Minireview

Runaway repeats force expansion of the *Phytophthora infestans* genome

Mark Gijzen

Address: Agriculture and Agri-Food Canada, 1391 Sandford Street, London, Ontario N5V 4T3, Canada. Email: mark.gijzen@agr.gc.ca

Abstract
Sequencing of the genome of the potato late blight pathogen *Phytophthora infestans* provides insight into genome structure and evolution within this genus of plant pathogenic oomycetes.

Parasitic organisms that rely on hosts to complete their life cycles face unstable and rapidly changing selective pressures, as their own reproduction affects the health and success of their host. Late blight disease of potatoes and tomatoes is caused by the plant pathogen *Phytophthora infestans*. The recent analysis of the *P. infestans* genome sequence by Haas *et al.* [1] has provided new answers as to how this pathogen adapts to its host, and why its 240 Mb genome is unusually large among the species within its assemblage.

The completion of the *P. infestans* genome sequence represents a milestone because of the importance of this organism to history and the study of plant disease. As every student of plant pathology knows, *P. infestans* caused the Irish potato famine of the 1840s. The rapid spread of the disease after it was accidentally introduced to Europe from the Americas resulted in utter destruction of the potato crop. It remains the most spectacular example of a catastrophic plant disease epidemic. Anton de Bary described the life cycle of *P. infestans* some 140 years ago, making it among the first plant pathogenic microorganisms to be well characterized. But this is not just a history lesson, as *P. infestans* and other oomycetes continue to pose problems as plant pathogens and invasive organisms around the globe.

Oomycetes are somewhat obscure, fungus-like organisms that capture public attention when they destroy crops and gardens or invade new landscapes. Oomycetes are filamentous microorganisms that resemble fungi in their morphology and lifestyle. Many oomycetes are parasites of plants or animals and cause disease, whereas others are saprophytic and free-living. Oomycetes and fungi are similar in form but have evolved independently, thus illustrating the concept of convergent evolution. These filamentous organisms share so many similarities in their appearance and mode of life that it seems natural to assume a close ancestry and to classify them together, as all textbooks did until recently. But it was also clear that oomycetes and fungi have fundamental differences, such as in the make-up of their cell walls, the flagellation of their spores and their methods of sexual reproduction.

DNA-sequence-based phylogenies have drastically reorganized the classification of eukaryotes and deeply separated the fungal and oomycete lineages [2]. True fungi are unikonts (any flagellated cells have only one flagellum) and are related to animals, whereas oomycetes are heterokonts (flagellated cells generally have two flagella, of different types) and belong to the chromalveolates. From a phylogenetic perspective the two groups of organisms are about as different as they could be (Figure 1). Despite this early bifurcation, which occurred at the initial stages of eukaryote radiation hundreds of millions of years ago, the fungi and the oomycetes have completed an evolutionary journey that has taken them to the same destination. This is especially true for the plant pathogenic types, which have grown to resemble each other in appearance and biology to a remarkable degree. Their filamentous and osmotrophic growth mode, with dissemination by spores, and their sub-specialization into necrotrophic, biotrophic and hemi-biotrophic pathogenic types, are the basic common elements that define plant pathogenic fungi and oomycetes.

