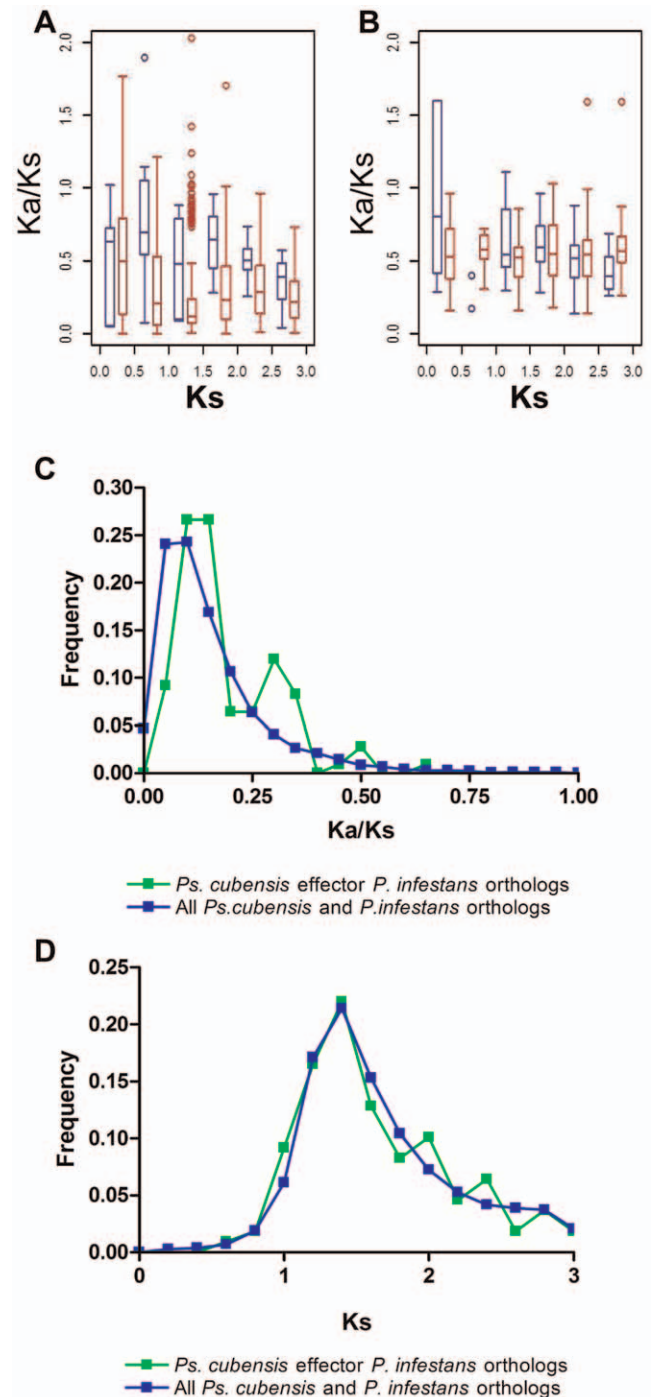


Figure 1. Strength of purifying selection on *Pseudoperonospora cubensis* effectors. Comparison of selective constraints on effector paralogs (blue) and all other genes (red) in (A) *Ps. cubensis* and (B) *P. infestans*. Frequency distributions (C) of ω , the ratio between the nonsynonymous substitution rate (Ka) and the synonymous substitution rate (Ks) of *Ps. cubensis* and *P. infestans* sequence pairs. Distributions of Ks values (D) of *Ps. cubensis* and *P. infestans* sequence pairs. Green symbols indicate *Ps. cubensis* effector – *P. infestans* non-effector ortholog pairs. Blue symbols represent other orthologous gene pairs between the two species.
doi:10.1371/journal.pone.0034701.g001

sequence. Alternatively, it is also possible that pathogen effectors, even those having undergone ancient duplication events, experience some degree of continuous positive selection.

Relationship between *Ps. cubensis* effectors and their *P. infestans* orthologs

Subsequent *in silico* analysis of candidate *Ps. cubensis* effectors and comparisons to annotated genes in *P. infestans* revealed that there were a number of orthologs between *Ps. cubensis* effector proteins and both effector and non-effector genes in *P. infestans*. We hypothesized that this scenario (i.e., effector with non-effector ortholog) may provide a foundation for the analysis of the evolution of effectors from non-effector proteins. Therefore, we identified orthologous pairs of predicted *Ps. cubensis* effector proteins and their non-effector counterparts in *P. infestans* considering sequence similarity and synteny (see Materials and Methods). With this approach, 11,601 orthologous gene pairs were identified between *Ps. cubensis* and *P. infestans* for comparison. Of 271 *Ps. cubensis* effector sequences, 109 had a predicted ortholog in *P. infestans* (Table S2). As shown in Figure 1C, the *Ps. cubensis* effector *P. infestans* (*PscE-Pi*) ortholog pairs have significantly higher ω values as compared to the baseline pairs (Kolmogorov-Smirnov test, $p < 7.9e-06$), consistent with what was found with the effector paralogs (Figure 1A). Additionally, the distribution of ω for the *PscE-Pi* pairs appears multi-modal. Given that the first effector ortholog peak (at $\omega \sim 0.15$) is mostly overlapping with that of the other orthologs, these effector paralogs are more highly conserved. The second peak at $\omega \sim 0.3$ likely indicates the presence of a group of effectors that are more quickly evolving (Figure 1C). However, we cannot rule out the possibility that these peaks are present simply due to the small effector ortholog sample size.

To determine if the overall higher ω value among effector orthologs is an artifact due to mis-identification of orthologous genes, we examined if putative effectors, as well as the other orthologs, have similar “age”. As shown in Figure 1D, the distributions of Ks values for the effector and the other orthologs are highly similar and are statistically indistinguishable. Thus, mis-identified orthologous pairs likely do not significantly impact our findings.

Signal peptide conservation among ortholog pairs

Signal peptides are essential components of oomycete effector proteins, as they are required for translocation of the protein from the pathogen haustorium to the extrahaustorial matrix prior to uptake by the host cell membrane [3]. As such, all 109 of the *Ps. cubensis* effector sequences in the *PscE-Pi* dataset are predicted to have signal peptides (Figure S1). However, only 71 (65%) of the corresponding *P. infestans* orthologs were predicted to be secreted proteins. For comparison, predictions of signal peptides were made for 10,383 of the 11,601 ortholog pairs. Of these, there were 688 (6.63%; *Psc-sec/Pi-sec*) ortholog pairs where both members were predicted to have signal peptides, 428 (4.12%; *Psc-sec/Pi-non*) pairs where the *Ps. cubensis* protein was predicted to be secreted and the *P. infestans* ortholog was not, and 622 (5.99%; *Psc-non/Pi-sec*) where the *Ps. cubensis* sequence did not have a predicted signal peptide and its corresponding *P. infestans* sequence was predicted to be secreted. Additionally, there were 8,645 (83.3%; *Psc-non/Pi-non*) ortholog pairs where neither member was predicted to be secreted. For statistical analysis, the *Psc-sec/Pi-sec* and *Psc-sec/Pi-non* datasets from the *Ps. cubensis* effector-*P. infestans* orthologs were compared to their respective genome-wide datasets. Using the Fisher’s exact test, no significant difference ($p = 0.5354$) was identified between the two datasets, indicating that presence of signal peptide prediction is not a good indicator of potential selection for effector peptide evolution.

Identification of *Pseudoperonospora cubensis* effector *PscRXLR1*

Using the RXLR effector identification pipeline [28], we previously identified 61 candidate effector protein sequences from a draft genome assembly of *Ps. cubensis* [24]. Initial analysis of these sequences using the Basic Alignment Analysis Search Tool (BLAST) against the proteome of *P. infestans* revealed that only 7 of these sequences aligned with annotated proteins within the *P. infestans* genome database; moreover, only 1 of these was predicted to be a secreted RXLR effector. Of these sequences, one (contig01871_F1) had 75% amino acid identity to *P. infestans* protein PITG_17484, a putative member of the drug/metabolite transporter (DMT) superfamily (CLO184; Figure 2). Additional cloning via 3’ RACE PCR and subsequent analysis revealed that the *Ps. cubensis* coding sequence, hereafter referred to as *Ps. cubensis* RXLR protein 1 (*PscRXLR1*), appeared significantly shorter (i.e., 132 amino acids), compared to its *P. infestans* ortholog PITG_1784 (i.e., 332 amino acids). This apparent truncation in *PscRXLR1* results in a protein coding sequence lacking the EamA functional domain (PF00892; formerly called DUF6) found in members of the DMT family [30].

