



Genomic, Network, and Phylogenetic Analysis of the Oomycete Effector Arsenal

 Jamie McGowan,  David A. Fitzpatrick

Department of Biology, Genome Evolution Laboratory, Maynooth University, Maynooth, Co. Kildare, Ireland

ABSTRACT The oomycetes are a class of microscopic, filamentous eukaryotes within the stramenopiles-alveolate-*Rhizaria* (SAR) supergroup and include ecologically significant animal and plant pathogens. Oomycetes secrete large arsenals of effector proteins that degrade host cell components, manipulate host immune responses, and induce necrosis, enabling parasitic colonization. This study investigated the expansion and evolution of effectors in 37 oomycete species in 4 oomycete orders, including *Albuginales*, *Peronosporales*, *Pythiales*, and *Saprolegniales* species. Our results highlight the large expansions of effector protein families, including glycoside hydrolases, pectinases, and necrosis-inducing proteins, in *Phytophthora* species. Species-specific expansions, including expansions of chitinases in *Aphanomyces astaci* and *Pythium oligandrum*, were detected. Novel effectors which may be involved in suppressing animal immune responses in *Ap. astaci* and *Py. insidiosum* were also identified. Type 2 necrosis-inducing proteins with an unusual phylogenetic history were also located in a number of oomycete species. We also investigated the "RxLR" effector complement of all 37 species and, as expected, observed large expansions in *Phytophthora* species numbers. Our results provide in-depth sequence information on all putative RxLR effectors from all 37 species. This work represents an up-to-date *in silico* catalogue of the effector arsenal of the oomycetes based on the 37 genomes currently available.

IMPORTANCE The oomycetes are a class of microscopic, filamentous eukaryotes and include ecologically significant animal and plant pathogens. Oomycetes secrete large arsenals of effector proteins that degrade host cell components, manipulate host immune responses, and induce necrosis, enabling parasitic colonization. In this study, we catalogued the number and evolution of effectors in 37 oomycete species whose genomes have been completely sequenced. Large expansions of effector protein families in *Phytophthora* species, including glycoside hydrolases, pectinases, and necrosis-inducing proteins, were observed. Species-specific expansions were detected, including chitinases in *Aphanomyces astaci* and *Pythium oligandrum*. Novel effectors which may be involved in suppressing animal immune responses were identified in *Ap. astaci* and *Py. oligandrum*. Type 2 necrosis-inducing proteins with an unusual phylogenetic history were also located. This work represents an up-to-date *in silico* catalogue of the effector arsenal of the oomycetes based on the 37 genomes currently available.

KEYWORDS comparative genomics, effectors, evolution, oomycota, secretome, similarity network

The oomycetes are a class of diverse eukaryotic microorganisms which includes some of the most devastating pathogens of plants, mammals, fish, and fungi (1). Previously, they were thought to be fungi, due to the similar morphologies, ecological roles, and modes of nutrition and filamentous growth (2). Molecular analyses have

Received 11 September 2017 **Accepted** 2 November 2017 **Published** 22 November 2017

Citation McGowan J, Fitzpatrick DA. 2017. Genomic, network, and phylogenetic analysis of the oomycete effector arsenal. *mSphere* 2:e00408-17. <https://doi.org/10.1128/mSphere.00408-17>.

Editor Aaron P. Mitchell, Carnegie Mellon University

Copyright © 2017 McGowan and Fitzpatrick. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to David A. Fitzpatrick, david.fitzpatrick@nuim.ie.

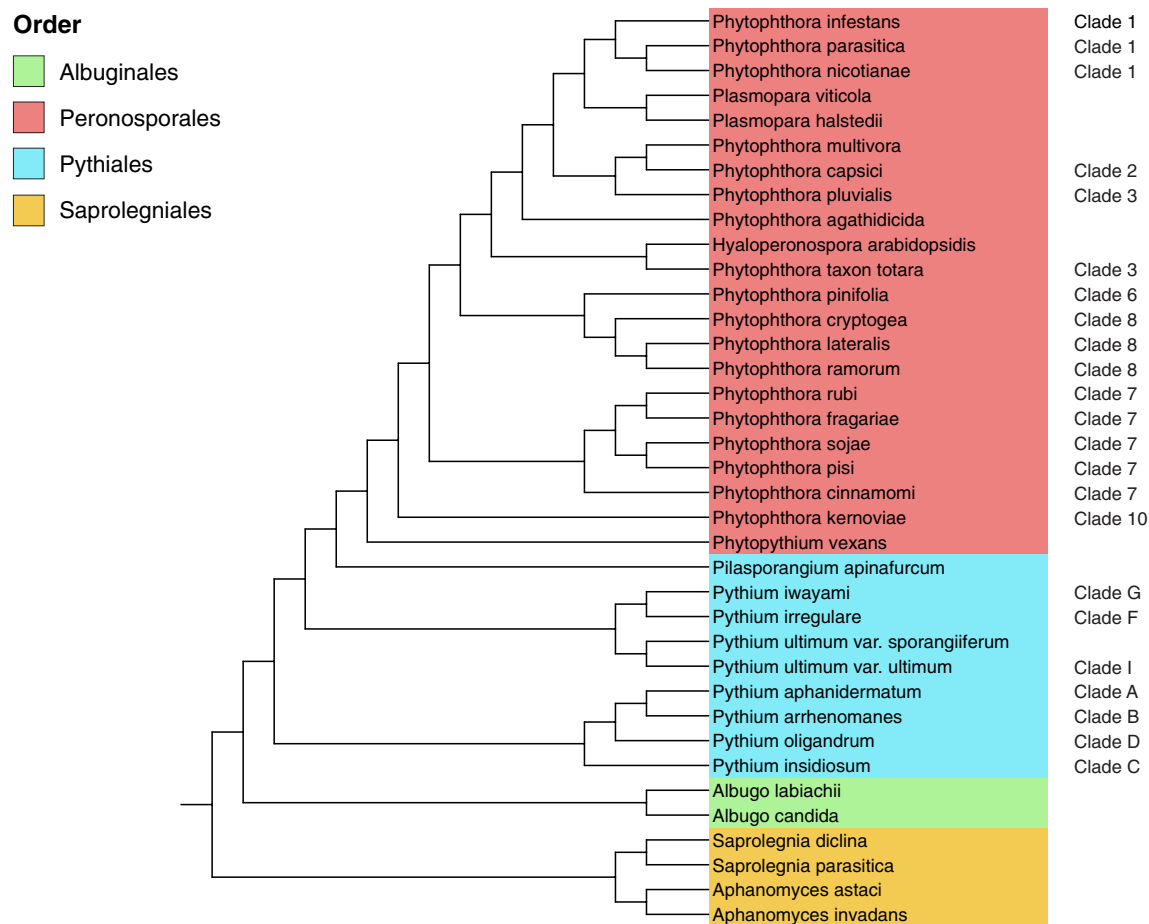


FIG 1 Phylogeny of the 37 oomycete species from 4 oomycete orders considered in this study. Data for *Phytophthora* clades as designated by Blair et al. (10) and *Pythium* clades as designated by de Cock et al. (19) are indicated in red and blue, respectively. (Adapted from reference 11.

placed the oomycetes into the *Stramenopiles-Alveolata-Rhizaria* (SAR) eukaryotic supergroup with close relationships to the diatoms and brown algae (3). Within the oomycete class, there are a number of highly diverse orders, including the *Saprolegniales*, *Peronosporales*, *Albuginales*, and *Pythiales*, that exhibit different lifestyles and can have either very specific or very broad host ranges.

More than 60% of known oomycetes are pathogens of plants (4) and have a devastating effect on many agriculturally important crops and ornamental plants. Members of the *Saprolegniales* order predominantly exhibit saprotrophic lifestyles and include the *Aphanomyces* animal and plant pathogens (5) as well as the fish-pathogenic *Saprolegnia* genus, known as “cotton molds” (6, 7). The *Peronosporales* order consists largely of phytopathogens and includes the hemibiotrophic genus *Phytophthora* (the “plant destroyers”). *Phytophthora* species include the notorious phytopathogen *Phytophthora infestans*, which is the causative agent of late potato blight, a disease reported to cause billions of dollars worth of damage worldwide annually (8). *Phytophthora* is the largest genus of the *Peronosporales* order and is divided into 10 phylogenetic clades (namely, clades 1 to 10) (9, 10) (Fig. 1). Also included in the *Peronosporales* are the genera *Hyaloperonospora* and *Plasmopara*, which are closely related to *Phytophthora* species (11) (Fig. 1). These two genera contain species that cause downy mildew in a number of economically important plants (12–14). In contrast to *Phytophthora* species, they are obligate biotrophs that are completely dependent on their host for survival. Other obligate oomycete biotrophs include the *Albugo* species (“white blister rusts”), which are members of the *Albuginales* order (15, 16) (Fig. 1). The *Pythiales*

order includes the genus *Pythium*, members of which are necrotrophs that cause root rot in many terrestrial plants. Exceptions include *Pythium insidiosum*, a pathogen causing pythiosis in mammals (17), and *Pythium oligandrum*, a pathogen of other oomycetes and fungi (18). *Pythium* species are divided into 10 clades (namely, clades A to J) (19) (Fig. 1).

Oomycetes are notorious for secreting a large arsenal of effector proteins (20). Effectors facilitate infection by manipulating host cell components, exploiting host nutrients, triggering defense responses, and inducing necrosis (21). Oomycete effectors can be categorized into two classes (the apoplastic and cytoplasmic classes) depending on where they localize. Apoplastic effectors are secreted by the pathogens and exert their pathogenic activity in the host's extracellular environment (22). Oomycete apoplastic effectors include a large number of hydrolytic enzymes which are involved in the degradation of host cell components, enabling penetration of host cells. These include cutinases, glycoside hydrolases (GHs), pectinases, and proteases, among other enzymes. Some oomycete species, such as *Phytophthora*, also encode members of extracellular toxin families such as necrosis-inducing proteins (NLPs) and Pcf family toxins (23). Host species are known to secrete protective proteases in an effort to degrade pathogen effectors; for example, P69B and P69C are subtilisin-like serine proteases secreted by tomatoes in response to *Phytophthora* species protease (24). To counteract this, *Phytophthora* species secrete protease inhibitors to block the host defense (25).

In contrast to apoplastic effectors, cytoplasmic effectors are secreted and translocated into the host cell, where they exhibit their pathogenic activity. Two types of cytoplasmic effectors dominate the oomycete secretome—"RxLR" effectors and Crinklers. RxLR effectors (RxLRs) are so named because they contain a highly conserved RxLR motif in their N-terminal domain (8, 26). This motif is followed by a downstream "EER" motif in many RxLRs. Studies have shown that the RxLR motif acts as a translocation signal, marking the protein for trafficking into the host cell (27). The mechanisms of the RxLR motif are thought to be similar to those of the "Pexel" translocation motif found in effectors of the malaria parasite *Plasmodium falciparum* (27–29). Some RxLRs can enter host cells in a manner independent of any additional pathogen-encoded machinery (30). RxLRs have been described as members of a rapidly evolving superfamily in which all members are related and share a common ancestor (31). They have very conserved N termini and divergent C termini, although conserved WY folds have been observed in some (8, 32). Genes encoding RxLR effectors are mainly found in gene-sparse regions of the genome which contain a high frequency of transposons (33). This could account for the rapid evolution of RxLR effectors. Large expansions of RxLR effectors have been observed in *Phytophthora* species, with some species reportedly encoding several hundred putative RxLRs (8, 26).

Other well-characterized oomycete effectors include Crinkler proteins (CRNs), named for their crinkling and necrosis-inducing activity, which are composed of a highly conserved N-terminal domain containing a signal peptide and an "LxLFLAK" motif which mediates translocation into the host cell (34). The end of the N terminus is marked by a highly conserved "HVLVxxP" motif, which separates the N terminus and C terminus. They are modular, rapidly evolving proteins that consist of a diverse collection of C-terminal domains (8, 23). The CRNs of some oomycetes carry a modified version of the "LxLFLAK" motif (15). CRNs have been reported to localize to, target, and accumulate in host nuclear components (34, 35). CRNs are thought to be a more ancient class of cytoplasmic effectors than RxLRs, as they have been found to be distributed across a wide range of oomycete orders (23), including *Albuginales* (15, 16), *Peronosporales* (8, 26, 36), and *Pythiales* (37, 38).

