

Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire

C André Lévesque, Henk Brouwer, Liliana Cano, John P Hamilton, Carson Holt, Edgar Huitema, Sylvain Raffaele, Gregg P Robideau, Marco Thines, Joe Win, et al.

▶ **To cite this version:**

C André Lévesque, Henk Brouwer, Liliana Cano, John P Hamilton, Carson Holt, et al.. Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire. Genome Biology, 2010, 11 (7), 10.1186/gb-2010-11-7-r73 . hal-01605006

HAL Id: hal-01605006
https://hal.science/hal-01605006

Submitted on 14 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Lévesque et al. Genome Biology 2010, 11:R73
http://genomebiology.com/2010/11/7/R73

Genome Biology

RESEARCH Open Access

Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire

C André Lévesque[1,2], Henk Brouwer[3†], Liliana Cano[4†], John P Hamilton[5†], Carson Holt[6†], Edgar Huitema[4†], Sylvain Raffaele[4†], Gregg P Robideau[1,2†], Marco Thines[7,8†], Joe Win[4†], Marcelo M Zerillo[9†], Gordon W Beakes[10], Jeffrey L Boore[11], Dana Busam[12], Bernard Dumas[13], Steve Ferriera[12], Susan I Fuerstenberg[11], Claire MM Gachon[14], Elodie Gaulin[13], Francine Govers[15,16], Laura Grenville-Briggs[17], Neil Horner[17], Jessica Hostetler[12], Rays HY Jiang[18], Justin Johnson[12], Theerapong Krajaejun[19], Haining Lin[5], Harold JG Meijer[15], Barry Moore[6], Paul Morris[20], Vipaporn Phuntmart[20], Daniela Puiu[12], Jyoti Shetty[12], Jason E Stajich[21], Sucheta Tripathy[22], Stephan Wawra[17], Pieter van West[17], Brett R Whitty[5], Pedro M Coutinho[23], Bernard Henrissat[23], Frank Martin[24], Paul D Thomas[25], Brett M Tyler[22], Ronald P De Vries[3], Sophien Kamoun[4], Mark Yandell[6], Ned Tisserat[9], C Robin Buell[5*]

Abstract

Background: *Pythium ultimum* is a ubiquitous oomycete plant pathogen responsible for a variety of diseases on a broad range of crop and ornamental species.

Results: The *P. ultimum* genome (42.8 Mb) encodes 15,290 genes and has extensive sequence similarity and synteny with related *Phytophthora* species, including the potato blight pathogen *Phytophthora infestans*. Whole transcriptome sequencing revealed expression of 86% of genes, with detectable differential expression of suites of genes under abiotic stress and in the presence of a host. The predicted proteome includes a large repertoire of proteins involved in plant pathogen interactions, although, surprisingly, the *P. ultimum* genome does not encode any classical RXLR effectors and relatively few Crinkler genes in comparison to related phytopathogenic oomycetes. A lower number of enzymes involved in carbohydrate metabolism were present compared to *Phytophthora* species, with the notable absence of cutinases, suggesting a significant difference in virulence mechanisms between *P. ultimum* and more host-specific oomycete species. Although we observed a high degree of orthology with *Phytophthora* genomes, there were novel features of the *P. ultimum* proteome, including an expansion of genes involved in proteolysis and genes unique to *Pythium*. We identified a small gene family of cadherins, proteins involved in cell adhesion, the first report of these in a genome outside the metazoans.

Conclusions: Access to the *P. ultimum* genome has revealed not only core pathogenic mechanisms within the oomycetes but also lineage-specific genes associated with the alternative virulence and lifestyles found within the pythiaceous lineages compared to the Peronosporaceae.

Background

Pythium is a member of the Oomycota (also referred to as oomycetes), which are part of the heterokont/chromist clade [1,2] within the 'Straminipila-Alveolata-Rhizaria' superkingdom [3]. Recent phylogenies based on multiple protein coding genes indicate that the oomycetes, together with the uniflagellate hyphochytrids and the flagellates *Pirsonia* and *Developayella*, form the sister clade to the diverse photosynthetic orders in the phylum Ochrophyta [2,4]. Therefore, the genomes of the closest relatives to *Pythium* outside of the oomycetes available to date would be those of the diatoms *Thalassiosira* [5] and *Phaeodactylum* [6], and the phaeophyte algae *Ectocarpus* [7].

* Correspondence: buell@msu.edu
† Contributed equally
[5]Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA

BioMed Central

© 2010 Lévesque et al.; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pythium is a cosmopolitan and biologically diverse genus. Most species are soil inhabitants, although some reside in saltwater estuaries and other aquatic environments. Most *Pythium* spp. are saprobes or facultative plant pathogens causing a wide variety of diseases, including damping-off and a range of field and post-harvest rots [8-12]. *Pythium* spp. are opportunistic plant pathogens that can cause severe damage whenever plants are stressed or at a vulnerable stage. Some species have been used as biological control agents for plant disease management whereas others can be parasites of animals, including humans [13-15]. The genus *Pythium*, as currently defined, contains over a hundred species, with most having some loci sequenced for phylogeny [16]. *Pythium* is placed in the Peronosporales *sensu lato*, which contains a large number of often diverse taxa in which two groups are commonly recognized, the paraphyletic Pythiaceae, which comprise the basal lineages of the second group, the Peronosporaceae.

The main morphological feature that separates *Pythium* lineages from *Phytophthora* lineages is the process by which zoospores are produced from sporangia. In *Phytophthora*, zoospore differentiation happens directly within the sporangia, a derived character or apomorphism for *Phytophthora*. In *Pythium*, a vesicle is produced within which zoospore differentiation occurs [12]; this is considered the ancestral or plesiomorphic state. There is a much wider range of sporangial shapes in *Pythium* than is found in *Phytophthora* (see [17] for more detailed comparison). Biochemically, *Phytophthora* spp. have lost the ability to synthesize thiamine, which has been retained in *Pythium* and most other oomycetes. On the other hand, elicitin-like proteins are abundant in *Phytophthora* but in *Pythium* they have been mainly found in the species most closely related to *Phytophthora* [18-20]. Many *Phytophthora* spp. have a rather narrow plant species host range whereas there is little host specificity in plant pathogenic *Pythium* species apart from some preference shown for either monocot or dicot hosts. Gene-for-gene interactions and the associated cultivar/race differential responses have been described for many *Phytophthora* and downy mildew species with narrow host ranges. In constrast, such gene-for-gene interactions or cultivar/race differentials have never been observed in *Pythium*, although single dominant genes were associated with resistance in maize and soybean against *Pythium inflatum* and *Pythium aphanidermatum*, respectively [21,22], and in common bean against *P. ultimum* var. *ultimum* (G Mahuku, personal communication). Lastly, in the necrotroph to biotroph spectrum, some *Pythium* spp. are necrotrophs whereas others behave as hemibiotrophs like *Phytophthora* spp. [23].

P. ultimum is a ubiquitous plant pathogen and one of the most pathogenic *Pythium* spp. on crop species [13]. It does not require another mating type for sexual reproduction as it is self-fertile - that is, homothallic - but outcrossing has been reported [24]. *P. ultimum* is separated into two varieties: *P. ultimum* var. *ultimum* is the most common and pathogenic group and produces oospores but very rarely sporangia and zoospores, whereas *P. ultimum* var. *sporangiiferum* is a rare and less pathogenic group that produces both oospores and sporangia [12]. The isolate (DAOM BR144 = CBS 805.95 = ATCC 200006) reported in this study belongs to *P. ultimum* var. *ultimum* and was found to be the most representative strain [16,25,26]. We use *P. ultimum* to refer to *P. ultimum* var. *ultimum* unless stated otherwise.

In this study, we report on the generation and analysis of the full genome sequence of *P. ultimum* DAOM BR144, an isolate obtained from tobacco. The genomes of several plant pathogenic oomycetes have been sequenced, including three species of *Phytophthora* (*Ph. infestans*, *Ph. sojae*, and *Ph. ramorum* [27,28]), allowing the identification and improved understanding of pathogenicity mechanisms of these pathogens, especially with respect to the repertoire of effector molecules that govern the outcome of the plant-pathogen interaction [27-30]. To initially assess the gene complement of *P. ultimum*, we generated a set of ESTs using conventional Sanger sequencing coupled with 454 pyrosequencing of *P. ultimum* (DAOM BR144) hyphae grown in rich and nutrient-starved conditions [31]. These transcriptome sequence data were highly informative and showed that *P. ultimum* shared a large percentage of its proteome with related *Phytophthora* spp. In this study, we report on the sequencing, assembly, and annotation of the *P. ultimum* DAOM BR144 genome. To gain insight into gene function, we performed whole transcriptome sequencing under eight growth conditions, including a range of abiotic stresses and in the presence of a host. While the *P. ultimum* genome has similarities to related oomycete plant pathogens, its complement of metabolic and effector proteins is tailored to its pathogenic lifestyle as a necrotroph.

Results and discussion

Sequence determination and gene assignment

Using a hybrid strategy that coupled deep Sanger sequencing of variable insert libraries with pyrosequencing, we generated a high quality draft sequence of the oomycete pathogen *P. ultimum* (DAOM BR144 = CBS 805.95 = ATCC 200006). With an N_{50} contig length of 124 kb (1,747 total) and an N_{50} scaffold length of 773,464 bp (975 total), the *P. ultimum* assembly

represents 42.8 Mb of assembled sequence. Additional metrics on the genome are available in Additional file 1.

P. ultimum, *Ph. sojae* and *Ph. ramorum* differ in mating behaviour: *P. ultimum and Ph. sojae* are homothallic while *Ph. ramorum* is heterothallic. The outcrossing preference in *Ph. ramorum* is reflected in the 13,643 single nucleotide polymorphisms identified in this species versus 499 found in the inbreeding *Ph. sojae* [27]. Although the *Ph. sojae* genome size is twice that of *P. ultimum*, a large number (11,916) of variable bases (that is, high quality reads in conflict with the consensus) were present within the DAOM BR144 assembly, indicating that the *in vitro* outcrossing reported for *P. ultimum* [24] might be common in nature.

The final genome annotation set (v4) contained 15,297 genes encoding 15,329 transcripts (15,323 protein coding and 6 rRNA coding) due to detection of alternative splice forms. Global analysis of the intron/exon structure revealed that while there are examples of intron-rich genes in the *P. ultimum* genome, the majority of genes tend to have few introns, with an average 1.6 introns occurring per gene that are relatively short (average intron length 115 bp), consistent with that of *Ph. infestans* (1.7 introns per gene, 124 bp average intron length). Coding exons in the *P. ultimum* genome tend to be relatively long when compared to other eukaryotes [32-40], having an average length of 498 bp, with 38.9% of the *P. ultimum* genes encoded by a single exon. This is comparable to that observed in *P. infestans*, in which the average exon is 456 bp with 33.1% encoding single exon genes.

In eukaryotic genomes such as that of *Arabidopsis thaliana* and human, 79% and 77% of all genes contain an InterPro domain, respectively. In comparison, only 60% of all *P. ultimum* genes contain an InterPro protein domain, which is comparable to that observed with *Phytophthora* spp. (55 to 66%). This is most likely attributable to the higher quality annotation of the human and *Arabidopsis* proteomes and, potentially, the lack of representation of oomycetes in protein databases.

Earlier transcriptome work with strain DAOM BR144 involved Sanger and 454 pyrosequencing of a normalized cDNA library constructed from two *in vitro* growth conditions [31]. When mapped to the DAOM BR144 genome, these ESTs (6,903 Sanger- and 21,863 454-assembled contigs) aligned with 10,784 gene models, providing expression support for 70.5% of the gene set. To further probe the *P. ultimum* transcriptome and to aid in functional annotation, we employed mRNA-Seq [41] to generate short transcript reads from eight growth/treatment conditions. A total of 71 million reads (2.7 Gb) were mapped to the DAOM BR144 genome and 11,685 of the 15,297 loci (76%) were expressed based on RNA-Seq data. Collectively, from the Sanger, 454, and Illumina transcriptome sequencing in which eight growth conditions, including host infection, were assayed, transcript support was detected for 13,103 genes of the 15,291 protein coding genes (85.7%). When protein sequence similarity to other annotated proteins is coupled with all available transcript support, only 190 of the 15,291 protein coding genes lack either transcript support or protein sequence similarity (Table S1 in Additional file 2).

Repeat content in DAOM BR144

In total, 12,815 repeat elements were identified in the genome (Table S2 in Additional file 2). In general, the relatively low repeat content of the *P. ultimum* genome (approximately 7% by length) is similar to what would be expected for small, rapidly reproducing eukaryotic organisms [42,43]. While the repeat content is much lower than that of the oomycete *Ph. infestans* [28], the difference is likely due to the presence of DNA methylases identified by protein domain analyses in the *P. ultimum* genome, which have been shown to inhibit repeat expansion [44]. Interestingly, the oomycete *Ph. infestans* lacks DNA methylase genes, the absence of which is believed to contribute to repeat element expansion within that organism, with repeats making up > 50% of the genome [27,28,45].

Mitochondrial genome

The *P. ultimum* DAOM BR144 mitochondrial genome is 59,689 bp and contains a large inverted repeat (21,950 bp) that is separated by small (2,711 bp) and large (13,078 bp) unique regions (Figure S1 in Additional file 3). The *P. ultimum* DAOM BR144 mitochondrion encodes the same suite of protein coding (35), rRNA (2), and tRNA (encoding 19 amino acids) genes present in other oomycetes such as *Phytophthora* and *Saprolegnia* [46-48]. However, the number of copies is different due to the large inverted repeat as well as some putative ORFs that are unique to *P. ultimum* (Additional file 1). No insertions of the mitochondrial genome into the nuclear genome were identified.

Proteins involved in plant-pathogen interactions

Comparative genome analyses can reveal important differences between *P. ultimum* and the Peronosporaceae that may contribute to their respective lifestyles, that is, the non-host specific *P. ultimum* and the host specific *Phytophthora* spp. We utilized two approaches to probe the nature of gene complements within these two clades of oomycetes. First, using the generalized approach of examining PANTHER protein families [49], we identified major lineage-specific expansions of gene families. Second, through targeted analysis of subsets of the *P. ultimum* proteome, including the secretome, effectors, proteins involved in carbohydrate metabolism, and pathogen/microbial-associated molecular patterns

Table 1 Major lineage-specific gene family expansions leading to differences in the P. ultimum gene complement compared to *Phytophthora*

Biological process	Comparison to *Phytophthora*	Protein family expansions (number of genes in *P. ultimum*/*Ph. ramorum*)
Proteolysis	Over-represented	HECT E3 ubiquitin ligase (56/28)
		Subtilisin-related serine protease S8A (43/7)
		Trypsin-related serine protease S1A (17/31)
		Pepsin-related aspartyl protease A1 (25/15)
		Metalloprotease M12 (12/3)
Intracellular signaling cascade	Under-represented	PTHR23257 S/T protein kinase (78/158)
		PTHR22985 S/T protein kinase (23/51)
		PTHR22982, CaM kinase (50/85)
		Phospholipase D (9/18)
Sulfur metabolism	Under-represented	Sulfatase (7/14)
		Cysteine desulfurylase (4/11)
		Sulfate transporter (10/18)
Water transport	Under-represented	Aquaporin (11/35)

(PAMPs or MAMPs; for review see [50]), we revealed commonalities, as well as significant distinct features, of *P. ultimum* in comparison to *Phytophthora* spp.

