USING GENOMIC DATA TO UNDERSTAND ANTHROPOGENIC INFLUENCES ON OOMYCETE AND *PHYTOPHTHORA* COMMUNITIES, AND THE EVOLUTION OF AN ALIEN INVASIVE SPECIES RESPONSIBLE FOR SUDDEN OAK DEATH, *PHYTOPHTHORA RAMORUM*.

by

ANGELA DALE

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Forestry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2018

Abstract

Emerging *Phytophthora* pathogens, often introduced, represent a threat to natural ecosystems. *Phytophthora* species are known for rapid adaptation and hybridization, which may be facilitated by anthropogenic activities. Little is known about natural *Phytophthora* and oomycete populations, or mechanisms behind rapid adaptation. We surveyed oomycete and *Phytophthora* communities from southwest B.C. under varying anthropogenic influences (urban, interface, natural) to determine effects on diversity, introductions and migration. We used DNA meta-barcoding to address these questions on oomycetes. We then focused on *Phytophthora*, adding baiting and culturing methods, and further sub-dividing urban sites into agricultural or residential. Finally, we studied an alien invasive species, *Phytophthora ramorum* responsible for sudden oak death, and how it overcame the invasion paradox, limited to asexual reproduction and presumed reduced adaptability. Anthropogenic activities increase oomycete and *Phytophthora* diversity. Putative introduced species and hybrids were more frequent in urban sites. Migration is suggested by shared species between urban and interface sites, and two known invasive species found in natural and interface sites. Different anthropogenic activities influence different communities. Abundance increased for some species in either residential or agricultural sites. Two hybrids appear to be spreading in different agricultural sites. In the invasive *Phytophthora ramorum*, mitotic recombination drives diversification of the four lineages (NA1, NA2, EU1 and EU2), generating runs of homozygosity. One genome region, enriched in putative plant pathogenicity genes and transposons, was fixed in NA1 and present in eight EU1 individuals, but affecting the opposite alleles. Longer lesions during initial colonization in inoculated larch and Douglas fir logs suggested a fitness advantage in these EU1 individuals.

Mitotic recombination breakpoints were associated with transposons and low gene density. Non-core and lineage-specific genomic regions were enriched in putative plant pathogenicity genes and transposons. Gene loss was observed in the EU2 non-core genome affecting all effectors. A two-speed genome, where regions enriched in transposons and plant pathogenicity genes evolve faster, appears to drive non-core genome divergence, and mitotic recombination resulting in population evolution. This may explain invasion success and adaptability in *Phytophthora* pathogens. These results highlight the importance of anthropogenic activities in the emergence of forest diseases.

Lay Summary

International trade can facilitate the transfer and introduction of non-native plant and forest pathogens into new environments, resulting in forest disease outbreaks. Urban activities like gardening or agriculture may also spread plant pathogens. This study investigated the effects of these activities by comparing a group of organisms found in soil and water in urban and natural environments. Several species from this group are forest and crop pathogens. We determined that populations of these organisms in urban areas were more diverse and had more introduced species. Natural areas close to urban centers were also more diverse and could act as a bridge for migration of introduced species into natural areas. Since these organisms are highly adaptable, we investigated DNA characteristics that may facilitate adaptation in an introduced invasive pathogen. DNA changes have led to increased population diversity which may explain the success of this pathogen in new environments.

Preface

Chapters 2 and 3 of this dissertation are an original intellectual product of the author, Angela Dale, with some collaborative input as described below. Chapter 4 was conducted as an international collaboration using data generated for the Genome Canada funded Taiga project. I, Angela Dale, was the lead investigator on this paper, working under the guidance of the supervisory author Dr. Richard Hamelin. I was responsible for the research questions, hypotheses, and direction of the paper, with input from my supervisor and collaborators, especially Dr. Nicolas Feau, and Dr. Sydney Everhart. Original manuscript composition was by me, and the version presented here was reviewed and edited by me as well as by Dr. Nicolas Feau and my supervisor, Dr. Richard Hamelin. Contributions of collaborators and the author to specific analyses are as described below. Drafts of each of the sections completed by collaborators (methods, results and figures) were given to me and edited for inclusion into both the main document and the supplementary materials by me, Dr. Nicolas Feau, and Dr. Richard Hamelin. The supplementary materials (Appendix B) describe the details of the work performed, including the work performed by the author, Angela Dale. However, it was not included in the main body of the document due to the format required for the type of journal where we are planning to submit the work. All the work described in chapters 2, 3 and 4 was conducted by me, except where indicated in the following sections.

Chapter 2: All the work described was conducted by me, except for the following. The samples for DNA meta-barcoding were prepared (DNA extraction, PCR amplified, barcoded, checked, pooled) by me and sent to the laboratory of Dr. Jean Bérubé of the Canadian Forest Service

where they were cleaned and quantified prior to being sequenced by the McGill platform as described in the dissertation.

Chapter 3: The chi-square independence test and heat maps of UniFrac distance were done by Dr. Nicolas Feau.

Chapter 4: Genome sequencing: DNA extraction was done under the supervision of Dr. Guillaume Bilodeau at the CFIA. Genome sequencing was done at the Genome Sciences Centre under the supervision of Dr. Steve Jones.

Extraction of a SNP set: SNP extraction, filtering and testing filtering parameters was done by the author, Angela Dale with help/contributions from Dr. Nicolas Feau and Dr. Sydney Everhart.

De novo genome assemblies were done by Dr. Braham Dhillon and Dr. Nicolas Feau.

Phylogenetic analyses: Neighbor joining phylogenetic analysis was done by the author, Angela Dale. Maximum Likelihood phylogenetic reconstruction was done by Javier F. Tabima under the supervision of Dr. Niklaus Grünwald. Divergence time between lineages was done by Dr. Nicolas Feau.

Detection of runs of homozygosity (ROH) was done by the author, Angela Dale, with help in development of a python script by Dr. Nicolas Feau. Determining the effects on protein content

from conversion to homozygosity was done by the author Angela Dale, with help from Javier F. Tabima who reconstructed each DNA strand in regions with ROH.

Identification and characterization of a pathogenicity hotspot was done by the author, Angela Dale with advice from Dr. Nicolas Feau.

Effects of ROH on phenotype was done by the author, Angela Dale, with contributions from collaborators. Dr. Clive Brasier provided advice on the test design, and testing on Rhododendron leaves was done at the CFIA under the direction of Dr. Guillaume Bilodeau. Testing on Larch and Douglas-fir log sections was done by the author, Angela Dale, using the facilities at FPInnovations.

Genome-level copy number variation was done by Dr. Nicolas Feau.

Mitotic recombination breakpoints were identified by the author, Angela Dale. Surrounding gene content, gene density and comparison with random datasets was done by Angela Dale.

OrthoMCL analysis and evolution of gene family size was done by Dr. Nicolas Feau.

Core and non-core genome analysis was done by Dr. Nicolas Feau. An original version of this work was done by both the author, Angela Dale, and Dr. Nicolas Feau. At the request of the author, Dr. Nicolas Feau extracted the non-core genome regions which were then analyzed for gene content and compared between lineages by the author, Angela Dale. A final, more robust

analysis was done by Dr. Nicolas Feau to use for the final version to submit for publication. This also included work on lineage-specific proteins and evolution of gene family size done by Dr. Nicolas Feau.

Genes encoding effectors: RxLR effector extraction was done by Dr. Sydney Everhart. CRN or Crinkler effector extraction was done by Dr. Nicolas Feau. Extraction of a random protein set and the CEGMA gene set, protein filtering, conversion to DNA and analysis of positive selection was done by Dr. Nicolas Feau. Statistical comparisons between datasets was done by the author, Angela Dale.

This collaboration was led by Dr. Richard Hamelin of UBC. In addition to the author, Angela Dale, the collaborators at UBC included Dr. Nicolas Feau and Dr. Braham Dhillon. Members in addition to the UBC team: Dr. Niklaus Grünwald, Dr. Sydney Everhart and Javier F. Tabima of Oregon State University and the USDA Horticultural Crops Research Unit; Dr. Guillaume Bilodeau from the Canadian Food Inspection Agency; Dr. Steve Jones of the Genome Sciences Centre; Dr. Jean Bérubé of the Canadian Forest Service; Dr. Clive Brasier of The Forestry Commision, UK (Alice Holt Lodge).