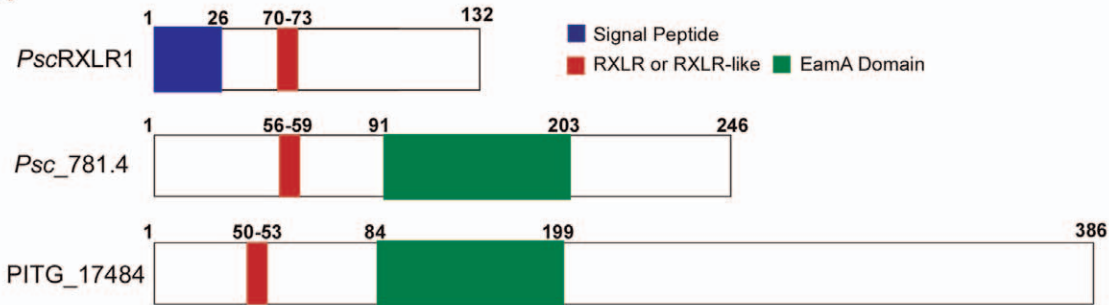
SignalP analysis of *PscRXLR1* identified a putative 26 amino acid signal peptide at the N-terminus of the protein (Figure 2). Based on the conserved features and domain organization of oomycete effectors, the presence of a canonical RXLR motif was identified at amino acid position 70 (Figure 2). However, unlike several previously characterized oomycete effector proteins, *PscRXLR1* does not contain an EER motif, which has also been implicated in oomycete effector translocation into the host cell [12,24,27]. The *P. infestans* ortholog, PITG_17848, while not having a predicted signal peptide, does contain an RXLR-like motif (i.e., RFMR; Figure 2A). To eliminate the possibility that PITG_17848 was mis-annotated and did in fact contain a signal peptide upstream of the predicted coding sequence, the region 500 bp upstream of the annotated PITG_17848 sequence was examined and a canonical signal peptide sequence was not identified. We therefore concluded that PITG_17848 is not an RXLR effector protein.

The absence of a predicted signal peptide in PITG_17848 suggests that *PscRXLR1* may have evolved this function independently. To address this possibility, and to further explore the ancestral function of these proteins, orthologous *PscRXLR1* sequences in additional plant oomycete pathogen species were identified, including *P. sojae*, *P. ramorum*, and *Py. ultimum*. Not surprisingly, the sequences from *P. sojae* and *P. ramorum* were more similar to *PscRXLR1* than those from *Py. ultimum* (76% and 72%, respectively, compared to 59%; Figure S2). Additionally, while none of these orthologs had predicted signal peptides, they did contain EamA functional domains, indicating that they were also members of the DMT superfamily (Figure S2). Phylogenetic analysis to infer evolutionary relationships between *PscRXLR1* and orthologs from *P. sojae*, *P. ramorum*, and *Py. ultimum* supported these observations (Figure 2C).

Functional validation of the *PscRXLR1* signal peptide

A primary characteristic of oomycete effector proteins is signal peptide-mediated secretion from the haustoria into the extrahaustorial matrix (EHM) prior to translocation into the host cell [3]. *PscRXLR1* contains a 26 amino acid signal peptide as predicted by SignalP 3.0 ([31]; HMM Probability, 0.966), whereas its closest *P. infestans* ortholog, PITG_17484, does not have a predicted signal peptide. To determine if the predicted signal peptide from *PscRXLR1* is functional, we used a yeast signal trap assay based on the requirement of invertase secretion for yeast

A



B

PscRXLR1 1 MVWLQLKKSGLGFTMSLSAVYGAVYAAANSVPAGKIDSGKKAMRHLENLPLLLASDSLES
Psc_781.4 1 -----MSLSAVYGAVYAAANSVPAGKIDSGKKAMRHLENLPLLLASDSLES
PITG_17484 1 -----MELPAVYGAVYASATSSAP----AGKKAIRHLENLPLLVASDSLES

PscRXLR1 61 VSTEGKWLRFRLRQAIMRSIANRIAGIILVSTSAFLASCIATLVKEDAVKLAPVEILFWR
Psc_781.4 47 VSTEGKWLRFRLRQAIMRSIANRIAGIILVSTSAFLASCIATLVKEDAVKLAPVEILFWR
PITG_17484 43 MTEGKWLRFMRSAVVRSVANRIAGIILVGTSAFLASCIATLVKEDTIKLSAIEALFWR

PscRXLR1 121 SLVSWLLTLVSS*-----
Psc_781.4 107 SLVSWLLTLVAITTTGVKTRIKKENYRPIVLRSETGCIATTLTIIMLQELAVSNALATY
PITG_17484 103 SLVSWELTVAAIATTTSTKMRVKKENRPIITLRCVFGCISTTLTIQVLEKLAVSNAATY

PscRXLR1 -----
Psc_781.4 167 FSPLLAFAFAARFLKEKPKLFAVACSVMCVIGAVLVVRRPVFLFGKSGSTDAWYRSMTS
PITG_17484 163 VSPLLAFAFAATFLKEKPGVETAVCSALCVAGAVLVVRRPAFLFGKSGSTDAKWYRSMAS

PscRXLR1 -----
Psc_781.4 227 FVTSYLFEGSLAIGCATIVVFMQAGAYVSLRSLQKVPHLVVMHYIYVITTLVSLP*----
PITG_17484 223 FVTSNLFEGSLAVGCAVVVAFMQAGAYVSLRSLHKVEYLVVMQYIYETMTLAAALAMLGI

PscRXLR1 -----
Psc_781.4 -----
PITG_17484 283 QHGKFKAGTSLETWGAIVGTGALAFVEQLFLTRGFQFDGAGVLAATRLLHVSCEFAWGV

PscRXLR1 -----
Psc_781.4 -----
PITG_17484 343 LLGTALNPWSAGGAGVTAAGVLFALRRVHTHWAARRSLRRILQ*

C

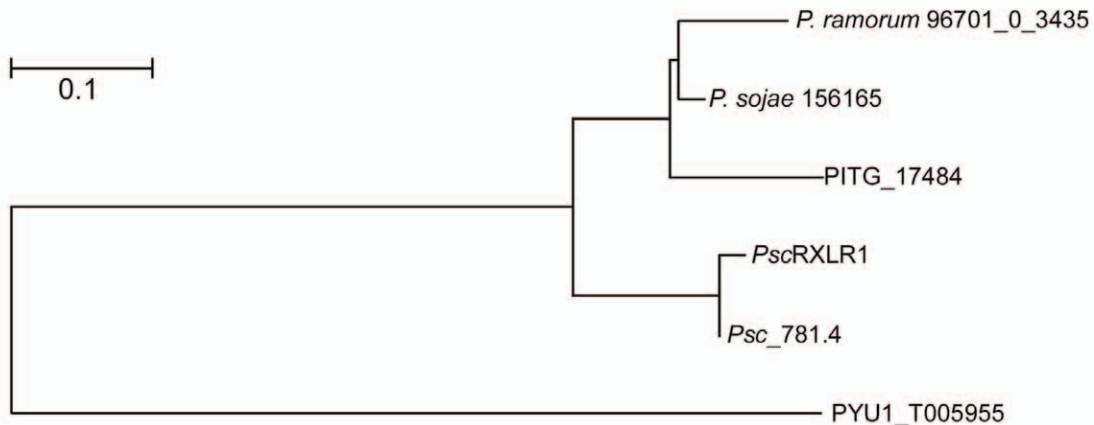


Figure 2. *PscRXLR1* encodes a RXLR-containing effector protein with homology to a non-effector protein in *Phytophthora infestans*. (A) Schematic representations of *PscRXLR1*, *Psc_781.4* and PITG_17484 from *P. infestans*. (B) Multiple sequence alignment of *PscRXLR1*, *Psc_781.4* and PITG_17484. Protein sequences were aligned using CLUSTALW and shaded for consensus using BOXSHADE. The asterisk indicates a stop codon. (C) The full length protein sequences of *PscRXLR1*, *Psc781.4* and their orthologs from *P. infestans* (PITG_17484), *Phytophthora ramorum* (*P. ramorum* 96701 0 3435), *Phytophthora sojae* (*P. sojae* 156165) and *Pythium ultimum* (PYU_T005955) were aligned using Muscle and evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [1] using MEGA5 [2]. 500 bootstrap runs were performed. doi:10.1371/journal.pone.0034701.g002