Over-represented gene families

Several families involved in proteolysis were over-represented in *P. ultimum* compared to *Phytophthora* spp. (Table 1). This is primarily due to a massive expansion of subtilisin-related proteases (PTHR10795) in *P. ultimum* following the divergence from ancestors of *Phytophthora*. With regard to the total complement of serine proteases, the subtilisin family expansion in *P. ultimum* is somewhat counterbalanced by the trypsin-related serine protease family, which has undergone more gene duplication events in the *Phytophthora* lineage than the *Pythium* lineage. The metalloprotease M12 (neprolysin-related) family has also undergone multiple expansions, from one copy in the stramenopile most recent common ancestor, to three in the oomycete most recent common ancestor (and extant *Phytophthora*), then up to 12 in *P. ultimum* (data not shown).

E3 ligases are responsible for substrate specificity of ubiquitination and subsequent proteolysis, and secreted E3 ligases have been shown to act as effectors for pathogens by targeting host response proteins for degradation [51,52]. The HECT E3 family of ubiquitin-protein ligases (PTHR11254) apparently underwent at least two major expansions, one in the oomycete lineage after the divergence from diatoms and another in the *P. ultimum* lineage (Figure S2 in Additional file 3; Table 1). Most of the expansion in the *P. ultimum* lineage appears to be derived from repeated duplication of only two genes that were present in the *Pythium-Phytophthora* common ancestor. This expanded subfamily is apparently orthologous to the *UPL1* and *UPL2* genes from *A. thaliana*. Of the 56 predicted HECT E3 ligases in the *P. ultimum* genome (that had long enough sequences for phylogenetic analysis), 16 are predicted by SignalP [53] to have *bona fide* signal peptides, and another 10 have predicted signal anchors, a substantially larger number than reported for other oomycete genomes [54].

Under-represented gene families

Several gene families are significantly under-represented in the *P. ultimum* genome compared to *Phytophthora* (Table 1) and it appears that these are mostly due to expansions in the *Phytophthora* lineage rather than losses in the *Pythium* lineage, though the relatively long distance to the diatom outgroup makes this somewhat uncertain. These include the aquaporin family (PTHR19139), the phospholipase D family (PTHR18896; Additional file 1), four families/subfamilies of intracellular serine-threonine protein kinases, and three families involved in sulfur metabolism (sulfatases (PTHR10342), cysteine desulfurylases (PTHR11601) and sulfate transporters (PTHR11814)).

The P. ultimum secretome

As oomycete plant pathogens secrete a variety of proteins to manipulate plant processes [30,55], we predicted and characterized in detail the soluble secreted proteins of *P. ultimum*. The secretome of *P. ultimum* was identified by predicting secreted proteins using the PexFinder algorithm [56] in conjunction with the TribeMCL protein family clustering algorithm. The *P. ultimum* secretome is composed of 747 proteins (4.9% of the proteome) that can be clustered into 195 families (each family contains at least 2 sequences) and 127 singletons (Table S3 in Additional file 2; selected families are shown in Figure S3 in Additional file 3). Of these, two families and one singleton encode transposable-element-related proteins that were missed in the repeat masking process. The largest family contains 77 members, mostly ankyrin repeat containing proteins, of which only 3 were predicted to have a signal peptide. Notable families of

secreted proteins include protease inhibitors (serine and cysteine), NPP1-like proteins (toxins), cellulose-binding elicitor lectin (CBEL)-like proteins with carbohydrate binding domains, elicitins and elicitin-like proteins, secreted E3 ubiquitin ligases (candidate effectors), cell-wall degrading enzymes, lipases, phospholipases, potential adhesion proteins, highly expanded families of proteases and cytochrome P450 (Table 2), and several families of 'unknown' function. A subset (88 proteins) of the secretome showed exclusive similarity to fungal sequences yet are absent in other eukaryotes (Table S4 in Additional file 2; see Table S1 in [57] for a list of organisms). These may represent shared pathogenicity proteins for filamentous plant pathogens, such as peroxidases (Family 68), CBEL-like proteins (Family 8), and various cell wall degrading enzymes and other hydrolases.

RXLR effectors

Many plant pathogens, especially biotrophic and hemibiotrophic ones, produce effector proteins that either enter into host cells or are predicted to do so [27,58,59]. The genomes of *Ph. sojae*, *Ph. ramorum* and *Ph. infestans* encode large numbers (370 to 550) of potential effector proteins that contain an amino-terminal cell-entry domain with the motifs RXLR and dEER [28,29], which mediate entry of these proteins into host cells in the absence of pathogen-encoded machinery [60,61]. RXLR-dEER effectors are thought, and in a few cases shown, to suppress host defense responses, but a subset of these effectors can be recognized by plant immune receptors resulting in programmed cell death and disease resistance. To search for RXLR effectors in the genome of *P. ultimum*, we translated all six frames of the genome sequence to identify all possible small proteins, exclusive of splicing. Among these, a total of 7,128 translations were found to contain an amino-terminal signal peptide based on SignalP prediction. We then used the RXLR-dEER Hidden Markov Model (HMM) [29] to search the translations for candidate effectors and, as a control, the same set of translations following permutation of their sequences downstream of the signal peptide (Figure 1a). Only 35 sequences with significant scores were found in the non-permuted set while an average of 5 were found in 100 different permuted sets. In comparison to the *Ph. ramorum* secretome, 300 hits were found without permutation. Examination of the 35 significant sequences revealed that most were members of a secreted proteinase family [62] in which the RXLR motif was part of a conserved subtilisin-like serine protease domain of 300 amino acids in length, and thus unlikely to be acting as a cell entry motif. A string search was then performed for the RXLR motif within the amino terminus of each translation, 30 to 150 residues from the signal peptide. In this case, the number of hits was not significantly different between the real sequences and the permuted sequences. The same result was obtained with the strings RXLX and RX[LMFY][HKR] (Figure 1b). HMMs have been defined to

Table 2 Protein families implicated in plant pathogenesis: *P. ultimum* versus *Phytophthora* spp. or diatoms

	P. ultimum	*Ph. infestans*	*Ph. sojae*	*Ph. ramorum*	*Thalassiosira pseudonana* (diatom)	*Phaeodactylum tricornutum* (diatom)
ABC transporters[a]	140	137	141	135	57	65
Aspartyl protease families A1, A8[b]	29	16	16	18	ND	8
Crinklers (CRN-family)[a]	26	196	100	19	0	0
Cutinase[c]	0	4	13	4	0	ND
Cysteine protease families C1, C2, C56[a]	42	38	33	42	ND	11
Cytochrome P450s[b]	41	28	31	31	ND	10
Elicitin-like proteins[d]	24	40	57	50	0	0
Glycoside hydrolases[c]	180	277	301	258	59	ND
Lipases[d]	31	19	27	17	22	17
NPP1-like proteins (necrosis-inducing proteins)[d]	7	27	39	59	0	0
PcF/SCR-like[d]	3	16	8	1	0	0
Pectin esterases[c]	0	13	19	11	0	ND
Polysaccharide lyases[c]	29	67	54	49	0	ND
Phospholipases[d]	20	36	31	28	18	11
Protease inhibitors, all[d]	43	38	26	18	11	5
RXLR effectors[a]	0	563	350	350	0	0
Serine protease families S1A, S8, S10[b]	85	60	63	57	ND	31

[a]Data from manual curation/analyses. [b]Data from PANTHER family analyses (MEROPS classification). [c]Data from CAZy. [d]Data from analysis of TRIBEMCL families. ND, not determined.

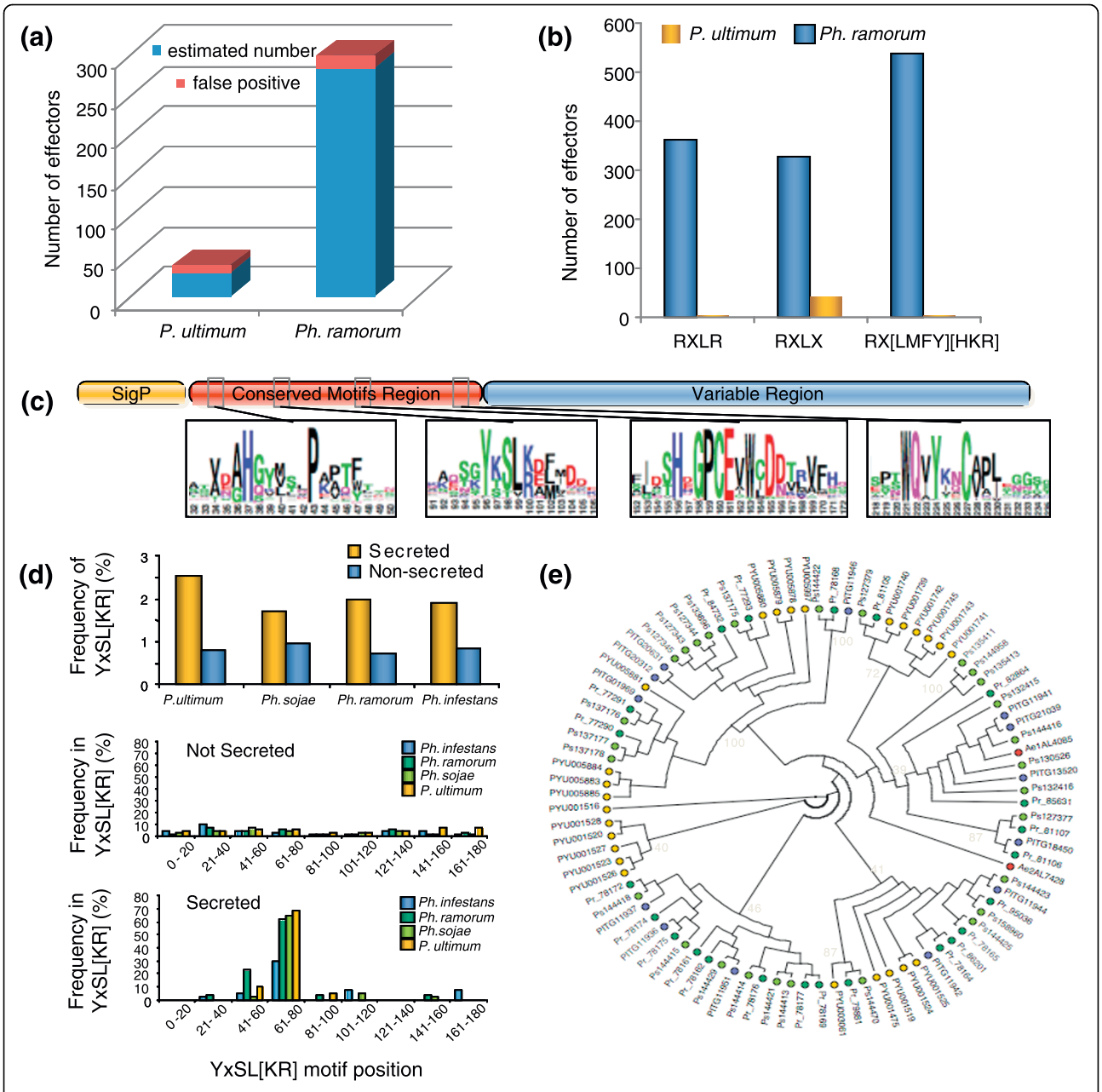


Figure 1 An original repertoire of candidate effector proteins in *P. ultimum*. **(a)** The number of candidate RXLR effectors estimated by Hidden Markov Model (HMM) searches of predicted proteins with amino-terminal signal peptides. The numbers of false positives were derived from HMM searches of the permutated protein sequences. **(b)** The number of candidate RXLR effectors discovered by motif searching. The search was performed on the total set of six-frame translated ORFs from the genome sequences that encode proteins with an amino-terminal signal peptide. The motif RXLR and two more degenerate motifs, RXLX or RX[LMIFY][HKR], were required to occur within 100 amino acids of the amino termini. **(c)** The typical architecture of a YxSL[RK] effector candidate inferred from 91 sequences retrieved from *P. ultimum*, three *Phytophthora* genomes and *A. euteiches*. **(d)** The YxSL[RK] motif is enriched and positionally constrained in secreted proteins in *P. ultimum* and *Phytophthora* spp. The top graph compares the abundance of YxSL[RK]-containing proteins among secreted and non-secreted proteins from four oomycete genomes. The middle and bottom graphs show the frequency of the YxSL[RK] motif among non-secreted and secreted proteins, respectively, according to its position in the protein sequence. **(e)** Cladogram based on the conserved motifs region of the 91YxSL[KR] proteins, showing boostrap support for the main branches.

identify carboxy-terminal motifs conserved in about 60% of RXLR-dEER effectors [29,63]. Searching the secretome and the permutated secretome with this HMM also identified no significant numbers of candidate effectors (data not shown). Blast searches with the most conserved *Phytophthora* effectors likewise produced no hits.

Based on synteny analysis of surrounding genes, a small number of *Phytophthora* effectors share conserved genomic positions [27]. Synteny analysis (see below) was used to identify the corresponding positions in the *P. ultimum* genome, but no predicted secreted proteins were found in those positions in the *P. ultimum* genome. A paucity of predicted RXLR effector sequences was reported previously in the transcriptome of *P. ultimum* [31]; the one candidate noted in the transcriptome sequence dataset has proven to be a false positive, matching the negative strand of a conserved transporter gene in the genome sequence. Therefore, we conclude that the *P. ultimum* genome lacks RXLR effectors that are abundant in other oomycetes, although this analysis does not rule out the possible presence of other kinds of effectors (see below). Nonetheless, the lack of RXLR effectors in *P. ultimum* is consistent with the absence of gene-for-gene interactions, all known instances of which in *Phytophthora* spp. involve RXLR effectors with avirulence activities.