The core oomycete genome versus the plastic genome

Plant pathogenic oomycetes cause diseases on all types of plants, in all sorts of environments, around the world. They include soil-, water-, and air-borne pathogens that attack roots or aerial plant parts. Typically, spores attach to host plants and then grow into host tissues where they proliferate and produce more spores. Within the oomycetes, there are some 80 members of the genus *Phytophthora* that are particularly destructive. Whole-genome sequencing of oomycete species began with *Phytophthora sojae* (soybean root rot pathogen; 95 Mb genome) and *Phytophthora ramorum* (sudden oak death pathogen; 65 Mb genome) because they have smaller and simpler genomes than the 240 Mb *P. infestans* [3,4]. In fact, the *P. infestans* genome is an outlier among oomycetes, being about three-fold larger than the genomes of most
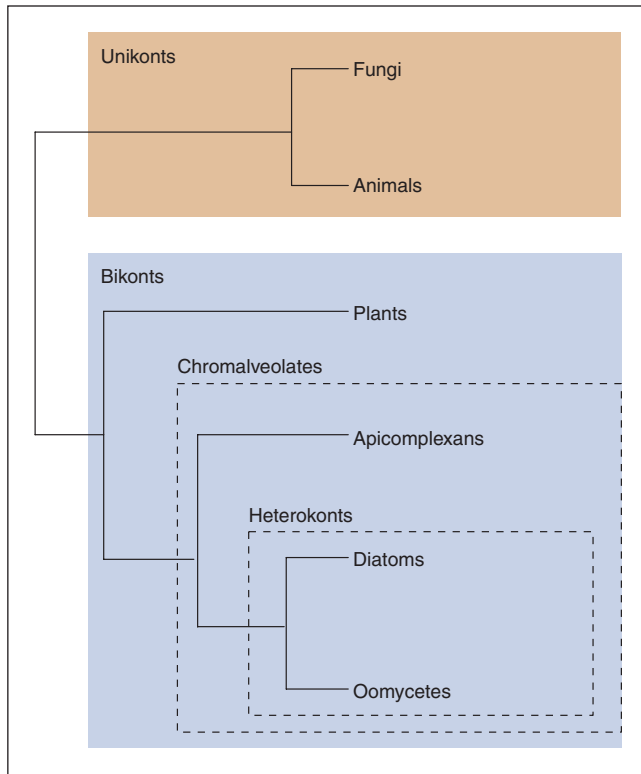
Figure 1

A current model of eukaryotic phylogeny showing selected groups of organisms. *Phytophthora infestans* is an oomycete.

Figure 2

Genome sizes of three different *Phytophthora* species, showing the content of non-repetitive and repetitive DNA.

other species in this class. The completion of the *P. infestans* genome has now shown how repeat-driven expansion has enlarged its size and differentially affected its structure (Figure 2). Three-quarters of the *P. infestans* genome is repetitive DNA and much of this appears to have recently arisen through an explosive spread of transposons [1]. Nonetheless, there remains a high degree of conserved synteny among all three genomes, and a core set of some 9,500 orthologous genes is present in each of the species [1]. The total number of genes annotated for each genome is similar, ranging from 14,451 (*P. ramorum*) to 16,988 (*P. sojae*) and 17,797 (*P. infestans*).

It is clear that the expansion of the *P. infestans* genome is not due to a wholesale duplication of the genome or large segments of it; rather, it is a consequence of the proliferation of transposable elements. The expansion of the genome is also not uniform. Overall, the genome shows a bimodal pattern whereby regions of low repeat content and high gene density with conserved synteny (the 'core genome') are interrupted by repeat-rich regions of low gene density where synteny has broken down (the 'plastic genome'). This pattern has been noted before, in the *P. ramorum* and *P. sojae* genomes, but it is so exaggerated
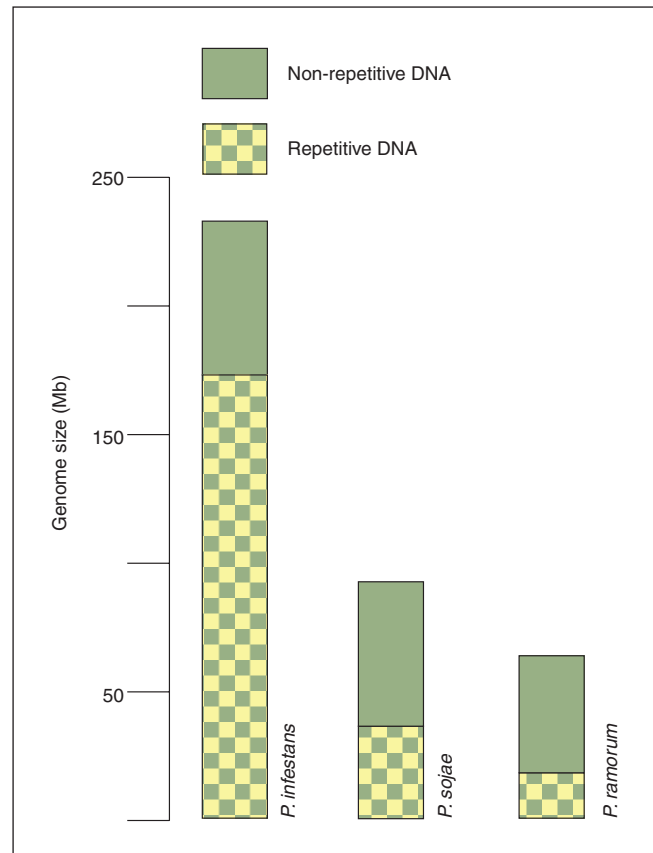
in *P. infestans* that the genome has become a caricature of features for its genus [1,3].