In this report, we catalogue the effector repertoire among 590,896 protein coding genes from 37 publically available genome sequences for the oomycete class, including *Albugo*, *Aphanomyces*, *Hyaloperonospora*, *Phytophthora*, *Phytophythium*, *Pilasporangium*, *Plasmopara*, *Pythium*, and *Saprolegnia* species (Table 1). Numerous bioinformatic techniques were employed to identify and catalogue putative proteins which may be involved in pathogenesis. A mix of network and phylogenetic methods was utilized to

TABLE 1 Taxonomic and genomic information for the 37 oomycete species analyzed here^a

Species name	Clade	Order	Host	Size (Mb)	Genes	Candidate effectors	Reference
<i>Al. candida</i>	n/a	Albuginales	Plants	34.40	13,310	162	15
<i>Al. laibachii</i>	n/a	Albuginales	Plants	32.76	13,804	143	16
<i>H. arabidopsidis</i>	n/a	Peronosporales	Plants	78.89	14,321	292	36
<i>Ph. agathidicida</i>	Clade 5	Peronosporales	Plants	37.23	14,110*	573	97
<i>Ph. capsici</i>	Clade 2	Peronosporales	Plants	64.00	19,805	755	96
<i>Ph. cinnamomi</i>	Clade 7	Peronosporales	Plants	53.96	12,942*	459	97
<i>Ph. cryptogea</i>	Clade 8	Peronosporales	Plants	63.80	11,876*	430	97
<i>Ph. fragariae</i>	Clade 7	Peronosporales	Plants	73.68	13,361*	368	98
<i>Ph. infestans</i>	Clade 1	Peronosporales	Plants	228.50	17,797	1,249	8
<i>Ph. kernoviae</i>	Clade 10	Peronosporales	Plants	43.00	10,650	329	99
<i>Ph. lateralis</i>	Clade 8	Peronosporales	Plants	43.17	11,635	400	100
<i>Ph. multivora</i>	Clade 2	Peronosporales	Plants	40.32	15,006*	637	97
<i>Ph. nicotianae</i>	Clade 1	Peronosporales	Plants	76.50	10,521	559	101
<i>Ph. parasitica</i>	Clade 1	Peronosporales	Plants	82.39	27,942	983	INRA-310 v. 3
<i>Ph. pinifolia</i>	Clade 6	Peronosporales	Plants	94.60	19,533*	358	97
<i>Ph. pluvialis</i>	Clade 3	Peronosporales	Plants	53.62	18,426*	632	95
<i>Ph. pisi</i>	Clade 7	Peronosporales	Plants	58.81	15,495*	596	PRJEB6298
<i>Ph. ramorum</i>	Clade 8	Peronosporales	Plants	65.00	15,743	762	26
<i>Ph. rubi</i>	Clade 7	Peronosporales	Plants	74.65	15,462*	311	PRJNA244739
<i>Ph. sojae</i>	Clade 7	Peronosporales	Plants	82.60	26,584	1,181	26
<i>Ph. taxon totara</i>	Clade 3	Peronosporales	Plants	55.58	16,691*	585	95
<i>Pl. halstedii</i>	n/a	Peronosporales	Plants	75.32	15,469	435	14
<i>Pl. viticola</i>	n/a	Peronosporales	Plants	74.73	12,048*	331	PRJNA329579
<i>P. vexans</i>	n/a	Peronosporales	Plants	33.90	11,958	265	38
<i>Pi. apinafurcum</i>	n/a	Pythiales	Plants	37.58	13,184*	388	PRJDB3797
<i>Py. aphanidermatum</i>	Clade A	Pythiales	Plants	34.26	12,312	348	38
<i>Py. arrhenomanes</i>	Clade B	Pythiales	Plants	44.67	13,805	354	38
<i>Py. insidiosum</i>	Clade C	Pythiales	Animals	53.23	19,290*	558	102
<i>Py. irregular</i>	Clade F	Pythiales	Plants	42.97	13,805	357	38
<i>Py. iwayami</i>	Clade G	Pythiales	Plants	43.20	14,875	330	38
<i>Py. oligandrum</i>	Clade D	Pythiales	Fungi	35.90	14,292*	413	103
<i>Py. ultimum</i> var. <i>sporangiferum</i>	Clade I	Pythiales	Plants	37.70	14,096	255	38
<i>Py. ultimum</i> var. <i>ultimum</i>	Clade I	Pythiales	Plants	42.80	15,323	386	37
<i>Ap. astaci</i>	n/a	Saprolegniales	Animals	75.84	26,259	582	APO3 v.2
<i>Ap. invadans</i>	n/a	Saprolegniales	Animals	71.40	20,816	290	9901 v.2
<i>S. diclina</i>	n/a	Saprolegniales	Animals	62.88	18,229	378	PRJNA168273
<i>S. parasitica</i>	n/a	Saprolegniales	Animals	53.13	20,121	385	56

^aProtein counts manually generated from assembly data are marked with an asterisk. References are to the genome publications where possible or otherwise to NCBI BioProject identifiers or to the Broad Institute strain identifier and assembly version. Species are colored by order as follows: green, *Albuginales*; red, *Peronosporales*; blue, *Pythiales*; orange, *Saprolegniales*. (Adapted from reference 11).

analyze their evolutionary history. Our results have identified novel effector families that appear to be unique to particular oomycete lineages, including *Ap. astaci* proteins, which might have the potential to cleave immunoglobulin. Consistent with previous studies, we have identified a significant expansion of effectors in *Phytophthora* species, including glycoside hydrolases and necrosis-inducing proteins. We have detected expansions of chitin degrading enzymes in *Ap. astaci* and *Py. oligandrum*. We have also identified multiple type 2 necrosis-inducing proteins in a number of oomycete species.

RESULTS AND DISCUSSION

The oomycete secretome. Oomycete pathogens secrete effector proteins and degradative enzymes to facilitate host colonization through altered physiology (38). Using the available proteome data, we undertook *in silico* prediction analyses to determine the number of proteins in each species that may be secreted. Our data set consisted of the predicted proteomes of 37 oomycete species. This included 18 *Phytophthora*, 8 *Pythium*, 2 *Albugo*, 2 *Aphanomyces*, 2 *Plasmopara*, 2 *Saprolegnia*, 1 *Hyaloperonospora*, 1 *Pilaspangium*, and 1 *Phytophythium* species (Table 1).

Proteins predicted to contain signal peptides were located using SignalP v3. SignalP v3 was chosen over earlier and later versions of the software as previous studies have found v3 the most sensitive in detecting oomycete signal peptides (39). Proteins that contained a transmembrane domain after the signal peptide cleavage were discarded as these proteins are not likely to be secreted and are instead retained in the plasma membrane.

A previous analysis of 13 stramenopiles (including 11 species in our data set) by Adhikari et al. found that between 6.19% and 10.34% of the proteins in each species were secreted (38). Our analysis showed that of the 590,896 proteins tested, 5.25% (30,996) are predicted to be secreted, from a relative low of 2.11% (291) in the obligate biotroph *Al. laibachii* to a relative high of 7.93% (834) in the necrotroph *Ph. nicotianae* (see Table S1 in the supplemental material). We observed that in all cases, the percentages of secreted proteins in our analysis differed from what was observed in the analysis by Adhikari et al. These differences cannot be accounted for by differences in the number of proteins per species as the numbers were consistent between the two analyses, as were the methodologies used. For example, Adhikari et al. predicted that 10.34% of the *Ph. ramorum* genome is secreted compared to 7.65% in our analysis (Table S1). Similarly, there are large discrepancies between the percentages of secreted proteins for *Ph. sojae* (10.24% versus 6.24%) and for *Py. irregular* (6.95% versus 5.03%) and the largest discrepancy was seen in the comparisons for *H. arabidopsidis* (8.81% versus 3.68%). For transparency, all scripts used to annotate secreted proteins are publically available (see Materials and Methods).

Our results show there is significant a difference ($P < 0.01$ [chi square test]) between two of the closely related *Saprolegniales* species (*Ap. astaci* and *Ap. invadans*) in the number of predicted proteins, where 4.26% (1,119 predicted proteins) of the *Ap. astaci* proteome is predicted to be secreted compared to 2.93% (609 predicted proteins) in *Ap. invadans* (Table S1). The predicted proteomes for those two species differ in size (26,529 versus 20,816 predicted proteins), indicating that an expansion in secreted proteins is partially responsible for the differences observed in the sizes of the proteomes.

Overall, the correlation between the number of predicted secreted proteins and the overall number of proteins per species was shown to be moderate ($R^2 = 0.455$) and significant ($P < 0.00001$ [Pearson correlation test]). The correlation between the number of predicted secreted proteins and genome size was shown to be weak ($R^2 = 0.1065$), however, and not significant ($P > 0.05$).

Secretome enrichment analysis. We have documented which biological functions are enriched in the predicted secretomes of individual species. This was achieved by comparing the frequency of Gene Ontology (GO) terms and Pfam domains in the secretome to the nonsecreted proportion of the proteome using the Fisher exact test corrected for false-discovery rate (FDR) (40). InterProScan was used to functionally annotate all proteins with GO terms and Pfam domains (41).

Comparing the putative secretomes of the 37 species in this analysis, we saw that the number of Pfam domains enriched in the secretome relative to the nonsecreted portion of the proteome varied from a low of 2 domains in *Albugo candida* to a high of 56 in *Ap. astaci* (Table S2). The enrichment analysis showed that the elicitor Pfam domain (PF00964) is enriched in all 37 species (Fig. 2 and Table S2). Elicitors are structurally conserved extracellular proteins in *Phytophthora* and *Pythium* species (37). They are known to bind lipids and sequester sterols from plants, thereby overcoming the inability of *Phytophthora* and *Pythium* species to synthesize sterols (42). Similarly, the serine protease inhibitor Kazal-like domain (PF07648) is enriched in 33 of the 37 species (Fig. 2 and Table S2). The Kazal-like domain has been implicated in the infection process of *Phytophthora* species and acts as an apoplastic effector (24). Another widely distributed domain that is enriched in a large number of oomycete secretomes (33 of 37) is that corresponding to the cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins (CAP) (Fig. 2 and Table S2). CAP-domain-containing secreted proteins are produced by nonvertebrate eukaryotes and prokaryotes and have been implicated in both virulence and immunity functions (43, 44); however, little is currently known about the molecular mode of action of such proteins. Of the 37 species analyzed, 31 showed enrichment for chymotrypsin (PF00089), which most probably has a role to play in extracellular proteolysis. Similarly, 30 of the 37 species showed enrichment for the PAN/Apple domain (PF14295). Previous analyses had shown that

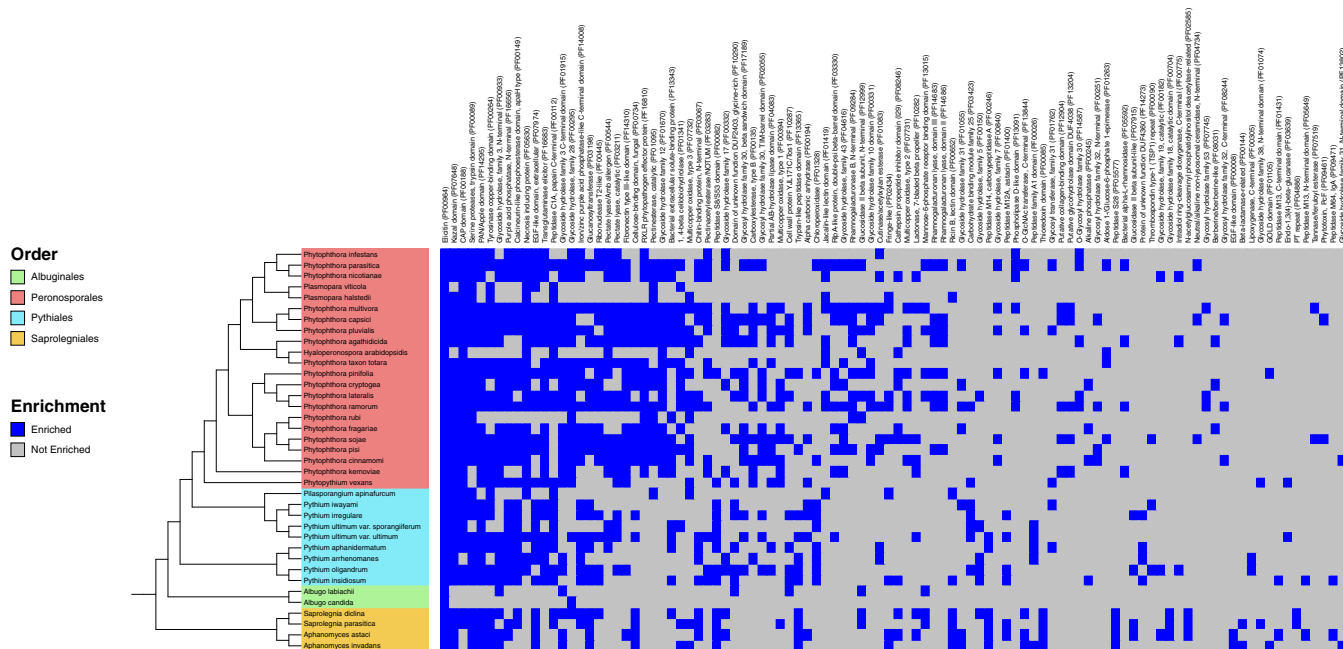


FIG 2 Heat map of enriched Pfam terms in oomycete secretomes. Terms are ordered with respect to the number of the times that they were observed. Only terms listed for two or more species are shown. Terms that were statistically enriched are colored blue in the heat map. Gray coloring indicates that the term was not enriched.

carbohydrate-binding module (CBM)-containing proteins that recognize and bind saccharide ligands from *Ph. parasitica* are associated with two PAN/Apple domains. The PAN/Apple domain is known to interact with both proteins and carbohydrates (45). Knockdown of *Ph. parasitica* CBM affects adhesion to cellulose substrates, including plant cell walls (46). Domains involved in the possible degradation of plant cell walls were also found to be enriched in the secretomes of many of the species investigated. For example, pectin degradation (PF00544, PF03283, and PF03211), glycoside hydrolases (PF00933, PF00915, PF00295, PF17189, and PF02055) and cellulose binding (PF00374) are all enriched across a range of species (Fig. 2 and Table S2). Well-known effectors, including necrosis-inducing protein (PF05630), are found to be enriched in *Phytophthora* and *Pythiales* species, while the RxLR phytopathogen domain (PF16810) is enriched in the secretome of the majority of *Phytophthora* species (Fig. 2 and Table S2).