CRN protein repertoire

In *Phytophthora* spp. the *Crinkler* (*crn*) gene family encodes a large class of secreted proteins that share a conserved amino-terminal LFLAK domain, which has been suggested to mediate host translocation and is followed by a major recombination site that forms the junction between the conserved amino terminus and diverse carboxy-terminal effector domains [28]. In sharp contrast to the RXLR effectors, the CRN protein family appears conserved in all plant pathogenic oomycete genomes sequenced to date. BLASTP searches of 16 well-defined amino-terminal domains from *Ph. infestans* against the *P. ultimum* predicted proteome identified 18 predicted proteins within *P. ultimum* (BLAST cutoff of 1×10^{-10}; Table S5 in Additional file 2). Examination of protein alignments revealed considerable conservation of the *P. ultimum* LFLAK domain. We used *P. ultimum* CRN sequence alignments to build an HMM and through HMM searches identified two additional predicted proteins with putative LFLAK-like domains. We assessed the distribution of candidate CRN proteins within *P. ultimum* families and identified six additional candidates in Family 64. Further examination of candidates confirmed the presence of LFLAK-like domains (Table S5 in Additional file 2). Surprisingly, only 2 (approximately 7.5%) of the 26 predicted CRN proteins were annotated as having signal peptides (Table S5 in Additional file 2). Two additional CRNs (PYU1_T003336 and PYU1_T002270) have SignalP v2.0 HMM scores of 0.89 and 0.76, respectively, which although below our stringent cutoff of 0.9 may still suggest potential signal peptides. Several of the remaining genes have incomplete ORFs and gene models, suggesting a high frequency of CRN pseudogenes as previously noted in *Ph. infestans* [28]. All 26 amino-terminal regions were aligned to generate a sequence logo. These analyses revealed a conserved LxLYLAR/K motif that is shared amongst *P. ultimum* CRN proteins (Figure S4 in Additional file 3) and is followed by a conserved WL motif. The LxLYLAR/K motif is closely related to the F/LxLYLALK motif found in *Aphanomyces euteiches* [64]. Consistent with results obtained in other oomycete genomes, we found that the LxLYLAR/K motif was located between 46 and 64 amino acids after the methionine, followed by a variable domain that ended with a conserved motif at the proposed recombination site (HVLVxxP), reflecting the modular design of CRN proteins in the oomycetes (Figure S4 in Additional file 3). This recombination site, which is characteristic for the DWL domain, was found highly conserved in 11 of the putative *P. ultimum CRN* genes, consisting of an aliphatic amino acid followed by a conserved histidine, another three aliphatic amino acids, two variable amino acids and a conserved proline. In a phylogenetic analysis, these 11 genes were predominantly placed basal to the validated CRNs from *Phytophthora* (Figure S5 in Additional file 3). Although the CRN-like genes in *Pythium* are more divergent than the validated CRNs of *Phytophthora* (Figure S5 in Additional file 3), both the recombination site and the LxLYLAR/K-motif, which is a modification of the prominent LxLFLAK-motif present in most *Phytophthora* CRNs, show a significant degree of conservation, highlighting that the CRN family, greatly expanded in *Phytophthora* [28], had already evolved in the last common ancestor of *P. ultimum* and *Phytophthora*.

A novel family of candidate effectors

In the absence of obvious proteins with an amino-terminal RXLR motif, we used other known features of effectors to identify candidate effector families in *P. ultimum*. *Ph. infestans* RXLR effectors are not only characterized by a conserved amino-terminal translocation domain but also by their occurrence in gene-sparse regions that are enriched in repetitive DNA [28]. Based on the length of the flanking non-coding regions, the distribution of *P. ultimum* genes is not multimodal as was observed in *Ph. infestans* (Figure S6 in Additional file 3). However, relative to the rest of the genes, *P. ultimum* secretome genes more frequently have long flanking non-coding regions (Figure S7 in Additional file 3). In addition, the secretome genes show a higher proportion of closely related paralogs, suggesting recent duplications in *P. ultimum* (Figure S7 in Additional file 3) and indicating that the secretome genes may have

distinct genome organization and evolution as noted in *Phytophthora* spp. [28,57]. Using genome organization properties to identify families of secreted proteins in *P. ultimum* that could correspond to novel effector candidates, we sorted the 194 secretome families based on highest rate of gene duplication, longest flanking non-coding region, and lowest similarity to *Ph. infestans* proteins (see Figure S8 in Additional file 3 for examples). One relatively large family of secreted proteins, Family 3, stood out because it fulfilled the three criteria and included proteins of unknown function. BlastP similarity searches identified similar sequences only in oomycete species (*Phytophthora* spp. and *A. euteiches*). Furthermore, of the 44 family members in *P. ultimum* for which transcripts could be detected, 32 (73%) were induced more than 2-fold during *Arabidopsis* infection compared to mycelia, with 5 members induced more than 40-fold. In total, we identified a set of 91 predicted secreted proteins with similarity to Family 3 proteins from the various oomycete species (Additional file 4). Multiple alignments of these proteins, along with motif searches, identified a YxSL[RK] amino acid motif (Figure 1c). This motif is at least two-fold enriched in secreted proteins compared to non-secreted proteins in four oomycete species (Figure 1d). In addition, the YxSL[RK] motif is positionally constrained between positions 61 and 80 in secreted oomycete proteins only (Figure 1d). The 91 YxSL[RK] proteins show a modular organization with a conserved amino-terminal region, containing four conserved motifs, followed by a highly variable carboxy-terminal region (Figure 1c; Figure S9 in Additional file 3) as reported for other oomycete effectors [30]. Phylogenetic analyses of the YxSL[RK] family revealed four main clades and suggest an expansion of this family in *Phytophthora* spp. (Figure 1e).

The YxSL[RK] motif appears to be a signature for a novel family of secreted oomycete proteins that may function as effectors. It is intriguing that the YxSL[RK] motif shares some similarity in sequence and position with the canonical RXLR motif, a resemblance increased by the fact that the variable amino acid is a basic amino acid (lysine) in 28 out of the 91 family members. Whether the YxSL[RK] motif defines a host-translocation domain as noted for RXLR effectors remains to be determined.

Detection of P. ultimum by the host

Detection of pathogens through the perception of PAMPs/MAMPs leads to the induction of plant immune responses (for review, see [50]). Oomycetes produce various and specific molecules able to induce defense responses like elicitins (for review, see [65]), but only two oomycete cell-surface proteins containing a MAMP have been characterized: a transglutaminase [66] and a protein named CBEL [67]. Genes encoding both of these cell-surface proteins were detected in *P. ultimum* (Additional file 1), suggesting that *P. ultimum* produces typical oomycete MAMPs, which can be efficiently perceived by a wide range of plant species. The occurrence of PAMPs/MAMPs in *P. ultimum* suggests that this pathogen must have evolved mechanisms to evade PAMP-triggered immunity. This could occur through a necrotrophic mechanism of infection or using the candidate effector proteins described above.

Metabolism of complex carbohydrates

A total of 180 candidate glycoside hydrolases (GHs) were identified in *P. ultimum* using the CAZy annotation pipeline [68]. This number is apparently similar to those reported previously for *Ph. ramorum* (173), *Ph. sojae* (190), and *Ph. infestans* (157) [27,28]. However, when the CAZy annotation pipeline was applied to *Ph. sojae*, *Ph. ramorum* and *Ph. infestans*, 301, 258 and 277 GHs were found, respectively, nearly twice the number present in *P. ultimum* (Table 2). Among these we identified putative cellulases belonging to families GH5, GH6 and GH7. All six GH6 candidate cellulases harbor secretion signals. Only one GH6 protein contains a CBEL domain at the carboxyl terminus. Three contain a transmembrane domain and one contains a glycosylphosphatidylinisotol anchor, features suggesting that these proteins may be targeting the oomycete cell wall rather than plant cell walls. The *P. ultimum* strain studied here could not grow when cellulose was the sole carbon source (Table 3; Figure S10 in Additional file 3).

Cutinases are a particular set of esterases (CAZy family CE5) that cleave cutin, a polyester composed of hydroxy and hydroxyepoxy fatty acids that protects aerial plant organs. No candidate cutinases could be found

Table 3 Growth comparison of *P. ultimum* DAOM BR 144 on different carbon sources and the pH of the medium after 7 days

	DAOM BR144	
Carbon source	Mycelium density	pH on day 7
No carbon	-	5.1
25 mM D-glucose	+++	2.9
25 mM D-fructose	+++	2.9
25 mM D-xylose	-	5
25 mM L-arabinose	-	5
25 mM cellobiose	+++	4
25 mM sucrose	+++	3.2
1% cellulose	-	5.2
1% birch wood xylan	-	4.7
1% soluble starch	+++	3.5
1% citrus pectin*	+	5

The symbols indicate poor growth (+), moderate growth (++), good growth (+++), very good growth (++++), or growth less than or equal to the no-carbon medium (-). The data are the average of the two duplicates used for this experiment.

in the *P. ultimum* genome. Cutinase activity was reported in culture filtrates of *P. ultimum*, but its growth was not supported on apple cutin [69] and low levels of fatty acid esterase were detected in *P. ultimum* only in 21-day-old culture [70]. The absence of recognizable cutinases suggests these enzymes are not critical for penetration and infection by *P. ultimum*, which attacks young, non-suberized roots and penetrates tissues indirectly through wounds. This contrasts with the number of putative cutinases identified in several *Phytophthora* spp. [27,71-73], which presumably promote penetration of leaf and stem tissues that are protected by a thick cuticle or colonization of heavily suberized root and bark tissue.

The xylan degrading capacity of *P. ultimum* appears to be limited, if not totally absent. No members of the GH10 and GH11 families encoding endoxylanases essential for xylan degradation could be found. Furthermore, families involved in the removal of xylan side chains or modifications such as GH67, CE3, and CE5 are absent while families CE1 and CE2 contain only a limited number of members. The lack of significant xylan digestion was confirmed by the absence of growth when xylan was used as a carbon source (Table 3; Figure S10 in Additional file 3), consistent with previous work on *P. ultimum* and other *Pythium* spp. [70].

Pectinases play a key role in infection by *Pythium* spp. [74]. Twenty-nine candidate pectin/pectate lyases (PL1, PL3 and PL4 families) are present in *P. ultimum* while the genomes of *Phytophthora* spp. [27,28] encode even larger PL families (Table 2). In *P. ultimum*, the set of pectin lyases is complemented by 11 pectin hydrolases from family GH28, several of which having been functionally characterized in various *Phytophthora* spp. [75-78]. *P. ultimum* lacks pectin methylesterases as well as genes encoding family GH88 and GH105 enzymes and therefore cannot fully saccharify the products of pectin/pectate lyases, consistent with previous reports of incomplete pectin degradation and little or no galacturonic acid production during *P. ultimum* infection of bentgrass [79]. The data from the carbon source utilization experiment (Table 3; Figure S10 in Additional file 3) show only limited growth on medium with citrus pectin as the sole carbon source.

We also observed that the *P. ultimum* genome encodes candidate GH13 α-amylases, GH15 glucoamylase and a GH32 invertase, suggesting that plant starch and sucrose are targeted. The growth data confirm these observations, with excellent growth on soluble starch and sucrose (Figure S10 in Additional file 3).

The CAZy database also contains enzymes involved in fungal cell wall synthesis and remodeling. Cell walls of oomycetes differ markedly from cell walls of Fungi and consist mainly of glucans containing β-1,3 and β-1,6 linkages and cellulose [80-82]. The *P. ultimum* genome encodes four cellulose synthases closely related to their orthologs described for *Ph. infestans* [82]. The genome also specifies a large number of enzyme activities that may be involved in the metabolism of β-1,3- and β-1,6-glucans (Additional file 1), as well as a large set of candidate β-1,3-glucan synthases likely involved in synthesis of cell wall β-glucans and in the metabolism of mycolaminaran, the main carbon storage compound in *Phytophthora* and *Pythium* spp. [81,83,84].

Reponses to fungicide

Metalaxyl and its enantiopure R form mefenoxam have been used widely since the 1980s for the control of plant diseases caused by oomycetes [17,85]. The main mechanism of action of this fungicide is selective inhibition of ribosomal RNA synthesis by interfering with the activity of the RNA polymerase I complex [86]. *P. ultimum* DAOM BR144 is sensitive to mefenoxam at concentrations higher than 1 μl/l (data not shown) and 45 genes were expressed five-fold or more when *P. ultimum* was exposed to it (Table S6 in Additional file 2). Active ABC pump efflux systems are important factors for drug and antifungal resistance in Fungi and oomycetes [87-91]. Although the substrates transported by ABC proteins cannot be predicted on the basis of sequence homology, it is clear that these membrane transporters play a key role in the adaptation to environmental change. Three pleiotropic drug resistance proteins (ABC, subfamily G) were strongly up-regulated (> 27-fold) in response to mefenoxam. These genes arose from a tandem duplication event but remain so similar that it is possible that only one of these genes is actually up-regulated under these conditions due to our inability to uniquely map mRNA-seq reads when there are highly similar paralogs. A fourth gene and a member of the multidrug resistance associated family was also up-regulated more than nine-fold. Notably, the ABC transporters in *P. ultimum* that were up-regulated are distinct from those that were up-regulated in *Ph. infestans* in response to metalaxyl [92], indicating that a unique set of ABC transporters may be involved in the response to the fungicide in *P. ultimum*. Three genes coding for E3 ubiquitin-protein ligase were more than 18-fold up-regulated in response to mefenoxam compared to the control, but not in the other tested conditions. Ubiquitin/proteasome-mediated proteolysis is activated in response to stress - such as nutrient limitation, heat shock, and exposure to heavy metals - that may cause formation of damaged, denatured, or misfolded proteins [93,94]. Thus, increased expression of these enzymes in *P. ultimum* exposed to mefenoxam might be related to decreased synthesis of rRNA and expression of aberrant proteins.

Comparative genomics

Zoospore production

P. ultimum does not typically exhibit release of zoospores from sporangia in culture [12] but zoospore release directly from aged oospores has been reported [95]. Comparative genomics with well studied whiplash flagellar proteins from the green algae *Chlamydomonas reinhardtii* and other model organisms indicates that indeed *P. ultimum* does have the necessary genetic complement for flagella. Orthologs of tinsel flagellar mastigoneme proteins have also been identified in *P. ultimum* through comparison to those studied in *Ochromonas danica*, a unicellular member of the Straminipila kingdom. Overall, approximately 100 putative whiplash and tinsel flagellum gene orthologs were identified in *P. ultimum* (Table S7 in Additional file 2) with corresponding orthologs present in *Ph. infestans*, *Ph. sojae*, and *Ph. ramorum*. Expression of flagellar orthologs was observed in 8 growth conditions used in whole transcriptome sequencing, although 14 putative flagellar orthologs for axonemal dynein and kinesin and intraflagellar transport did not show expression in any condition.