growth on media with raffinose as the sole carbon source [32]. This assay has been used previously to confirm predicted signal peptide sequences in candidate effector proteins from both *P. infestans* and *Ps. cubensis* [24,33]. As shown in Figure 3A, pSUC2-*PscRXLR1* (column 4) is able to grow on medium containing only raffinose (YPRAA), indicating that the signal peptide of *PscRXLR1* is sufficient for secretion of invertase. As a second confirmation of signal peptide function, we monitored the reduction of 2,3,5-triphenyltetrazolium chloride (TTC) to the red-colored compound triphenylformazan. Again, pSUC2-*PscRXLR1* (column 4) was confirmed as having a functional signal peptide. In contrast, neither the control yeast strains (i.e., YPK12 [column 1] or pSUC2 [column 2]), nor the pSUC2-PITG_17484 (column 3) yeast strain containing a PITG_17484-invertase fusion construct, grew on YPRAA, nor were they able to reduce TTC. Our positive control, PcQNE-SP1 (5) was, as shown previously [24], both able to grow on YPRAA medium and reduce TTC. These data support the annotation of *PscRXLR1* as a secreted RXLR effector protein and confirm that its *P. infestans* ortholog is a non-secreted protein.

PscRXLR1 and PITG_1784 localize to the plant plasma membrane

In planta localization of effector proteins has been successfully used to guide functional analysis and to infer the function itself [34,35,36,37]. To identify a possible role for *PscRXLR1* in the pathogenicity of *Ps. cubensis*, and to provide clues as to its putative role *in planta*, we investigated the localization of *PscRXLR1*. To this end, a C-terminal CFP-fusion protein (i.e., *PscRXLR1*-CFP) was transiently expressed in *Nicotiana benthamiana*, and protein localization visualized using laser scanning confocal microscopy. Based on the similarity of *PscRXLR1* with members of the DMT superfamily in *Phytophthora* spp. and *Py. ultimum*, *PscRXLR1* was predicted to localize to the plant plasma membrane, despite the absence of a predicted transmembrane domain when analyzed using InterProScan. To confirm this, *PscRXLR1*-CFP was co-expressed with a YFP-tagged construct encoding the aquaporin gene *AtPIP2A*, a marker for plasma membrane localization [38]. As predicted, *PscRXLR1*-CFP co-localized with *AtPIP2A*-YFP (Figure 3B), confirming that *PscRXLR1* localizes to the plasma membrane *in planta*. Additionally, a C-terminal CFP fusion was made to the *P. infestans* ortholog PITG_17484, which was also confirmed to be plasma membrane localized (Figure 3B).

PscRXLR1, but not its *P. infestans* ortholog, elicits a rapid cell death response when expressed in *Nicotiana benthamiana*

The obligate nature of a plant pathogen often presents challenges towards functional characterization of virulence and susceptibility within their respective host(s). To circumvent this limitation, transient heterologous systems have been developed and have proved valuable in their functional characterization [34,39,40,41,42]. To investigate the activity of *PscRXLR1* *in planta*, we utilized heterologous expression in *N. benthamiana* as means to characterize and determine the function of *PscRXLR1*. As shown in Figure 3C, expression of *PscRXLR1* resulted in leaf

chlorosis throughout the entire infiltration zone by 2 dpi, followed by browning and initiation of necrosis at 4 dpi. By 7 dpi, the zone of infiltration was completely dehydrated. In comparison, neither infiltration with PITG_17484, nor pMDGFP, resulted in any detectable cell death-type phenotype in *N. benthamiana* leaves at 4 dpi (Figure 3D; Figure S3).

PscRXLR1 expression is induced during *Ps. cubensis* infection of cucumber

The function of pathogen effector molecules is to enhance the virulence of the pathogen during its lifecycle, as well as to dampen host defense responses activated during infection. In this regard, the temporal expression of effector molecules during infection and pathogen development often signals critical stages in the host-pathogen interaction. Expression of *PscRXLR1* mRNA was measured using quantitative reverse transcription (qRT)-PCR following infection of *Ps. cubensis* on the susceptible cucumber cultivar 'Vlaspik'. As shown in Figure 4, expression of *PscRXLR1* was significantly ($p < 0.001$) induced during infection, beginning at 1 dpi and continuing through 4 dpi, as compared to the basal expression level in sporangia. Induction of gene expression at 1 dpi corresponds with zoospore encystment in the stomata, the first stage of pathogen entry into the host (Figure 4B, left panel). Subsequent expression observed through 4 dpi corresponds with hyphal penetration through the stomata, growth throughout the mesophyll, and formation of haustoria (Figure 4B, center and right panels). This pattern of expression supports a potential role for *PscRXLR1* in initial establishment of the infection possibly through dampening host defense responses. Additionally, this pattern is consistent with the expression patterns observed in other oomycete plant pathogen effectors, further supporting the prediction of *PscRXLR1* as an effector protein with a role in infection and disease development.

PscRXLR1 is a splice variant of *Psc_781.4*

Automated annotation of the Illumina-generated *Ps. cubensis* assembly described in this study resulted in *Psc_781.4*, a gene model at the *PscRXLR1* locus that more closely mirrored PITG_17484 than our prediction for *PscRXLR1* and what was obtained through molecular cloning (Figure 5), with the primary difference between the two predictions being that intron 1 is either spliced in *Psc_781.4*, or retained in *PscRXLR1* (Figure 5A). Empirical whole transcriptome sequence data (RNA-seq) from *Ps. cubensis* sporangia (unpublished results) provides support for both isoforms at this locus. When the first intron is retained, a stop codon is also brought into frame, yielding a truncated transcript (i.e., putative effector *PscRXLR1*), and subsequently, a smaller protein, which, as described in Figure 2, lacks the EamA functional domain (Figure 5A).

Based on our *in silico* predictions we confirmed which gene model, or both, was represented *in vivo*. Using an RT-PCR-based approach, we were able to amplify both splice variants from purified sporangia (SP), as well as from infected leaf material harvested at 4 and 8 days post-inoculation (dpi) (Figure 5B), suggesting that both isoforms are present throughout the infection process. As an added control, the resultant PCR products were