With respect to GO term secretome enrichment within our data set, we observed results that corroborated our Pfam enrichment analysis. For example, enriched GO terms associated with plant cell wall degradation such as pectin activity (GO: 0030570 and GO: 0030599), cellulose activity (GO: 0030248 and GO: 0008810), polygalacturonase activity (GO: 004650), glucan catabolism (GO: 0009251), cellulose catabolism (GO: 0030245), cellulose metabolism (GO: 0030243), and carbohydrate binding (GO: 0030246) are widely distributed. Cell wall organization (GO: 0071555), modification (GO: 0042545), and biogenesis (GO: 0071554) are all enriched in the majority of *Phytophthora* species (see Table S2 and Fig. S1 in the supplemental material). Hydrolase activities acting on glycosyl bonds (GO: 0004553, GO: 0016787, and GO: 0016798) were found to be enriched in all 37 species (Table S2 and Fig. S1). Other terms that are ubiquitously enriched include defense response (GO: 0006952), pathogenesis (GO: 0009405), interspecies interactions (GO: 0044419), and multiorganism processes (GO: 0051704) (Table S2 and Fig. S1).

Previous researchers have undertaken similar analyses of some of the species within our data set, including *Ph. infestans* (47) and six *Pythium* species (38). We found broad agreement between our results and those previously reported for the *Ph. infestans* analysis. For example, we also observed enrichment in GO terms associated with

carbohydrate metabolic processes (GO: 0005975 and GO: 0016052), sugar metabolism (GO:0006040), sugar binding (GO: 0030246), sugar modification (GO: 0008810, GO: 0004650, GO: 0030570, and others), pathogenesis (GO: 0009405), defense response (GO: 0006952) proteolysis (GO: 0006508), and serine peptidase activity (GO: 0004252 and GO: 0008236) (Table S2). Similarly, we also observed Pfam domains associated with pectin degradation (PF03283, PF00544, and PF03211), elicitors (PF00964), Kazal-type domains (PF07648), and necrosis-inducing protein (PF05630) (Table S2). With respect to our results and those previously reported for *Pythium* species, the level of agreement is not as strong. We did observe enrichment in GO terms associated with pathogenesis, proteolysis, carbohydrate metabolic process (GO: 0006508), hydrolyase activity (GO: 0004553), and glycosyl hydrolyase activity (GO: 16798) but failed to detect enrichment in terms such nucleotide binding (GO: 000166), integral to membrane (GO: 0016021), transmembrane transport (GO: 0055085), and RNA processing (GO: 0006396) as previously reported (38).

The oomycete effector arsenal. We set out to investigate the abundance of effectors in oomycete species. The list of effectors considered in our analysis (Table S3) is based on a number of previous studies which described pathogenic proteins from oomycete species, including *Plasmopara* (14), *Phytophthora* (26), and *Pythium* (37) species. Any protein identified as having a pathogenicity-related domain was classified as a putative effector. These were combined with our secretome analysis to determine whether or not the effector was predicted to be secreted. The overall effector content of each species was quantified to detect expansions. The following counts excluded RxLR effectors as they are treated in more depth in a following section.

In total, 13,751 proteins were identified as having domains that could be implicated in pathogenicity (Table S3). Overall, 6,250 (~45%) of these were predicted to be secreted by SignalP (Table S3). Our results show that *Phytophthora* proteomes generally contain the highest frequency of effectors, with *Ph. sojae* (733 effectors) and *Ph. infestans* (646 effectors) possessing the largest arsenals of effectors, representing 2.76% and 3.63% of their total proteomes, respectively (Table S3). *Albugo* species were found to contain the smallest number of effectors. A trend was identified whereby hemibiotrophic species, such as *Phytophthora*, typically possess more effectors than saprotrophic species (members of the *Saprolegniales* order) and necrotrophs, such as *Pythium* species. Three species of obligate biotrophs, including *Al. candida*, *Al. laibachii*, and *H. arabidopsidis*, feature the lowest numbers of effectors, 156, 131, and 197, respectively (Table S3). Exceptions to this trend exist, most notably, *Ap. astaci* (554 effectors), *Py. insidiosum* (494 effectors), and *Py. oligandrum* (359 effectors), all of which contain large repertoires of effectors compared to other closely related oomycete species with similar lifestyles. For example, *Ap. astaci* was found to contain 554 effectors, in comparison to the closely related *Ap. invadans* species, which contained 272 (Table S3). A number of oomycete effectors are discussed in more detail in the following sections.

Necrosis-inducing proteins (NLPs). Necrosis-inducing proteins (NLPs) are apoplastic effectors found in bacteria, fungi, and oomycetes (48). The mechanisms by which NLPs act are not fully understood, but they are known to induce necrosis, trigger ethylene accumulation, and elicit immune responses in dicotyledons (49). A number of NLPs have previously been reported to be noncytotoxic but to act instead as microbe-associated molecular patterns (MAMPs) which are recognized by the plant hosts and result in the activation of the plant immune system (50).

Our InterProScan analysis detected 771 proteins with signatures of NLPs; 499 (67%) of these NLPs were predicted to contain signal peptides by SignalP (Table S3). Our results show that NLPs are absent from *Albugo*, *Aphanomyces*, and *Saprolegnia* species but are present in all *Peronosporales* and *Pythiales* species. Most *Pythium* species have fewer than seven copies. *Py. insidiosum* is the only *Pythium* species in our data set to lack NLPs. NLPs are highly expanded in *Phytophthora* species. In particular, *Ph. ramorum* has 69 copies, *Ph. parasitica* has 74 copies, and *Ph. sojae* has 80 copies (Table S3).

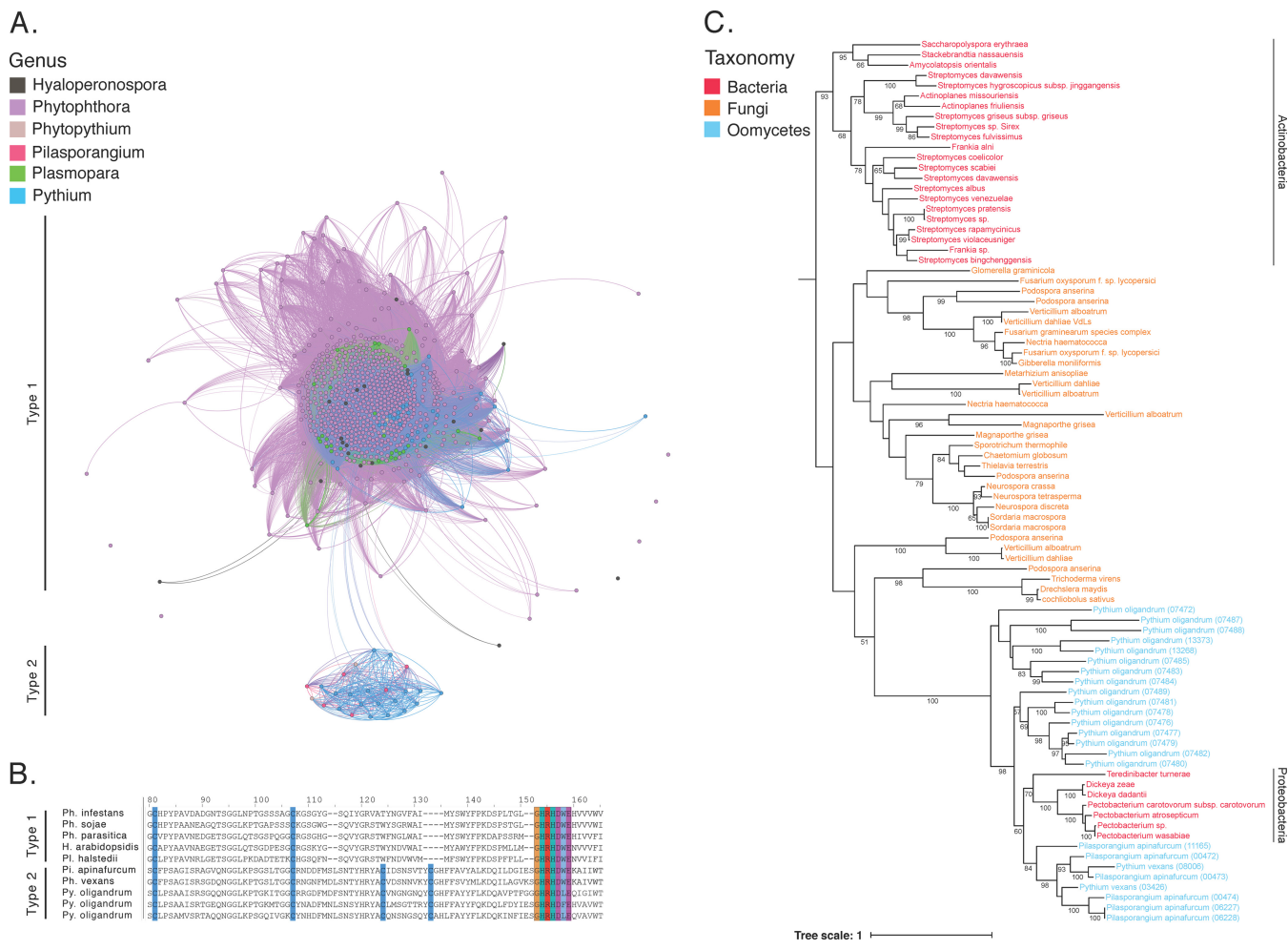


FIG 3 Analysis of oomycete necrosis-inducing proteins (NLP). (A) Homology network of 771 oomycete NLPs. Identified type 2 NLPs are arranged in a small, strongly connected cluster of 25 proteins from *Phytophthora vexans*, *Pilasporangium apinafurcum*, and *Pythium oligandrum*. (B) Multiple-sequence alignment of five type 1 NLPs and five type 2 NLPs. The two types share a conserved “GHRHDWE” NLP motif. Type 2 NLPs have an additional pair of cysteine residues. (C) Maximum-likelihood phylogeny of 87 NLPs containing 29 bacterial proteins, 33 fungal proteins, and 25 oomycete proteins. Bootstrap values greater than 50% are shown.

NLPs can be divided into two types: type 1 NLPs and type 2 NLPs (51). The two types share a conserved amino acid motif, “GHRHDWE.” They are distinguished by the presence of pairs of cysteine residues. Type 1 NLPs have one pair of conserved cysteines, while type 2 NLPs have two pairs of conserved cysteine residues (49). Each pair of cysteines could potentially form disulfide bridges, providing additional stability. Type 1 NLPs are found in bacteria, fungi, and oomycetes. Type 2 NLPs have been located in bacteria and fungi but were originally thought to be absent from oomycetes (49, 51). However, work by Horner et al. has shown that *Py. oligandrum* contains a type 2 NLP with similarity to a homolog from the proteobacterial plant pathogen *Pectobacterium atrosepticum* (52).

To further investigate the oomycete NLPs in our data set, we constructed a homology network of all 771 NLPs in our data set (Fig. 3A). An interactive version of the network is available online at <https://oomycetes.github.io>. In the network, the degree of a node is the number of edges it has connecting it to other nodes; therefore, the degree is the number of homologs that the protein has in our network. The overall NLP network (Fig. 3A) has an average degree value of 522.5, revealing that many of these NLPs are homologous to each other. 740 NLPs are grouped in a large, dense cluster (Fig. 3A). The network included six singletons that did not share significant sequence similarity with any other protein in the network. They included one *H. arabidopsidis*, one

Ph. cryptogea, one *Ph. kernoviae*, and three *Ph. parasitica* proteins. Interestingly, a smaller cluster of 25 NLPs was identified (Fig. 3A). Only five of these proteins were homologous to other proteins outside this cluster; this lack of homology is illustrated by the few edges seen between both clusters (Fig. 3A). The 25 proteins included 17 *Py. oligandrum*, 6 *Pilasporangium apinafurcum*, and 2 *Phytophthium vexans* putative NLPs. One of the *Py. oligandrum* proteins is an ortholog of the type 2 NLP (GenBank accession number [EV243877](https://www.ncbi.nlm.nih.gov/nuccore/EV243877)) previously reported by Horner et al. (52). A multiple-sequence alignment of a selection of proteins from both clusters was carried out (Fig. 3B). Inspection of this alignment revealed no significant sequence similarity between the two clusters except for the presence of the conserved “GHRHDWE” motif and conserved cysteine residues (Fig. 3B). The NLPs from the large cluster contained two conserved cysteine residues, indicating that they are type 1 NLPs. Proteins from the smaller cluster of 25 proteins contained four conserved cysteine residues, indicating that they are type 2 NLPs. Apart from the shared NLP motif and conserved cysteine residues, there is no significant sequence similarity between the two types of NLPs.