Cadherins, an animal gene family found in oomycetes

Perhaps the most remarkable discovery relative to gene family expansion is that there are four *P. ultimum* genes that encode cadherins. Previously, members of this gene family have only been found in metazoan genomes (and the one fully sequenced genome from the clade of nearest relatives, the choanoflagellate *Monosiga brevicollis*). Cadherins are cell adhesion proteins that presumably evolved at the base of the clade containing metazoans and choanoflagellates [96]. Cadherin-related proteins are encoded in several bacterial genomes, but these bacterial proteins lack important calcium ion-binding motifs (the LDRE and DxND motifs) found in the extracellular (EC) repeat domains of 'true' cadherins [97]. The cadherin genes in *P. ultimum* do contain these motifs, and this is therefore the first report of true cadherins in a genome outside the metazoans/choanoflagellates. In metazoans, but not in choanoflagellates, some cadherins also contain an intracellular catenin-binding domain (CBD) that connects intercellular binding via EC domains to intracellular responses such as cytoskeletal changes. A search of predicted gene models with the PANTHER HMMs for cadherins (PTHR10596) identified two genes containing cadherin EC domains in the *Ph. infestans* genome, but none in the *Ph. ramorum*, *Ph. sojae* and *Phaeodactylum tricornutum* genomes. The identification of cadherin EC domains in both *P. ultimum* and *Ph. infestans* led us to postulate that such genes may also exist in other *Phytophthora* genomes that were not found in the original analysis of these genomes. Indeed, a TBLASTN search of genomic DNA using the predicted *P. ultimum* cadherin domain-containing proteins identified one putative cadherin-containing ORF in the *Ph. sojae* genome and four in the *Ph. ramorum* genome. The *P. ultimum* cadherin genes contain between 2 and 17 full-length cadherin EC domains, as predicted by the Pfam database [98] at the recommended statistical significance threshold, and likely a number of additional cadherin domains that have been truncated and/or have diverged past this similarity threshold. The genes from the *Phytophthora* genomes each contain between one and seven intact cadherin EC domains, though we did not attempt to construct accurate gene models for the *Phytophthora* genes. None of the oomycete cadherins appear to have the catenin-binding domain, nor do these genomes appear to encode a β-catenin gene, so like in *M. brevicollis*, the β-catenin-initiated part of the classical metazoan cadherin pathway appears to be absent from oomycetes.

In order to explore the evolution of these domains in the oomycetes, we performed a phylogenetic analysis. The first (amino-terminal) cadherin EC domain has been used to explore gene phylogeny among the cadherins [96,99], and to facilitate comparison we used both neighbor joining [100] and maximum likelihood (using the PhyML program [101,102]) to estimate a phylogenetic tree for these same sequences together with all of the intact cadherin domains from the *P. ultimum* and *Ph. infestans* genomes (Figure 2). To generate a high-quality protein sequence alignment for phylogeny estimation, we used the manual alignment of Nollet *et al*. [99] as a 'seed' for alignment of other sequences using MAFFT [102]. We found that all of the oomycete domains fall within a single clade. However, this clade is broad and also contains several cadherins from the choanoflagellate *M. brevicollis*, as well as some of the more divergent metazoan cadherins (Cr-2 and Cr-3 subfamilies). In general, the branches in this clade are very long, making phylogenetic reconstruction somewhat unreliable (all branches with bootstrap values > 50% are marked with a circle in Figure 2). Nevertheless, most of the cadherin domains found in *P. ultimum* are reliably orthologous to domains in one or more *Phytophthora* species, suggesting descent from a common ancestor by speciation. The most notable example is for the genes PITG_09983 and PYU1_T011030, in which a region spanning three consecutive EC repeats appears to have been inherited by both species from that common ancestor (apparently followed by substantial duplication and rearrangement of individual cadherin domains). These repeats are also apparently orthologous to repeats in both *Ph. sojae* and *Ph. ramorum*. The oomycete cadherins may have been initially obtained either vertically (by descent from the common ancestor with metazoans) or horizontally (by transfer of metazoan DNA long after divergence). No cadherins have been found in genomes

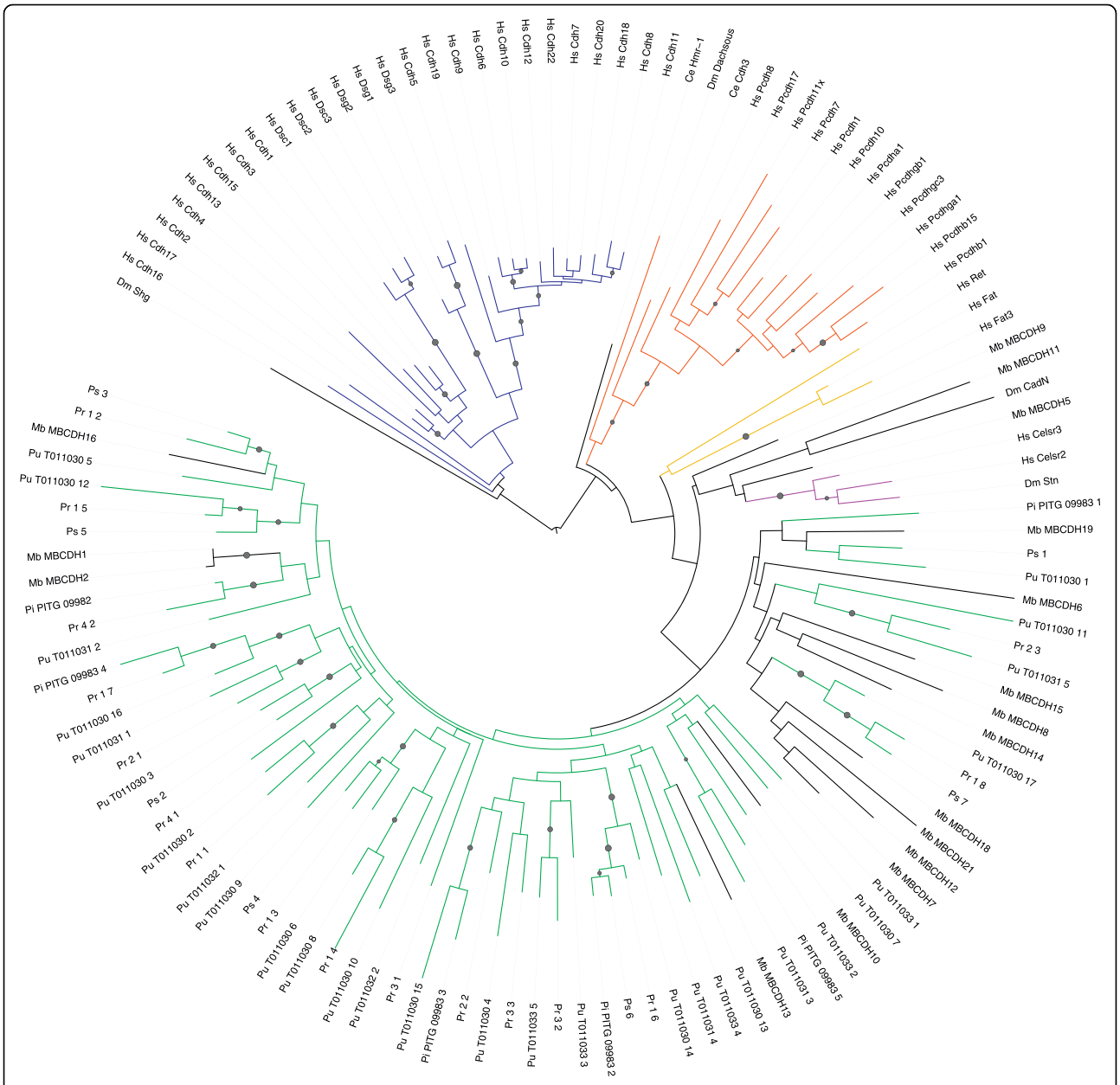


Figure 2 Phylogenetic tree of the cadherin family, showing all members of the novel oomycete subfamily (green) and their relationships to representative metazoan and choanoflagellate cadherins. The major clades of cadherins [96] are colored: C-1 (blue), Cr-1a and Cr-1b (red), C-2 (purple), and Cr-3 (orange). Most of the oomycete cadherins fall within a fairly distinct subfamily (green), though this subfamily has many long branches and also includes some cadherins from the choanoflagellate *M. brevicollis* (labeled starting with 'MB') that are also highly diverged from other cadherins. Reliable branches (bootstrap > 50%) are labeled with a circle. All full-length oomycete cadherin domains are shown, from *P. ultimum* (labeled starting with 'Pu' and ending with the number of the repeat relative to the amino terminus), *Ph. infestans* (labeled starting with 'Pi'), *Ph. sojae* (Ps) and *Ph. ramorum* (Pr). Other cadherins are from the human genome ('Hs') unless labeled starting with 'Dm' (*Drosophila melanogaster*) or 'Ce' (*Caenorhabditis elegans*). The figure was drawn using the iTOL tool [143].

sequenced from other clades more closely related to either oomycetes (for example, diatoms and alveolates) or the metazoan/choanoflagellates (for example, Fungi and amoebozoa). This means that, if cadherins were present in the most recent common ancestor of oomycetes and metazoans, these genes must have been lost independently in all of these other diverging lineages. Given the data currently available, it is more probable that at least one horizontal cadherin gene transfer event occurred from a choanoflagellate or metazoan to an oomycete ancestor, prior to the divergence of *Pythium* from *Phytophthora*. The source of the metazoan DNA

may have been a host of the ancestral oomycete, or possibly introduced by a virus. Nevertheless, the subsequent preservation of cadherin domains in at least two lineages of oomycetes over a substantial period of time suggests that the genes are likely to perform an important function, which remains to be explored.

Synteny with other oomycete plant pathogens

A phylogenetic approach (PHRINGE [103]) was used to identify *P. ultimum* proteins orthologous to proteins encoded in the genomes of *Ph. infestans*, *Ph. sojae*, and *Ph. ramorum*. Of the 15,322 proteins predicted from the *P. ultimum* genome sequence, 12,230 were identified as orthologous to a protein in at least one *Phytophthora* genome sequence. A total of 11,331 proteins were identified as orthologs common to all three *Phytophthora* spp., and of these, 8,504 had identifiable orthologs in *P. ultimum*. PHRINGE was also used to examine the conservation of gene order (synteny) between the *Phytophthora* and *P. ultimum* genomes. As previously described [27], the gene order of orthologs is very highly conserved among *Phytophthora* spp. In *P. ultimum* the ortholog content was very similar between broad regions of the *P. ultimum* and *Phytophthora* genomes, but the local gene order was greatly rearranged, primarily as a result of inversions. Only short runs of up to 10 orthologs were found to be collinear, whereas runs of more than 100 could be identified between the *Phytophthora* spp. Figure 3 shows an example spanning a well-assembled region of the *Ph. infestans*, *Ph. ramorum* and *P. ultimum* genome sequences. In *Ph. ramorum*, the region spans 1.18 Mb and 383 predicted genes and in *P. ultimum* the region spans 1.15 Mb and 435 predicted genes. Of these genes, 286 are identified as orthologs. In the *Ph. ramorum* sequence there are an additional 38 genes with orthologs in *Ph. infestans* but not in *P. ultimum*. Due to a much larger number of repeat sequences, and expanded gene numbers, the corresponding region in *Ph. infestans* spans 2.38 Mb and 499 predicted genes, but the order of the orthologous genes is highly conserved with that of *Ph. ramorum*. The *Ph. sojae* genome shows similar conservation of gene order in this region but for simplicity is not shown.

Conclusions

Analysis of the *P. ultimum* genome sequence suggests that not all oomycete plant pathogens contain a similar 'toolkit' for survival and pathogenesis. Indeed, *P. ultimum* has a distinct effector repertoire compared to *Phytophthora* spp., including a lack of the hallmark RXLR effectors, a limited number of *Crinkler* genes, and a novel YxSL[RK] family of candidate effectors. The absence of any convincing RXLR effectors from the predicted proteome of *P. ultimum*, first noted by Cheung *et al*. [31] and rigorously confirmed here, provides a striking contrast to the *Phytophthora* genomes. RXLR effectors are also absent from the proteome of *A. euteiches*, a member of the Saprolegniales, which was predicted from an EST collection [64]. It is possible that RXLR effectors are confined to oomycete pathogens in the family Peronosporaceae, and represent an adaptation to facilitate biotrophy. The absence of RXLR effectors from *P. ultimum* (and possibly all other species of the genus) may be functionally associated with the very broad host range of *Pythium* pathogens. It also correlates with the lack of gene-for-gene resistance against *Pythium* and the fact that *Pythium* pathogens are generally restricted to necrotrophic infection of seedlings, stressed plants, and plant parts (for example, fruit) with diminished defenses against infection. In contrast to the RXLR effectors, the genome of *P. ultimum* does encode

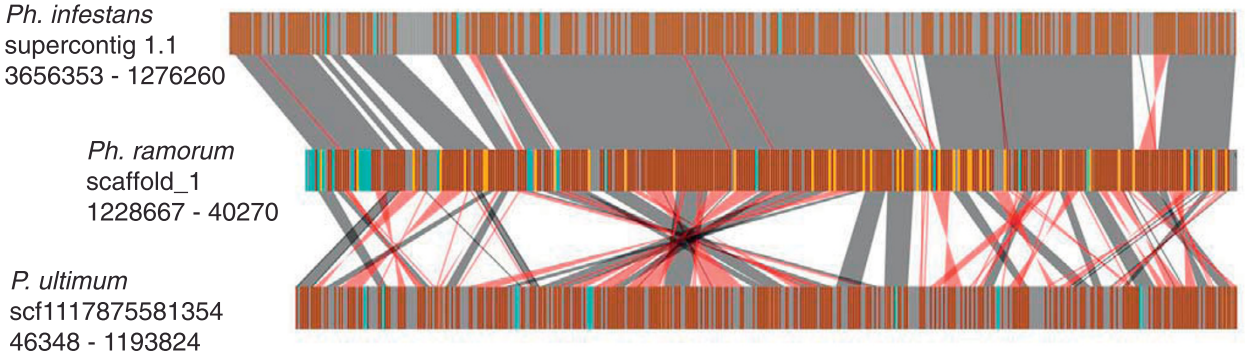


Figure 3 Rearrangements in gene order in the *P. ultimum* genome relative to *Phytophthora* genomes. Vertical brown bars indicate orthologs shared among *P. ultimum, Ph. infestans* and *Ph. ramorum*. Gold bars indicate orthologs shared only between *Ph. infestans* and *Ph. ramorum*. Turquoise bars indicate genes with orthologs in other regions of the compared genomes (that is, non-syntenic orthologs). Grey bars indicate genes without orthologs. Gray and red shaded connections indicate blocks of syntenic orthologs with the same or opposite relative transcriptional orientations, respectively. Non-coding regions of the genome are not depicted.

members of the Crinkler class of effectors, albeit not at the numbers present in *Phytophthora* genomes. These effectors may also enter host cells, and can trigger cell death [28]. They are also found in *A. euteiches* [64] and may represent a basal family of effectors that contribute to necrotrophy. This study uncovered a third family of secreted proteins conserved across all oomycetes sequenced so far with characteristics that suggest they might act inside host cells. These characteristics include high sequence variability, small size, hydrophilic nature, and a conserved RXLR-like motif (YxSL[KR]) with several family members specifically and highly expressed during infection. However, as yet no experimental data support this hypothesis.