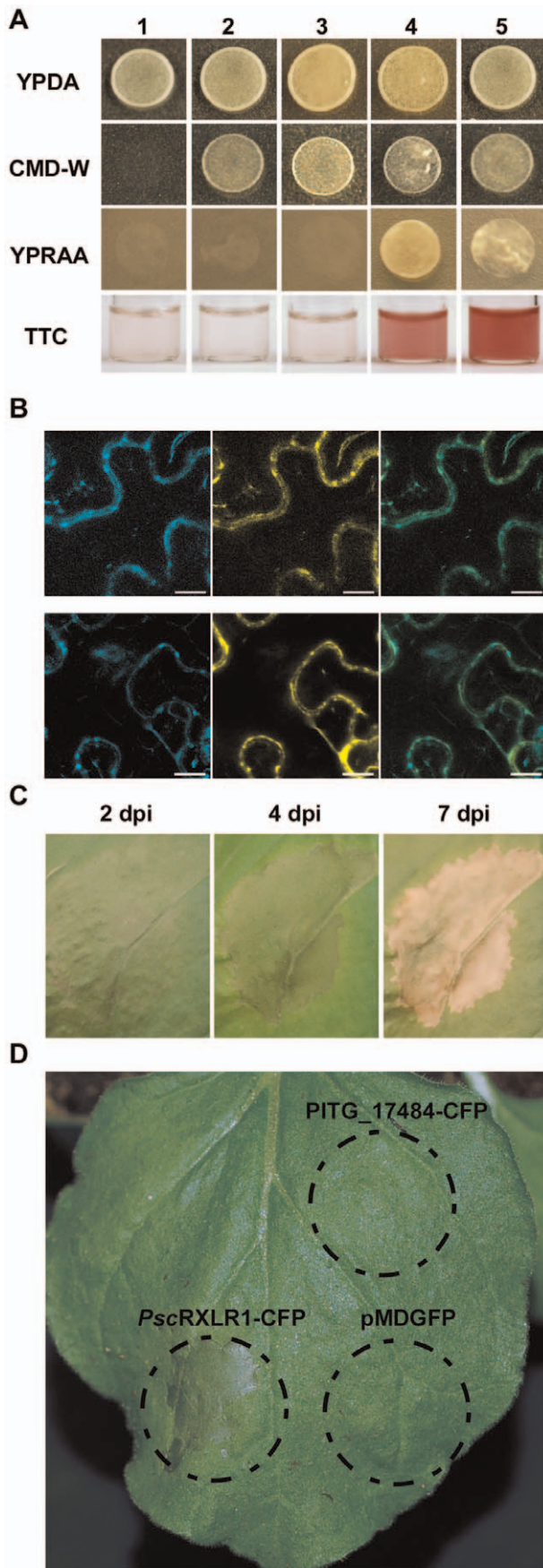


Figure 3. Functional characterization of *PscRXLR1* and *PITG_17484*. (A) *PscRXLR1* has a functional signal peptide validated by the yeast signal trap assay. Yeast strains were grown on media with raffinose (YPRAA) as the sole carbon source or in the presence of 2,3,5-triphenyltetrazolium chloride (TTC). Yeast strains YTK12 and YTK12:pSUC2 EV both lack a functional invertase gene and cannot grow on YPRAA medium or reduce TTC to red formazan. The functional signal peptide of *PscRXLR1*, when fused in-frame to the mature yeast invertase (pSUC2-*PscRXLR1*), allows for secretion of invertase, resulting in growth on YPRAA medium, as well as reduction of TTC to red formazan. *PITG_17484*, as predicted, does not have a functional signal peptide (pSUC2-*PITG_17484*). (B) Both *PscRXLR1_CFP* (top row) and *PITG_17484_CFP* (bottom row) co-localize with a plasma membrane-specific *AtPIP2A_YFP* marker. Left panels: C-CFP fusion protein only. Center panels: *AtPIP2A_YFP*. Right panels: Merge of CFP and YFP images. Scale bar = 10 μ m. (C, D) Heterologous expression of *PscRXLR1* induces chlorosis and necrosis in *Nicotiana benthamiana*. *Agrobacterium tumefaciens* C58-C1 expressing *PscRXLR1_CFP*, *PITG_17484_CFP*, or an *AtPIP2A_YFP* construct were infiltrated into *N. benthamiana*. The chlorosis and necrosis phenotype of *PscRXLR1_CFP* infiltrated leaves (C) was photographed at 2, 3, and 7 days post-inoculation (dpi). Leaf areas infiltrated with either *PITG_17484_CFP* or *AtPIP2A_YFP* (D) as designated by the dash line circles did not show any phenotype at 3 dpi. doi:10.1371/journal.pone.0034701.g003

cloned and sequenced to confirm that they corresponded to the appropriate splice variant (Figure S4). Additional functional analysis of *Psc_781.4* confirmed transient expression in *N. benthamiana* does not elicit a cell death response *in planta*, indicating that it likely has no virulence function in *Ps. cubensis* (Figure S5). In

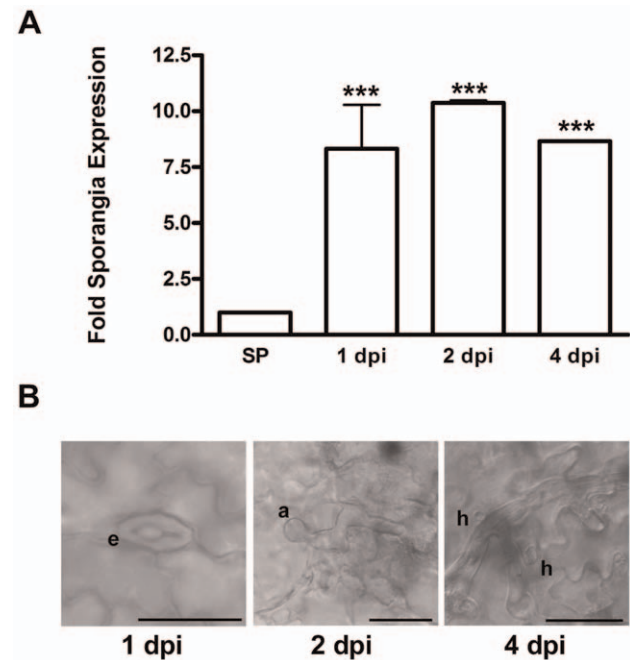


Figure 4. *PscRXLR1* mRNA expression is up-regulated during *Pseudoperonospora cubensis* infection of cucumber. (A) Expression levels of *PscRXLR1* in sporangia (SP) and at 1, 2, and 4 days post-inoculation (dpi). Expression is displayed as fold sporangia expression and all time points are significantly different compared to SP control (***) indicates $p < 0.001$) using Two-way ANOVA. Error bars represent the standard error of the mean of 2 technical replicates from each of 2 biological replicates. (B) Differential interference (DIC) microscopy images of stages of *Ps. cubensis* infection on cucumber where e = encysted zoospore; a = appressorium; and h = haustorium. Scale bars = 25 μ m. doi:10.1371/journal.pone.0034701.g004

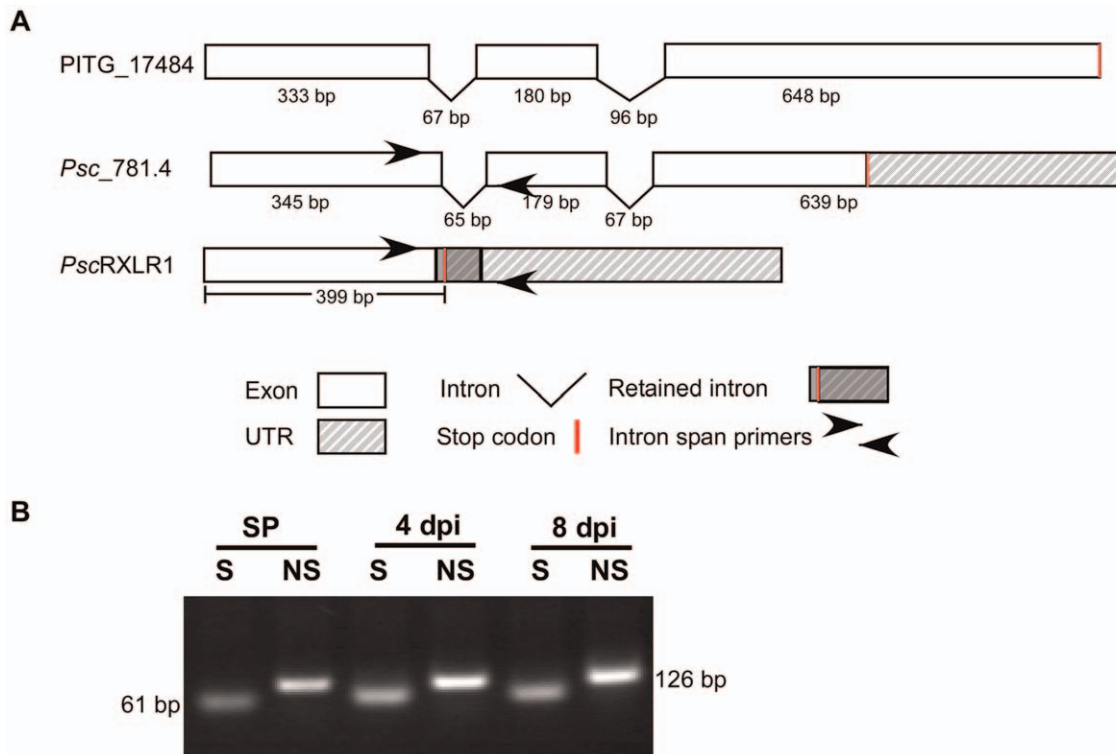


Figure 5. *PscRXLR1* is a splice variant of *Psc_781.4*. (A) Schematic representation of intron and exon structures of PITG_17484, *Psc_781.4*, and *PscRXLR1*. (B) RT-PCR analysis of alternative splicing in *Psc_781.4* in sporangia (SP) and at 4 and 8 days post-inoculation (dpi). RT-PCR products were subcloned into the pGEM vector and DNA was amplified by PCR using intron spanning primers as shown in (A). S, transcript with spliced intron 1. NS, non-spliced transcript.

doi:10.1371/journal.pone.0034701.g005

total, these independent methods confirm our conclusion that *Psc_781.4* is alternatively spliced leading to generation of a functional RXLR effector protein.