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to offer my sincere thanks and appreciation to my supervisor, Dr. Richard Hamelin for his support, guidance, inspiration, and the opportunity to do this work. Also for all of the great scientific discussions; not only did they offer inspiration and helped to define my work, but they were moments of fun and laughter that created a great team environment that I will miss working in.

A special thanks to Dr. Nicolas Feau. It has been an amazing and inspirational journey. Thank you for your encouragement, guidance, advice, and for all the fun and laughter we shared over the past six years. Thank you for encouraging me to step outside of my comfort zone, to learn to code with Python, and to look beyond the surface of the data for the good stuff. Thank you for listening to my endless stream of thoughts and for help in turning my ideas into something concrete. And overall, thank you for your help in producing this work.

A special thanks to FPInnovations for the opportunity to do this work and for the support provided during the project. A special thanks to my colleagues for the inspiration, moral support, help and guidance over the years.

Thank you to my committee members, Dr. Kermit Ritland, Dr. Sue Grayston, and Dr. William Mohn for their guidance, inspiration and for their time.

A big thank you to all of the members of the UBC Hamelin lab, whose support was instrumental during the project, especially during the field season and in processing the hundreds of samples.

This project was done as part of the Taiga project, made possible by funding and collaboration from Genome Canada, Genome British Columbia, The University of British Columbia, University of Laval, FPInnovations, The Canadian Food Inspection Agency, Natural Resources Canada, Canadaôs Michael Smith Genome Sciences Centre, and Boreal Genomics. Many thanks to the contributors for the opportunity to do this work.

A special thank you to my family and to my parents. Your never-ending support and your faith in me were instrumental in me being here today. Thank you for helping me through the tough times and for celebrating during the good times.

I dedicate this thesis to my children, Matthew, Nathan and Noah, to my sister, Lorelei, my niece Jorden, and to my parents. You are my inspiration. You make me strong and you continue to bring happiness into my life. You have taught me that you can do anything you put your mind to.

Introduction: Phytophthora, the plant destroyer genus

The *Phytophthora* de Bary genus, first described in the late 1800s by Heinrich Anton de Bary , is considered one of the most destructive groups of plant pathogens in history. *Phytophthora infestans* (Mont.) de Bary caused the notorious Irish potato famine which resulted in more than (De Bary, 1876) one million deaths and the migration of more than two million people. *Phytophthora* species infect a wide range of hosts including economically valuable tree species, both angiosperm and gymnosperm, and many important agricultural species. Many known species of *Phytophthora* are invasive and have been introduced through nursery and agricultural trade.

In the United States, the recently emerged *Phytophthora ramorum* Werres, De Cock & Man in 't Veld has killed large numbers of oak trees along the West Coast and the range of the pathogen is increasing despite efforts to eradicate outbreaks (Kanaskie *et al.*, 2007). The primary range of the epidemic is over 750 km from southern California to southern Oregon (Meentemeyer *et al.*, 2008). Over 230 000 oak and tanoak were estimated to have been killed by 2005 in the Big Sur ecosystem of California, the heart of the epidemic (Meentemeyer *et al.*, 2008). *Phytophthora ramorum* is believed to have been spread via the nursery trade and is invasive in California, Oregon and the UK (Grünwald *et al.*, 2012). The origin of the pathogen remains unknown (Rizzo & Garbelotto, 2003; Grünwald *et al.*, 2012). In the UK the pathogen has also rapidly spread despite intensive control measures and has recently jumped hosts from the previously reported known angiosperm hosts to Japanese larch (*Larix kaempferi* (Lamb.) Carr.) (Brasier & Webber, 2010). This is the first time that the pathogen has been reported to cause mortality and

widespread damage on conifers (Brasier & Webber, 2010). Previously, some conifer species were shown to be susceptible hosts in laboratory studies (Chastagner *et al.*, 2004; Hansen *et al.*, 2005; Denman *et al.*, 2005). Infections were observed in heavily infected areas on some conifer species including Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), California red fir (*Abies magnifica* A. Murray bis), and coast redwood (*Sequoia sempervirens* (Lamb.) Endl.) (Davidson *et al.*, 2002; Maloney *et al.*, 2002; Brasier & Webber, 2010; Chastagner & Riley, 2010). These cases were associated with high inoculum loads, and infection appears to be limited to needles, shoots, and branches. The recent jump to a new conifer host and subsequent spread on a landscape scale raises concern in western North America both for forest health due to the relative abundance of susceptible conifers in forests, and for the forest industry that relies heavily on some of these species.

Other *Phytophthora* species attacking trees have also spread widely. *Phytophthora lateralis* Tucker & Milbrath, also an invasive pathogen, infects Port Orford Cedar (*Chamaecyparis lawsoniana* (A.Murray) Parl) and has spread throughout the natural range of the tree (Hansen, 2008). The epidemic has affected both the wood export market as this is a valued species in foreign markets as well as the nursery trade as it is also a valued horticultural species (Hansen, 2008). *Phytophthora cinnamomi* Rands has also been introduced and spread to many countries worldwide. *Phytophthora cinnamomi* has a very broad host range, from grasses and ferns to several different tree species. Many of the species affected by this pathogen are important horticultural and forest species and *P. cinnamomi* has been responsible for widespread epidemics in several forest species including Chestnut in the USA, cork oak in Europe, and in Australia the pathogen affects many species and is causing extensive damage to Jarrah forests and heathlands

(Hansen, 2008). The devastating effects of invasive *Phytophthora* species have resulted in trade restrictions being implemented in many countries in order to decrease the probability of further spread of these pathogens. Worldwide, millions of dollars have been spent on quarantine efforts, attempted control, and eradication programs.

In addition to economic costs, non-indigenous pathogens can also have wide ranging effects on ecosystems (Loo, 2009). The loss of a host species can change the dynamics of an ecosystem, especially when that species is a dominant one (Ellison *et al.*, 2005). For example there can be changes in microclimate, hydrology, and nutrient fluxes (Ellison *et al.*, 2005; Loo, 2009), and important food sources and habitat may be lost for some wildlife (Ellison *et al.*, 2005) . Population reduction of Port-Orford cedar in Oregon and California by the pathogen *P. lateralis* could result in significant ecological changes including changes in soil calcium content, loss of habitat structure, as well as loss of stream structure and stability along riparian zones (Hansen *et al.*, 2000; Loo, 2009). Extensive mortality of tanoak (*Lithocarpus densiflorus* (Hook. & Arn.) Rehd.) in Oregon and California forests due to *P. ramorum* increases fire risk in those forests, may affect animals dependent on the acorns as a food source, and could cause extended periods of chaparral vegetation in some areas (Hansen, 2008).

Introducing species to new environments and other anthropogenic activities such as agricultural activities may also have an effect on the evolutionary processes occurring in a community. For example bringing different species together could increase opportunities for hybridization between closely related species. Agricultural systems may also increase opportunities for these events as host plants are in high abundance and in close proximity to one another allowing for

increased contact between pathogens. *Phytophthora* x*alni* Brasier & S.A. Kirk 2004 is a hybrid species currently spreading throughout European ecosystems and responsible for alder decline. It has been suggested that it emerged after the introduction of *Phytophthora uniformis* Brasier & S.A. Kirk 2004 from North America followed by hybridization with *Phytophthora* x*multiformis* Brasier & S.A. Kirk 2004 (Ioos *et al.*, 2006; Aguayo *et al.*, 2016).

The international trade of plants and trees has increased the global movement of plant and forest pathogens. Understanding the effect of this movement and of other forestry and agricultural activities on pathogen, population, and community evolution is important in order to mitigate negative effects on natural ecosystems. Improving our understanding of how activities alter pathogen communities and how that leads to disease can help in the development of management strategies to mitigate the effects of those activities. For instance if nursery activities lead to evolution of new species and spread into natural systems, then measures need to be taken to reduce or prevent pathogen escape from these systems.

So far the interest in *Phytophthora*s has stemmed from the extensive damage that members of this genus have inflicted on crops and forests. However little is known about the natural roles and communities of *Phytophthora*, and the oomycetes in general. Likewise, we have a poor understanding of the influence of anthropogenic activities on natural communities of oomycetes. This review will first give a general overview of the oomycetes, followed by a review of *Phytophthora* communities and the niches they occupy. The second part of the review will look at *Phytophthora* biology, genomics and evolutionary processes, specifically the importance of

interspecific hybridization. The third part of the review will look at focus on an example of an introduced and emerging pathogen, *P. ramorum*.