The repertoire of metabolic genes within the *P. ultimum* genome reflects its pathogenic lifestyle (Figure 4). *P. ultimum* is an opportunistic pathogen of young seedlings and plant roots with little or no cuticle or heavily suberized tissue, consistent with lack of cutinase encoding genes. It is a poor competitor against secondary invaders of damaged plant tissues and soil organisms with better saprobic ability [13]. The *P. ultimum* genome contains a suite of GHs that fits well with an organism in this ecological niche. The genome encodes cellulases and pectinases that facilitate initial penetration and infection of the host, but it does not appear to use these plant polysaccharides as a major carbon source in culture and it lacks the ability to effectively degrade other complex polysaccharides such as xylan [70] (Table 3) and chitin [74,104]. As a primary pathogen that usually initiates infection, *P. ultimum* probably has first-hand access to easily degradable carbohydrates such as starch and sucrose. Following depletion of these carbon sources, it appears to focus on quick reproduction and production of survival structures [13] rather than switching its metabolism to the more difficult carbon sources such as plant cell wall polysaccharides. Intriguingly, the arsenal of *P. ultimum* enzymes targeting plant carbohydrates is strikingly similar to that found in the genome of the root-knot nematode *Meloidogyne incognita* [105], a root pathogen that also lacks xylanases yet has a strong pectin degrading capacity. In summary, access to the *P. ultimum* genome sequence has reinforced earlier hypotheses on pathogenesis and survival mechanisms in oomycete plant pathogens and has advanced our understanding of events at the plant-pathogen interface, especially during necrotrophy.

Materials and methods

Sequencing, assembly and autoclosure of *P. ultimum* DAOM BR144

P. ultimum (DAOM BR144 = CBS 805.95 = ATCC 200006) was sequenced using a whole-genome shotgun sequencing approach. Sequencing of three Sanger libraries generated 263,715 quality filtered reads (281,088 attempted reads). Three full runs of 454 FLX standard pyrosequencing [106] generated 1,296,941 reads. These were assembled by a 'shredding' pipeline [107] that generates pseudo-Sanger reads from the contigs of a Newbler assembly [106] of 454 reads and then assembles all of the reads with Celera Assembler [108]. This yielded 2,659 contigs in 714 scaffolds with an N_{50} contig size of 40,520 bp. The 1,945 intra-scaffold gaps were subjected to AutoClosure, an in-house pipeline that automates primer design, template re-array, and reaction orders. This produced 6,468 reads, of which 5,014 passed quality filtering. Subsequently, the Celera Assembler software was modified to accept 454 reads without shredding [109] and Celera Assembler 5.2 was run on the Sanger shotgun, 454 shotgun, and Sanger AutoClosure reads together. Contigs for the mitochondrial genome were identified and annotated separately with 16,277 sequences assembled for a greater than 200-fold coverage. The whole genome shotgun project has been deposited at NCBI [GenBank:ADOS00000000] along with the 454 reads [SRA:SRX020087], the mitochondrial genome [GenBank:GU138662], and the Sanger reads (NCBI Trace Archive under species code 'PYTHIUM ULTIMUM DAOM BR144'). The version described in this paper is the first version [WGS: ADOS01000000].

Genome annotation

The *P. ultimum* genome annotations were created using the MAKER program [110]. The program was configured to use both spliced EST alignments as well as single exon ESTs greater than 250 bp in length as evidence for producing hint-based gene predictions. MAKER was also set to filter out gene models for short and partial gene predictions that produce proteins with fewer than 28 amino acids. The MAKER pipeline was set to produce *ab initio* gene predictions from both the repeat-masked and unmasked genomic sequence using SNAP [111], FGENESH [112], and GeneMark [113]. Hint-based gene predictions were derived from SNAP and FGENESH.

The EST sequences used in the annotation process were derived from Sanger and 454 sequenced *P. ultimum* DAOM BR144 ESTs [31] considered together with ESTs from dbEST [114] for *Aphanomyces cochlioides*, *Phytophthora brassicae*, *Phytophthora capsici*, *Phytophthora parasitica*, *Ph. sojae*, *Ph. infestans*, and *Pythium oligandrum*. Protein evidence was derived from the UniProt/Swiss-Prot protein database [115,116] and from predicted proteins for *Ph. infestans* [28], *Ph. ramorum* [27], and *Ph. sojae* [27]. Repetitive elements were identified within the MAKER pipeline using the Repbase repeat library [117] and RepeatMasker [45] in

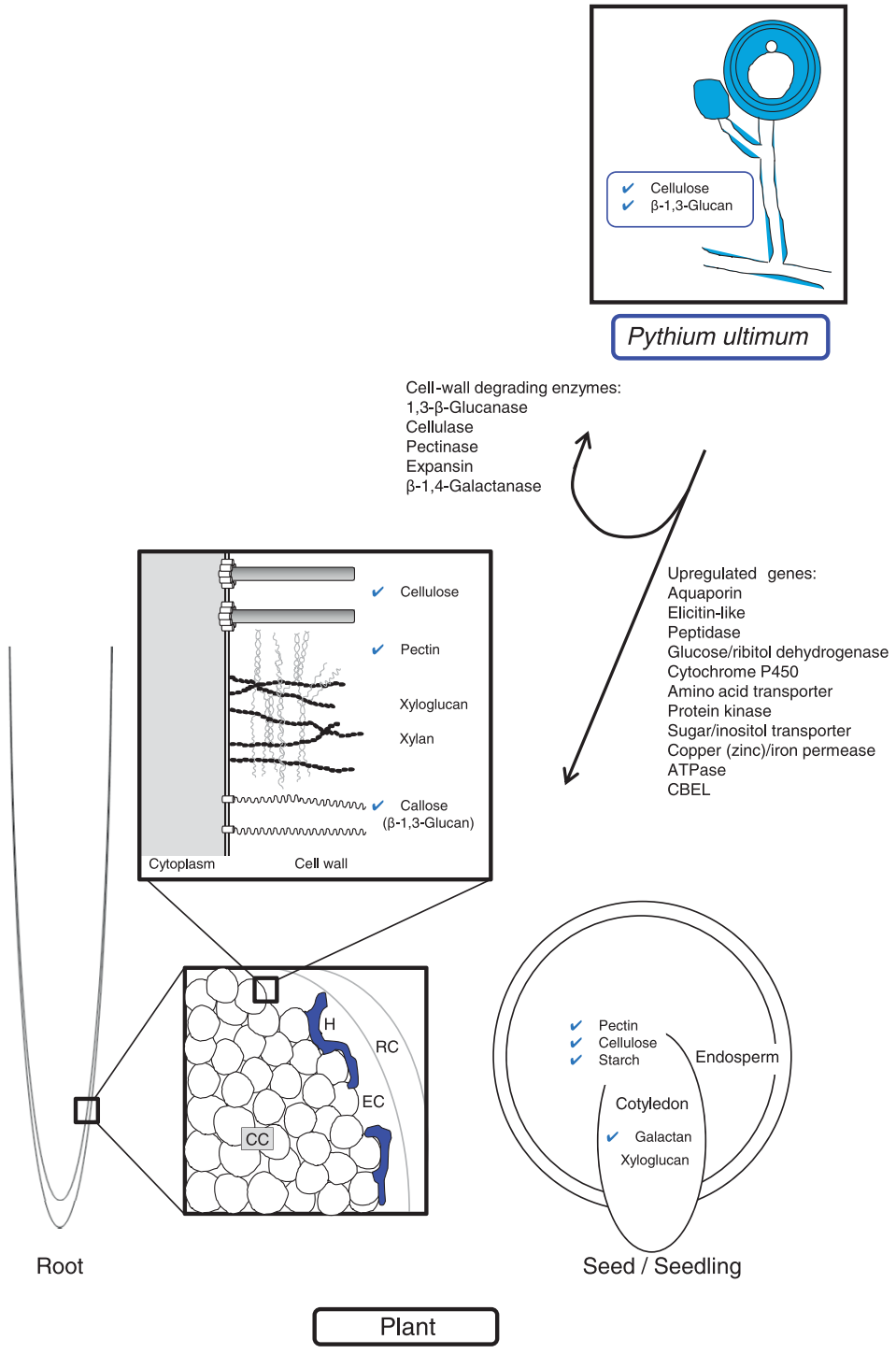


Figure 4 The *P. ultimum* genome contains genes encoding enzyme activities necessary for the degradation of plant cell wall polysaccharides and storage sugars (blue ticks indicate the polysaccharides targeted by *P. ultimum* enzymes). Some activities are equally relevant for *P. ultimum*'s own cell wall metabolism. Degradation of the plant cell wall relies essentially on the action of cellulases and pectinases. Significantly, the absence of identified enzymes with xylanase, pectin methylesterase or cutinase activities is in agreement with previous studies of *P. ultimum* and other *Pythium* spp. [70,104,144]. For *Pythium*'s pathogenic action, penetration is primarily limited to wounded tissue, or to young roots and germinating seedlings with little or no suberized tissue. Penetration and root rot, for some *Pythium* spp., is limited to the first layers of cells (RC and EC) [104]. Other genes, including those coding for transporters, elicitin-like, and stress proteins, were upregulated when *P. ultimum* was grown in contact with *A. thaliana* seeds. CC, cortical cells; EC, epidermal cells; RC, root cap; H, hyphae. (Figure adapted from [104,144-146].)

conjunction with a MAKER internal transposable element database [118] and a *P. ultimum* specific repeat library prepared for this work (created using PILER [119] with settings suggested in the PILER documentation). *Ab initio* gene predictions and hint-based gene predictions [110] were produced within the MAKER pipeline using FGENESH trained for *Ph. infestans*, GeneMark trained for *P. ultimum* via internal self-training, and SNAP trained for *P. ultimum* from a conserved gene set identified by CEGMA [110].

Following the initial MAKER run, a total of 14,967 genes encoding 14,999 transcripts were identified, each of which were supported by homology to a known protein or had at least one splice site confirmed by EST evidence. Additional *ab initio* gene predictions not overlapping a MAKER annotation were scanned for protein domains using InterProScan [120-122]. This process identified an additional 323 gene predictions; these were added to the annotation set, producing a total of 15,290 genes encoding 15,322 transcripts (referred to as v3). Selected genes within the MAKER produced gene annotation set were manually annotated using the annotation-editing tool Apollo [123]. The final annotation set (v4) contained 15,297 genes encoding 15,329 transcripts, including six rRNA transcripts.

Putative functions were assigned to each predicted *P. ultimum* protein using BLASTP [124] to identify the best homologs from the UniProt/Swiss-Prot protein database and/or through manual curation. Additional functional annotations include molecular weight and isoelectric point (pI) calculated using the pepstats program from the EMBOSS package [125], subcellular localization predicted with TargetP using the non-plant network [126], prediction of transmembrance helices via TMHMM [127], and PFAM (v23.0) families using HMMER [128] in which only hits above the trusted cutoff were retained. Expert annotation of carbohydrate-related enzymes was performed using the Carbohydrate-Active Enzyme database (CAZy) annotation pipeline [68].

Transcriptome sequencing

Eight cDNA libraries were constructed to assess the expression profile of *P. ultimum*. Initially, plugs of 10% V8 agar containing *P. ultimum* strain DAOM BR144 were incubated for 1 day in yeast extract broth (YEB; 30 g/l sucrose, 1 g/l KH_2PO_4, 0.5 g/l $MgSO_4 \cdot 7H_2O$, 0.5 g/l KCl, 10 mg/l $FeSO_4 \cdot 7H_2O$, 1 g/l yeast extract) at 25°C with shaking (200 rpm). Approximately 50 mg of hyphae growing out of the plugs were then transferred to flasks containing media for the various expression assays. Mycelium was grown under the following conditions: 1, nutrient-rich YEB medium for 3 days at 25°C with shaking (200 rpm) and nutrient-starved Plich medium (S Kamoun, unpublished) for 10 days at 25°C in standing culture, as previously described [31]; 2, YEB medium under hypoxic conditions (oxygen concentration of 0.2%) for 1 and 3 days in standing liquid culture at 25°C; 3, YEB medium for 2 days at 25°C with shaking (200 rpm) followed by the addition of 1 and 100 μl/l of the fungicide mefenoxam (Subdue MAXX™, Novartis Crop Production, Greensboro, NC, USA) and subsequent incubation for an additional 0.25, 3 and 6 hours at the same temperature and with agitation; 4, YEB medium for the same time periods but without the addition of mefenoxam was included (mefenoxam control); 5, YEB medium for 2 days at 25°C with shaking (200 rpm) followed by a cold stress of 0°C with shaking (200 rpm) for 0.25, 3 and 6 hours; or 6, YEB medium for 2 days at 25°C with shaking (200 rpm) followed by exposure to heat stress of 35°C for 0.25, 3 and 6 hours; 7, YEB medium for 2 days at 25°C followed by exposure to 25°C for 0.25, 3 and 6 hours (temperature control); 8, 0.1% V8-juice medium containing surface-sterilized *A. thaliana* ecotype Columbia Col-0 seeds. Approximately 200 seeds were placed in the liquid medium at 25°C with shaking (200 rpm) for 1, 2 and 7 days. Mycelium of *P. ultimum* was then added and allowed to grow in contact with the seeds for 3 days.

For each condition listed above, mycelium was harvested, macerated in liquid nitrogen and RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) as described [31]. RNA was treated with DNAse (Promega RQ1 RNase-Free DNase, Madison, WI, USA) and 10 μg RNA was used to construct cDNA using the mRNA-Seq Sample Prep Kit (Illumina, San Diego, CA, USA), which was sequenced with Illumina Genome Analyzer (GA) II using version 3 sequencing reagents for 41 cycles. Base calling was carried out using the Illumina GA pipeline v1.4.

For each library, the filtered reads from the Illumina GA II pipeline were mapped using Tophat, a splice-site-aware short read mapper that works in conjunction with Bowtie short read aligner [129]. Reads were deposited in the NCBI Short Read Archive [SRA:SRP002690]. The minimum and maximum intron sizes were 5 bp and 15 kbp, respectively, for each Tophat run. The final annotation GFF3 file was provided to Tophat and expression values were calculated using reads per kilobase of exon model per million mapped reads (RPKM) [130]. The minimum RPKM for all eight conditions was 0, the median RPKM ranged from 5 to 8, while the maximum RPKM ranged from 10,182 in the YEP+Plich library to 32,041 from the 35°C temperature treatment. Using a RPKM value of 2.5 (approximately half of the median RPKM of all genes in each library) as a cutoff for expression, loci with differential expression in treatment versus control were identified (comparisons: hypoxia

versus YEB+Plich; *Arabidopsis* versus YEB+Plich; mefenoxam treatment versus mefenoxam control; heat treatment versus temperature control; cold treatment versus temperature control; Table S6 in Additional file 2). The fold changes were calculated for loci with RPKM ≥ 2.5 in both treatment and control samples. Loci with control values < 2.5 RPKM but with expression in treatment conditions were flagged as 'U' as a true ratio could not be calculated. Loci with treatment values < 2.5 RPKM but with expression in control conditions were flagged as 'D' as a true ratio could not be calculated. Loci with RPKM < 2.5 in both treatment and control samples were flagged as 'N'.