Accelerated change pervades the effector repertoire

A closer look at the contrasting segments that make up the core and plastic genomes of *P. infestans* illustrates how changes in genomic structure are shaped by the sharpest biological challenges and selective pressures that an organism faces. Genes encoding disease effector proteins are at the forefront of evolution for pathogens, just as genes controlling immunity are among the most dynamic in their hosts. Pathogen effectors comprise a diverse collection of secreted proteins that are delivered to extra- and intracellular host targets in order to foster susceptibility and disease. Effectors are crucial to pathogens because they enable growth in the host organism. For species of *Phytophthora*, numerous types of effectors have been described or proposed. Examples of extracellular or apoplastic effectors include enzymes that digest host cells and mobilize nutrients [5], inhibitors of host enzymes that attack pathogen components [6], and toxins that cause

necrosis of host tissues [7]. Recent studies have shown how the *P. sojae* and *P. ramorum* genomes have an expanded collection of polysaccharide-degrading enzymes in comparison to other non-plant-pathogenic chromalveolates [8]. The phytotoxic necrosis and ethylene inducing-like protein (NLP) family is also inordinately large and among the chromalveolates is only present in plant pathogenic oomycete species [1]. Current theories suggest that NLPs are cytolytic toxins and virulence factors that contribute to pathogen growth in the host [8].

Secreted pathogen effectors that enter host cells are also known. The RXLR effectors, for example, named after the Arg-X-Leu-Arg pattern (where X denotes any amino acid) that occurs downstream of the signal peptide, are know to be intracellular effectors. Conserved amino acid sequence motifs, or higher-order protein structural signals, somehow direct the uptake or trafficking of these effectors across host plasma membranes [9,10].

It is not known what functions RXLR effectors perform but there is evidence from studies on *P. infestans* and *P. sojae* that they can suppress host defenses and accelerate infection [11,12]. In fact, host immune systems have honed in on RXLR effectors for surveillance, as cues to activate rapid and robust defenses that are self-destructive but that limit pathogen growth and halt disease through a process known as effector-triggered immunity. So far, several different RXLR effectors that activate effector-triggered immunity are known from three species of oomycete plant pathogens - *P. infestans*, *P. sojae*, and *Hyaloperonospora arabidopsidis* (downy mildew pathogen of *Arabidopsis*) [13-20]. Any pathogen molecule that triggers immunity in the host is a dead-end liability and subject to immense selective pressures. Thus, pathogen effectors that are recognized by host immune systems evolve rapidly and acquire gain-of-virulence mutations in a type of molecular arms race.

It is also intuitive that pathogens will build effector arsenals that are diverse and redundant so that particular ones could be dispensable or re-adaptable should they come under host surveillance. Analysis of the *P. infestans* genome indicates this to be the case. There are 563 predicted RXLR effectors but only 16 of these are part of the core genome with identifiable orthologs in both *P. sojae* and *P. ramorum* [1]. The RXLR effectors constitute a huge and fast-evolving class of proteins in all three *Phytophthora* species sequenced to date. Previous studies on *P. sojae* and *P. ramorum* have suggested that the RXLR effector class arose from a single ancestor [21] but Haas *et al.* [1] favor the view that there are distinct families of effectors with separate origins.