Chapter 1: Literature Review

Oomycetes

The oomycetes are commonly known as watermolds. They are part of the Superphylum heterokonts (also known as Stramenopiles) (Baldauf, 2003). They are superficially similar to fungi and share some key characteristics such as a mycelial growth habit, spore production, infection strategies and they are able to reproduce both sexually and asexually (Latijnhouwers *et al.*, 2003). However, they are phylogenetically distinct from fungi (Baldauf, 2003) and differ in some key characteristics such as the presence of cellulose as the major constituent of their cell walls, instead of chitin, and they are diploid during the vegetative state (Latijnhouwers *et al.*, 2003).

The oomycetes are a widespread group of organisms containing both saprobes and pathogens; however they are mostly known as plant pathogens. They are responsible for infamous crop diseases such as the potato late blight that was responsible for the Irish potato famine. In a ranking of the top ten plant pathogenic oomycetes, *Phytophthora* species (foliar blights, root and stem rots) made up the majority of the species (Kamoun *et al.*, 2015). Other important species are the downy mildews such as *Hyaloperonospora arabidopsidis* and *Plasmopara viticola*, as well as the damping off and root rot pathogen, *Pythium ultimum*, and the white rust, *Albugo candida* (Kamoun *et al.*, 2015). However oomycetes can also infect other organisms including animals and fish. Some of the most important diseases of fish are from the family

Saprolegniaceae (Noga, 1993; Torto-Alalibo *et al.*, 2005; Phillips *et al.*, 2008), including *Saprolegnia parasitica* which is especially important on salmon and trout species (Torto-Alalibo *et al.*, 2005). *Pythium* is often considered an important disease of young plants in nurseries; however *Pythium insidiosum* is an important pathogen of mammals, including humans, causing the disease Pythiosis (Phillips *et al.*, 2008; Gaastra *et al.*, 2010).

Oomycetes display a range of lifestyles from obligate biotroph to necrotroph (Fawke *et al.*, 2015). They can share an intimate relationship with a single host as in the case of obligate biotrophs, or they can exhibit a generalist lifestyle, infecting hundreds or thousands of different host species as in the forest pathogens *P. ramorum* or *P. cinnamomi* (Hardham, 2005; Cahill *et al.*, 2008; Grünwald *et al.*, 2012). Many are common inhabitants of aquatic and terrestrial environments. Natural dispersal is through wind and water (Fawke *et al.*, 2015); however their frequent association with plant material, soil and roots makes them very amenable to human dispersal through trade, agricultural and horticultural activities (Desprez-Loustau *et al.*, 2010). In general, we have a much better understanding of the species that cause important diseases with a large economic or ecological impact. Our knowledge of these microorganisms in natural habitats is more limited (Arcate *et al.*, 2006). Several recent studies of oomycetes in soil environments show a dominance of the genus *Pythium* in both agricultural and natural environments (Arcate *et al.*, 2006; Nelson & Karp, 2013; Sapkota & Nicolaisen, 2015; Riit *et al.*, 2016; Esmaeili Taheri *et al.*, 2017). In natural environments, oomycetes may play a role in driving succession (Dickie *et al.*, 2017), nutrient cycling (Marano *et al.*, 2016), and creating or maintaining forest diversity (Packer & Clay, 2000).

Phytophthora in natural ecosystems

Phytophthora species have long been known as pathogenic species associated with agricultural activities. However, studies done since the mid-1990s suggest that a diverse community of *Phytophthora* species exist in natural ecosystems (Hansen *et al.*, 2012a), many of which rarely cause noticeable disease or disturbance to the ecosystem (Hansen *et al.*, 2012a). In part because of the emergence of new destructive *Phytophthora* species, there has been increased interest in studying *Phytophthora* communities in forest soils and streams (Hwang *et al.*, 2008; Reeser *et al.*, 2010, 2011; Hüberli *et al.*, 2013; Oh *et al.*, 2013; Shrestha *et al.*, 2013; Brazee *et al.*, 2016; Stamler *et al.*, 2016), or to link tree declines with a possible *Phytophthora* pathogen (Greslebin *et al.*, 2005, 2007; Balci *et al.*, 2007; Vannini *et al.*, 2013). These studies have uncovered many new *Phytophthora* species. The genus now contains over 100 described species, half of which have been described only in the past ten years (Kroon *et al.*, 2012; Hansen *et al.*, 2012a). The genus has a global distribution with some widely distributed species and others with a limited distribution. *Phytophthora* species have been found in waterways, forest soils and more recently, are being discovered in the tree canopies of forests (Hansen *et al.*, 2012a, 2017).

Phytophthora communities from waterways appear to be dominated by species belonging to a monophyletic group designated as ITS clade 6, based on the sequence of the internal transcribed spacer of the ribosomal cluster (Jung *et al.*, 2011). Species from this group are rarely associated with disease (Brasier *et al.*, 2003). Several of the species appear to be saprophytic, often being found in riparian habitats or on submerged wood and plant material (Brasier *et al.*, 2003; Jung *et al.*, 2011). In North America *Phytophthora gonapodyides* (H.E. Petersen) Buisman,

Phytophthora lacustris Brasier, Cacciola, Nechwatal, Jung & Bakonyi, and *Phytophthora chlamydospora* (Brasier *et al.*, 2003) are the species most frequently recovered in waterways (Hwang *et al.*, 2008; Hulvey *et al.*, 2010; Reeser *et al.*, 2011; Hansen *et al.*, 2012b). Although these species have a rather ubiquitous distribution in the northern hemisphere, they are limited in the southern hemisphere where other species dominate (Jung *et al.*, 2011).

Soil communities show similar trends as waterway communities, except that invasive pathogens appear to be more prevalent. In the soil of declining oak forests in the Eastern US, nine different *Phytophthora* species were found but only *P. cinnamomi* was associated with disease symptoms (Balci *et al.*, 2007). *Phytophthora cinnamomi* was the most frequently isolated species in forest soils in the eastern US, and the second most frequently isolated species in Oregon after *P. ramorum* (Balci *et al.*, 2007; Reeser *et al.*, 2010). Both of these species are believed to be introduced (Rizzo *et al.*, 2002, Hansen, 2008, Hansen *et al.*, 2012a). Their presence in such high frequency in natural forest soils highlights the importance of invasive species in *Phytophthora* communities.

The forest canopy represents another niche where *Phytophthora* are found that has only recently been explored (Hansen *et al.*, 2012a). *Phytophthora ramorum* is an aerially dispersed pathogen infecting foliage and stems of its hosts. Since its discovery, monitoring activities in the US have uncovered additional *Phytophthora* species in the forest canopy. *Phytophthora nemorosa* E. M. Hansen and Reeser and *Phytophthora pseudosyringae* T. Jung & Delatour were frequently found in California and Oregon forest canopies (Wickland *et al.*, 2008; Reeser *et al.*, 2010). In Oregon, *P. gonapodyides* and *P. chlamydospora* were also present in soils and canopy drip, but in lower frequency than in waterways (Reeser *et al.*, 2010) suggesting not only a global distribution in

waterways for these two species, but possibly also in soils and forest canopies. Aerial dispersal can greatly increase the efficiency and length of dispersal and can generate rapid spread across and even between continents (Brown & Hovmøller, 2002). This can be a serious problem for epidemics in plantations where a pathogen can quickly destroy an entire stand.

Influence of anthropogenic activities on natural communities

Some *Phytophthora* species, such as *Phytophthora cryptogea* Pethybr. & Laff. or *Phytophthora citricola* Sawada, are found globally. Whether this ubiquitous distribution is natural or has taken place through trade and their association with the agriculture and food industry is not known. Hansen *et al.* (2006) report finding *P. citricola, Phytophthora cambivora* (Petri) Buisman, *Phytophthora syringae* (Berk.) Kleb. and *Phytophthora cactorum* (Lebert & Cohn) J. Schr. only occasionally in forest systems despite their dominance in urban and nursery settings. This pattern of distribution suggests that movement through trade has resulted in the ubiquitous nature of these agricultural pathogens, followed by subsequent escape or migration into forests. How much agricultural practices have changed the natural species composition of *Phytophthora* or of oomycetes in general is largely unknown. In the eastern US, Hulvey *et al.* (2010) compared isolates found in nurseries and in natural waterways and found three species common to streams and nurseries: *P. citricola*, *Phytophthora hydropathica* Hong, C. and Gallegly, M. and *Phytophthora citrophthora* (R.E. Sm. & E.H. Sm.) Leonian (the *P. citricola* complex has since been revised and *P. citricola* may be the wrong name). For the species from the *P. citricola* complex, an identical DNA fingerprint in nursery and stream samples suggest an escape (Hulvey *et al.*, 2010).