We set out to further investigate the evolutionary history of the putative oomycete type 2 NLPs. BLASTp searches of the 25 type 2 NLPs against the NCBI databases revealed top hits with proteobacterial species. To reconstruct the evolutionary history of these proteins, we used the 25 type 2 NLPs as bait sequences in a BLASTp homology search (E value cutoff of 10^{-20}) against a local database of 8,805,033 proteins, with broad taxon sampling across prokaryotes and eukaryotes (53). This search identified 87 homologous proteins, including 25 oomycete proteins (included in our data set), 33 fungal proteins, and 29 bacterial proteins. We aligned and manually edited the resulting homologs to give a final alignment of 460 amino acids and generated a maximum-likelihood phylogeny with PhyML using the Whelan and Goldman (WAG) model of substitution (Fig. 3C). This phylogeny places all fungal proteins in a single clan (41% bootstrap support). All oomycete proteins were located in a single clan with 100% bootstrap support (Fig. 3C). Within this large oomycete-containing clan, there was also a sister group clan of *Proteobacteria* proteins with 70% bootstrap support (Fig. 3C). There was also a completely separate actinobacterial clan with 93% bootstrap support (Fig. 3C).

The phylogenetic distribution of these type 2 NLPs is interesting. All of the fungal homologs are from the subphylum *Pezizomycotina*. Two bacterial phyla are represented, the *Actinobacteria* and the *Proteobacteria*. However, the *Proteobacteria* proteins are inferred to be more closely related to their oomycete homologs than to the actinobacterial homologs (Fig. 3C). One possible scenario is that horizontal gene transfer (HGT) has occurred during the evolutionary history of these proteins. However, due to the patch phyletic distribution, we cannot confidently infer the direction of gene transfer or indeed whether HGT has definitely occurred.

Immunoglobulin A peptidases. Immunoglobulin A peptidases represent a family of hydrolytic enzymes that cleave immunoglobulin A (IgA) and have been implicated as important virulence factors in bacterial infections of humans (54). InterProScan analysis revealed 40 oomycete proteins containing “IgA peptidase M64” domains (Table S3), 21 (53%) of which were predicted to contain signal peptides (Table S3). Expansions of these proteins are present in the genomes of both *Ap. astaci* and *Py. insidiosum*. Ten proteins containing this domain were found in *Ap. astaci* (6 with signal peptides) and 6 in *Py. insidiosum* (4 with signal peptides) (Table S3). Other *Pythium* species in our data set contained between one and five copies of the protein. No proteins with this domain were found in any *Albugo*, *Phytophthora*, *Plasmopara*, or *Saprolegnia* species (Table S3). Therefore, within our data set, IgA peptidases are unique to *Aphanomyces* species and the *Pythiales* order.

We aligned and manually curated all 40 IgA peptidase-containing proteins using MUSCLE and JalView, respectively, to give a final alignment of 770 amino acids. A maximum-likelihood phylogeny of all IgA peptidase-containing proteins was generated using a WAG substitution model and 100 bootstrap replicates (Fig. 4). Our phylogenetic

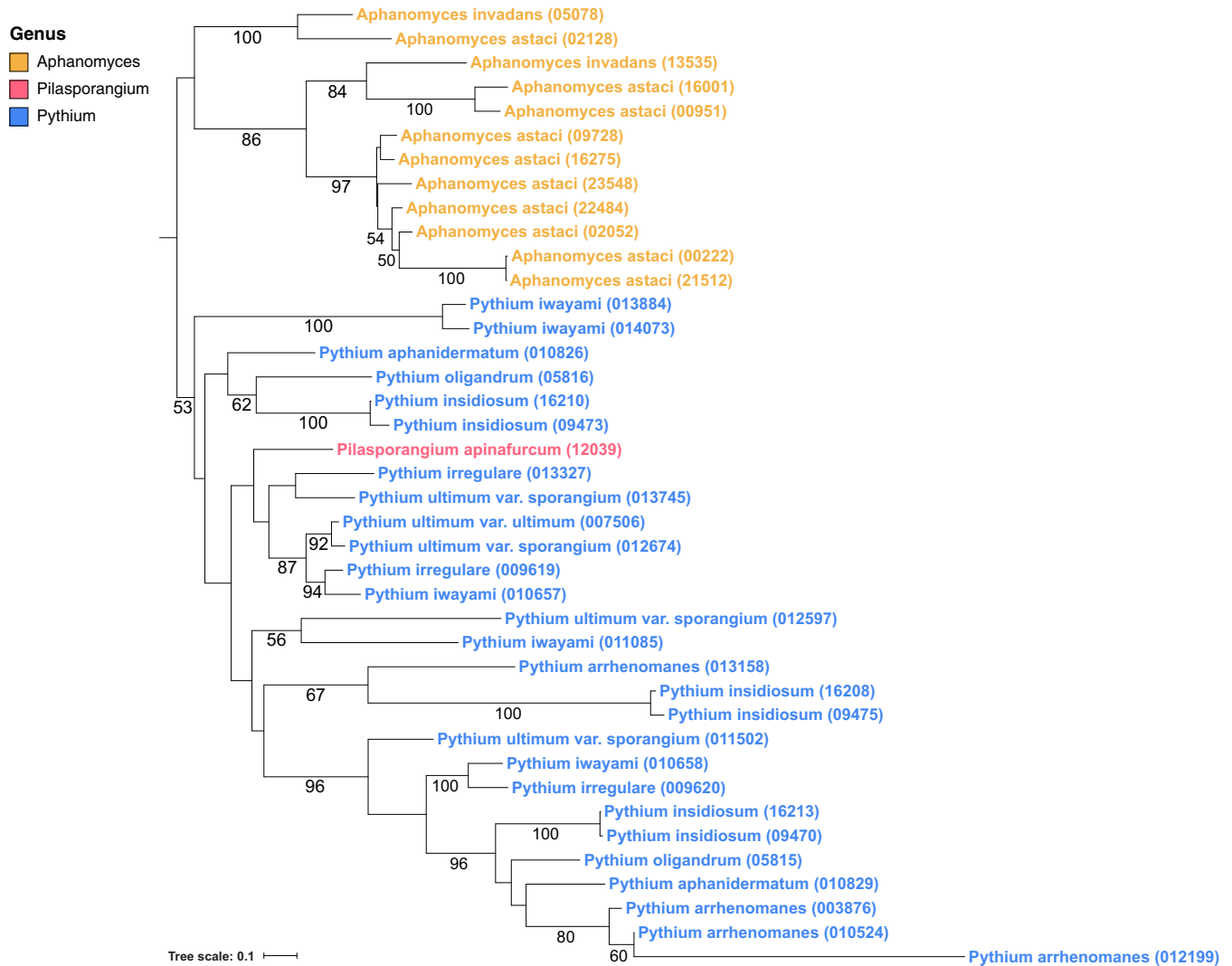


FIG 4 Maximum-likelihood phylogenetic reconstruction of oomycete IgA peptidases. Bootstrap values greater than 50% are shown.

reconstruction shows multiple species-specific duplication events which have led to the expansion of IgA peptidases in *Ap. astaci* and *Py. insidiosum* (Fig. 4). The IgA peptidases are grouped into two separate monophyletic clades. The first contains all *Pythiales* homologs, while the second contains all *Aphanomyces* homologs (Fig. 4).

Ap. astaci and *Py. insidiosum* are pathogens of crayfish and mammals, respectively (17, 55). It is tempting to speculate that these proteins possibly provide a defense mechanism for these species, allowing them to suppress the immune response of their animal hosts. However, it should be noted that a number of *Pythium* species with plant hosts were also shown to contain multiple copies of IgA peptidases, including five in *Py. iwayami* and four in both *Py. arrhenomanes* and *Py. ultimum* var. *sporangiferum* (Table S3).

Glycoside hydrolases. Glycoside hydrolases (GH), or glycosyl hydrolases, are hydrolytic proteins that cleave glycosidic bonds in complex sugars. They can be used to break down cellular components, for example, the cell membrane and cell wall. We set out to investigate the extent of expansions within the members of GH families across oomycete species. A total of 4,521 proteins with GH domains were located in our InterProScan analysis (Table S3 and S4). Expansions of GHs have been observed previously in oomycetes (38, 56). These were distributed across 47 different GH families. SignalP predicted secretion of 1,928 (42.6%) GHs. *Phytophthora* species possess the

highest number of GH proteins, with an average of 162, followed by *Aphanomyces* and *Saprolegnia* species, with averages of 106 and 98, respectively (Table S4). *Al. candida* and *Al. laibachii* feature smaller sets of GH proteins, with fewer than 60 each.

A number of GH families were found to be species or genus specific. For example, GH families 4, 26, 29, 42, and 48 are restricted to *Ph. kernoviae* whereas GH families 8, 24, and 44 are unique to *Ph. rubi*. Similarly, members of GH family 36 are found only in *Ph. kernoviae* and *Ph. rubi*. Members of GH families 20 and 27 are found only in members of the *Saprolegniales* order. GH families 18, 19, and 48 have known chitinase activity (57). Expansions of these families were revealed in *Ap. astaci*, *Py. oligandrum*, and *Saprolegnia* species. They are investigated in more detail below (see also Table S4).

Chitinases. Chitin is the second-most-abundant biopolymer in nature and is an important structural component of invertebrate exoskeletons and fungal cell walls (58). Chitinases are enzymes that degrade chitin. In fungi, chitinases play an important role in hyphal growth, spore germination, and cell wall remodeling (59). Oomycete cell walls, however, contain no or very small amounts of chitin (60). Thus, it would appear that chitinases encoded by oomycete species play roles primarily in the degradation of exogenous chitin, for example, in the breakdown of chitin in host cell walls.

Several GH families are known to have chitinase activity (57). These include GH families 18, 19, 23, and 48. A total of 281 proteins in our data set were reported to have putative chitinase activity in our InterProScan analysis; 112 (40%) of these are predicted to be secreted (Table S3). The chitinases identified were spread across three GH families, 18, 19, and 48. No members of GH family 23 were located in our data set. Only one protein was identified as a member of GH family 48; it belonged to *Ph. kernoviae*. However, it was not predicted to be secreted.

A large expansion of 50 chitinases was detected in *Ap. astaci* (Table S3). This is significantly more than the expansion seen with any other oomycete; 38 of these proteins belong to GH family 18, and the remaining 12 belong to GH family 19 (Table S4). *Ap. astaci* is a pathogen of crayfish, growing on and within the crayfish cuticle (55), which is composed of chitin. This expansion of chitinases may reveal a successful adaptation of *Ap. astaci* to its crayfish host, allowing the pathogen to penetrate through the chitin layers of crayfish cuticles. In agreement with previous work (56), our analysis also revealed large numbers of chitinases in the genome of *S. diclina* and *S. parasitica* (25 and 23, respectively) (Table S3). Additionally, large numbers of chitinases were located in *Al. laibachii* and *Py. oligandrum*. A total of 16 chitinases were identified in *Al. laibachii* and 15 in *Py. oligandrum* (Table S3). This is particularly significant for *Py. oligandrum*, which is a pathogen of fungi (18). Some fungal cell walls are composed of up to 20% chitin (61). Possessing a large repertoire of chitinase enzymes may prove useful for the pathogen for breaking down host fungal cells. Other oomycetes in the data set typically had fewer than 10 chitinases.

We constructed a homology network of all 281 oomycete chitinases in our data set to investigate their evolutionary history (Fig. S2). An interactive version of the network is available online at <https://oomycetes.github.io>. In the networks, nodes represent proteins and edges represent sequence similarity between two proteins. Our network consists of a number of disconnected clusters (no edges/homology to other clusters). We identified one singleton (a node with no edges, i.e., a protein with no homologs), the GH family 48 protein belonging to *Ph. kernoviae*. All GH family 19 proteins were placed into a single, strongly connected cluster (Fig. S2). No *Albugo* proteins were found in this cluster. GH family 18 proteins were divided into five clusters. *Saprolegnia* proteins dominated the network; in particular, proteins of *Ap. astaci* made up almost 18% of the network. Thus, our network analysis shows that oomycete chitinases are divided into a number of subclasses that are not homologous. This is consistent with the division of chitinases into 5 classes (62).

Our analysis also revealed 31 *Ap. astaci* proteins with N-terminal chitin-binding domains (Table S3); 25 (81%) of these were predicted to be secreted (Table S3). Other oomycetes contained fewer than 10 of these proteins, with *Al. candida*, *Al. laibachii*, *Ap.*

invadans, and *Ph. rubi* containing no proteins with this domain. BLAST homology searches of these proteins against the NCBI databases did not reveal any hits with annotated proteins. Their function is unknown; it is possible that the N-terminal domain facilitates attachment to the chitin exoskeleton of its crayfish host.