It is debatable whether the RXLR effectors are monophyletic, but this is certainly not the case for another class of proposed intracellular effectors, called the crinklers (CRNs). The CRNs are secreted proteins that were identified on the basis of their ability to cause cell death and necrosis when expressed in plant cells. The 196 predicted CRN genes of *P. infestans* can be aligned into distinct families that arose through recombination and domain swapping [1]. Like the RXLR effectors, the CRNs appear to rely on separate conserved signaling domains within the amino-terminal region of the protein for secretion and delivery into host cells, although this needs to be proved by further experimentation. The CRN genes also populate the repeat-rich plastic genome of *P. infestans*, which has diverged and expanded in comparison to those of *P. sojae* and *P. ramorum*.

The proliferation of transposons and the spread of repetitive sequences in *P. infestans* have apparently aided sequence-exchange mechanisms that rely on homologous recombination. These processes can also cause the expansion of gene families through tandem duplication and amplification of sequence segments. Clusters of CRN genes in the *P. infestans* genome may have arisen this way. Likewise, analysis of the *P. sojae* genome has shown how multiple copies of identical or near-identical RXLR effector genes occur in tandem arrays that display copy-number variation among pathogen strains [20], or that show evidence of ongoing sequence exchanges within an array [19].

The post-genomic era for *Phytophthora*

Genetic research on oomycetes has always been challenging because it is difficult to perform sexual crosses and follow segregation. Moreover, for a long time these diploid organisms were misclassified and misunderstood. The application of molecular and DNA sequencing analyses has altered our view of oomycetes and provided new insight into the evolution of these organisms and how they have excelled as plant pathogens [22]. Parasites often follow a path of retrograde evolution, where structures and functions are lost over time as they become dispensable or redundant with functions provided by their hosts. Plant pathogenic oomycetes display many such trends, having lost photosynthetic capacity and associated functions and developmental programs. Recent studies using the *P. sojae* and *P. ramorum* sequences reveal a large number of gene fusions compared with other eukaryotes and suggest that the core genome has been under pressure for consolidation [23]. This is in contrast to the expansionary forces that dominate genome regions harboring disease effector genes.

The *P. infestans* genome appears to be exceptionally fluid, perhaps because we have caught it in a period of rapid change. To maintain the size and structure of a genome, the spread of repetitive elements, such as transposons, must be kept in check. For *P. infestans*, the unleashing of transposons has apparently provided the genetic means for recombination and adaptation that has enabled its recent evolution. Nevertheless, the highly conserved core genome

appears somehow protected or insulated in this hyper-changing environment, although this insulation may simply be imposed through the forces of natural selection. Like any good scientific study, the analysis of the *P. infestans* genome raises as many questions as it answers. How did transposons come to infest the genome of *P. infestans* to such a degree? What are the costs and consequences of carrying such a load of repetitive DNA? What are the probable future trajectories for such a genome? How exactly do the RXLR and CRN effectors operate? What other effector types lurk in the vast pool of predicted genes with no functional annotation? Where do we look in the *P. infestans* genome to find the key to controlling this pathogen and others like it? These questions and countless others are on the agenda as we enter the post-genomic era for *Phytophthora* research.

Acknowledgements

References

1. 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.

2. Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J: **Phylogenomics reshuffles the eukaryotic supergroups.** *PLoS ONE* 2007, **2:**e790.

3. 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.

4. Govers F, Gijzen M: ***Phytophthora* genomics: The plant destroyers' genome decoded.** *Mol Plant Microbe Interact* 2006, **19:**1295-1301.

5. 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.

6. Damasceno CM, Bishop JG, Ripoll DR, Win J, Kamoun S, Rose JK: **Structure of the glucanase inhibitor protein (GIP) family from *Phytophthora* species suggests coevolution with plant endo-beta-1,3-glucanases.** *Mol Plant Microbe Interact* 2008, **21:**820-830.