Phytophthora in British Columbia

Natural *Phytophthora* communities have not been surveyed in British Columbia and Canada. Similar to other countries, it is likely that BC forests harbour both native and introduced *Phytophthora* species. Some common forest pathogens have been found in irrigation water, gardens and parks and may have made their way into forests as well. In a study from BC in the 1960s, some common agricultural pathogens were found in irrigation water from orchards in the Okanagan and Similkameen valleys in the southern part of the province (McIntosh, 1966). Several irrigation sources sampled identified *P. cactorum* as the most common contaminant. In addition, *P. cambivora*, *P. citricola*, and *Phytophthora megasperma* Drechsler as well as some unidentified species were found. Identifications were made from morphological characteristics and there was some level of uncertainty associated with the identifications (McIntosh, 1966); furthermore many species have since been re-evaluated using molecular techniques, and isolates that were previously identified as one species, for example *P. megasperma*, turned out to be different species. In 1965, Atkinson looked at the causal agent of root rot in *C. lawsoniana* in nurseries and three gardens in southern BC in the Fraser Valley and in Victoria. He found *P. lateralis*, *P. cinnamomi*, and an unidentified *Phytophthora* species (Atkinson, 1965). More recently, *P. lateralis* has been found associated with diseased ornamental *C. lawsoniana* in the city of Victoria (Gorman, 1993) where infected trees had to be eradicated from parks and boulevards (Michelle Gorman, personal communication). *Phytophthora ramorum* has also been found in BC nurseries (Bilodeau *et al.*, 2007). There are likely other pathogens present in agricultural systems that may have forest hosts as well, so there is a potential that these species

may have made their way into forest systems affecting the natural *Phytophthora* community, and the natural host community as well.

Oomycete and *Phytophthora* monitoring

Oomycete and *Phytophthora* monitoring or surveying have long been conducted using stream baiting methods or by directly isolating the pathogen from infected plant tissues. Stream baiting involves using selected bait materials placed in some sort of container or bag which is anchored in a waterway allowing water to pass through the trap. The motile zoospores of oomycetes swim to the bait tissues, encyst and colonize. After a couple of days to weeks, the pathogen can be isolated from the necrotic tissues. Many variations of this technique have been used, ranging from orchard fruits placed in tin cans with holes punched into the sides for surveys of species found in orchard irrigation water (McIntosh, 1966), to leaves of target species placed in mesh bags and floated in streams to gain a general idea of the species present in an ecosystem (Hulvey *et al.*, 2010; Reeser *et al.*, 2011; Hüberli *et al.*, 2013; Shrestha *et al.*, 2013; Brazee *et al.*, 2016). Rhododendron leaves are now commonly utilized as bait to attract a wide variety of species, but other plant tissues are sometimes used to target species that are more host specific or to match host and pathogen species that are found in particular ecosystems. For instance, Stamler and group used poplar and willow leaves as bait in stream surveys in the Southwestern USA as those were the common tree species in the ecosystems being surveyed (Stamler *et al.*, 2016). *Quercus robur* leaves were used as bait in *Phytophthora* surveys of Oak systems to target species potentially involved in Oak decline (Balci *et al.*, 2007). Many alternative techniques are now being commonly utilized to increase sampling efficiency. For instance baiting in the lab, where

soil and water samples are processing in the lab rather than *in situ* (Hansen *et al.* 2017; Balci *et al.*, 2007; Oh *et al.* 2013). This has allowed surveying to be conducted on soil samples as well as water. Soil samples are placed in containers, flooded with water, and bait leaves are floated on top to attract the pathogen. Filtering is another method where water samples are passed through a filter and the pathogen can either be isolated from the filters, or directly identified using molecular techniques (Balci *et al.*, 2007; Hwang *et al.*, 2008; Reeser *et al.*, 2011; Rollins *et al.*, 2016). Finally, direct identification and characterization of oomycete and *Phytophthora* communities is being done using next generation sequencing techniques (Vannini *et al.* 2013; Català *et al.*, 2015, 2017; Sapkota and Nicolaisen 2015; Prigigallo *et al.*, 2016; Riit *et al.*, 2016; Burgess *et al.*, 2017; Dickie *et al.*, 2017; Esmaeili Taheri *et al.*, 2017)(Català *et al.*, 2015, 2017; Burgess *et al.*, 2017b).

Oomycete and *Phytophthora* species require water and favorable environmental conditions for dissemination of the motile zoospores. Like many other microorganisms, environmental conditions can dictate timing for various life cycle stages. Many survey and monitoring studies are conducted over different time frames to capture a wider sampling of the species present (Balci *et al.*, 2007; Reeser *et al.*, 2011). Time of year affected the diversity and species detected in streams over a year-long survey conducted in the USA (Hwang *et al.* 2008). Sampling in spring versus fall resulted in twice as many species in one study on *Phytophthora* in Oak stands (Balci *et al.*, 2007). Selection of monitoring methods, season, habitat and geographic location can all affect results in *Phytophthora* or oomycete community studies which may complicate disease management or underestimate natural community diversity.

Phytophthora biology

Life cycle, reproduction and dispersal

Phytophthora can be either homothallic (possessing the ability to self) or heterothallic (requiring individuals with different mating types for reproduction) (Brasier, 1992); however some outcrossing or self-fertilizing can be found in either mating system (Brasier, 1992; Goodwin, 1997). *Phytophthora* have a mixed mating system, possessing both a sexual and an asexual phase. Chlamydospores and sporangia are spores produced during the asexual phase and oospores during the sexual phase. Sporangia are released under favorable conditions which could be during periods of rainfall or extended leaf wetness in the rainy seasons for aerial species (Davidson *et al.*, 2005; Hansen *et al.*, 2008, 2012a), or in times of extended saturation in soils (Hansen *et al.*, 2008). These spores and can either directly germinate on the host surface or germinate and release zoospores, a type of spores with flagella, which can swim to host tissue by chemotaxis. Once in contact with host tissue the zoospores encyst and initiate infection. Chlamydospores are asexual resting spores that can remain dormant in host tissue or soil and germinate under suitable conditions (Grünwald *et al.*, 2012). During the sexual phase, the sexual reproductive organs (oogonia and antheridia) are produced, in the presence of the opposite mating type for heterothallic species, and fusion of the organs results in the production of an oospore (Grünwald *et al.*, 2012). Oospores are thought to be resting spores that can withstand extreme conditions (Latijnhouwers *et al.*, 2003; Grünwald *et al.*, 2012).

Infection biology

Phytophthora are hemibiotrophs possessing both biotrophic and necrotrophic phases (Latijnhouwers *et al.*, 2003). They colonize and interact with living plant tissue and in the later necrotrophic phase when the plant cells die, they feed on the dead cells (Latijnhouwers *et al.*, 2003). Once established on the host, *Phytophthora* produce appressoria that penetrate the host where the pathogen can then grow within the intercellular matrix (Kamoun, 2006; Birch *et al.*, 2006). Once inside the host *Phytophthora* secrete proteins called effectors that manipulate host structure or function (Huitema *et al.*, 2004; Kamoun, 2006). Effectors can be secreted into the plant apoplast or inside host cells through infection structures such as haustoria or through infection vesicles (Kamoun, 2006).

Apoplastic effectors include several classes of proteins involved in infection or in blocking plant defense processes enabling pathogens to overcome their hosts. Some proteins trigger host cell death such as the Nep1-like proteins (NLP) and small cysteine-rich proteins. The small cysteine rich proteins are specific to oomycetes (Jiang & Tyler, 2012) whereas the NLP proteins have been found in oomycetes, bacteria and fungi (Pemberton & Salmond, 2004). *Phytophthora* species contain much larger NLP proteins families in comparison to other closely related oomycetes (Jiang & Tyler, 2012).

Many effectors are under diversifying selection, most likely to evade host recognition, and are poorly conserved. However elicitins are a highly conserved group of proteins under purifying selection that are unique to *Phytophthora* and *Pythium* (Panabières *et al.*, 1997; Jiang *et al.*, 2006). Elicitins induce a host response called the hypersensitive response in plants (Jiang *et al.*,

2006; Kamoun, 2006). Jiang *et al.* (2006) found 17 distinct clades in the elicitin gene family, 13 of which included proteins from three or more *Phytophthora* species suggesting that divergence in the gene family pre-dates species divergence.