Proteases. A large number of proteins with hydrolytic activity were reported, including glycoside hydrolases, proteases, and pectin modifying proteins (Table S3). These are thought to be involved in the degradation of host cells (8). Several protease families were found to make up a large part of oomycete secretomes, including aspartyl proteases, papain family cysteine proteases, subtilases, trypsins, and trypsin-like proteases. Members of the *Saprolegniales* order contained the largest number of proteases, with each member possessing over 150 proteins with predicted protease activity (Table S3). *Pythium* species also had a large number of proteases, which may be attributable to their highly damaging necrotrophic lifestyles. *Pythium* species had an average of 139 proteases, higher than the *Phytophthora* average of 88 (Table S3). A large number of these proteases were reported to be secreted in our SignalP analysis, indicating that they may be involved in the breakdown of host cells. For example, *Saprolegnia* species have an average of 86 secreted proteases, *Pythium* species have an average of 55, and *Phytophthora* have an average of 24 (Table S3). Obligate biotrophs such as *Al. laibachii*, *Al. candida*, and *H. arabidopsidis* feature the fewest proteases, with each species possessing fewer than 60 proteases. Expansions of aspartyl proteases were detected in *Py. insidiosum* (95 proteases), *Ph. ramorum* (70 proteases), and *Ph. sojae* (68 proteases). *Ap. astaci* contains 161 proteins with predicted trypsin or trypsin-like activity, the largest number of any species in our data set. SignalP predicted 93 (58%) of these to be secreted (Table S3). Every other oomycete had fewer than 61 copies. Our analysis also identified an expansion of proteins with subtilase domains in *Aphanomyces*, *Pythium*, and *Saprolegnia* species. Again, *Ap. astaci* contained a large repertoire of 94 proteins with subtilase domains compared to *Phytophthora* species, all of which has fewer than 15 subtilases (Table S3). Our results show that the *Ap. astaci* proteome harbors vast number of proteases, significantly more than any other oomycete. The large arsenal of hydrolytic enzymes in *Ap. astaci*, and in other *Saprolegniales* members, may play an important role in the degradation of host cells.

Oomycete pectin modifying proteins. Pectin is a major component of plant cell walls, making up to 35% of primary walls in higher plants, and plays important roles in plant defense, development, and growth (63). A total of 1,048 oomycete proteins were found to contain domains that are involved in the modification or degradation of pectin, including 693 pectate lyases, 226 pectin esterases, and 129 pectin acetyl-esterases (Table S3). Pectate lyases are involved in the cleavage of pectin and result in fruit softening and rot via degradation of the plant cell wall (64). Proteins with pectate lyase domains were abundant in *Phytophthora* and *Pythium* species (Table S3). They are completely absent in *Al. laibachii*, *Ap. astaci*, and *Ap. invadans*. One copy was found in each of *Al. candida*, *S. diclina*, and *S. parasitica*, but the *S. diclina* copy was not predicted to be secreted. Pectinesterases, or pectin methylesterases, catalyze the de-esterification of pectin and are used by plants in a wide range of biological processes, including cell wall remodeling, fruit ripening, pollen growth, and root development (65). However, they can also be exploited by pathogens to invade plant tissues (66). Our results show that pectin esterases were present only in members of the *Peronosporales* order in our data set (Table S3). The majority of these proteins contain signal peptides (Table S3), indicating that they are effectors that potentially cause damage to the cell walls of their hosts. Similarly, pectin acetyl-esterases can be exploited by pathogens to catalyze the deacetylation of pectin, making the pectin backbone more accessible to pectin-degrading enzymes such as pectate lyases (67). Proteins with the pectin acetyl-esterase domain were found in *Phytophthora*, *Phytophythium*, *Pilasporangium*, *Plasmopara*, and *Pythium* species (Table S3). Species with more pectate lyases typically had more pectin acetyl-esterases, and the correlation was strong and significant ($R^2 = 0.600$, $P < 0.00001$ [Pearson correlation test]).

Cutinases. Cutin is one of the main components of plant cuticles (68). The plant cuticle acts as a physical barrier, allowing plant cells to tolerate external environmental stresses and also protecting interior plant tissues from invading pathogens (69). Cutinases are extracellular enzymes that hydrolyze cutin and have been identified in bacteria, fungi, and oomycetes (70). They can be used by plant pathogens to compromise the structural integrity of the plant cuticle, allowing them to penetrate and infect inner plant tissues. Enzymatic digestion of cutin has been proven to be an essential initial step in the infection process of plant pathogens (71).

Overall, we have identified 122 proteins with cutinase domains in our data set, 79 (65%) of which were predicted to be secreted (Table S3). Cutinases were completely absent in *Aphanomyces* and *Saprolegnia* species. *Pythium* species have previously been reported to lack cutinases (72). We found this to be true for the majority of *Pythium* species in our data set; however, we identified nine cutinases in *Py. aphanidermatum* and seven cutinases in *Py. arrhenomanes* (Table S3). Both *Albugo* species in our data set also contained multiple copies. *Ph. pinifolia* was the only *Peronosporales* member to lack cutinases.

Toxin families. We identified a number of toxin families in our data set, including necrosis-inducing proteins (see previous section) and members of the phytotoxic protein family (PcF). Relative to the level seen with other effectors, the overall number of PcFs detected in our data set was low. PcFs are known to induce necrosis (73). In total, 31 proteins with signatures of PcF proteins were identified and were unique to several members of the *Peronosporales*; 19 (61%) of these were predicted to be secreted (Table S3). *Ph. infestans* contains 14 PcF proteins, 5 of which are predicted to be secreted (Table S3). *Ph. capsici* contains six PcF proteins; *Ph. sojae* has four; *Ph. parasitica* has two; and *Ph. multivora*, *Ph. pisi*, and *Ph. ramorum* have one copy each. All of these are predicted to be secreted (Table S3). The proteome of *Ph. lateralis* was also reported to have one PcF protein, but it is not predicted to be secreted. *H. arabidopsidis* was the only non-*Phytophthora* species reported to have a PcF protein (Table S3). However, it was not predicted to be secreted. This finding suggests that PcF proteins are unique to *Phytophthora* and closely related species, suggesting that they may have arisen in the last common ancestor of the *Peronosporales* order.

Crinklers. A combination of regular expression searches and hidden Markov models was utilized to identify Crinkler effectors (CRNs). After manual inspection and removal of suspected false positives, a total of 899 CRNs were identified in our data set. Our results highlight that *Phytophthora* species possess large expansions of CRNs, more than the species of any other genera. In particular, we found 91, 92, and 167 CRNs in *Ph. capsici*, *Ph. sojae*, and *Ph. infestans*, respectively (Table S3). *Pilastrangium apinafurcum* and *Pl. viticola* also feature large numbers of CRNs relative to other species, having 52 and 65 CRNs each, respectively. CRN numbers were sparse in the *Albuginales* and *Saprolegniales* orders, with some species (*Ap. invadans*, *S. diclina*, and *S. parasitica*) possessing only one copy (Table S3). Thus, it would appear that CRNs play an important role in plant infection for *Phytophthora* species. Only 177 CRNs were predicted to be secreted in our SignalP analysis (Table S3). However, previous reports have indicated that a large number of CRNs may be secreted via unconventional protein secretion systems that cannot be predicted *in silico* (74).

Oomycete protease inhibitors. Plants and plant pathogens are constantly undergoing an evolutionary arms race with one another (75). Production and secretion of proteases by plants to degrade pathogen effectors are among the examples of this. To counteract this, plant pathogens, including oomycetes, have coevolved to secrete protease inhibitors (4). These protease inhibitors block the defensive mechanism of plant proteases. We have identified a number of secreted oomycete protease inhibitors. The most abundant were Kazal-type protease inhibitors (507 in total; 355 were predicted to be secreted) and cathepsin propeptide inhibitors (155 in total; 95 were predicted to be secreted) (Table S3). Counts of cathepsin propeptide inhibitors were consistent across our oomycete data set, with most species possessing fewer than 5

copies (Table S3). Kazal-type protease inhibitors were more abundant and not evenly distributed; for example, expansions were recorded in several species, including *Py. insidiosum* (43 copies), *Ph. sojae* (39 copies), and *Ph. infestans* (33 copies). Most other species contained fewer than 15 Kazal-type protease inhibitors. *Albugo* species have fewer than five copies each (Table S3).

RxLR effectors. Numerous *in silico* analyses performed using various bioinformatics strategies have been employed in previous studies to identify candidate RxLR effectors in oomycetes. The most liberal, first described by Win et al., is a strategy in which all possible open reading frames (ORFs) are examined for the presence of a signal peptide within the first 30 amino acid residues followed by an RxLR motif between residues 30 and 60 (Win method) (76). Extensions to this method have been developed (27, 76) with searches for a downstream EER motif (or a loose match); the EER motif is present in numerous validated RxLR effectors (regular expression or Regex) (27). Hidden Markov model (HMM) profiles derived from alignments of RxLR-EER effectors have also been implemented successfully (27). The initial genome analyses that described the *Ph. infestans* RxLR complement utilized all three approaches described above as well as additional criteria such as exhibiting sequence homology to previously known RxLRs or belonging to protein families where the majority of proteins are deemed putative RxLRs based on Win, Regex, or HMM criteria. Depending on the method or combination of methods utilized, it is possible to detect a broad range of potential RxLR effectors (8). As RxLR effectors are most abundant and have been best characterized in *Phytophthora* species (8), a *Phytophthora*-biased approach was taken to identify candidate oomycete RxLR effectors. To be classified as a putative RxLR effector, proteins or ORFs had to meet one of the RxLR criteria in analyses performed using the Win method, HMMsearch, Regex, or the Homologous method (see Materials and Methods).

The four RxLR criteria were tested on the predicted proteomes of each of the 37 oomycete species in our data set. Utilizing predicted proteins adds an additional criterion layer but may potentially miss open reading frames that were missed by gene callers during annotation. Additional criteria utilized by others such as belonging to a TribeMCL cluster (8) were not considered due to the associated computational costs. It is obvious that some of the criteria described above may detect a large number of false positives. For example, a number of the 563 ORFs counted as putative RxLRs in *Ph. infestans* were found by Blast homology alone and do not contain a RxLR domain; using these as “bait” in a BlastP search against other oomycetes will locate non-RxLR domain-containing homologs. However, we have noticed that when the RxLR repertoires of model genomes such as *Ph. infestans*, *Ph. sojae*, and *Ph. ramorum* are referenced in the literature, the researchers normally consider homologs of known RxLRs located via BlastP alone to represent putative RxLRs. Our analysis extends this as we use all proteins from these model genomes as bait sequences in our analysis. We also searched for the presence of repeating sequence motifs termed “W,” “Y,” and “L” that are found toward the C terminus of a number of *Phytophthora* cytoplasmic effectors (31). These domains form an alpha-helical fold known as the WY fold that may provide structural flexibility (32). However, due to the sequence divergence observed in *Phytophthora* RxLRs, it seems that alternative folds most likely exist (77).

Putative RxLRs along with sequence information and the criteria that were satisfied in classifying the proteins as putative RxLRs are listed in Table S6. For completeness and to allow comparisons to previous analyses, we also searched all putative ORFs that were nonoverlapping and more than 70 amino acids in length using the criteria described above (Table S7). We do not discuss the results of these counts here and instead concentrate on counts related to predicted protein coding genes. The purpose of this is to maintain consistency, as our secretome analysis also considered putative protein coding genes and did not look at all possible open reading frames.

In total, 4,131 proteins in our data set matched one or more of the RxLR criteria, ranging from a low of 6 proteins in *Al. candida* to a high of 603 in *Ph. infestans* (Table S5). Unsurprisingly, the vast majority (3,600 or 87%) of the 4,131 proteins are

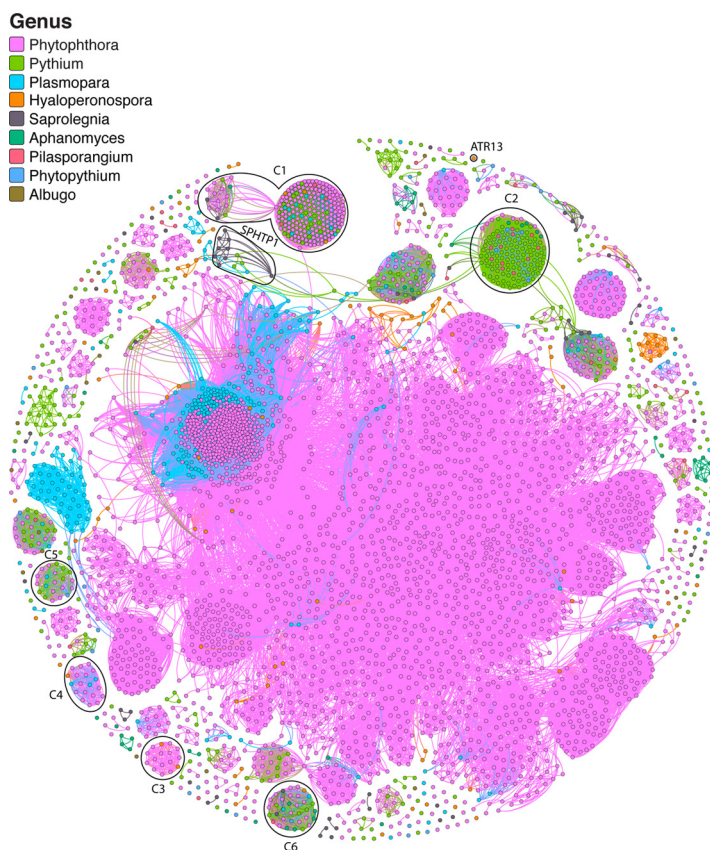


FIG 5 Homology network of 4,131 putative oomycete RxLR effectors. Each protein is represented by a node. An edge joining two nodes represents sequence similarity shared by the two proteins (E value cutoff of 10^{-10}). Many disconnected clusters of RxLR effectors are represented, indicating no significant homology between clusters. Nodes are colored by genus. Clusters discussed in the text are labeled. An interactive version of this network is available online at <https://oomycetes.github.io/rxlr.html>.

located in *Peronosporales* species (Table S5). A homology network of the RxLR-like proteins was generated to investigate the evolutionary relationships within the putative RxLRs (Fig. 5). An interactive and searchable version of the network is available online at <https://oomycetes.github.io>. The online version of the network permits users to query the network based on the protein identifier (ID) and to retrieve sequence information as well as performing BlastP searches against the NCBI database. Furthermore, users can filter proteins based on the genus or the RxLR criteria used (Win, Regex, HMM, or BLAST). Proteins can also be viewed based on the presence or absence of the WY fold. The network is composed of 4,131 nodes; each node corresponds to an individual protein and 184,302 edges; edges link nodes if they are homologous. The putative RxLRs were clustered into 357 connected components (Fig. 5). The nodes had an average degree of 89 (i.e., on average, each protein in the network has 89 homologs). The fact that the network is comprised of a number of disconnected clusters (individual clusters of connected nodes that are not connected to other clusters) shows there are no significant sequence similarities between some clusters (Fig. 5).