Discussion

In this study, we describe a candidate RXLR-type effector from *Ps. cubensis* that results from a splice variant of a putative multi-drug transporter protein, and additionally expands the scope of our understanding of the function and evolutionary history of the *Ps. cubensis* effector repertoire. While *Ps. cubensis* is an oomycete pathogen of worldwide economic importance, insight into the mechanism(s) underlying its virulence and pathogenicity remain limited [43]. A recent study has provided a foundation for investigating the genetic basis for virulence and pathogenicity in *Ps. cubensis* through generation of a large scale genomic dataset [24]. We build upon this previous work using a combination of *in silico* analyses, gene expression studies, and cell biology to functionally characterize *PscRXLR1* and establish a potential role in promoting *Ps. cubensis* infection and proliferation.

Alternative splicing has been previously described in oomycetes pathogens of plants; specifically related to the family 5 endoglucanases (EGL5) from *P. sojae* [21], as well as in gene families such as Crinklers (CRNs), protein kinases, and transcription factors [22]. In *P. sojae*, EGL5 proteins have a role in infection of soybean and are highly up-regulated during the early stages of infection. As part of these analyses, four different mechanisms of alternative splicing were described: intron skip, exon skip, alternative donor site, and alternative acceptor site, with intron skip, where the intron is retained, being the most commonly observed mechanism

[22]. In agreement with this previous observation, we propose that the *PscRXLR1* transcript is generated *via* a retained intron from *Psc_781.4*, yielding an RXLR-type effector. From an evolutionary standpoint, alternative-splicing functions to expand the capacity of an organism's proteome, thus enabling the generation of multiple functional isoforms from a single coding sequence. Over time, new isoforms may be maintained if they have a beneficial function [44], or lost, if their function is not critical to the lifecycle of the organism. In the case of plant pathogens, this process could potentially serve an adaptive role to allow for generation of isoforms of "housekeeping" type genes that gain new function(s), potentially to promote virulence or infection. Alternatively, the *PscRXLR1* splice variant could represent a step in evolutionary time, as it is generated from the same coding sequence as *Psc_781.4* and maintained in the coding repertoire, but has not been duplicated or retained as a separate sequence.

Like other oomycete effectors characterized to date, *PscRXLR1* has a functional 26 amino acid signal peptide necessary for secretion from the haustorium into the extrahaustorial matrix prior to translocation into the host cytoplasm. Interestingly, *PscRXLR1* is also localized to the host plasma membrane, despite the lack of a predicted transmembrane domain or localization signal (Figure 3). To further examine this, we surveyed the genomes of additional oomycete plant pathogens for orthologous sequences. We hypothesize that if functional characterization data for orthologs in any of these better characterized species (e.g., *P. infestans*, *P. ramorum* and *P. sojae*) existed, it might provide insight into both the function and conservation of *PscRXLR1*. Through BLAST analysis of the *P. infestans*, *P. sojae*, *P. ramorum* and *Py. ultimum* genomes, we identified orthologous sequences in each of

the pathogens, all of which were annotated as members of the Drug/Metabolite Transporter (DMT) superfamily [30]. The DMT superfamily encompasses 19 families; the orthologs described here are members of the EamA family, named for the O-acetylserine/cysteine export gene in *Escherichia coli* [45]. While *PscRXLR1* is lacking the EamA transmembrane domains that are characteristic of these transporter proteins, our data clearly demonstrated plasma membrane localization.

Monitoring the expression of both pathogen and host genes during infection can provide insight into the interplay between resistance and susceptibility [46]. Using qRT-PCR, we demonstrated that expression of *PscRXLR1* mRNA is up-regulated during the early stages of infection of cucumber. While we were unable to distinguish between isoforms, expression was induced nearly 10-fold at 1, 2 and 4 days post-inoculation (dpi), corresponding with zoospore encystment, appressorium formation and penetration, and proliferation and haustoria establishment, respectively (Figure 5). Several effector proteins from *P. infestans* have also been demonstrated to have distinct temporal patterns of expression, and are often expressed during the pre- and early stages of infection, representative of the biotrophic phase of the *P. infestans* life cycle [12]. Based on the robust induction of *PscRXLR1* mRNA during the early stages of infection, as well as the aggressive nature of the necrosis-inducing activity observed in *N. benthamiana*, we hypothesize that the expression pattern of *PscRXLR1* could support a role in the initial infection process, possibly through dampening of host defense responses. Indeed, effectors from other oomycete plant pathogens, including *PcQNE* from *Ps. cubensis* and members of the CRN family from *P. infestans*, have also been shown to elicit similar phenotypes when transiently expressed in *N. benthamiana* [3,24,47], supporting the classification of *PscRXLR1* as an effector protein with a putative role in virulence.

In the current study, ongoing analysis of the *Ps. cubensis* genome has expanded the candidate effector complement of *Ps. cubensis* to 271 sequences, revealing significant variation in the conserved translocation motif. While previous analyses revealed a near equal distribution of RXLR:QXLR motifs in *Ps. cubensis*, our current work, based on a higher coverage draft genome sequence and predicted protein sequences rather than open reading frames (ORFs), predicts sequences with 20 different amino acid possibilities at the R1 position of the XXLR motif. Of these 20 predicted R1 substitutions, 19 have expression support (Table S1). While all 20 R1 substitutions have yet to be functionally validated, it is not surprising that *Ps. cubensis* effectors may in fact utilize a more diverse set of translocation motifs compared to the *Phytophthora* spp., given its obligate lifestyle and relatively narrow host range. Among *Phytophthora* spp., the conservation of the RXLR motif is well-established, yet there are additional classes of oomycete effectors, such as the CRN family, that appear to utilize disparate translocation motifs [34,47]. Moreover, analysis of the *Py. ultimum* genome has identified an additional predicted translocation motif, YxSL[RK] [9]. Indeed, divergence of transport signal sequences is even more pronounced between oomycetes and the true fungi, which have no obvious conserved motifs that could function in transport and show high degrees of variation even within the same species [48]. For example, the effectors AvrM and AvrL567 from *Melampsora lini*, an obligate rust fungi with a similar lifestyle to *Ps. cubensis*, rely on unique N-terminal sequences for uptake [48]. These sequences, while different in regards to sequence similarity from the RXLR motif observed in *Phytophthora* spp., are similar in that they feature positively charged residues, implying that secondary protein structure may be a factor contributing to uptake of these proteins.

Both *M. lini* and *Ps. cubensis* are obligate biotrophs with specific host ranges, which may have influenced the evolution of their effector repertoires to select for unique translocation motifs compared to those found in *Phytophthora* spp.

Preliminary analysis of the *Ps. cubensis* effector repertoire reveals minimal orthology with annotated effector proteins from *P. infestans*, similar to what has been observed when comparing the effector complement from *P. infestans* with *P. ramorum*, *P. sojae* or *H. arabidopsidis* [8,27,28]. Through extensive analysis using both evolutionary and comparative genomics, *Phytophthora* RXLR effector genes have been shown to be undergoing accelerated rates of birth and death evolution as well as both widespread gene duplication and loss events [7,8]. As such, among the predicted RXLR effector genes from *P. infestans*, *P. sojae*, and *P. ramorum*, there are few genes with high degrees of sequence similarity or 1:1:1 orthology [7]. Similarly, the same phenomenon was observed in comparing candidate effector sequences from *Ps. cubensis* to those of *P. infestans*. Of 271 predicted *Ps. cubensis* sequences, less than half (41%) of these had significant similarity ($e\text{-value} < 1e^{-5}$) to predicted *P. infestans* proteins, and only 3 of these sequences had similarity to annotated *P. infestans* RXLR effector proteins. These results indicate that the effector repertoire *Ps. cubensis* utilizes to promote its virulence and pathogenicity on its hosts is, as could be predicted, very different than that utilized by *P. infestans*, and likely the other *Phytophthora* spp. as well. This is likely due to differing selective pressures on *Ps. cubensis* resulting from host specificity as well as differences in lifestyle between the two pathogens (i.e., obligate vs. non-obligate).