The cytoplasmic effectors include the RXLR effectors and the CRN or crinklers and necrosis inducing effectors. RXLR effectors appear to suppress the host immune response that often results in host cell death (Jiang & Tyler, 2012). The originally identified CRN effectors cause cell death and leaf crinkling in tomato and *Nicotiana* species (Torto *et al.*, 2003; Kamoun, 2006). The C-terminus of the proteins are highly divergent both from their homologs in other *Phytophthora* species as well as between paralogs within species (Jiang *et al.*, 2008; Haas *et al.*, 2009; Schornack *et al.*, 2010). Jiang *et al.* (2008) found a large super family of RXLR effectors (avirulence homolog or Avh genes) that they believe has evolved from a single gene in a common ancestor of *Phytophthora sojae* Kaufm. & Gerd. and *P. ramorum*. Many RXLR proteins are under positive selection (Win *et al.*, 2007; Jiang *et al.*, 2008), and the majority of the genes are in frequently rearranged non-conserved regions of the genome, and do not form a homologous cluster based on sequence homology (Jiang *et al.*, 2008).

Phytophthora genomics

Whole genome sequencing has allowed for some important observations and inferences on the evolution of the *Phytophthora* genus as well as individual species within the genus. Genome size varies widely in the *Phytophthora* species sequenced to date. *Phytophthora infestans* has experienced large-scale genome expansion and is approximately 240 Mb, whereas the genomes

of other species are much smaller at 95 Mb for *P. sojae* and 65 Mb for *P. ramorum* (Tyler *et al.*, 2006; Haas *et al.*, 2009). *Phytophthora infestans* has a much higher amount of repetitive DNA and transposable elements, accounting for approximately 74% of the genome (Haas *et al.*, 2009). Seidl et al. (2012) show large number of gene gains and duplications in the evolutionary history of the oomycetes and of the *Pythium* ï *Peronosporales* clade, and a higher amount of duplications leading to the *Phytophthora* genus. There are large expansions of the protein groups involved in host pathogen interactions in oomycetes (Seidl *et al.*, 2012) and there is copy number variation in genes between strains of a species, for example in *P. sojae* (Qutob *et al.*, 2009). The differences in copy numbers of a virulence or avirulence gene can result in virulence differences between strains (Qutob *et al.*, 2009). Copy number variation has been found in *Phytophthora* genomes, however has not been reported in fungal genomes (Raffaele & Kamoun, 2012).

Genome organization likely plays an important role in *Phytophthora* evolution. Comparative genome analyses of *Phytophthora* species show a distinct pattern of gene rich regions interspersed by gene sparse regions (Haas *et al.*, 2009; Raffaele *et al.*, 2010a). Gene dense regions have short intergenic distances and gene order is highly conserved. In gene sparse regions, intergenic regions are larger and gene order is not conserved (Haas *et al.*, 2009; Raffaele *et al.*, 2010a). This pattern is most pronounced in *P. infestans* which has much larger intergenic distances and a much higher repeat content compared to *P. ramorum* and *P. sojae* (Haas *et al.*, 2009). The gene dense regions are rich in core orthologous genes and the gene sparse regions contain effector genes. In a core set of 7,113 orthologous genes sharing a 1:1:1 relationship between *P. infestans*, *P. sojae*, and *P. ramorum*, approximately 90% were located in the gene dense, conserved regions (Haas *et al.*, 2009). These included genes involved in cellular processes

such as DNA replication, transcription, and translation. In contrast, effector or virulence genes are commonly found in the gene sparse regions (Raffaele *et al.*, 2010b). This pattern was also found by Raffaele *et al.* (2010a) in a genome comparison of the clade 1c species ï *P. infestans*, *Phytophthora ipomoeae* Flier & Grünwald, *Phytophthora mirabilis* Galindo & H.R. Hohl, and *Phytophthora phaseoli* Thaxt. Raffaele *et al.* (2010a) also found signs of accelerated evolution in the genes located in repeat rich, gene-sparse regions compared to those located in the gene-dense regions. Genes showed signs of positive selection and there were more structural polymorphisms in these regions. Many of the genes found in the repeat rich, gene-sparse regions were plant-induced genes, including many effectors Raffaele *et al.* (2010a). Haas *et al.* (2009) hypothesize that the mobile elements and dynamic nature of the repeat rich, gene-sparse regions could cause more recombination and gene loss and gene gain events in RXLR effector genes and may be a mechanism involved in the speed of evolution of *P. infestans*.

In addition to faster evolution in regions of the genome, genes involved in host interactions evolve more rapidly relative to the rest of the genome. The genomes of *Phytophthora ramorum* and *P. sojae* were sequenced by Tyler *et al.* (2006) and comparisons revealed rapid expansion and diversification of gene families associated with pathogenicity, in particular with the RXLR gene family. In addition they found a faster rate of evolution in the collective group of secreted proteins (secretome) relative to the proteome (Tyler *et al.*, 2006).

Emergence and evolution of new pathogens

Pathogen evolution can happen through many processes such as clonal speciation, host jumps, horizontal gene transfer and hybridizations (Brasier, 1995; Stukenbrock & McDonald, 2008; Stukenbrock, 2013, 2016). Some processes are slow and can take many generations to result in speciation; this is the case of allopatric speciation in a stable habitat. However, some events can result in the rapid evolution of new species. Sudden changes in a population or habitat can result in increased selection causing periods of rapid evolution (Brasier, 1995). There can be several causes of sudden change resulting in increased selection including massive habitat changes, species introductions, and host jumps. Species that are introduced from their native range to a new environment often confront new conditions and hosts. Successful introduced species adapt to the new environment and hosts and may result in clonal speciation or they may hybridize with existing species to create hybrid species able to establish in the new environment (Brasier, 1995, 2001; Stukenbrock & Bataillon, 2012; Gladieux *et al.*, 2014).

Hybridization

Hybridization may provide an opportunity for the evolution of a new species. In many cases, if hybrids are able to form, they introgress back to one or both parental species and do not persist as an independent species (Hegarty, 2012). However in some cases hybrids are able to occupy a new niche different from that occupied by the parental species. Transgressive segregation and increases in genetic variation or genome plasticity can generate novel gene combinations creating hybrid individuals that are pre-adapted for novel habitats (Rieseberg *et al.*, 1999; Stukenbrock & Bataillon, 2012; Hegarty, 2012; Callaghan & Guest, 2015; Stukenbrock, 2016).

The novel niche and gene combinations can lead to rapid evolution, hybrid speciation or the emergence of new diseases (Rieseberg *et al.*, 1999; Stukenbrock & Bataillon, 2012; Hegarty, 2012; Callaghan & Guest, 2015; Stukenbrock, 2016).

Several examples exist of pathogens that have emerged on new hosts or in new environments as a result of hybridization. *Phytophthora andina* Adler & Flier, a hybrid between *P. infestans* and an unknown *Phytophthora* parent (Goss *et al.*, 2011a), is an emerging pathogen of *Solanum* species in the Andes. Although both pathogens infect *Solanum* hosts, the species they infect are different and there is little overlap (Oliva *et al.*, 2010; Goss *et al.*, 2011a) suggesting that hybridization could have led to a host range expansion or shift (Goss *et al.*, 2011a). *Phytophthora* x*alni* was first discovered on Alder in Europe (Brasier *et al.*, 1995). The pathogen resembled *P. cambivora* but was suggested to be a group of heteroploid hybrids (Brasier *et al.*, 1999). Three hybrid sub-species were described as *P. alni* subsp. *alni*, *P. alni* subsp. *uniformis* Brasier & S.A. Kirk, and *P. alni* subsp. *multiformis* Brasier & S.A. Kirk (Brasier *et al.*, 2004); and subsequently reclassified as *P.* x*alni*, *P. uniformis*, and *P.* x*multiformis* (Husson *et al.*, 2015). *Phytophthora* x*alni* is widespread across most of Europe, and it is also the most aggressive pathogen of Alder (Brasier *et al.*, 2004). It is a hybrid between *P. uniformis*, and *P.* x*multiformis* (Ioos *et al.*, 2006). *Phytophthora uniformis* may have been generated by an ancient hybridization event between *P. cambivora* and another *Phytophthora* species and *P.* x*multiformis* could have been generated interspecific hybridization of two unknown species (Ioos *et al.*, 2006; Husson *et al.*, 2015). *Phytophthora* x*alni* likely hybridized with *P. uniformis* after it was introduced from North America (Aguayo *et al.*, 2016). It appears that these *Phytophtora* species are specific to

Alder, and also that Alder is not a host of *P. cambivora* (Petri) Buisman (Brasier *et al.*, 1999) suggesting that hybridization has led to a host jump.