A significant proportion (1,852 or 44.8%) of the 4,131 proteins have only one line of evidence labeling them as RxLRs. For example, 1,045 (25.3%) were located based on BlastP homology alone (therefore, the RxLR domain is absent), 683 (16.5%) based on the Win method alone, 25 (0.6%) based on the HMM alone, and 99 (2.4%) on Regex alone (Fig. 6 and Fig. S3). Of the 1,045 proteins that were located by sequence homology exclusively (no evidence of RxLR motif), 981 were found in species from the *Peronosporales* order and may not be functional RxLRs, although 287 of these homologs were found to contain the WY fold (Table S6). Conversely, 967 (23.4%) of all proteins

C6 contains 33 RxLRs, with proteins from every genus in our data set except *Albugo*. All of these proteins have homology to reference RxLRs (Table S6) and are classified RxLRs based on the Win or Regex criteria. The group forms a clique (a subnetwork where each member is connected to every other member), showing that every protein in the group is homologous to every other protein. On average, these proteins have sequence similarity of 52%. The position of the RxLR motif is ubiquitously conserved (not shown). However, the majority of these proteins contain a KDEL endoplasmic reticulum (ER) retention motif at the C terminus as previously reported for the *Py. oligandrum* protein represented in this cluster (52). Proteins with this ER motif are not predicted to be secreted and therefore likely represent false positives (52). Overall, 75 of the 4,131 putative RxLRs contain the KDEL motif (or a variation) (Table S6).

Of the 4,131 putative RxLRs, 3,011 were located based on Win, Regex, or HMM results, meaning that 1,120 have been labeled due to homology to a reference RxLR alone, or they may also contain a KDEL retention motif and so may not be secreted. Ignoring these 1,120 proteins, we observe that *Ph. infestans* is predicted to contain 470 RxLRs, a figure that is consistent with the figure of 563 putative RxLRs reported based on analysis of ORFs (8). Furthermore, 95 of these 563 ORFs were labeled based on homology alone, therefore giving an overall number of 468 if non-RxLR-containing homologs are excluded. Similarly, the comparison between our *Ph. sojae* RxLR prediction number and the corresponding number reported from previous studies (338 versus 312) is consistent.

Conclusions. The first oomycete genome sequences were published in 2009 (8), and at the time of writing there were 37 oomycete genome assemblies publically available. Due to their importance as pathogens of economically important crops and animals, along with the ongoing advances and reductions in costs associated with next-generation sequencing technologies, this number is expected to increase dramatically over the coming years. In this analysis, we performed an inventory of known oomycete effectors in all available genome sequences. As well as quantifying their occurrences, we have in a number of cases also investigated their evolutionary history.

This genome-wide survey provides an up-to-date inventory of previously described effectors in the class *Oomycota*. It is by no means a complete list, and we are cognizant that many additional effectors will be characterized in the coming years, especially with improved host-pathogen interaction omics studies. However, it does provide the current overview of the arsenal of known effectors in this economically important class of animal and plant pathogens.

We have also examined the presence and absence of glycoside hydrolases and found a diverse range of families across the oomycote class. The majority (54%) of these are secreted. Glycoside hydrolase families 18, 19, and 48 all have chitinase activity. A homology network (Fig. S2) showed that there is no sequence homology between family members 18 and 19. Interestingly, family 18 can be subdivided into 4 distinct clusters, indicating that while family members have the same enzymatic function, they do not share sequence similarity, confirming that family 18 has different subclasses. Our analysis has also detected the presence of immunoglobulin A peptidases in *Pythiales* species and *Aphanomyces* species (Fig. 4). Some of these species are animal pathogens, and it is possible that their presence may be important in suppressing the immune response of their animal hosts.

We have also catalogued the putative RxLR repertoire of all 37 oomycete species. Our results are consistent with those of other studies in showing that *Phytophthora* species have undergone expansions in these proteins. We are aware that this analysis may be reporting false positives, particularly for species outside the *Peronosporales* order. However, a number of previously described RxLR effectors such as *H. arabidopsisidis* ATR1 and ATR13 were detected, as was SPHTP1 *S. parasitica*. We are also aware that we may be underreporting the number of RxLRs, as we did not analyze all possible open reading frames and instead concentrated on predicted protein coding genes; this decision was taken because our secretome analysis also used protein coding genes.

Furthermore, using open reading frames could in itself lead to the reporting of false positives by reporting pseudogenes or noncoding regions of the genome. The confirmation of all putative RxLRs is beyond the scope of this *in silico* catalogue; however, detailed sequence information on all proteins is provided in the supplemental material.

MATERIALS AND METHODS

Data set assembly. The predicted proteomes for 23 oomycete species were obtained from public databases (Table 1). Predicted proteomes for another 14 oomycete species (10 *Phytophthora* species and *Pl. viticola*, *Pi. apinafurcum*, *Py. insidiosum*, and *Py. oligandrum*) were generated from publically available data using AUGUSTUS (82) (Table 1). Templates for *ab initio* protein prediction were generated for AUGUSTUS using assembly and expressed sequence tag (EST) data from a number of reference oomycete species. *Ph. capsici* was used as a reference for *Ph. agathidicida*, *Ph. multivora*, *Ph. pluvialis*, and *Ph. taxon totara*. *Ph. sojae* was used as a reference for *Ph. cinnamomi*, *Ph. cryptogea*, *Ph. fragariae*, *Ph. pinifolia*, *Ph. pisi*, and *Ph. rubi*. *Pl. halstedii* was used as a reference for *Pl. viticola*. *Py. ultimum* var. *sporangiiiferum* was used as a reference for the two *Pythium* species. GeneMark-ES (83) was used in addition to AUGUSTUS for predicting proteins of *Pi. apinafurcum*. The final data set contained 590,896 proteins from 37 predicted oomycete proteomes (Table 1). All proteomes used in this analysis as well as pipelines and scripts for identification of effectors, secretomes, and RxLRs are available at <https://github.com/oomycetes/oomycetes.github.io/tree/master/SupplementaryMaterial>.

Identification of putative effectors. InterProScan 5 (41) was run on all 590,896 predicted oomycete proteins in our data set. Any proteins reported by InterProScan as having a Pfam domain that could be implicated in pathogenesis were classified as potential effectors (see Table S3 in the supplemental material). The list of pathogenic Pfam domains considered was based on a number of previous analyses (14, 26, 37).

Identification of putatively secreted proteins. Transmembrane domain prediction was carried out for all 590,896 proteins using THMM (84), and signal peptides were predicted using SignalP v3 (85). Proteins that had an HMM S probability value of ≥ 0.9 , an NN Y_{\max} score of ≥ 0.5 , and an NN D score of ≥ 0.5 with predicted localization "Secreted" and no transmembrane domain after the signal peptide cleavage site were considered to be putatively secreted.

Enrichment analyses. Enrichment for particular Pfam and GO terms was undertaken in Blast2GO (86) by comparing the frequencies of GO terms and Pfam domains in the secretome to those in the nonsecreted proportion of the proteome using Fisher's exact test corrected for false-discovery rate (FDR) (40). Only Pfam domains or GO terms with enrichment *P* values of < 0.05 are reported.

Identification of putative RxLR effectors. Proteins were classified as putative RxLR effectors if they satisfied one of the following four criteria. (i) In the Win method, the protein must contain a signal peptide in residues 1 to 30 followed by an RxLR motif within residues 30 to 60 (76). (ii) In the HMMsearch method, hidden Markov model analysis was performed for all proteins predicted to be secreted to detect the RxLR motif using the "cropped.Hmm" HMM profile constructed by Whisson et al. (27). This accounts for variations in the RxLR and EER motifs. Matches with a bit score of > 0 were retained. (iii) In the Regex method, the protein must contain a signal peptide between residues 10 to 40 and an RxLR motif within the following 100 residues followed by the EER motif within 40 residues downstream of the RxLR motif (allowing replacements of E to D and R to K) (27). The regular expression used was as follows: $^{\wedge}\{10,40\}\{1,96\}R.LR.\{1,40\}[ED][ED][KR]$. (iv) In the Homologous method, the set of 1,207 putative *Phytophthora* RxLR effectors was downloaded for *Ph. infestans*, *Ph. ramorum*, and *Ph. sojae* (8). These were used as reference RxLRs in the RxLR homology search. Proteins located via a BlastP search (*E* value cutoff, 10^{-20}) corresponding to a reference *Phytophthora* RxLR were considered putative RxLRs.

A HMM search was run on all candidate RxLRs to determine if they contain the WY fold using the HMM method developed by Boutemy et al. (77).

Identification of Crinkler effectors. String searches were performed to account for variations in the Crinkler "LxLFLAK" motif. First, a search was carried out with the regular expression $^{\wedge}\{30,70\}LFLA[RK]$. All hits were aligned using MUSCLE (v3.8.31) (87). A hidden Markov model (HMM) was built for the alignment using HMMER (3.1) (88). The HMM was searched against our entire data set using hmmsearch to identify homologs. A second string search was carried out using the regular expression $^{\wedge}\{30,70\}LYLA[RK]$. Again, hits from this search were aligned using MUSCLE and an HMM was built and searched against our data set. The two results were merged as our candidate Crinkler effector set. These candidates were manually inspected, and any proteins that did not contain an obvious "LxLFLAK-like" motif were excluded.

Generation of homology networks. Homology networks were generated for a number of protein families. In each instance, an all-versus-all BLASTp search (89) was run against each member of the family with an *E* value cutoff of 10^{-10} for the chitinase and NLP networks and an *E* value cutoff of 10^{-5} for the RxLR network. Each protein was represented by a node in the network. Two proteins were connected by an undirected edge if they were identified as homologous in our all-versus-all BLASTp search. The network was visualized and annotated in Gephi (90) and arranged using the Fruchterman-Reingold layout (91). Network statistics were calculated within Gephi. Online interactive versions of all networks are available at <https://oomycetes.github.io>. Protein/node information is available for viewing in the network. The networks can be filtered to hide/show particular proteins by protein ID, species, genus, or order.

Maximum-likelihood phylogenetic reconstruction of effector families. For the NLP phylogeny, the 25 type 2 NLPs were used as bait sequences in a BLASTp (89) homology search (*E* value cutoff of

10^{-20}) against a local database of 8,805,033 proteins (53). All homologs were aligned using MUSCLE (87) and manually edited, giving a final alignment of 460 amino acids. ModelGenerator inferred that the optimum model of substitution was the WAG model of substitution (92). A maximum-likelihood phylogeny was reconstructed in PhyML using this model, and 100 bootstrap replicates were undertaken (93). The final tree was visualized and annotated with iTOL (94).

For the IgA peptidase-containing proteins, all 40 were aligned using MUSCLE and edited to give a final alignment of 770 amino acids. ModelGenerator inferred that the optimum model of substitution was the WAG model of substitution. A maximum-likelihood phylogeny of all IgA peptidase-containing proteins was generated using a WAG substitution model in PhyML, and 100 bootstrap replicates were undertaken. The final tree was visualized and annotated with iTOL (94).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSphere.00408-17>.

- FIG S1, PDF file, 0.8 MB.
- FIG S2, PDF file, 0.2 MB.
- FIG S3, JPG file, 2 MB.
- TABLE S1, XLSX file, 0.04 MB.
- TABLE S2, XLSX file, 0.3 MB.
- TABLE S3, XLSX file, 0.1 MB.
- TABLE S4, XLSX file, 0.1 MB.
- TABLE S5, XLSX file, 0.05 MB.
- TABLE S6, XLSX file, 1.2 MB.
- TABLE S7, XLSX file, 2.7 MB.

ACKNOWLEDGMENTS

We acknowledge the DJEI/DES/SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support. We also thank Charley McCarthy for the provision of data sets.

J.M. is funded by a postgraduate scholarship from the Irish Research Council, Government of Ireland (grant number GOIPG/2016/1112).