Identification of secreted proteins and effector families

The secretome of *P. ultimum* was identified using SignalP V2.0 program following the PexFinder algorithm as described previously [56]. In addition, sequences that were predicted to contain transmembrane domains or organelle targeting signals were omitted from the secretome. Each sequence in the secretome was searched against two 'Darwin' databases [57] that were compiled from > 50 eukaryote whole proteomes from major phylogenetic branches using BLASTP with an E-value cutoff of 1×10^{-3}. One database contained sequences only from Fungi and the other contained the sequences from other organisms excluding the Fungi and oomycetes. Protein sequences of the secretome were clustered into families along with their related non-secretory proteins by using the TRIBE-MCL algorithm [131] using BLASTP with an E-value cutoff of 1×10^{-10}. Each family was named according to the existing annotation of the member sequences. Families and singletons were searched against Pfam-A release 24.0 using the HMMER3 beta 3 hmmsearch with trusted cutoffs to detect any transposable-element-related proteins that may have been missed in the repeat masking process. Families or singletons where at least 50% of the members matched transposon-associated Pfam domains were manually curated to identify and exclude true transposon-related sequences from the secretome.

For the analysis of genome organization, *P. ultimum* predicted genes were binned according to the length of their flanking non-coding regions (FIRs). FIRs were computed using predicted gene coordinates on scaffolds. Binning according to 5' FIRs and 3' FIRs was performed along the x-aixs and y-axis, respectively, using conditional counting functions. Logarithmic size was chosen for the bins in order to allow a maximum dispersion of the values. A color code was used to represent the number of genes or average values in bins. Average values were computed for bins containing a minimum of three genes.

Motif searches were done using the MEME [132] prediction server with default parameters except the following: min width = 4; max width = 12; min sites = 10. Sequences with homology to gene models in oomycetes genomes were identified by BLAST analysis against the NR database and aligned using MUSCLE [133]. For phylogenetic inference of the *CRN* genes, alignments were done using RevTrans [134] with the dialign-T algorithm. Molecular phylogenetic reconstructions were done using RAxML [135] version 7.0. Sequence logos were constructed on the basis of the RevTrans alignment using WebLogo [136].

Comparative genomics analyses

In order to find substantial expansions and contractions of gene families observed in other eukaryotes, we used the PANTHER Classification System [49,137,138]. We first scored all predicted proteins from the *P. ultimum* genome against the PANTHER HMMs, and created a tab-delimited file with two columns: the *P. ultimum* protein identifier and the PANTHER HMM identifier from the top-scoring HMM (if E-value < 0.001). We created similar files for three *Phytophthora* genomes (*Ph. infestans*, *Ph. ramorum*, and *Ph. sojae*), and a diatom genome (*P. tricornutum*) for comparison. We removed protein families of probable viral origin or transposons (PTHR19446, PTHR10178, PTHR11439, PTHR23022, PTHR19303). This left 7,762 *P. ultimum* proteins in PANTHER families, 8,169 from *Ph. infestans*, 7,667 from *Ph. ramorum* and 7,701 from *Ph. sojae*. We then uploaded the tab-delimited files to the PANTHER Gene List Comparison Tool [137,139] and analyzed the list for under- and over-representation of genes with respect to molecular functions, biological processes, and pathways. For each class that was significantly different (Bonferroni-corrected $P < 0.05$) between *P. ultimum* and all of the *Phytophthora* genomes, we determined the protein family expansions or contractions that made the biggest contributions to these differences (Table 1). Finally, we determined likely gene duplication and loss events that generated the observed protein family expansions and contractions by building phylogenetic trees of each of these families using the 48 genomes included in the trees on the PANTHER website [140], in addition to the five stramenopile genomes above (*P. ultimum*, *Ph. infestans*, *Ph. ramorum*, *Ph. sojae*, *P. tricornutum*). Phylogenetic trees were constructed using the GIGA algorithm [141], which infers the timing of likely gene duplication events relative to speciation events, allowing the reconstruction of ancestral genome content and lineage-specific duplications and losses. Using v3 of the annotation (MAKER output without manual curation), *P. ultimum* genes orthologous to genes in *Ph. infestans*, *Ph. sojae* and *Ph. ramorum* were identified using PHRINGE ('Phylogenetic

Resources for the Interpretation of Genomes') [103] in which the evolutionary relationships among all oomycete protein families are reconstructed.

Carbohydrate utilization

Growth was compared on different media. Carbon sources were added to Minimal Media (MM; per liter: 0.5 g KH_2PO_4, 0.5 g K_2HPO_4, 4 × 10^{-4} g $MnSO^4$, 4 × 10^{-4} g $ZnSO_4$, 1.05 g NH_4Cl, 6.8 ml 1M $CaCl_2{\cdot}2H_2O$, 2 ml 1M $MgSO_4{\cdot}7H_2O$, 4 × 10^{-3} g $FeSO_4$ and 1% (w/v) agarose) at the following concentrations: 1% (w/v) for cellulose, soluble starch, citrus pectin and birchwood xylan and 25 mM for D-glucose, D-fructose, D-xylose, cellobiose, sucrose and L-arabinose. The pH of the medium was adjusted to 6.0 and the medium was autoclaved at 121°C for 20 minutes. $CaCl_2$, $MgSO_4$, and monosaccharides were autoclaved separately from the rest of the medium and $FeSO_4$ was sterile filtered (Whatman 0.2 μm millipore filter, Dassel, Germany). All of these components were added to the autoclaved medium before it solidified. The growth of *P. ultimum* DAOM BR144 was compared on the different media mentioned above; Minimal Media without a carbon source was used as the negative control in this experiment. The strain was initially grown on Potato Carrot Agar [142]. A small agar plug containing mycelium (1 mm diameter) was transferred from the edge of a vigorously growing 1-day-old colony to the center of the Petri dishes with the different media. The cultures were incubated in the dark at 21°C. Mycelium density and colony diameter were measured daily for the first 5 days and again after 7 days. Colony morphology pictures were taken, and pH was measured after 7 days. The growth test was conducted twice for each strain.

Additional material

Additional file 1: Supplemental methods and results. Additional details on sequencing methods and analysis results citing methods or data from [147-166].

Additional file 2: Supplemental Tables S1 to S11 providing detailed lists and analyses.

Additional file 3: Supplemental Figures S1 to S16, supporting data analyses.

Additional file 4: Multiple sequence alignment of oomycete proteins with similarity to *P. ultimum* Family 3 proteins. Predicted secreted proteins (91) with similarity to Family 3 proteins from various oomycete species were aligned demonstrating the YxSL[KR] motif.

Abbreviations

bp: base pairs; CBEL: cellulose-binding elicitor lectin; CRN" Crinkler; EC: extracellular; EST: expressed sequence tag; FIR: flanking non-coding region; GH: glycoside hydrolase; HMM: Hidden Markov Model; MAMP: microbial-associated molecular pattern; ORF: open reading frame; PAMP: pathogen-associated molecular pattern; RPKM: reads per kilobase of exon model per million mapped reads; YEB: yeast extract broth.

Acknowledgements

Funding for the work was provided by the US Department of Agriculture (USDA) National Institute of Food and Agriculture Microbial Genome Sequencing Program to CRB and NT (2007-35600-17774 and 2007-35600-18886). The MAKER genome annotation pipeline is funded by NIH/NHGRI 5R01HG004694 to MY. Analysis of the genome was funded in part by grants to BMT from the National Research Initiative of the USDA CSREES #2007-35600-18530 and from the National Science Foundation #MCB-0731969. CAL and GPR are supported by the NSERC Discovery and Network programs. CMMG is the recipient of a Marie Curie Intra-European Fellowship (MIEF-CT-2006-022837) and a European Reintegration Grant (PERG03-GA-2008-230865) from the European Commission. HB is supported by the Dutch Ministry of Agriculture, Nature and Food Quality through a FES program. MT acknowledges support by the Max-Planck Society, the German Science Foundation (DFG), the Landesstiftung Baden-Württemberg and the LOEWE - Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz research program of Hesse's Ministry of Higher Education, Research, and the Arts. An allocation of computer time from the Center for High Performance Computing at the University of Utah is gratefully acknowledged. The computational resources for annotating the genome were provided by the National Institutes of Health (grant # NCRR 1 S10 RR17214-01) on the Arches Metacluster, administered by the University of Utah Center for High Performance Computing. We want to thank the Canadian Collection of Fungal Cultures (CCFC/DAOM) for supplying and maintaining the culture for this study. We acknowledge the assistance of Nicole Désaulniers for culturing and DNA extractions as well as Jason Miller and Brian Walenz of the J Craig Venter Institute in assembly of the *P. ultimum* genome. We want to thank George Mahuku for sharing results on resistance of beans before publication.

Author details

[1]Agriculture and Agri-Food Canada, 960 Carling Ave, Ottawa, ON, K1A 0C6, Canada. [2]Department of Biology, Carleton University, Ottawa, ON, K1S 5B6, Canada. [3]CBS-KNAW, Fungal Biodiversity Centre, Uppsalalaan 8, Utrecht, 3584 CT, The Netherlands. [4]The Sainsbury Laboratory, Norwich, NR4 7UH, UK. [5]Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA. [6]Eccles Institute of Human Genetics, University of Utah, 15 North 2030 East, Room 2100, Salt Lake City, UT 84112-5330, USA. [7]Biodiversity and Climate Research Centre, Georg-Voigt-Str 14-16, D-60325, Frankfurt, Germany. [8]Department of Biological Sciences, Insitute of Ecology, Evolution and Diversity, Johann Wolfgang Goethe University, Siesmayerstr. 70, D-60323 Frankfurt, Germany. [9]Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523-1177, USA. [10]School of Biology, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK. [11]Genome Project Solutions, 1024 Promenade Street, Hercules, CA 94547, USA. [12]J Craig Venter Institute, 9704 Medical Center Dr., Rockville, MD 20850, USA. [13]Surfaces Cellulaires et Signalisation chez les Végétaux, UMR5546 CNRS-Université de Toulouse, 24 chemin de Borde Rouge, BP42617, Auzeville, Castanet-Tolosan, F-31326, France. [14]Scottish Association for Marine Science, Oban, PA37 1QA, UK. [15]Laboratory of Phytopathology, Wageningen University, NL-1-6708 PB, Wageningen, The Netherlands. [16]Centre for BioSystems Genomics (CBSG), PO Box 98, 6700 AB Wageningen, The Netherlands. [17]Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD, UK. [18]The Broad Institute of MIT and Harvard, Cambridge, MA 02141, USA. [19]Department of Pathology, Faculty of Medicine-Ramathibodi Hospital, Mahidol University, Rama 6 Road, Bangkok, 10400, Thailand. [20]Department of Biological Sciences, Bowling Green State University, Bowling Green, OH 43403, USA. [21]Department of Plant Pathology and Microbiology, University of California, Riverside, CA 92521, USA. [22]Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Washington Street, Blacksburg, VA 24061-0477, USA. [23]Architecture et Fonction des Macromolecules Biologiques, UMR6098, CNRS, Univ. Aix-Marseille I & II, 163 Avenue de Luminy, 13288 Marseille, France. [24]USDA-ARS, 1636 East Alisal St, Salinias, CA, 93905, USA. [25]Evolutionary Systems Biology, SRI International, Room AE207, 333 Ravenswood Ave, Menlo Park, CA 94025, USA.

Authors' contributions

CAL, NT, and CRB directed the project, performed analyses, and drafted the manuscript. GWB drafted the manuscript. HB, LC, EH, SR, GR, MT, JW, JLB, BD, SIF, CMMG, EG, FG, LG-B, NH, RHYJ, TK, HJGM, PM, VP, ST, SW, PvW, PMC, BH, FM, PDT, BMT, RPDV, and SK participated in the genome analysis and drafted the manuscript. JPH, CH, BRW, and MY conducted genomic/transcriptomic analyses and drafted the manuscript. MMZ isolated RNA, constructed cDNA libraries, participated in the genome analysis and drafted the manuscript. DB, JJ, HL, BM, DP, and JES conducted genomic/transcriptomic analyses. SF, JH, and JS performed genome sequencing. All authors read and approved the final manuscript.