Hybridization can also be an important source of genetic exchange between closely related species. Introgression through backcrossing of hybrid species with parental species can introduce new genes from one parental species to the other. In Europe, the Dutch elm pathogen *Ophiostoma novo-ulmi* Brasier, which was initially introduced as a few clones of only one mating type which were spreading asexually, acquired the second mating type gene as well as other pathogenicity genes through introgression with *Ophiostoma ulmi* (Buisman) Nannf. (Brasier, 2012). The new genes and ability to sexually recombine would have increased the long term evolutionary potential of the pathogen evident by the sudden reduced impact of a virus that had been spreading through the pathogen population (Brasier, 2012). In cases where hybrids are unable to produce an F2 generation, they may still be able to backcross with parental species allowing gene flow back to the parental populations. In crosses between *Phytopthora capsici* Leonian and *Phytophthora tropicalis* (Aragaki & J.Y. Uchida) hybrids were unable to reproduce with each other however they were able to reproduce with the *P. capsici* parent (Donahoo & Lamour, 2008).

Influence of disturbances on hybridization

Hybridization between species has often been linked to human activities or disturbances in natural habitats (Callaghan & Guest, 2015; Burgess, 2015) . There are many examples of hybridization when an organism is introduced into a region that a closely related species inhabits.

Hubbs (1955), in his work with hybridization in fish species, made the important observation that the scarcity of one species and abundance of another closely related species often led to hybridization due to difficulties of the rare species in finding the proper mate. Introduced species hybridizing with native species may often be due to a lack of reproductive opportunities with conspecific mates in the introduced individuals. Hybridization can also occur as a result of human activities, habitat disturbance, or environmental changes if the changes bring together or increase the frequency of species that were once separated spatially or temporally (Hegarty, 2012). Nursery systems have very different environmental conditions than natural systems. The constant influx of plant material often brings pathogens together that would normally be separated in a natural habitat. In addition, there can be large numbers of host plants all within close proximity and abundant watering which are conditions that are conducive to pathogen development and that are usually not associated with natural habitats. These conditions may increase the opportunities for hybridization. There are several examples of hybridization of *Phytophthora* species in nurseries. *Phytophthora alni* subsp. *alni* is believed to have originated from a hybridization event in nurseries. Hybridization appears to be relatively common between *Phytophthora nicotianae* Breda de Haan and *P. cactorum* in hydroponic cultures in the Netherlands (Man in ά Veld *et al.*, 1998; Bonants *et al.*, 2000). Hybrids were recovered from different hosts in different years and molecular analyses suggest that the hybrids were generated on multiple occasions. Some of the hosts that the hybrids were isolated from were not host species of the parental species suggesting a possible shift in host range in the hybrids (Bonants *et al.*, 2000). Hybrids of *P. cactorum* and *Phytophthora hedraiandra* De Cock & Man in 't Veld, believed to be introduced, were also found in nurseries in the Netherlands as well as in a few parks suggesting movement through infected nursery material (Man in ά Veld *et al.*, 2007).

These hybrids were also able to infect numerous hosts that the parental species could not infect (Man in ά Veld *et al.*, 2007). These examples suggest that nursery environments may be conducive to hybridization events.

Hybridization in *Phytophthora*s appears to be relatively frequent, and may be an important process in the evolution and adaptability of these species (Bertier *et al.*, 2013; Burgess, 2015). In some environments, hybridization appears to be a natural and frequent process (Burgess, 2015). How anthropogenic activities or niche preferences have influenced these processes is also unknown; however the more frequent finding of natural hybrids in nursery settings suggests that there is an influence.

Phytophthora ramorum: An emerging pathogen introduced through nursery trade

Successful establishment into a new habitat does not necessarily require a pathogen to develop novel adaptations. If the environment or hosts are similar enough the pathogen may be able to establish and successfully spread with little to no adaptive changes. In other cases the pathogen may be able to establish, but selection pressures may be high favoring certain genotypes over others and may lead to clonal speciation (Brasier, 1995). *Phytophthora ramorum* is an invasive pathogen that has been introduced to North America and Europe and is spreading asexually throughout the range where it was introduced.

History of *P. ramorum*

Phytophthora ramorum was discovered in the mid-1990s in both Europe (Werres *et al.*, 2001) and the US (Rizzo *et al.*, 2002). In the US, the pathogen was most likely first introduced to a nursery in California (Mascheretti *et al.*, 2008), but spread to adjacent forests where it caused extensive mortality of coast live oak (*Quercus agrifolia* Nee.) and tanoak before it was noticed (Rizzo *et al.*, 2002; Mascheretti *et al.*, 2008). It is now found in natural forest ecosystems extending from California through Oregon (Rizzo & Garbelotto, 2003) and has been most recently isolated from waterways in Washington State as well as in some south eastern states (Chastagner *et al.*, 2009). *Phytophthora ramorum* has been raising concerns due to the impact on the nursery trade and on valued oak species in the United States, and more recently on commercially important Japanese larch in Europe. In 2005, an estimated 20% of the oak and tanoak trees in the Big Sur Ecosystem of Coastal California had been killed by *P. ramorum* (Meentemeyer *et al.*, 2008).

In Europe, *P. ramorum* was first found on woody ornamentals in Germany and the Netherlands (Werres *et al.*, 2001). It is now found throughout many European countries (Grünwald *et al.*, 2012) in nurseries and gardens largely affecting *Rhododendron* species with limited infection on trees (Brasier & Webber, 2010); however the pathogen has recently jumped to Japanese larch plantations in the UK where it is causing large-scale mortality. By early summer 2010, 1,900 hectares of Japanese larch had been affected by *P. ramorum* (Brasier & Webber, 2010).

In Canada, the pathogen has been found in several nurseries in British Columbia (Bilodeau *et al.*, 2007). Plants originating from infected nursery beds were shipped between British Columbia and

West Coast US states several times, indicating the possibility for movement of the pathogen between the two countries (Goss *et al.*, 2009). *Phytophthora ramorum* has not been reported in natural forests in Canada, however it has been found in some residential locations where eradication is underway (Kliejunas, 2011).

Biology and Life Cycle of *P. ramorum*

The life cycle of *P. ramorum* follows the general life cycle of *Phytophthora* species (Grünwald *et al.*, 2008) except that *P. ramorum* affects the above ground parts of the host species instead of the roots (Rizzo & Garbelotto, 2003; Hansen *et al.*, 2008). Sporangia are produced on infected leaf tissue and released during periods of rainfall or periods of extended leaf wetness in the rainy seasons (Davidson *et al.*, 2005; Hansen *et al.*, 2008). Sporangia are most commonly dispersed by rain droplets (Davidson *et al.*, 2005; Hansen *et al.*, 2008) and wind driven rain usually within zero to ten meters and in some cases up to 15 m (Davidson *et al.*, 2005); however dispersal (longer range dispersal) and infection patterns (top down infection in canopies) suggest that air dispersal is also an important means of spread (Hansen *et al.*, 2008). Spores can also be dispersed through waterways; however this is thought to be a secondary source of infection suggested by the distribution and proximity of new infections (Hansen *et al.*, 2008; Grünwald *et al.*, 2012). Zoospores are produced and released from the sporangia (Davidson *et al.*, 2005) and it is believed that once in contact with a suitable host the zoospores encyst and penetrate the host (Grünwald *et al.*, 2008). Direct germination has not been documented from the sporangia; however this could occur similar to other *Phytophthora* species (Grünwald *et al.*, 2008). *Phytophthora ramorum* also produces chlamydospores in infected plant tissue (Werres *et al.*,

2001; Davidson *et al.*, 2005). They can remain dormant in plant tissue or plant debris in the soil and cause infection when in contact with a host under suitable conditions (Shishkoff, 2007). *Phytophthora ramorum* is heterothallic, requiring both mating types for sexual reproduction to occur (Werres *et al.*, 2001), however in the current known range sexual reproduction has not been observed (Goss *et al.*, 2009; Prospero *et al.*, 2009). Oospores are produced under laboratory conditions when the two mating types are grown together (Werres *et al.*, 2001; Brasier & Kirk, 2004) suggesting that this part of the general *Phytophthora* life cycle should hold true for *P. ramorum*.