REFERENCES

1. Beakes GW, Glockling SL, Sekimoto S. 2012. The evolutionary phylogeny of the oomycete “fungi.” *Protoclasma* 249:3–19. <https://doi.org/10.1007/s00709-011-0269-2>.
2. Richards TA, Dacks JB, Jenkinson JM, Thornton CR, Talbot NJ. 2006. Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. *Curr Biol* 16:1857–1864. <https://doi.org/10.1016/j.cub.2006.07.052>.
3. Burki F. 2014. The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb Perspect Biol* 6:a016147. <https://doi.org/10.1101/cshperspect.a016147>.
4. Thines M, Kamoun S. 2010. Oomycete-plant coevolution: recent advances and future prospects. *Curr Opin Plant Biol* 13:427–433. <https://doi.org/10.1016/j.pbi.2010.04.001>.
5. Diéguez-Urbeondo J, García MA, Cerenius L, Kozubíková E, Ballesteros I, Windels C, Weiland J, Kator H, Söderhäll K, Martín MP. 2009. Phylogenetic relationships among plant and animal parasites, and saprotrophs in Aphanomyces (oomycetes). *Fungal Genet Biol* 46:365–376. <https://doi.org/10.1016/j.fgb.2009.02.004>.
6. Hulvey JP, Padgett DE, Bailey JC. 2007. Species boundaries within Saprolegnia (Saprolegniales, Oomycota) based on morphological and DNA sequence data. *Mycologia* 99:421–429.
7. Banfield MJ, Kamoun S. 2013. Hooked and cooked: a fish killer genome exposed. *PLoS Genet* 9:e1003590. <https://doi.org/10.1371/journal.pgen.1003590>.
8. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AMV, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JIB, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, Fischbach MA, Fugelstad J, Gilroy EM, Gnerre S, Green PJ, Grenville-Briggs LJ, Griffith J, Grünwald NJ, Horn K, Horner NR, Hu C-H, Huitema E, Jeong D-H, Jones AME, Jones JDG, Jones RW, Karlsson EK, Kunjeti SG, Lamour K, Liu Z, et al. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461:393–398. <https://doi.org/10.1038/nature08358>.
9. Cooke DE, Drenth A, Duncan JM, Wagels G, Brasier CM. 2000. A molecular phylogeny of *Phytophthora* and related oomycetes. *Fungal Genet Biol* 30:17–32. <https://doi.org/10.1006/fgbi.2000.1202>.
10. Blair JE, Coffey MD, Park SY, Geiser DM, Kang S. 2008. A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. *Fungal Genet Biol* 45:266–277. <https://doi.org/10.1016/j.fgb.2007.10.010>.
11. McCarthy CGP, Fitzpatrick DA. 2017. Phylogenomic reconstruction of the oomycete phylogeny derived from 37 genomes. *mSphere* 2:e00095-17. <https://doi.org/10.1128/mSphere.00095-17>.
12. Coates ME, Beynon JL. 2010. *Hyaloperonospora arabidopsidis* as a pathogen model. *Annu Rev Phytopathol* 48:329–345. <https://doi.org/10.1146/annurev-phyto-080508-094422>.
13. Gascuel Q, Martinez Y, Boniface MC, Vear F, Pichon M, Godiard L. 2015. The sunflower downy mildew pathogen *Plasmopara halstedii*. *Mol Plant Pathol* 16:109–122. <https://doi.org/10.1111/mpp.12164>.
14. Sharma R, Xia X, Cano LM, Evangelisti E, Kemen E, Judelson H, Oome S, Sambles C, van den Hoogen DJ, Kitner M, Klein J, Meijer HJG, Spring O, Win J, Zipper R, Bode HB, Govers F, Kamoun S, Schornack S, Studholme DJ, Van den Ackerveken G, Thines M. 2015. Genome analyses of the sunflower pathogen *Plasmopara halstedii* provide insights into effector evolution in downy mildews and *Phytophthora*. *BMC Genomics* 16:741. <https://doi.org/10.1186/s12864-015-1904-7>.
15. Links MG, Holub E, Jiang RHY, Sharpe AG, Hegedus D, Beynon E, Sillito D, Clarke WE, Uzuhashi S, Borhan MH. 2011. De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. *BMC Genomics* 12:503. <https://doi.org/10.1186/1471-2164-12-503>.
16. Kemen E, Gardiner A, Schultz-Larsen T, Kemen AC, Balmuth AL, Robert-

- Seilianantz A, Bailey K, Holub E, Studholme DJ, MacLean D, Jones JDG. 2011. Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. *PLoS Biol* 9:e1001094. <https://doi.org/10.1371/journal.pbio.1001094>.
17. Gaastra W, Lipman LJA, De Cock AWAM, Exel TK, Pegge RBG, Scheurwater J, Vilela R, Mendoza L. 2010. *Pythium insidiosum*: an overview. *Vet Microbiol* 146:1–16. <https://doi.org/10.1016/j.vetmic.2010.07.019>.
 18. Benhamou N, le Floch G, Vallance J, Gerbore J, Grizard D, Rey P. 2012. *Pythium oligandrum*: an example of opportunistic success. *Microbiology* 158(Pt 11):2679–2694.
 19. de Cock AWAM, Lodhi AM, Rintoul TL, Bala K, Robideau GP, Abad ZG, Coffey MD, Shahzad S, Lévesque CA. 2015. Phytophythium: molecular phylogeny and systematics. *Persoonia* 34:25–39. <https://doi.org/10.3767/003158515X685382>.
 20. Wawra S, Belmonte R, Löblich L, Saraiva M, Willems A, van West P. 2012. Secretion, delivery and function of oomycete effector proteins. *Curr Opin Microbiol* 15:685–691. <https://doi.org/10.1016/j.mib.2012.10.008>.
 21. Kamoun S. 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu Rev Phytopathol* 44:41–60. <https://doi.org/10.1146/annurev.phyto.44.070505.143436>.
 22. Giraldo MC, Valent B. 2013. Filamentous plant pathogen effectors in action. *Nat Rev Microbiol* 11:800–814. <https://doi.org/10.1038/nrmicro3119>.
 23. Jiang RHY, Tyler BM. 2012. Mechanisms and evolution of virulence in oomycetes. *Annu Rev Phytopathol* 50:295–318. <https://doi.org/10.1146/annurev-phyto-081211-172912>.
 24. Tian M, Huitema E, da Cunha L, Torto-Alalibo T, Kamoun S. 2004. A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. *J Biol Chem* 279:26370–26377. <https://doi.org/10.1074/jbc.M400941200>.
 25. Tian M, Win J, Song J, van der Hoorn R, van der Knaap E, Kamoun S. 2007. A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apolipase protease. *Plant Physiol* 143:364–377. <https://doi.org/10.1104/pp.106.090050>.
 26. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee MK, McDonald WH, Medina M, Meijer HJ, Nordberg EK, Maclean DJ, Ospina-Giraldo MD, Morris PF, Phuntumart V, Putnam NH, Rash S, Rose JK, Sakihama Y, Salamov AA, Savidor A, Scheuring CF, Smith BM, Sobral BW, Terry A, Torto-Alalibo TA, Win J, Xu Z, Zhang H, et al. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313:1261–1266. <https://doi.org/10.1126/science.1128796>.
 27. Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, Gilroy EM, Armstrong MR, Grouffaud S, van West P, Chapman S, Hein I, Toth IK, Pritchard L, Birch PR. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450:115–118. <https://doi.org/10.1038/nature06203>.
 28. Bhattacharjee S, Hiller NL, Liolios K, Win J, Kanneganti TD, Young C, Kamoun S, Haldar K. 2006. The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PLoS Pathog* 2:e50. <https://doi.org/10.1371/journal.ppat.0020050>.
 29. Haldar K, Kamoun S, Hiller NL, Bhattacharjee S, van Ooij C. 2006. Common infection strategies of pathogenic eukaryotes. *Nat Rev Microbiol* 4:922–931. <https://doi.org/10.1038/nrmicro1549>.
 30. Dou D, Kale SD, Wang X, Jiang RHY, Bruce NA, Arredondo FD, Zhang X, Tyler BM. 2008. RXLR-mediated entry of *Phytophthora sojae* effector *Avr1b* into soybean cells does not require pathogen-encoded machinery. *Plant Cell* 20:1930–1947. <https://doi.org/10.1105/tpc.107.056093>.
 31. 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. <https://doi.org/10.1073/pnas.0709303105>.
 32. Win J, Krasileva KV, Kamoun S, Shirasu K, Staskawicz BJ, Banfield MJ. 2012. Sequence divergent RXLR effectors share a structural fold conserved across plant pathogenic oomycete species. *PLoS Pathog* 8. <https://doi.org/10.1371/journal.ppat.1002400>.
 33. Anderson RG, Deb D, Fedkenheuer K, McDowell JM. 2015. Recent progress in RXLR effector research. *Mol Plant Microbe Interact* 28: 1063–1072. <https://doi.org/10.1094/MPMI-01-15-0022-CR>.
 34. Schornack S, van Damme M, Bozkurt TO, Cano LM, Smoker M, Thines M, Gaulin E, Kamoun S, Huitema E. 2010. Ancient class of translocated oomycete effectors targets the host nucleus. *Proc Natl Acad Sci U S A* 107:17421–17426. <https://doi.org/10.1073/pnas.1008491107>.
 35. Stam R, Jupe J, Howden AJM, Morris JA, Boevink PC, Hedley PE, Huitema E. 2013. Identification and characterisation CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS One* 8:e59517. <https://doi.org/10.1371/journal.pone.0059517>.
 36. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, Thines M, Ah-Fong A, Anderson R, Badejoko W, Bittner-Eddy P, Boore JL, Chibucos MC, Coates M, Dehal P, Delehaunty K, Dong S, Downton P, Dumas B, Fabro G, Fronick C, Fuerstenberg SI, Fulton L, Gaulin E, Govers F, Hughes L, Humphray S, Jiang RHY, Judelson H, Kamoun S, Kyung K, Meijer H, Minx P, Morris P, Nelson J, Phuntumart V, Qutob D, Rehmany A, Rougon-Cardoso A, Ryden P, Torto-Alalibo T, Studholme D, Wang Y, Win J, Wood J, Clifton SW, Rogers J, Van den Ackerveken G, Jones JDG, et al. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330:1549–1551. <https://doi.org/10.1126/science.1195203>.
 37. Lévesque CA, Brouwer H, Cano L, Hamilton JP, Holt C, Huitema E, Raffaele S, Robideau GP, Thines M, Win J, Zerillo MM, Beakes GW, Boore JL, Busam D, Dumas B, Ferriera S, Fuerstenberg SI, Gachon CMM, Gaulin E, Govers F, Grenville-Briggs L, Horner N, Hostetler J, Jiang RHY, Johnson J, Krajaeun T, Lin H, Meijer HJG, Moore B, Morris P, Phuntumart V, Puiu D, Shetty J, Stajich JE, Tripathy S, Wawra S, van West P, Whitty BR, Coutinho PM, Henrissat B, Martin F, Thomas PD, Tyler BM, De Vries RP, Kamoun S, Yandell M, Tisserat N, Buell CR. 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol* 11: R73. <https://doi.org/10.1186/gb-2010-11-7-r73>.
 38. 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:e75072. <https://doi.org/10.1371/journal.pone.0075072>.
 39. Sperschneider J, Williams AH, Hane JK, Singh KB, Taylor JM. 2015. Evaluation of secretion prediction highlights differing approaches needed for oomycete and fungal effectors. *Front Plant Sci* 6:1168. <https://doi.org/10.3389/fpls.2015.01168>.
 40. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300.
 41. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30: 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
 42. Ponchet M, Panabières F, Milat M-L, Mikes V, Montillet JL, Suty L, Triantaphyllides C, Tirilly Y, Blein JP. 1999. Are elicitors cryptograms in plant-oomycete communications? *Cell Mol Life Sci* 56:1020–1047. <https://doi.org/10.1007/s000180050491>.
 43. Gibbs GM, Roelants K, O'Bryan MK. 2008. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—roles in reproduction, cancer, and immune defense. *Endocr Rev* 29:865–897. <https://doi.org/10.1210/er.2008-0032>.
 44. Schneiter R, Di Pietro A. 2013. The CAP protein superfamily: function in sterol export and fungal virulence. *Biomol Concepts* 4:519–525. <https://doi.org/10.1515/bmc-2013-0021>.
 45. Zhou H, Casas-Finet JR, Coats RH, Kaufman JD, Stahl SJ, Wingfield PT, Rubin JS, Bottaro DP, Byrd RA. 1999. Identification and dynamics of a heparin-binding site in hepatocyte growth factor. *Biochemistry* 38: 14793–14802. <https://doi.org/10.1021/bi9908641>.
 46. Gaulin E, Jauneau A, Villalba F, Rickauer M, Esquerré-Tugayé MT, Bottin A. 2002. The CBEL glycoprotein of *Phytophthora parasitica* var. *nicotianae* is involved in cell wall deposition and adhesion to cellulosic substrates. *J Cell Sci* 115:4565–4575. <https://doi.org/10.1242/jcs.00138>.
 47. Raffaele S, Win J, Cano LM, Kamoun S. 2010. Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*. *BMC Genomics* 11:637. <https://doi.org/10.1186/1471-2164-11-637>.
 48. Feng BZ, Zhu XP, Fu L, Lv RF, Storey D, Tooley P, Zhang XG. 2014. Characterization of necrosis-inducing NLP proteins in *Phytophthora capsici*. *BMC Plant Biol* 14:126. <https://doi.org/10.1186/1471-2229-14-126>.
 49. Oome S, Van den Ackerveken G. 2014. Comparative and functional analysis of the widely occurring family of Nep1-like proteins. *Mol Plant*