In this study, we identified minimal conservation between the predicted *Ps. cubensis* effector complement and effectors from *P. infestans*. We hypothesize that the identification and analysis of effector to non-effector relationships among oomycete plant pathogens is a valid measure to assess conservation and rates of evolution. Additionally, with the identification of *PscRXLR1*, a splice variant of a non-effector gene, we posit that these types of analyses as well as a more thorough analysis of alternative splicing may provide a preliminary baseline to not only investigate evolutionary differences among oomycete plant pathogens, but to also infer the relationship between effector repertoire and the host-pathogen specificity and lifestyle. We have used several criteria (i.e., prediction of selection pressure, secretion, etc.) to identify and analyze the relationship between predicted *Ps. cubensis* effectors and their orthologs in *P. infestans*. As observed for other effector proteins, some *Ps. cubensis* effectors may have experienced stronger positive selection than most other proteins within the genome. Interestingly, in addition to varying significantly from the genome average, the distribution of ω for the *PscE-Pi* pairs has two distinct peaks, representing groups of effectors under different levels of selection pressure. Thus, it appears that aside from acting as effectors during infection, some of these slower evolving genes may have additional, “housekeeping” roles that are yet to be uncovered. Despite computational evidence indicating that these slower evolving genes are likely effectors, their role(s) in pathogenesis remain to be established.

Materials and Methods

Ps. cubensis culture and growth conditions

Ps. cubensis was maintained on *Cucumis sativus* cv. ‘Vlaspik’ as previously described [24]. Cucumber plants were grown at 22°C with a 12 h light/dark photoperiod. For *Ps. cubensis* inoculation, sporangia were collected from heavily sporulating leaves by washing with cold sterile distilled water and collecting sporangia in a centrifuge tube. Sporangia were enumerated with a

hemocytometer and suspended to a concentration of 1×10^5 sporangia/ml in sterile distilled water. The underside of fully expanded 2nd or 3rd true leaves of 4-week-old cucumber plants were spray-inoculated, until run-off, with the suspension, and incubated for 24 h at 100% humidity in the dark. After 24 h, inoculated plants were moved to a growth chamber (22°C with a 12 h light/dark photoperiod).

DNA and RNA Extraction

Genomic DNA of *Ps. cubensis* was isolated from sporangia of isolate MSU-1 using the DNeasy Plant Mini kit (Qiagen, Germantown, MD) with slight modifications. Sporangia were washed from heavily sporulating leaves with sterile distilled water and filtered through a 40 µm nylon cell strainer to remove residual plant debris. The resultant sporangia suspension was centrifuged, and the supernatant decanted. Sporangia were suspended in buffer AP1 containing RNase and 5 µl of Proteinase K and incubated at 37°C for 20 min. 50 µl of 425–600 µm acid washed beads were added to the sporangia suspension and sporangia disrupted for 3 min using a vortex. Subsequent DNA extraction steps were performed according to manufacturer's instructions.

Ps. cubensis total RNA was isolated as follows: sporangia were collected as described above for DNA isolation, yet re-suspended in buffer RLT from the RNeasy Plant Mini Kit (Qiagen, Germantown, MD) and disrupted as above. RNA isolation was performed according to the manufacturer's instructions. RNA samples were treated with DNase (Promega, Madison, WI) prior to use.

Sequence, Assembly and Annotation of the *Ps. cubensis* Genome

Genomic DNA was isolated from *Ps. cubensis* MSU-1 as described above and libraries constructed using the Illumina Genomic DNA Sample kit (Illumina, San Diego, CA). Two separate paired-end libraries were end sequenced using an Illumina Genome Analyzer II (Illumina, San Diego, CA) at the UC-Davis Genome Center. The first library was sequenced with 84 bp reads and an insert size of 180 bp yielding 7.8 Gbp of sequence. The second library was sequenced with 100 bp reads and an insert size of 480 bp yielding 5.5 Gbp of sequence. Illumina reads were trimmed to 51 bp to remove low quality regions at the 3' end of the reads. Reads with more than one N base or a base with a quality score less than 20 were removed. The reads were then searched against the Cucumber genome assembly [49] with Bowtie v0.12.7 [50] and matching reads were removed; 4.5 Gbp of sequence was retained following trimming and cleaning the reads. The trimmed and cleaned reads were assembled using Velvet v1.0.14 [25]. Three Velvet runs were performed with hash lengths of 31, 41, and 51 and coverage cutoffs of 7, 3.6, and 2, respectively. A minimum contig size cutoff of 200 bp was used for all the assembly runs. The contigs from each Velvet run were then merged into one assembly using the Minimus2 pipeline (<http://sourceforge.net/apps/mediawiki/amos/index.php?title=Minimus2>). Contaminant-containing and mitochondrial contigs were removed; the final assembly contains 35,546 contigs with an N50 contig size of 4.0 Kbp; the total assembly is 64.4 Mbp. Reads were deposited in the Sequence Read Archive at the National Center for Biotechnology Information under study number SRP011018. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AHJF000000000. The version described in this paper is the first version, AHJF01000000.

The assembly was annotated using the MAKER [26] annotation pipeline. The FGENESH gene finder [51] was used

with the *Phytophthora* matrix to produce the initial gene calls for the pipeline. All transcript and protein sequences from sequenced oomycete genomes were provided to MAKER to improve the quality of the annotation. In total, 23,519 loci and 23,522 gene models were predicted. Putative functional annotation was assigned by searching the gene models against UniRef100 using BLASTX (cutoff: $1E-5$) and transferring the first hit with informative annotation.

Identification and Cloning of *PscRXLR1*, *Psc_781.4* and PITG_17484

Amplification of the coding sequence of *PscRXLR1* was performed using DNA primers that correspond to the open reading frame of *PscRxLR1* (Figure 5). Subsequent isolation and cloning of *PscRxLR1* was performed by PCR using gene-specific primers and genomic DNA from *Ps. cubensis* sporangia. Resultant amplicons were cloned into the TA cloning vector pGEM-T-Easy (Promega), generating pGEM_ *PscRXLR1*. To ensure identification of a complete coding sequence, as well as to verify the absence of introns in the sequence, 3' RACE (Rapid Amplification of cDNA Ends) was performed using the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA). Amplification of *Psc_781.4* was performed using 3' RACE as described above and the final coding sequence was amplified using gene specific primers (Figure 5). Fidelity of all sequences was confirmed by DNA sequencing using the ABI 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA).

P. infestans clone PITG_17484 was amplified from cDNA of *P. infestans* by PCR using gene-specific primers. Amplicons were subcloned into pGEM-T-Easy and sequences confirmed by sequencing.

DNA cloning and construct preparation

To validate the predicted signal peptide (SP) of *PscRXLR1*, the yeast signal trap assay was used [24,32,33]. A DNA fragment corresponding to the predicted 26 amino acid signal peptide (including start codon) was amplified by PCR using gene specific primers modified to include 5' *EcoR1* and 3' *XhoI* restriction sites. Resultant amplicons were cloned into TA cloning vector pGEM-T-Easy (Promega), to yield pGEM-*PscRXLR1*-SP. The plasmid pGEM-*PscRXLR1*-SP was digested with *EcoR1* and *XhoI*, the 84 bp SP fragment was gel purified, and ligated into the *EcoR1*/*XhoI* sites of the yeast signal trap vector pSUC2T7M13ROI [32], generating pSUC2-*PscRXLR1*.

The SP sequence of PITG_17484 was similarly amplified and cloned, generating pSUC2-PITG_17484. Plasmids pSUC2T7M13-ROI, pSUC2-*PscRXLR1* and pSUC2-PITG_17484 were transformed into the yeast *SUC2* minus strain YTK12 using the Frozen-EZ Yeast Transformation IITM kit (Zymo Research, Orange, CA) following the manufacturer's instructions. Transformants were selected on CMD-W plates (0.67% yeast nitrogen base without amino acids, 0.075% tryptophan dropout supplement, 2% sucrose, 0.1% glucose, and 2% agar) for 3 days at 28 °C. To confirm signal peptide function, YTK12 containing the pSUC2 constructs were grown on raffinose-containing YPRAA plates (1% yeast extract, 2% peptone, 2% raffinose, 2 µg/L Antimycin A, and 2% agar). Yeast strains were replicated onto YPDA plates (1% yeast, 2% peptone, 2% glucose, 0.003% adenine hemisulfate, and 2% agar) and CMD-W plates as equal viability controls. YTK12 without pSUC2 was used as a negative control. The detection of the secreted invertase activity with 2,3,5-triphenyltetrazolium chloride (TTC) was performed as described by Tian et al. [24].

For construction of plasmids used for localization and phenotype studies, the open reading frames of *PscRXLR1*, *Psc_781.4* and PITG_17484 (minus stop codons) were amplified and cloned into the Gateway entry vector pENTR/D-TOPO (Invitrogen, Carlsbad, CA), yielding pENTR-*PscRXLR1*-GW, pENTR-*Psc_781.4*-GW, and pENTR-PITG_17484-GW, respectively. The destination vector pVKH18En6gw-cCFP [24] was used to create the C-terminal CFP fusions, using heterologous recombination *via* LR Clonase, as per the manufacturer's instructions (Invitrogen).