Phytophthora ramorum causes two distinct disease types: a bole canker type or a foliar blight (Rizzo *et al.*, 2002; Grünwald *et al.*, 2008). The canker type occurs on some oak species, tanoaks European beech (*Fagus sylvatica* L.) and on Japanese larch. The cankers spread under the bark, killing the phloem and girdling the tree (Grünwald *et al.*, 2008) and the infection may also spread to the xylem (Brown & Brasier, 2007). The foliar blight occurs on a much broader range of hosts, over 100 different species (Grünwald *et al.*, 2012), ranging from *Rhododendron* and *Viburnum* to conifer species including Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) (Grünwald *et al.*, 2008). The foliar blight is characterized by leaf spots, or shoot tip dieback, and wilting (Grünwald *et al.*, 2008). *Phytophthora ramorum* sporulation differs on the various hosts, ranging from no sporulation occurring in the cankers on *Quercus* species to prolific sporulation on the foliar tissue in some species such as California bay laurel (*Umbellularia californica* (Hook. & Arn.) Nutt.) (Davidson *et al.*, 2005).

Population Characteristics and Migration

Three distinct lineages of *P. ramorum* have been described in the known range of the pathogen, and a fourth newly described lineage (EU2) that has been found in Europe (Van Poucke *et al.*, 2012). NA1 is the dominant lineage in the United States (Ivors *et al.*, 2006; Mascheretti *et al.*, 2008), and has been found in BC nurseries as well (Goss *et al.*, 2011b). NA2 is the dominant lineage in British Columbia nurseries but has only been found in Washington and California nurseries in the United States, but not in forest stands (Goss *et al.*, 2011b). The EU1 lineage was first found in Europe but has since spread to North America where it is found in nurseries in Canada and the US (Goss *et al.*, 2011b). To date, all of the NA1 and NA2 isolates that have been found are the A2 mating type and EU1 has mostly been of the A1 mating type; however three isolates found in Belgium belonging to the EU1 lineage possessed the A2 mating type (Vercauteren *et al.*, 2011). Analysis on the evolutionary history of NA1, NA2 and EU1 using nuclear sequence data suggests that the three lineages diverged approximately 165 000 to 500 000 years ago and remained isolated since their divergence (Goss *et al.*, 2009); however the nuclear genes studied and the genome show high levels of heterozygosity suggesting that the founder population was sexually reproducing (Tyler *et al.*, 2006; Goss *et al.*, 2009). It is likely that the three lineages were initially introduced independently in North America and Europe (Goss *et al.*, 2009) and that the current populations were generated by the introduction of only one or a few different isolates (Goss *et al.*, 2009; Vercauteren *et al.*, 2010). Some differences were found in laboratory tests in the aggressiveness of the three lineages on *Rhododendron* leaves between, with NA2 and EU1 showing higher levels of aggressiveness than NA1 (Elliott *et al.*, 2011).

Overall genetic diversity was low in all three lineages of *P. ramorum*, reminiscent of clonally expanding populations (Prospero *et al.*, 2009; Vercauteren *et al.*, 2010; Goss *et al.*, 2011b). In the NA1 and EU1 lineages there were a high number of genotypes; however they were based on only a few microsatellite loci and genotypes were closely related to one parent genotype (Prospero *et al.*, 2009; Vercauteren *et al.*, 2010). Very little genotypic diversity has been reported in the NA2 lineage; however this is based on microsatellite markers developed for the NA1 and EU1 lineages which could underestimate diversity in NA2 (Goss *et al.*, 2011b). Differences between the different genotypes in all lineages can be explained by microsatellite mutation suggesting microsatellite mutation as the dominant process generating new genotypes rather than sexual recombination (Prospero *et al.*, 2007, 2009; Vercauteren *et al.*, 2010). The absence of sexual recombination could be from a lack of opportunity due to the presence of only one or the other mating type (Prospero *et al.*, 2009), a partially dysfunctional reproductive system suggested by the difficulty in obtaining successful pairings *in vitro* (Brasier & Kirk, 2004; Garbelotto *et al.*, 2006), or incompatibility between the three lineages due to an extended period of divergence (Garbelotto *et al.*, 2006; Goss *et al.*, 2009). In Belgium where the two mating types were found in the EU1 lineage, no signs of recombination were detected in the overall population (Vercauteren *et al.*, 2010).

Population genetic studies suggest that the nursery trade is the predominant method of dispersal shaping the population structure of *P. ramorum*. Prospero *et al.* (2009) looked at genetic structure and migration patterns of *P. ramorum* in California, Oregon and Washington in both forests and nurseries using microsatellite markers. They found low overall genetic differentiation. Genetic distances were small between nursery populations from Oregon,

Washington and California, and there were very large geographical distances between populations with shared genotypes, much larger than the distances found for natural dispersal. In addition, there was no correlation between geographic distance and the number of shared genotypes with some of the populations sharing the highest numbers of genotypes separated by the largest geographic distances. Their data was consistent with trace back and trace forward surveys showing that infected plants were moved back and forth between the nurseries where the *P. ramorum* infections were found (Prospero *et al.*, 2009). Other studies looking at the inter-continental migration patterns of *P. ramorum* have also highlighted the importance of the nursery trade in pathogen spread. There is continental migration from Europe to North America (Goss *et al.*, 2011b). Shared genotypes were found in both continents; however they were found in Europe in years prior to being found in North America. It is unclear if migration from Europe occurred simultaneously into Canada and the US, or via secondary spread from Canada to the US or from the US to Canada. Movement of NA1 could be in a northward direction from the US to British Columbia suggested by a staggered temporal occurrence of one particular genotype, the lower occurrence of that lineage in BC, and trace forward and back surveys also showing infected material moving north (Goss *et al.*, 2011b). The higher occurrence of the NA2 lineage in BC and the more infrequent occurrence in the US could indicate a southward migration from BC to the US for that lineage.

Studies in US forests and nurseries suggest that migration of *P. ramorum* genotypes occur from nurseries to forests (Prospero *et al.*, 2007; Mascheretti *et al.*, 2008). Higher genetic diversity was found in California nurseries than in forests using AFLP markers (Ivors *et al.*, 2004) and microsatellite markers (Ivors *et al.*, 2006) and forest populations appear to have been started

from the introduction of a single isolate (Ivors *et al.*, 2006). In California the lack of genetic structure between some forest sites and nurseries reflects human mediated spread from nurseries to forests rather than natural spread (Mascheretti *et al.*, 2008). In Oregon, population studies suggest the epidemic was started in Oregon nurseries and forests by introductions from outside nurseries and genotype flow does not occur from forests to nurseries (Prospero *et al.*, 2007).

Center of Origin

The origin of *P. ramorum* is currently unknown. Evidence suggests that in Europe and in North America the pathogen is introduced. The range of *P. ramorum* is limited in relation to its hosts (Rizzo & Garbelotto, 2003) and the dominance of single genotypes in California forests and in Europe as well as a clonal population structure suggests that the pathogen has been introduced to these locations (Rizzo & Garbelotto, 2003; Ivors *et al.*, 2006; Vercauteren *et al.*, 2010). In addition, it would be expected that both mating types would be found in high frequencies in a pathogenôs native range, which is not the case for *P. ramorum*, even though data suggests that these lineages were from a sexually reproducing parent population.

Disease management

The impacts of wide scale forest disease epidemics can be quite severe both economically and ecologically. Economic impacts can include costs from activities such as monitoring pathogen spread, enacting quarantine regulations and eradicating disease outbreaks. Income is lost due to the loss of marketable species or restrictions imposed on traded commodities. *Phytophthora ramorum* has had wide ranging impacts at many levels (Grünwald *et al.*, 2012). Economic losses

have had a substantial effect on the nursery industry where multiple important species are hosts for *P. ramorum*. Losses result from the destruction of large numbers of infected plants, to the loss of trade opportunities due to quarantine regulations. In some nurseries, these losses have resulted in bankruptcy (Grünwald *et al.*, 2012). Control efforts can cost governments millions of dollars. For example, the cost of eradication efforts for *P. ramorum* in Curry County, Oregon state between 2001 and 2008 was approximately 4.3 million US dollars (Kanaskie *et al.*, 2010).