Received: 17 February 2010 Revised: 2 May 2010
Accepted: 13 July 2010 Published: 13 July 2010

References

1. Cavalier-Smith T, Chao EE: **Phylogeny and megasystematics of phagotrophic heterokonts (kingdom Chromista).** *J Mol Evol* 2006, **62**:388-420.
2. Riisberg I, Orr RJ, Kluge R, Shalchian-Tabrizi K, Bowers HA, Patil V, Edvardsen B, Jakobsen KS: **Seven gene phylogeny of heterokonts.** *Protist* 2009, **160**:191-204.
3. Burki F, Shalchian-Tabrizi K, Pawlowski J: **Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes.** *Biol Lett* 2008, **4**:366-369.
4. Tsui CK, Marshall W, Yokoyama R, Honda D, Lippmeier JC, Craven KD, Peterson PD, Berbee ML: **Labyrinthulomycetes phylogeny and its implications for the evolutionary loss of chloroplasts and gain of ectoplasmic gliding.** *Mol Phylogenet Evol* 2009, **50**:129-140.
5. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou SG, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, *et al*: **The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism.** *Science* 2004, **306**:79-86.
6. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, *et al*: **The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes.** *Nature* 2008, **456**:239-244.
7. Cock JM, Sterck L, Rouze P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury JM, Badger JH, Beszteri B, Billiau K, Bonnet E, Bothwell JH, Bowler C, Boyen C, Brownlee C, Carrano CJ, Charrier B, Cho GY, Coelho SM, Collen J, Corre E, Da Silva C, Delage L, Delaroque N, Dittami SM, Doulbeau S, Elias M, Farnham G, *et al*: **The Ectocarpus genome and the independent evolution of multicellularity in brown algae.** *Nature* **465**:617-621.
8. Cook RJ, Sitton JW, Waldher JT: **Evidence for *Pythium* as a pathogen of direct-drilled wheat in the Pacific Northwest.** *Plant Dis* 1980, **64**:102-103.
9. Larkin RP, English JT, Mihail JD: **Effects of infection by *Pythium* spp. on the root system morphology of alfalfa seedlings.** *Phytopathology* 1995, **85**:430-435.
10. Sumner DR, Gascho GJ, Johnson AW, Hook JE, Threadgill ED: **Root diseases, populations of soil fungi, and yield decline in continuous double-crop corn.** *Plant Disease* 1990, **74**:704-710.
11. Snowdon AL: **A color atlas of post-harvest diseases & disorders of fruits and vegetables.** *CRC Press* 1990, **I**:302.
12. van der Plaats-Niterink AJ: **Monograph of the genus *Pythium*.** *Studies Mycol* 1981, **21**:242.
13. Martin FN, Loper JE: **Soilborne plant diseases caused by *Pythium* spp.: ecology, epidemiology, and prospects for biological control.** *Crit Rev Plant Sci* 1999, **18**:111-181.
14. Saunders GA, Washburn JO, Egerter DE, Anderson JR: **Pathogenicity of fungi isolated from field-collected larvae of the western treehole mosquito, *Aedes sierrensis* (Diptera: Culicidae).** *J Invertebr Pathol* 1988, **52**:360-363.
15. Mendoza L, Ajello L, McGinnis MR: **Infections caused by the oomycetous pathogen *Pythium insidiosum*.** *J Mycologie Medicale* 1996, **6**:151-164.
16. Lévesque CA, deCock WAM: **Molecular phylogeny and taxonomy of the genus *Pythium*.** *Mycol Res* 2004, **108**:1363-1383.
17. Erwin DC, Ribeiro OK: **Phytophthora *Diseases Worldwide*.** St Paul, MN: American Phytopathological Society 1996.
18. Cooke DE, Drenth A, Duncan JM, Wagels G, Brasier CM: **A molecular phylogeny of *Phytophthora* and related oomycetes.** *Fungal Genet Biol* 2000, **30**:17-32.
19. Panabières F, Ponchet M, Allasia V, Cardin L, Ricci P: **Characterization of border species among Pythiaceae: several *Pythium* isolates produce elicitins, typical proteins from *Phytophthora* spp.** *Mycol Res* 1996, **101**:1459-1468.
20. Jiang RH, Tyler BM, Whisson SC, Hardham AR, Govers F: **Ancient origin of elicitin gene clusters in *Phytophthora* genomes.** *Mol Biol Evol* 2006, **23**:338-351.
21. Rosso ML, Rupe JC, Chen P, Mozzoni LA: **Inheritance and genetic mapping of resistance to Pythium damping-off caused by *Pythium aphanidermatum* in 'Archer' soybean.** *Crop Sci* 2008, **48**:2215-2222.
22. Yang DE, Jin DM, Wang B, Zhang DS, Nguyen HT, Zhang CL, Chen SJ: **Characterization and mapping of Rpi1, a gene that confers dominant resistance to stalk rot in maize.** *Mol Genet Genomics* 2005, **274**:229-234.
23. Latijnhouwers M, de Wit PJGM, Govers F: **Oomycetes and fungi: similar weaponry to attack plants.** *Trends Microbiol* 2003, **11**:462-469.
24. Francis DM, St Clair DA: **Outcrossing in the homothallic oomycete, *Pythium ultimum*, detected with molecular markers.** *Curr Genet* 1993, **24**:100-106.
25. Barr DJS, Warwick SI, Desaulniers NL: **Isozyme variation, morphology, and growth response to temperature in *Pythium ultimum*.** *Can J Bot* 1996, **74**:753-761.
26. Martin F: **Phylogenetic relationships among some *Pythium* species inferred from sequence analysis of the mitochondrially encoded cytochrome oxidase II gene.** *Mycologia* 2000, **92**:711-727.
27. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorrance 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, *et al*: ***Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis.** *Science* 2006, **313**:1261-1266.
28. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AM, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JI, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, *et al*: **Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*.** *Nature* 2009, **461**:393-398.
29. Jiang HY, Tripathy S, Govers F, Tyler BM: **RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members.** *Proc Natl Acad Sci USA* 2008, **105**:4874-4879.
30. Kamoun S: **A catalogue of the effector secretome of plant pathogenic oomycetes.** *Annu Rev Phytopathol* 2006, **44**:41-60.
31. Cheung F, Win J, Lang J, Hamilton J, Vuong H, Leach J, Kamoun S, Levesque CA, Tisserat N, Buell CR: **Analysis of the *Pythium ultimum* transcriptome using Sanger and pyrosequencing approaches.** *BMC Genomics* 2008, **9**:542.
32. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, *et al*: **The sequence of the human genome.** *Science* 2001, **291**:1304-1351.
33. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, *et al*: **The genome sequence of *Drosophila melanogaster*.** *Science* 2000, **287**:2185-2195.
34. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K,

Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, *et al*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.

35. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, *et al*: **Initial sequencing and comparative analysis of the mouse genome.** *Nature* 2002, **420**:520-562.
36. *C. elegans* Sequencing Consortium: **Genome sequence of the nematode *C. elegans*: a platform for investigating biology.** *Science* 1998, **282**:2012-2018.
37. *Arabidopsis* Genome Initiative: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.** *Nature* 2000, **408**:796-815.
38. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG: **Life with 6000 genes.** *Science* 1996, **274**:546-567.
39. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, *et al*: **The genome sequence of the filamentous fungus *Neurospora crassa*.** *Nature* 2003, **422**:859-868.
40. Yandell M, Mungall CJ, Smith C, Prochnik S, Kaminker J, Hartzell G, Lewis S, Rubin GM: **Large-scale trends in the evolution of gene structures within 11 animal genomes.** *PLoS Comput Biol* 2006, **2**:e15.
41. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB: **Alternative isoform regulation in human tissue transcriptomes.** *Nature* 2008, **456**:470-476.
42. Devos KM, Brown JKM, Bennetzen JL: **Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*.** *Genome Res* 2002, **12**:1075-1079.
43. Eichler EE, Sankoff D: **Structural dynamics of eukaryotic chromosome evolution.** *Science* 2003, **301**:793-797.
44. Yoder JA, Walsh CP, Bestor TH: **Cytosine methylation and the ecology of intragenomic parasites.** *Trends Genet* 1997, **13**:335-340.
45. Judelson HS, Randall TA: **Families of repeated DNA in the oomycete *Phytophthora infestans* and their distribution within the genus.** *Genome* 1998, **41**:605-615.
46. Avila-Adame C, Gomez-Alpizar L, Zismann V, Jones KM, Buell CR, Beagle Ristaino J: **Mitochondrial genome sequences and molecular evolution of the Irish potato famine pahtogen, *Phytophthora infestans*.** *Curr Genet* 2005, **49**:39-46.
47. Grayburn WS, Hudspeth DS, Gane MK, Hudspeth ME: **The mitochondrial genome of *Saprolegnia ferax* : organization, gene content, and nucleotide sequence.** *Mycologia* 2004, **96**:981-989.
48. Martin FN, Bensasson D, Tyler BM, Boore JL: **Mitochondrial genome sequences and comparative genomics of *Phytophthora ramorum* and *P. sojae*.** *Curr Genet* 2007, **51**:285-296.
49. Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, Guo N, Muruganujan A, Doremieux O, Campbell MJ, Kitano H, Thomas PD: **The PANTHER database of protein families, subfamilies, functions and pathways.** *Nucleic Acids Res* 2005, **33**:D284-288.
50. Boller T, He SY: **Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens.** *Science* 2009, **324**:742-744.
51. Rosebrock TR, Zeng L, Brady JJ, Abramovitch RB, Xiao F, Martin GB: **A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity.** *Nature* 2007, **448**:370-374.
52. Zeng LR, Vega-Sanchez ME, Zhu T, Wang GL: **Ubiquitination-mediated protein degradation and modification: an emerging theme in plant-microbe interactions.** *Cell Res* 2006, **16**:413-426.
53. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
54. Birch PR, Armstrong M, Bos J, Boevink P, Gilroy EM, Taylor RM, Wawra S, Pritchard L, Conti L, Ewan R, Whisson SC, van West P, Sadanandom A, Kamoun S: **Towards understanding the virulence functions of RXLR effectors of the oomycete plant pathogen *Phytophthora infestans*.** *J Exp Bot* 2009, **60**:1133-1140.
55. Birch PR, Rehmany AP, Pritchard L, Kamoun S, Beynon JL: **Trafficking arms: oomycete effectors enter host plant cells.** *Trends Microbiol* 2006, **14**:8-11.
56. Torto TA, Li S, Styer A, Huitema E, Testa A, Gow NA, van West P, Kamoun S: **EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*.** *Genome Res* 2003, **13**:1675-1685.
57. Win J, Bos J, Krasileva K, Cano L, Chapparo-Garcia A, Ammar R, Staskawicz BJ, Kamoun S: **Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes.** *The Plant Cell* 2007, **19**:2349-2369.
58. Chisholm ST, Coaker G, Day B, Staskawicz BJ: **Host-microbe interactions: shaping the evolution of the plant immune response.** *Cell* 2006, **124**:803-814.
59. Panstruga R, Dodds PN: **Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens.** *Science* 2009, **324**:748-750.
60. Dou D, Kale SD, Wang X, Jiang RHY, Bruce NA, Arredondo FD, Zhang X, Tyler BM: **RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery.** *Plant Cell* 2008, **20**:1930-1947.
61. 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: **A translocation signal for delivery of oomycete effector proteins into host plant cells.** *Nature* 2007, **450**:115-118.
62. Bangyeekhun E, Cerenius L, Soderhall K: **Molecular cloning and characterization of two serine proteinase genes from the crayfish plague fungus, *Aphanomyces astaci*.** *J Invertebr Pathol* 2001, **77**:206-216.
63. Dou D, Kale SD, Wang X, Chen Y, Wang Q, Wang X, Jiang RHY, Arredondo FD, Anderson R, Thakur P, McDowell J, Wang Y, Tyler BM: **Carboxy-terminal motifs common to many oomycete RXLR effectors are required for avirulence and suppression of BAX-mediated programmed cell death by *Phytophthora sojae* effector Avr1b.** *Plant Cell* 2008, **20**:1118-1133.
64. Gaulin E, Madoui MA, Bottin A, Jacquet C, Mathe C, Couloux A, Wincker P, Dumas B: **Transcriptome of *Aphanomyces euteiches*: New oomycete putative pathogenicity factors and metabolic pathways.** *PLoS ONE* 2008, **3**:e1723.
65. Hein I, Gilroy EM, Armstrong MR, Birch PR: **The zig-zag-zig in oomycete-plant interactions.** *Mol Plant Pathol* 2009, **10**:547-562.
66. Brunner F, Rosahl S, Lee J, Rudd JJ, Geiler C, Kauppinen S, Rasmussen G, Scheel D, Nurnberger T: **Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases.** *EMBO J* 2002, **21**:6681-6688.
67. Gaulin E, Drame N, Lafitte C, Torto-Alalibo T, Martinez Y, Ameline-Torregrosa C, Khatib M, Mazarguil H, Villalba-Mateos F, Kamoun S, Mazars C, Dumas B, Bottin A, Esquerre-Tugaye MT, Rickauer M: **Cellulose binding domains of a *Phytophthora* cell wall protein are novel pathogen-associated molecular patterns.** *Plant Cell* 2006, **18**:1766-1777.
68. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics.** *Nucleic Acids Res* 2009, **37**:D233-D238.
69. Baker CJ, Bateman DF: **Cutin degradation by plant pathogenic fungi.** *Phytopathology* 1978, **68**:1577-1584.
70. Campion C, Massiot P, Rouxel F: **Aggressiveness and production of cell-wall degrading enzymes by *Pythium violae, Pythium sulcatum* and *Pythium ultimum*, responsible for cavity spot on carrots.** *Eur J Plant Pathol* 1997, **103**:725-735.
71. Belbahri L, Calmin G, Mauch F, Andersson JO: **Evolution of the cutinase gene family: evidence for lateral gene transfer of a candidate *Phytophthora* virulence factor.** *Gene* 2008, **408**:1-8.
72. Muñoz CI, Bailey AM: **A cutinase-encoding gene from *Phytophthora capsici* isolated by differential-display RT-PCR.** *Curr Genet* 1998, **33**:225-230.
73. Jiang RH, Tyler BM, Govers F: **Comparative analysis of *Phytophthora* genes encoding secreted proteins reveals conserved synteny and lineage-specific gene duplications and deletions.** *Mol Plant Microbe Interact* 2006, **19**:1311-1321.
74. Chérif M, Benhamou N, Bélanger RR: **Ultrastructural and cytochemical studies of fungal development and host reactions in cucumber plants infected by *Pythium ultimum*.** *Physiol Mol Plant Pathol* 1991, **39**:353-375.