Transient expression and localization in *N. benthamiana*

Infiltration and transient expression in *N. benthamiana* using *A. tumefaciens* was performed on 4–6 week old plants as described in Tian et al. [24]. *A. tumefaciens* strains were grown overnight at 28°C on Luria-Bertani (LB) plates containing 50 µg/mL rifampicin and 25 µg/mL kanamycin. *A. tumefaciens* clones were re-suspended in induction buffer (10 mM MES, pH 5.6, 10 mM MgCl₂, 150 µM acetosyringone) and incubated at room temperature, shaking in the dark, for 2 hours prior to infiltration. *A. tumefaciens* suspensions were infiltrated at a final concentration of OD₆₀₀ = 0.8.

A. tumefaciens-mediated transient expression in *N. benthamiana* for localization of *PscRXLR1*-CFP, *Psc_781.4*-CFP, and PITG_17484-CFP with *AtPIP2A*-YFP [38] was performed as described above. Visualization of fluorescently tagged proteins was observed using an Olympus Fluoview 1000 laser scanning confocal microscope. Images were adjusted for contrast in Canvas X (ACD Systems).

Quantitative real time PCR

First strand cDNA was synthesized from 1 µg total RNA using the first-strand cDNA synthesis kit (USB, Cleveland, OH). Quantitative RT-PCR was performed using a Mastercycler ep Realplex real-time PCR (Eppendorf, Hamburg, Germany) using HotStart SYBR Green qPCR Master Mix (2x) (USB), as previously described [52]. For amplification of *PscRXLR1* transcripts, gene specific primers were designed to amplify a 50 bp fragment (Forward: 5'-TGCGTAGCATCGCCAACCGA-3 and Reverse: 5'-TCTTGCCAGCTGCATCGCGA-3'). Primers specific for the *Ps. cubensis* internal transcribed spacer (ITS) region were used as an endogenous control. Cycle parameters were as follows: 95°C for 2 min, followed by 40 cycles of: 95°C (15 sec), 60°C (15 sec) and 72°C (45 sec). Fold expression was calculated based on expression in sporangia. Data were analyzed by two-way ANOVA using Prism 4 (GraphPad Software).

Splice variant analysis

Primers spanning the region of intron 1 were used (Figure S4) to amplify RT-PCR products from SP, 4, and 8 dpi cDNA samples and resultant products were cloned into the TA cloning vector pGEM-T-Easy. Fidelity of all sequences was confirmed by DNA sequencing as described above.

Ortholog identification and sequence analysis

Candidate effector proteins were identified from the predicted proteome of *Ps. cubensis* generated from the draft genome assembly using a modified RXLR effector identification pipeline [28]. Orthologs of *PscRXLR1* were identified in *P. infestans* (<http://www.broadinstitute.org/>), *P. sojae* (<http://genome.jgi-psf.org>), *P. ramorum* (<http://genome.jgi-psf.org>) and *Py. ultimum* (<http://pythium.plantbiology.msu.edu>) with BLAST [53]. Signal peptides were predicted using SignalP 3.0 ([54]; <http://www.cbs.dtu.dk/services/SignalP/>) and protein motifs identified using InterProScan ([55]; <http://www.ebi.ac.uk/Tools/pfa/iprscan/>). Amino

acid alignments were generated using ClustalW2 ([56]; <http://www.clustal.org>), and resultant figures generated using BoxShade v3.21 (http://www.ch.embnet.org/software/BOX_form.html). *P. infestans* effector sequences from Haas et al. [7] were used for analyses and *P. infestans* protein models were obtained from the *Phytophthora infestans* Sequencing Project (<http://www.broadinstitute.org/>).

Defining paralogs and orthologs and evolutionary rate estimates

Synonymous and non-synonymous substitution rates (K_s and K_a , respectively) were determined using the yn00 program in PAML [57]. Protein sequences were aligned first and “back-translated” to coding sequence alignments. Very few pairs had run errors (e.g., NAN in PAML output), and those with run errors were excluded. Sequence pairs that were too similar ($K_s \leq 0.005$) or too divergent ($K_s > 3$) were also excluded from further analysis. For each *Ps. cubensis* or *P. infestans* effector protein, the closest paralogous genes were identified using within-species BLAST searches and used for rate calculation. Rates between putative orthologs were calculated as well. Putative orthologs were identified globally between *Ps. cubensis*, *P. infestans* or *P. ultimum* by first determining pairwise sequence similarities between all genes in these species. For each *Ps. cubensis* protein, X, a protein in a second species, Y, is considered an ortholog if the following two conditions are met: 1) X is the reciprocal best match of Y and 2) X is located in a syntenic block where Y is found. Syntenic regions were established using Multiple Collinearity Scan [58], with $1e^{-5}$ as an alignment significance threshold, match size ≥ 5 , and average intergenic distance.

Molecular phylogenetic analysis

The full-length protein sequences of *PscRXLR1* and its orthologs were aligned using default parameters with MUSCLE [59]. The multiple sequence alignment was used to infer phylogenetic relationships between *PscRXLR1* and its orthologs using the Maximum Likelihood method, based on the JTT matrix-based model [60] with MEGA5 [61]. Bootstrap values (based on 500 replicates) for each node are given if $>25\%$.

Supporting Information

Figure S1 Signal peptide distribution among ortholog pairs. (A) Distribution of secreted or non-secreted proteins in the *Pseudoperonospora cubensis* – *Phytophthora infestans* ortholog baseline dataset. (B) Distribution of *P. infestans* orthologs of *Ps. cubensis* effectors that are predicted to be secreted. *Psc-sec* = *Ps. cubensis* secreted protein. *Psc-non* = *Ps. cubensis* non-secreted protein. *Pi-sec* = *P. infestans* secreted protein. *Pi-non* = *P. infestans* non-secreted protein. (TIF)

Figure S2 Relationship between *PscRXLR1* and oomycete orthologs. Alignment of *PscRXLR1*, *Psc_781.4*, PITG_17484 (*P. infestans*), PYU_T005955 (*Py. ultimum*), *P. ramorum* 96701_0_3435, and *P. sojae* 156165 amino acid sequences were generated using ClustalW and represented with BoxShade. *PscRXLR1* signal peptide is boxed in blue. The RXLR or RXLR-like domains are boxed in red. The green boxes represent the EamA domains found in each protein sequence. Stop codons are represented by asterisks. (PDF)

Figure S3 Heterologous expression of *PscRXLR1* specifically results in cell death in *Nicotiana benthamiana*. Infiltration of *PscRXLR1*-CFP with or without the plasma

membrane marker construct *AtPIP2A-YFP* results in chlorosis and necrosis 4 days post-inoculation (dpi). Circles mark the infiltration zones, visible at 0 dpi. Infiltration with *AtPIP2A-YFP* alone does not result in any observable phenotype in *N. benthamiana* leaves. (TIF)

Figure S4 Multiple sequence alignments of splice variant isoforms. (A) Alignment representing non-spliced, *PscRXLR1* isoform. (B) Alignment representing spliced isoform indicative of *Psc_781.4*. (PDF)

Figure S5 Heterologous expression of *Psc_781.4* in *Nicotiana benthamiana*. Infiltration and expression of *Psc_781.4* does not result in any observable phenotype in *N. benthamiana* leaves at 4 days post-inoculation (dpi). Circles mark the infiltration zones, visible at 9 dpi. (TIF)

Table S1 List of candidate effectors of *Pseudoperonospora cubensis*, their annotation, signal peptide predic-

tion, RXLR-like motif, amino acid sequence, and expression. (XLSX)

Table S2 List of *Pseudoperonospora cubensis* effector and *Phytophthora infestans* ortholog pairs. (XLSX)

Acknowledgments

We thank members of the Day lab for critical reading of the manuscript. Dr. Joe Win (The Sainsbury Laboratory) is gratefully acknowledged for his contributions in the preliminary analysis of genomic sequences from *Ps. cubensis*.

Author Contributions

Conceived and designed the experiments: EAS S-HS CRB BD. Performed the experiments: EAS CZ BNA. Analyzed the data: EAS CZ BNA JPH S-HS CRB BD. Contributed reagents/materials/analysis tools: EAS CZ BNA JPH. Wrote the paper: EAS BD.