7. Ottmann C, Luberacki B, Küfner I, Koch W, Brunner F, Weyand M, Mattinen L, Pirhonen M, Anderluh G, Seitz HU, Nürnberger T, Oecking C: **A common toxin fold mediates microbial attack and plant defense.** *Proc Natl Acad Sci USA* 2009, **106:**10359-10364.

8. Martens C, Vandepoele K, Van de Peer Y: **Whole-genome analysis reveals molecular innovations and evolutionary transitions in chromalveolate species.** *Proc Natl Acad Sci USA* 2008, **105:**3427-3432.

9. 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.

10. Dou DL, Kale SD, Wang X, Jiang RHY, Bruce NA, Arredondo FD, Zhang XM, 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.

11. Dou D, Kale SD, Wang X, Chen Y, Wang Q, Wang X, Jiang RH, Arredondo FD, Anderson RG, Thakur PB, McDowell JM, Wang Y, Tyler BM: **Conserved C-terminal motifs required for avirulence and suppression of cell death by *Phytophthora sojae* effector Avr1b.** *Plant Cell* 2008, **20:**1118-1133.

12. Bos JIB, Kanneganti TD, Young C, Cakir C, Huitema E, Win J, Armstrong MR, Birch PRJ, Kamoun S: **The C-terminal half of *Phytophthora infestans* RXLR effector *AVR3a* is sufficient to trigger *R3a*-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*.** *Plant J* 2006, **48:**165-176.

13. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, Beynon, JL: **Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew.** *Science* 2004, **306:**1957-1960.

14. Shan WX, Cao M, Dan LU, Tyler BM: **The *Avr1b* locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps1b*.** *Mol Plant Microbe Interact* 2004, **17:**394-403.

15. van Poppel P, Guo J, de Vondervoort P, Jung MWM, Birch PRJ, Whisson SC, Govers F: **The *Phytophthora infestans* avirulence gene *Avr4* encodes an RXLR- dEER effector.** *Mol Plant Microbe Interact* 2008, **21:**1460-1470.

16. Armstrong MR, Whisson SC, Pritchard L, Bos JI, Venter E, Avrova AO, Rehmany AP, Böhme U, Brooks K, Cherevach I, Hamlin N, White B, Fraser A, Lord A, Quail MA, Churcher C, Hall N, Berriman M, Huang S, Kamoun S, Beynon JL, Birch PR: **An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm.** *Proc Natl Acad Sci USA* 2005, **102:**7766-7771.

17. Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, Whisson SC, Kamoun S, Tyler BM, Birch PR, Beynon JL: **Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines.** *Plant Cell* 2005, **17:**1839-1850.

18. Vleeshouwers VG, Rietman H, Krenek P, Champouret N, Young C, Oh SK, Wang M, Bouwmeester K, Vosman B, Visser RG, Jacobsen E, Govers F, Kamoun S, Van der Vossen EA: **Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes.** *PLoS ONE* 2008, **3:**e2875.

19. Dong S, Qutob D, Tedman-Jones J, Kuflu K, Wang Y, Tyler BM, Gijzen M: **The *Phytophthora sojae* avirulence locus *Avr3c* encodes a multi-copy RXLR effector with sequence polymorphisms among pathogen strains.** *PLoS ONE* 2009, **4:** e5556.

20. Qutob D, Tedman-Jones J, Dong S, Kuflu K, Pham H, Wang Y, Dou D, Kale SD, Arredondo FD, Tyler BM, Gijzen M: **Copy number variation and transcriptional polymorphisms of *Phytophthora sojae* RXLR effector genes *Avr1a* and *Avr3a*.** *PLoS ONE* 2009, **4:**e5066.

21. Jiang RHY, 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.

22. Lamour K, Kamoun S (Eds): *Oomycete Genetics and Genomics. Diversity, Interactions, and Research Tools.* Hoboken, NJ: Wiley-Blackwell; 2009.

23. Morris PF, Schlosser LR, Onasch KD, Wittenschlaeger T, Austin R, Provart N: **Multiple horizontal gene transfer events and domain fusions have created novel regulatory and metabolic networks in the oomycete genome.** *PLoS ONE* 2009, **4:**e6133.