- Microbe Interact 27:1081–1094. <https://doi.org/10.1094/MPMI-04-14-0118-R>.
50. Oome S, Raaymakers TM, Cabral A, Samwel S, Böhm H, Albert I, Nürnberger T, Van den Ackerveken G. 2014. Nep1-like proteins from three kingdoms of life act as a microbe-associated molecular pattern in Arabidopsis. *Proc Natl Acad Sci U S A* 111:16955–16960. <https://doi.org/10.1073/pnas.1410031111>.
 51. Gijzen M, Nürnberger T. 2006. Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry* 67:1800–1807. <https://doi.org/10.1016/j.phytochem.2005.12.008>.
 52. Horner NR, Grenville-Briggs LJ, van West P. 2012. The oomycete *Pythium oligandrum* expresses putative effectors during mycoparasitism of *Phytophthora infestans* and is amenable to transformation. *Fungal Biol* 116:24–41. <https://doi.org/10.1016/j.funbio.2011.09.004>.
 53. McCarthy CGP, Fitzpatrick DA. 2016. Systematic search for evidence of interdomain horizontal gene transfer from prokaryotes to oomycete lineages. *mSphere* 1:e00195-16. <https://doi.org/10.1128/mSphere.00195-16>.
 54. Mistry D, Stockley RA. 2006. IgA1 protease. *Int J Biochem Cell Biol* 38:1244–1248. <https://doi.org/10.1016/j.biocel.2005.10.005>.
 55. Unestam T, Weiss DW. 1970. The host-parasite relationship between freshwater crayfish and the crayfish disease fungus *Aphanomyces astaci*: responses to infection by a susceptible and a resistant species. *J Gen Microbiol* 60:77–90. <https://doi.org/10.1099/00221287-60-1-77>.
 56. Jiang RHY, de Bruijn I, Haas BJ, Belmonte R, Löbach L, Christie J, van den Ackerveken G, Bottin A, Bulone V, Díaz-Moreno SM, Dumas B, Fan L, Gaulin E, Govers F, Grenville-Briggs LJ, Horner NR, Levin JZ, Mammella M, Meijer HJG, Morris P, Nusbaum C, Oome S, Phillips AJ, van Rooyen D, Rzeszutek E, Saraiva M, Secombes CJ, Seidl MF, Snel B, Stassen JHM, Sykes S, Tripathy S, van den Berg H, Vega-Arrequin JC, Wawra S, Young SK, Zeng Q, Dieguez-Uribeondo J, Russ C, Tyler BM, van West P. 2013. Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. *PLoS Genet* 9:e1003272. <https://doi.org/10.1371/journal.pgen.1003272>.
 57. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
 58. Seidl V. 2008. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biol Rev* 22:36–42. <https://doi.org/10.1016/j.fbr.2008.03.002>.
 59. Duo-Chuan L. 2006. Review of fungal chitinases. *Mycopathologia* 161:345–360. <https://doi.org/10.1007/s11046-006-0024-y>.
 60. Mérida H, Sandoval-Sierra JV, Diéguez-Uribeondo J, Bulone V. 2013. Analyses of extracellular carbohydrates in oomycetes unveil the existence of three different cell wall types. *Eukaryot Cell* 12:194–203. <https://doi.org/10.1128/EC.00288-12>.
 61. Bowman SM, Free SJ. 2006. The structure and synthesis of the fungal cell wall. *BioEssays* 28:799–808. <https://doi.org/10.1002/bies.20441>.
 62. Henrissat B, Bairoch A. 1993. New families in the classification of glycosylhydrolases based on amino acid sequence similarities. *Biochem J* 293:781–788. <https://doi.org/10.1042/bj2930781>.
 63. Mohnen D. 2008. Pectin structure and biosynthesis. *Curr Opin Plant Biol* 11:266–277. <https://doi.org/10.1016/j.pbi.2008.03.006>.
 64. Marin-Rodríguez MC, Orchard J, Seymour GB. 2002. Pectate lyases, cell wall degradation and fruit softening. *J Exp Bot* 53:2115–2119. <https://doi.org/10.1093/jxb/erf089>.
 65. Di Matteo A, Giovane A, Raiola A, Camardella L, Bonivento D, De Lorenzo G, Cervone F, Bellincampi D, Tsernoglou D. 2005. Structural basis for the interaction between pectin methylesterase and a specific inhibitor protein. *Plant Cell* 17:849–858. <https://doi.org/10.1105/tpc.104.028886>.
 66. Fries M, Ihrig J, Brocklehurst K, Shevchik VE, Pickersgill RW. 2007. Molecular basis of the activity of the phytopathogen pectin methylesterase. *EMBO J* 26:3879–3887. <https://doi.org/10.1038/sj.emboj.7601816>.
 67. Vercauteren I, de Almeida Engler J, De Groot R, Gheysen G. 2002. An Arabidopsis thaliana pectin acetyltransferase gene is upregulated in nematode feeding sites induced by root-knot and cyst nematodes. *Mol Plant Microbe Interact* 15:404–407. <https://doi.org/10.1094/MPMI.2002.15.4.404>.
 68. Schäfer W. 1993. The role of cutinase in fungal pathogenicity. *Trends Microbiol* 1:69–71. [https://doi.org/10.1016/0966-842X\(93\)90037-R](https://doi.org/10.1016/0966-842X(93)90037-R).
 69. Yeats TH, Rose JKC. 2013. The formation and function of plant cuticles. *Plant Physiol* 163:5–20. <https://doi.org/10.1104/pp.113.222737>.
 70. Belbahri L, Calmin G, Mauch F, Andersson JO. 2008. Evolution of the cutinase gene family: evidence for lateral gene transfer of a candidate *Phytophthora* virulence factor. *Gene* 408:1–8. <https://doi.org/10.1016/j.gene.2007.10.019>.
 71. Li D, Ashby AM, Johnstone K. 2003. Molecular evidence that the extracellular cutinase Pbc1 is required for pathogenicity of *Pyrenopeziza brassicae* on oilseed rape. *Mol Plant Microbe Interact* 16:545–552. <https://doi.org/10.1094/MPMI.2003.16.6.545>.
 72. Kamoun S, Furzer O, Jones JDG, Judelson HS, Ali GS, Dalio RJD, Roy SG, Schena L, Zambounis A, Panabières F, Cahill D, Ruocco M, Figueiredo A, Chen XR, Hulvey J, Stam R, Lamour K, Gijzen M, Tyler BM, Grünwald NJ, Mukhtar MS, Tomé DFA, Tör M, Van Den Ackerveken G, McDowell J, Daayf F, Fry WE, Lindqvist-Kreuzer H, Meijer HJG, Petre B, Ristaino J, Yoshida K, Birch PRJ, Govers F. 2015. The top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol* 16:413–434. <https://doi.org/10.1111/mpp.12190>.
 73. Orsomando G, Lorenzi M, Raffaelli N, Dalla Rizza M, Mezzetti B, Ruggieri S. 2001. Phytotoxic protein Pcf, purification, characterization, and cDNA sequencing of a novel hydroxyproline-containing factor secreted by the strawberry pathogen *Phytophthora cactorum*. *J Biol Chem* 276:21578–21584. <https://doi.org/10.1074/jbc.M101377200>.
 74. Meijer HJG, Mancuso FM, Espadas G, Seidl MF, Chiva C, Govers F, Sabido E. 2014. Profiling the secretome and extracellular proteome of the potato late blight pathogen *Phytophthora infestans*. *Mol Cell Proteomics* 13:2101–2113. <https://doi.org/10.1074/mcp.M113.035873>.
 75. Stahl EA, Bishop JG. 2000. Plant-pathogen arms races at the molecular level. *Curr Opin Plant Biol* 3:299–304. [https://doi.org/10.1016/S1369-5266\(00\)00083-2](https://doi.org/10.1016/S1369-5266(00)00083-2).
 76. Win J, Morgan W, Bos J, Krasilova KV, Cano LM, Chaparro-García A, Ammar R, Staskawicz BJ, Kamoun S. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 19:2349–2369. <https://doi.org/10.1105/tpc.107.051037>.
 77. Boutemy LS, King SRF, Win J, Hughes RK, Clarke TA, Blumenschein TMA, Kamoun S, Banfield MJ. 2011. Structures of *Phytophthora* RXLR effector proteins: a conserved but adaptable fold underpins functional diversity. *J Biol Chem* 286:35834–35842. <https://doi.org/10.1074/jbc.M111.262303>.
 78. Van West P, De Bruijn I, Minor KL, Phillips AJ, Robertson EJ, Wawra S, Bain J, Anderson VL, Secombes CJ. 2010. The putative RXLR effector protein SpHtp1 from the fish pathogenic oomycete *Saprolegnia parasitica* is translocated into fish cells. *FEMS Microbiol Lett* 310:127–137. <https://doi.org/10.1111/j.1574-6968.2010.02055.x>.
 79. Wawra S, Bain J, Durward E, de Bruijn I, Minor KL, Matena A, Löbach L, Whisson SC, Bayer P, Porter AJ, Birch PRJ, Secombes CJ, van West P. 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* 109:2096–2101. <https://doi.org/10.1073/pnas.1113775109>.
 80. Kale SD, Gu B, Capelluto DGS, Dou D, Feldman E, Rumore A, Arredondo FD, Hanlon R, Fudal I, Rouxel T, Lawrence CB, Shan W, Tyler BM. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* 142:284–295. <https://doi.org/10.1016/j.cell.2010.06.008>.
 81. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, Beynon JL. 2004. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science* 306:1957–1960. <https://doi.org/10.1126/science.1104022>.
 82. Stanke M, Morgenstern B. 2005. AUGUSTUS: a Web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res* 33:W465–W467. <https://doi.org/10.1093/nar/gki458>.
 83. Ter-Hovhannisyán V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res* 18:1979–1990. <https://doi.org/10.1101/gr.081612.108>.
 84. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>.
 85. Bendtsen JD, Nielsen H, Von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340:783–795. <https://doi.org/10.1016/j.jmb.2004.05.028>.

86. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>.
87. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
88. Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* 14: 755–763. <https://doi.org/10.1093/bioinformatics/14.9.755>.
89. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
90. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. *Abstr 3rd Int AAAI Conf Weblogs Soc Media*, p 361–362.
91. Fruchterman TMJ, Reingold EM. 1991. Graph drawing by force-directed placement. *J Softw Pract Exper* 21:1129–1164. <https://doi.org/10.1002/spe.4380211102>.
92. Keane TM, Creevey CJ, Pentony MM, Naughton TJM, McInerney JO. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol* 6:29. <https://doi.org/10.1186/1471-2148-6-29>.
93. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 2.0. *Syst Biol* 59: 307–321. <https://doi.org/10.1093/sysbio/syq010>.
94. Letunic I, Bork P. 2007. Interactive Tree of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128. <https://doi.org/10.1093/bioinformatics/btl529>.
95. Studholme DJ, McDougal RL, Sambles C, Hansen E, Hardy G, Grant M, Ganley RJ, Williams NM. 2016. Genome sequences of six phytophthora species associated with forests in New Zealand. *Genom Data* 7:54–56. <https://doi.org/10.1016/j.gdata.2015.11.015>.
96. Lamour KH, Stam R, Jupe J, Huitema E. 2012. The oomycete broad-host-range pathogen *Phytophthora capsici*. *Mol Plant Pathol* 13: 329–337. <https://doi.org/10.1111/j.1364-3703.2011.00754.x>.
97. Feau N, Taylor G, Dale AL, Dhillon B, Bilodeau GJ, Birol I, Jones SJM, Hamelin RC. 2016. Genome sequences of six *Phytophthora* species threatening forest ecosystems. *Genom Data* 10:85–88. <https://doi.org/10.1016/j.gdata.2016.09.013>.
98. Gao R, Cheng Y, Wang Y, Guo L, Zhang G. 2015. Genome sequence of *Phytophthora fragariae* var. *fragariae*, a quarantine plant-pathogenic fungus. *Genome Announc* 3:e00034-15. <https://doi.org/10.1128/genomeA.00034-15>.
99. Sambles C, Schlenzig A, O'Neill P, Grant M, Studholme DJ. 2015. Draft genome sequences of *Phytophthora kernoviae* and *Phytophthora ramorum* lineage EU2 from Scotland. *Genom Data* 6:193–194. <https://doi.org/10.1016/j.gdata.2015.09.010>.
100. Quinn L, O'Neill PA, Harrison J, Paskiewicz KH, McCracken AR, Cooke LR, Grant MR, Studholme DJ. 2013. Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. *FEMS Microbiol Lett* 344:179–185. <https://doi.org/10.1111/1574-6968.12179>.
101. Liu H, Ma X, Yu H, Fang D, Li Y, Wang X, Wang W, Dong Y, Xiao B. 2016. Genomes and virulence difference between two physiological races of *Phytophthora nicotianae*. *Gigascience* 5:3. <https://doi.org/10.1186/s13742-016-0108-7>.
102. Rujirawat T, Patumcharoenpol P, Lohnoo T, Yingyong W, Lerksuthirat T, Tangphatsomruang S, Suriyaphol P, Grenville-Briggs LJ, Garg G, Kittichotirat W, Krajaeun T. 2015. Draft genome sequence of the pathogenic oomycete *Pythium insidiosum* strain Pi-S, isolated from a patient with pythiosis. *Genome Announc* 3:e00574-15. <https://doi.org/10.1128/genomeA.00574-15>.
103. Berger H, Yacoub A, Gerbore J, Grizard D, Rey P, Sessitsch A, Compant S. 2016. Draft genome sequence of biocontrol agent *Pythium oligandrum* strain Po37, an oomycota. *Genome Announc* 4:e00215-16. <https://doi.org/10.1128/genomeA.00215-16>.