References

- De Wit PJGM, Mehrabi R, Van Den B, Harrold A, Stergiopoulos I (2009) Fungal effector proteins: past, present and future. *Mol Plant Pathol* 10: 735–747.
- Mansfield JW (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Mol Plant Pathol* 10: 721–734.
- Schornack S, Huitema E, Cano LM, Bozkurt TO, Oliva R, et al. (2009) Ten things to know about oomycete effectors. *Mol Plant Pathol* 10: 795–803.
- Phytopathology TPCRI (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Ann Rev Phytopathol* 44: 41–60.
- Chisholm ST, Coaker G, Day B, Staskawicz BJ (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124: 803–814.
- Jones JD, Dangl JL (2006) The plant immune system. *Nature* 444: 323–329.
- Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, et al. (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461: 393–398.
- Tyler BM (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313: 1261–1266.
- Levesque CA, Brouwer H, Cano L, Hamilton JP, Holt C, et al. (2010) Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol* 11: R73.
- Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, et al. (2010) Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330: 1549–1551.
- Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, et al. (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell* 17: 1839–1850.
- Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, et al. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450: 115–118.
- Bos JL, Kanneganti TD, Young C, Kadir C, Huitema E, et al. (2006) The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant J* 48: 165–176.
- Allen RL (2004) Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science* 306: 1957–1960.
- Armstrong MR, Whisson SC, Pritchard L, Bos JL, Venter E, et al. (2005) An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. *Proc Natl Acad Sci U S A* 102: 7766–7771.
- Shan W, Cao M, Leung D, Tyler BM (2004) The Avr1b locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps 1b. *Mol Plant-Microbe Interact* 17: 394–403.
- Qutob D, Tedman-Jones J, Dong S, Kufu K, Pham H, et al. (2009) Copy number variation and transcriptional polymorphisms of *Phytophthora sojae* RXLR effector genes Avr1a and Avr3a. *PLoS ONE* 4: e5066.
- Dong S, Qutob D, Tedman-Jones J, Kufu K, Wang Y, et al. (2009) The *Phytophthora sojae* avirulence locus Avr3c encodes a multi-copy RXLR effector with sequence polymorphisms among pathogen strains. *PLoS ONE* 4: e5556.
- Matlin AJ, Clark F, Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* 6: 386–398.
- Win J, Kanneganti TD, Torto-Alalibo T, Kamoun S (2006) Computational and comparative analyses of 150 full-length cDNA sequences from the oomycete plant pathogen *Phytophthora infestans*. *Fungal Genet Biol* 43: 20–33.
- Costanzo S, Ospina-Giraldo M, Deahl K, Baker C, Jones R (2007) Alternate intron processing of family 5 endoglucanase transcripts from the genus *Phytophthora*. *Curr Genetics* 52: 115–123.
- Shen D, Ye W, Dong S, Wang Y, Dou D (2011) Characterization of intronic structures and alternative splicing in *Phytophthora sojae* by comparative analysis of expressed sequence tags and genomic sequences. *Canadian J Microbiol* 57: 84–90.
- Savory EA, Granke LL, Quesada-Ocampo LM, Varbanova M, Hausbeck MK, et al. (2011) The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*. *Mol Plant Pathol* 12: 217–226.
- Tian M, Win J, Savory E, Burkhardt A, Held M, et al. (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol Plant-Microbe Interact* 24: 543–553.
- Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Gen Res* 18: 821–829.
- Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, et al. (2008) MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Gen Res* 18: 188–196.
- Jiang RHY, Tripathy S, Govers F, Tyler BM (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.
- Win J, Morgan W, Bos J, Krasileva KV, Cano LM, et al. (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 19: 2349–2369.
- Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290: 1151–1155.
- Jack DL, Yang NM, Saier MH, Jr. (2001) The drug/metabolite transporter superfamily. *Eur J Biochem* 268: 3620–3639.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783–795.
- Jacobs KA, Collins-Racie LA, Colbert M, Duckett M, Golden-Fleet M, et al. (1997) A genetic selection for isolating cDNAs encoding secreted proteins. *Gene* 198: 289–296.
- Oh SK, Young C, Lee M, Oliva R, Bozkurt TO, et al. (2009) In planta expression screens of *Phytophthora infestans* RXLR effectors reveal diverse phenotypes, including activation of the *Solanum tuberosum* disease resistance protein Rpi-blb2. *Plant Cell* 21: 2928.
- Schornack S, van Damme M, Bozkurt TO, Cano LM, Smoker M, et al. (2010) Ancient class of translocated oomycete effectors targets the host nucleus. *Proc Natl Acad Sci U S A* 107: 17421–17426.
- Gurlebeck D, Jahn S, Gurlebeck N, Szczesny R, Szurek B, et al. (2009) Visualization of novel virulence activities of the *Xanthomonas* type III effectors AvrBs1, AvrBs3 and AvrBs4. *Mol Plant Pathol* 10: 175–188.
- Jin P, Wood MD, Wu Y, Xie Z, Katagiri F (2003) Cleavage of the *Pseudomonas syringae* type III effector AvrRpt2 requires a host factor(s) common among eukaryotes and is important for AvrRpt2 localization in the host cell. *Plant Physiol* 133: 1072–1082.
- Shan L, Thara VK, Martin GB, Zhou JM, Tang X (2000) The *Pseudomonas* AvrPto protein is differentially recognized by tomato and tobacco and is localized to the plant plasma membrane. *Plant Cell* 12: 2323–2338.

38. Nelson BK, Cai X, Nebenführ A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J* 51: 1126–1136.
39. Chaparro-García A, Wilkinson RC, Gimenez-Ibanez S, Findlay K, Coffey MD, et al. (2011) The receptor-like kinase SERK3/BAK1 is required for basal resistance against the late blight pathogen *Phytophthora infestans* in *Nicotiana benthamiana*. *PLoS ONE* 6: e16608.
40. Win J, Kamoun S, Jones AM (2011) Purification of effector-target protein complexes via transient expression in *Nicotiana benthamiana*. *Methods Mol Biol* 712: 181–194.
41. Liu T, Ye W, Ru Y, Yang X, Gu B, et al. (2011) Two host cytoplasmic effectors are required for pathogenesis of *Phytophthora sojae* by suppression of host defenses. *Plant Physiol* 155: 490–501.
42. Halterman DA, Chen Y, Sopee J, Berduo-Sandoval J, Sanchez-Perez A (2010) Competition between *Phytophthora infestans* effectors leads to increased aggressiveness on plants containing broad-spectrum late blight resistance. *PLoS ONE* 5: e10536.
43. Savory EA, Granke LL, Quesada-Ocampo LM, Varbanova M, Hausbeck MK, et al. (2011) The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*. *Mol Plant Pathol* 12: 217–226.
44. Keren H, Lev-Maor G, Ast G (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* 11: 345–355.
45. Franke I, Resch A, Dassler T, Maier T, Bock A (2003) YfiK from *Escherichia coli* promotes export of O-acetylserine and cysteine. *J Bacteriol* 185: 1161–1166.
46. Chandran D, Tai YC, Hather G, Dewdney J, Denoux C, et al. (2009) Temporal global expression data reveal known and novel salicylate-impacted processes and regulators mediating powdery mildew growth and reproduction on Arabidopsis. *Plant Physiol* 149: 1435–1451.
47. Torto TA (2003) EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res* 13: 1675–1685.
48. Rafiqi M, Gan PH, Ravensdale M, Lawrence GJ, Ellis JG, et al. (2010) Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen. *Plant Cell* 22: 2017–2032.
49. Huang S, Li R, Zhang Z, Li L, Gu X, et al. (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41: 1275–1281.
50. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
51. Salamov AA, Solovyev VV (2000) Ab initio Gene Finding in Drosophila Genomic DNA. *Gen Res* 10: 516–522.
52. Knepper C, Savory EA, Day B (2011) Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. *Plant Physiol* 156: 286–300.
53. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc Acids Res* 25: 3389–3402.
54. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Prot* 2: 953–971.
55. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, et al. (2005) InterProScan: protein domains identifier. *Nuc Acids Res* 33: W116–120.
56. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
57. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comp App Biosci* 13: 555–556.
58. Tang H, Wang X, Bowers JE, Ming R, Alam M, et al. (2008) Unraveling ancient hexploidy through multiply-aligned angiosperm gene maps. *Gen Res* 18: 1944–1954.
59. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nuc Acids Res* 32: 1792–1797.
60. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comp App Biosci* 8: 275–282.
61. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* doi: 10.1093/molbev/msr121.