75. Torto TA, Rauser L, Kamoun S: **The pipg1 gene of the oomycete *Phytophthora infestans* encodes a fungal-like endopolygalacturonase.** *Curr Genet* 2002, **40**:385-390.
76. Wu CH, Yan HZ, Liu LF, Liou RF: **Functional characterization of a gene family encoding polygalacturonases in *Phytophthora parasitica*.** *Mol Plant Microbe Interact* 2008, **21**:480-489.
77. Yan HZ, Liou RF: **Cloning and analysis of pppg1, an inducible endopolygalacturonase gene from the oomycete plant pathogen *Phytophthora parasitica*.** *Fungal Genet Biol* 2005, **42**:339-350.
78. Götesson A, Marshall JS, Jones DA, Hardham AR: **Characterization and evolutionary analysis of a large polygalacturonase gene family in the oomycete plant pathogen *Phytophthora cinnamomi*.** *Mol Plant Microbe Interact* 2002, **15**:907-921.
79. Moore LD, Couch HB: **Influence of calcium nutrition on pectolytic and cellulolytic enzyme activity of extracts of highland bentgrass foliage blighted by *Pythium ultimum*.** *Phytopathology* 1968, **58**:833-838.
80. Bartnicki-Garcia S: **Cell wall chemistry, morphogenesis, and taxonomy of fungi.** *Annu Rev Microbiol* 1968, **22**:87-108.
81. Blaschek W, Käsbauer J, Kraus J, Franz G: ***Pythium aphanidermatum*: culture, cell-wall composition, and isolation and structure of antitumour storage and solubilised cell-wall (1->3), (1- >6)-beta-D-glucans.** *Carbohydr Res* 1992, **231**:293-307.
82. Grenville-Briggs LJ, Anderson VL, Fugelstad J, Avrova AO, Bouzenzana J, Williams A, Wawra S, Whisson SC, Birch PR, Bulone V, van West P: **Cellulose synthesis in *Phytophthora infestans* is required for normal appressorium formation and successful infection of potato.** *Plant Cell* 2008, **20**:720-738.
83. Zevenhuizen LP, Bartnicki-Garcia S: **Structure and role of a soluble cytoplasmic glucan from *Phytophthora cinnamomi*.** *J Gen Microbiol* 1970, **61**:183-188.
84. Wang MC, Bartnicki-Garcia S: **Distrubution of mycolaminarans and cell wall β-glucans in the life cycle of *Phytophthora*.** *Exp Mycol* 1980, **4**:269-280.
85. Cohen Y, Coffey MD: **Systemic fungicides and the control of oomycetes.** *Annu Rev Phytopathol* 1986, **24**:311-338.
86. Davidse L: **Phenylamide fungicides: Mechanism of action and resistance.** *Fungicide Resistance in North America* St Paul, Minnesota: American Phytopathological SocietyDelp CJ 1988, 63-65.
87. Bauer BE, Wolfger H, Kuchler K: **Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance.** *Biochim Biophys Acta* 1999, **1461**:217-236.
88. Cannon RD, Lamping E, Holmes AR, Niimi K, Baret PV, Keniya MV, Tanabe K, Niimi M, Goffeau A, Monk BC: **Efflux-mediated antifungal drug resistance.** *Clin Microbiol Rev* 2009, **22**:291-321.
89. Connolly MS, Sakihama Y, Phuntumart V, Jiang Y, Warren F, Mourant L, Morris PF: **Heterologous expression of a pleiotropic drug resistance transporter from *Phytophthora sojae* in yeast transporter mutants.** *Curr Genet* 2005, **48**:356-365.
90. Katzmann DJ, Hallstrom TC, Voet M, Wysock W, Golin J, Volckaert G, Moye-Rowley WS: **Expression of an ATP-binding cassette transporter-encoding gene (YOR1) is required for oligomycin resistance in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1995, **15**:6875-6883.
91. Kolaczkowski M, Kolaczowska A, Luczynski J, Witek S, Goffeau A: ***In vivo* characterization of the drug resistance profile of the major ABC transporters and other components of the yeast pleiotropic drug resistance network.** *Microb Drug Resist* 1998, **4**:143-158.
92. Judelson H, Senthil G: **Investigating the role of ABC transporters in multifungicide insensitivity in *Phytophthora infestans*.** *Mol Plant Pathol* 2006, **7**:17-29.
93. Goldberg AL: **Protein degradation and protection against misfolded or damaged proteins.** *Nature* 2003, **426**:895-899.
94. Staszczak M: **The role of the ubiquitin-proteasome system in the response of the ligninolytic fungus *Trametes versicolor* to nitrogen deprivation.** *Fungal Genet Biol* 2008, **45**:328-337.
95. Dreschler C: **Zoospore development from oospores of *Pythium ultimum* and *Pythium debaryanum* and its relation to rootlet-tip discoloration.** *Plant Disease Reporter* 1946, **30**:226-227.
96. Hulpiau P, van Roy F: **Molecular evolution of the cadherin superfamily.** *Int J Biochem Cell Biol* 2009, **41**:349-369.
97. Cao L, Yan X, Borysenko CW, Blair HC, Wu C, Yu L: **CHDL: a cadherin-like domain in Proteobacteria and Cyanobacteria.** *FEMS Microbiol Lett* 2005, **251**:203-209.
98. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A: **The Pfam protein families database.** *Nucleic Acids Res* 2008, **36**:D281-288.
99. Nollet F, Kools P, van Roy F: **Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.** *J Mol Biol* 2000, **299**:551-572.
100. Saitou N, Nei M: **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4**:406-425.
101. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, **52**:696-704.
102. Katoh K, Toh H: **Recent developments in the MAFFT multiple sequence alignment program.** *Brief Bioinform* 2008, **9**:286-298.
103. **PHRINGE..** [http://genomeprojectsolutions.com/PHRINGE_pipeline.html].
104. Boudjeko T, Andème-Onzighi C, Vicré M, Balangé A, Ndoumou D, Driouich A: **Loss of pectin is an early event during infection of cocoyam roots by *Pythium myriotylum*.** *Planta* 2006, **223**:271-282.
105. Abad P, Gouzy J, Aury JM, Castagnone-Sereno P, Danchin EG, Deleury E, Perfus-Barbeoch L, Anthouard V, Artiguenave F, Blok VC, Caillaud MC, Coutinho PM, Dasilva C, De Luca F, Deau F, Esquibet M, Flutre T, Goldstone JV, Hamamouch N, Hewezi T, Jaillon O, Jubin C, Leonetti P, Magliano M, Maier TR, Markov GV, McVeigh P, Pesole G, Poulain J, Robinson-Rechavi M, *et al*: **Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*.** *Nat Biotechnol* 2008, **26**:909-915.
106. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, *et al*: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**:376-380.
107. Goldberg SM, Johnson J, Busam D, Feldblyum T, Ferriera S, Friedman R, Halpern A, Khouri H, Kravitz SA, Lauro FM, Li K, Rogers YH, Strausberg R, Sutton G, Tallon L, Thomas T, Venter E, Frazier M, Venter JC: **A Sanger/ pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes.** *Proc Natl Acad Sci USA* 2006, **103**:11240-11245.
108. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC: **A whole-genome assembly of *Drosophila*.** *Science* 2000, **287**:2196-2204.
109. Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G: **Aggressive assembly of pyrosequencing reads with mates.** *Bioinformatics* 2008, **24**:2818-2824.
110. Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sanchez Alvarado A, Yandell M: **MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes.** *Genome Res* 2008, **18**:188-196.
111. Korf I: **Gene finding in novel genomes.** *BMC Bioinformatics* 2004, **5**:59.
112. Salamov AA, Solovyev VV: ***Ab initio* gene finding in *Drosophila* genomic DNA.** *Genome Res* 2000, **10**:516-522.
113. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M: **Gene identification in novel eukaryotic genomes by self-training algorithm.** *Nucleic Acids Res* 2005, **33**:6494-6506.
114. Boguski MS, Lowe TMJ, Tolstoshev CM: **dbEST - database for "expressed sequence tags".** *Nat Genet* 1993, **4**:332-333.
115. Bairoch A, Apweiler R: **The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000.** *Nucleic Acids Res* 2000, **28**:45-48.
116. UniProt Consortium: **The Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2007, **35**:D193-197.
117. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J: **Repbase Update, a database of eukaryotic repetitive elements.** *Cytogenet Genome Res* 2005, **110**:462-467.
118. Smith CD, Edgar RC, Yandell MD, Smith DR, Celniker SE, Myers EW, Karpen GH: **Improved repeat identification and masking in Dipterans.** *Gene* 2007, **389**:1-9.
119. Edgar RC, Myers EW: **PILER: identification and classification of genomic repeats.** *Bioinformatics* 2005, **21 Suppl 1**:i152-i158.
120. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MDR, Durbin R, Falquet L,

Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJA, Zdobnov EM: **The InterPro database, an integrated documentation resource for protein families, domains and functional sites.** *Nucleic Acids Res* 2001, **29**:37-40.

121. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R: **InterProScan: protein domains identifier.** *Nucleic Acids Res* 2005, **33**: W116-120.
122. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley R, Courcelle E, Durbin R, Falquet L, Fleischmann W, Gouzy J, Griffith-Jones S, Haft D, Hermjakob H, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Orchard S, Pagni M, Peyruc D, Ponting CP, *et al*: **InterPro: An integrated documentation resource for protein families, domains and functional sites.** *Brief Bioinform* 2002, **3**:225-235.
123. Lewis SE, Searle SM, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglir L, Birney E, Crosby MA, Kaminker JS, Matthews BB, Prochnik SE, Smithy CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME: **Apollo: a sequence annotation editor.** *Genome Biol* 2002, **3**:RESEARCH0082.
124. Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ: **Basic Local Alignment Search Tool.** *J Mol Biol* 1990, **215**:403-410.
125. Rice P, Longden I, Bleasby A: **EMBOSS: the European Molecular Biology Open Software Suite.** *Trends Genet* 2000, **16**:276-277.
126. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nat Protoc* 2007, **2**:953-971.
127. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**:567-580.
128. **HMMER..** [http://hmmer.org/].
129. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**:R25.
130. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5**:621-628.
131. Enright AJ, Van Dongen S, Ouzounis CA: **An efficient algorithm for large-scale detection of protein families.** *Nucleic Acids Res* 2002, **30**:1575-1584.
132. Bailey TL, Elkan C: **Fitting a mixture model by expectation maximization to discover motifs in biopolymers.** *Proc Int Conf Intell Syst Mol Biol* 1994, **2**:28-36.
133. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**:1792-1797.
134. Wernersson R, Pedersen AG: **RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences.** *Nucleic Acids Res* 2003, **31**:3537-3539.
135. Stamatakis A: **RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics* 2006, **22**:2688-2690.
136. **WebLogo3..** [http://weblogo.berkeley.edu].
137. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A: **PANTHER: a library of protein families and subfamilies indexed by function.** *Genome Res* 2003, **13**:2129-2141.
138. **PANTHER Database..** [http://www.pantherdb.org].
139. **PANTHER Classification System..** [http://www.pantherdb.org/tools/compareToRefListForm.jsp].
140. **PANTHER ftp site..** [ftp://ftp.pantherdb.org/genome/pthr7.0].
141. Thomas PD: **GIGA: a simple, efficient algorithm for gene tree inference in the genomic age.** *BMC Bioinformatics* **11**:312.
142. Gams KW, Hoekstra ES, Aptroot A: **CBS Course of Mycology.** Baarn, The Netherlands: Centraalbureau voor Schimmelcultures 1998.
143. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation.** *Bioinformatics* 2007, **23**:127-128.
144. Campion C, Vian B, Nicole M, Rouxel F: **A comparative study of carrot root tissue colonization and cell wall degradation by *Pythium violae* and *Pythium ultimum*, two pathogens responsible for cavity spot.** *Can J Microbiol* 1998, **44**:221-230.
145. Buckeridge MS, dos Santos HP, Tiné MA: **Mobilisation of storage cell wall polysaccharides in seeds.** *Plant Physiol Biochem* 2000, **38**:141-156.
146. Doblin MS, Vergara CE, Read S, Newbigin E, Bacic A: **Plant cell wall biosynthesis: making the bricks.** *The Plant Cell* Ithaca, New York: Blackwell Publishing Ltd - CRC PressRose JKC 2003, **8**:183-222.
147. Denisov G, Walenz B, Halpern AL, Miller J, Axelrod N, Levy S, Sutton G: **Consensus generation and variant detection by Celera Assembler.** *Bioinformatics* 2008, **24**:1035-1040.
148. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, *et al*: **The diploid genome sequence of an individual human.** *PLoS Biol* 2007, **5**:e254.
149. Brunner F, Wirtz W, Rose JK, Darvill AG, Govers F, Scheel D, Nürnberger T: **A beta-glucosidase/xylosidase from the phytopathogenic oomycete, *Phytophthora infestans*.** *Phytochemistry* 2002, **59**:689-696.
150. Chérif M, Benhamou N, Bélanger RR: **Occurence of cellulose and chitin in the hyphal walls of *Pythium ultimum*: a comparative study with other plant pathogenic fungi.** *CanJMicrobiol* 1993, **39**:213-222.
151. Cosgrove DJ: **Loosening of plant cell walls by expansins.** *Nature* 2000, **407**:321-326.
152. Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, Porcel BM, Couloux A, Aury JM, Segurens B, Poulain J, Anthouard V, Grossetete S, Khalili H, Coppin E, Dequard-Chablat M, Picard M, Contamine V, Arnaise S, Bourdais A, Berteaux-Lecellier V, Gautheret D, de Vries RP, Battaglia E, Coutinho PM, Danchin EG, Henrissat B, Khoury RE, Sainsard-Chanet A, Boivin A, Pinan-Lucarre B, *et al*: **The genome sequence of the model ascomycete fungus *Podospora anserina*.** *Genome Biol* 2008, **9**:R77.
153. Coutinho PM, Anderson MR, Kolenova K, vanKuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T, Seiboth B, Battaglia E, Aguilar-Osorio G, de Jong JF, Ohm RA, Aguilar M, Henrissat B, Nielsen J, Stålbrand H, de Vries RP: **Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*.** *Fungal Genet Biol* 2009, **46 Suppl 1**:S161-S169.
154. Agnihotri VP, Vaartaja O: **The influence of nitrogenous compounds on growth of 711-727 species.** *Can J Microbiol* 1967, **13**:1509-1519.
155. Gaulin E, Jauneau A, Villalba F, Rickauer M, Esquerré-Tugayé MT, Bottin A: **The CBEL glycoprotein of *Phytophthora parasitica* var-*nicotianae* is involved in cell wall deposition and adhesion to cellulosic substrates.** *J Cell Sci* 2002, **115**:4565-4575.
156. Torto-Alalibo T, Tian M, Gajendran K, Waugh M, van West P, Kamoun S: **Expressed sequence tags from the oomycete fish pathogen *Saprolegnia parasitica* reveals putative virulence factors.** *BMC Microbiol* 2005, **5**:46.
157. Tian M, Benedetti B, Kamoun S: **A second Kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato.** *Plant Physiol* 2005, **138**:1785-1793.
158. Song J, Win J, Tian M, Schornack S, Kaschani F, Ilyas M, van der Hoorn RA, Kamoun S: **Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3.** *Proc Natl Acad Sci USA* 2009, **106**:1654-1659.
159. Tian M, Huitema E, Da Cunha L, Torto-Alalibo T, Kamoun S: **A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B.** *J Biol Chem* 2004, **279**:26370-26377.
160. Tian M, Win J, Song J, van der Hoorn R, van der Knaap E, Kamoun S: **A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease.** *Plant Physiol* 2007, **143**:364-377.
161. Bouzidi MF, Parlange F, Nicolas P, Mouzeyar S: **Expressed sequence tags from the oomycete *Plasmopara halstedii*, an obligate parasite of the sunflower.** *BMC Microbiol* 2007, **7**:110.
162. Meijer HJ, Govers F: **Genomewide analysis of phospholipid signaling genes in *Phytophthora* spp.: novelties and a missing link.** *Mol Plant Microbe Interact* 2006, **19**:1337-1347.
163. Werck-Reichhart D, Feyereisen R: **Cytochromes P450: a success story.** *Genome Biol* 2000, **1, REVIEWS3003.**
164. Iordachescu M, Imai R: **Trehalose biosynthesis in response to abiotic stresses.** *J Integr Plant Biol* 2008, **50**:1223-1229.
165. Vogel G, Aeschbacher RA, Muller J, Boller T, Wiemken A: **Trehalose-6-phosphate phosphatases from *Arabidopsis thaliana*: identification by**

functional complementation of the yeast tps2 mutant. *Plant J* 1998, **13**:673-683.

166. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D: Genomic footprints of a cryptic plastid endosymbiosis in diatoms.. *Science* 2009, **324**:1724-1726.

doi:10.1186/gb-2010-11-7-r73
Cite this article as: Lévesque *et al.*: Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology* 2010 **11**:R73.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central