Early detection and eradication

Avoiding introduction and early eradication of pathogens such as *P. ramorum* are two important management options but are dependent on early detection. Avoiding introduction of harmful pathogens can be far less costly than controlling an established pathogen (Mack *et al.*, 2000). This has been attempted for various pathogens through quarantines, trade restrictions, and phytosanitary treatment requirements. If outbreaks are small and isolated then eradication is an option. However once a pathogen becomes established and widespread, eradication becomes difficult and costly (Mack *et al.*, 2000).

Effective quarantine measures rely on accurate identification of pathogens, which can be problematic for pathogens that are difficult to identify or isolate. Genomic data can be used to create powerful tools to aid in identifying, screening for, and understanding the biology and epidemiology of important forest pathogens. Genomic data has the advantage of using data from the whole genome of an organism rather than small fragments of DNA and can therefore be used to create more efficient, specific and sensitive detection tools. Part of developing effective tools

will be to understand the target species its close relatives in order to avoid detection methods that give false positives. For example molecular tools were developed for the detection of *P. lateralis* (Winton & Hansen, 2001) prior to the discovery of *P. ramorum*. These tools showed a positive detection signal for both species and were not specific to *P. lateralis* (Sutton *et al.*, 2008). A more complete understanding of *Phytopthora* genomics and the species in the genus will help to avoid these situations in the future. Understanding natural communities and the species present is also important for detection and monitoring for invasive pathogens. It would be ineffective to quarantine against pathogens that are natural in an ecosystem.

Controlling spread into natural environments

It is important to determine if anthropogenic activities allow the movement of non-indigenous pathogens, virulent strains or hybrid strains into the natural environment. And if they do, the particular activities that pose the greater risks should be identified. For example if pathogens are transferred into natural systems through irrigation water, better practices or control measures can be developed to treat irrigation water before it runs into natural systems. Similarly gaining a better understanding of conditions or activities conducive to pathogen evolution may enable the development of practices to reduce evolutionary opportunities. This may include simple tactics such as appropriate methods of dealing with diseased plant waste. In other cases, the knowledge may assist in targeting eradication programs to deal with the highest risk localities. If hybridization is a likely scenario between *Phytophthora* pathogens then it may be an option to direct eradication efforts to populations likely to hybridize, for instance, different lineages of *P. ramorum* or *P. ramorum* and *P. lateralis*.

Management of established invasive pathogens

If quarantine and eradication measures fail and invasive *Phytophthora*s become established in the environment then management efforts may need to be focused on controlling spread rather trying to halt spread. Spread of a disease can be slowed by limiting vehicle and recreational traffic in an area and by practicing sanitary procedures when leaving areas infected with a pathogen. These procedures have been used in Australia to prevent the spread of *P. cinnamomi* (Hardy *et al.*, 2001), in Europe to prevent the spread of *P. ramorum* (Forestry Commision Great Britian, 2012) and in the US to prevent the spread of *P. lateralis* (Hansen, 2008). Reducing inoculum sources and spread has also been attempted by removing infected host plants for example Port Orford cedar in the range of *P. lateralis* in Oregon (Hansen, 2008), and understory Rhododendron as well as whole Japanese larch stands to reduce *P. ramorum* spread in Europe (Forestry Commision Great Britian, 2012). In areas where the pathogen is present or has been eradicated resistant species can be planted in order to avoid new outbreaks and host species should be planted in areas less prone to disease or where spread is less likely (Hansen, 2008). Fungicides can be applied in nursery or agricultural settings but is not always economically feasible or allowed in many countries to control disease in natural forests. Fungicides have been used to control *P. cinnamomi* in Australia (Hardy *et al.*, 2001), and in California some oak trees on private lands are individually injected with a fungicide to control *P. ramorum* (Parke & Lucas, 2008); however this is not a viable option for controlling landscape level disease epidemics.

Objectives of study

The overall objective of this current study was to improve our understanding of *Phytophthora* and of oomycete communities in general, and to assess the influence of anthropogenic activities on the evolution of these communities and the emergence of pathogenic species. In chapter two DNA meta-barcoding was used to compare the oomycete communities under various anthropogenic influences, from urban to natural environments. The hypothesis was that oomycete communities in urban areas would have a higher diversity and more introduced species than communities from natural environments due to influences from horticulture, gardening and agriculture.

In chapter three traditional culture based methods were used to further sample and characterize the *Phytophthora* community under various anthropogenic influences. *Phytophthora* communities from residential and agricultural areas were also compared. The hypothesis was that *Phytophthora* communities in urban areas are more diverse, have more introduced species, and have more hybridization events among species. It was also hypothesized that communities under different anthropogenic influence would show a different species profile.

In chapter four we focused on the evolution of an invasive alien species with a known record of anthropogenic dispersal, *P. ramorum*, by studying the variation at the genome level within and between the clonal lineages, following their initial introduction and subsequent invasion. Specifically, we wanted to test the hypothesis that *P. ramorum* populations are highly adaptable

due to a rapidly evolving genome, and that populations would show a genomic signature of rapid evolution.

Chapter 2: Anthropogenic activities increase diversity of oomycete communities in urban and urban/natural interface environments and are pathway for potential new pathogens to natural ecosystems

Introduction

Anthropogenic activities are an important driver of the distribution and diversity of plant pathogens. One of the largest impacts is from international trade which has facilitated invasion of alien species into new regions, causing both widespread ecosystem damage and huge economic losses. There is a direct correlation between a country's gross domestic product (GDP) and the number of invasive plant pathogens introduced (Desprez-Loustau *et al.*, 2010). Although some of this relationship is likely due to more research and more reporting of plant pathogens in countries with more resources (Scott *et al.*, 2013; Hulbert *et al.*, 2017), these same countries are also likely involved in more trade. The global plant trade is one of the major pathways for movement of exotic pathogens into new territories (Brasier, 2008). Climate change could compound these problems by creating favorable conditions or by increasing the range and distribution of invasive species. Both global trade and climate change have been implicated in forest health changes (Pautasso *et al.*, 2010, 2015; Ramsfield *et al.*, 2016). Horticultural and agricultural activities could add to these problems by facilitating the emergence of pathogens and providing a bridge to natural ecosystems. Urban trees may provide an opportunity for establishment and subsequent dispersal for introduced pests and pathogens (Paap *et al.*, 2017). Agricultural ecosystems are less prone to environmental fluctuations, are more homogenous, and

undergo processes such as soil mixing, irrigation and fertilization which are not found in natural ecosystems and may all influence pathogen populations (Stukenbrock & McDonald, 2008).

The oomycetes are a group of eukaryotic microorganisms that resemble fungi and share a similar mycelial growth habit and infection strategies (Latijnhouwers *et al.*, 2003) but belong to a different major subgroup, the Heterokonts or Stramenopiles (Baldauf, 2003). They can be found in plant foliage and forest canopies, in soil and in aquatic habitats. Sporangia can be dispersed through wind or water, and can either directly infect host tissue, or they can produce motile zoospores which are released in the presence of water and swim to host material (Tyler, 2002). Two orders within the oomycetes, Peronosporales and Pythiales, contain genera comprising some of the worldôs most important plant pathogens: *Phytophthora* de Bary and *Pythium* Pringsh. Species in the genus *Phytophthora* are generally considered to be pathogens of plants and several species infect economically valuable tree species, both angiosperm and gymnosperm, and important agricultural species. Species in the genus *Pythium* include pathogens and saprotrophs that are widely distributed, diverse and are found in soil and water (Martin & Loper, 1999). Many *Pythium* species are also important crop pathogens and can cause extensive losses in agriculture.

Oomycetes are particularly prone to dissemination via the plant-for-planting trade since they can be propagated in the soil as resting spores or within plant tissues in an endophytic stage. DNA meta-barcoding of *Phytophthora* species in soil of nursery plants highlighted the role of nurseries and plant trade in the spread of *Phytophthora*: 25 *Phytophthora* phylotypes were found including two pathogens of conifers with a history of propagation outside nurseries: *Phytophthora*

ramorum, the causal agent of sudden oak death and *Phytophthora lateralis* the causal agent of Port Orford cedar root rot (Prigigallo *et al.*, 2016). Molecular detection of *Phytophthora* using qPCR on the soil and foliage of nursery plants also revealed a high frequency of *Phytophthora*: the soil of 87% of the plants tested positive and 70% of the asymptomatic plants tested positive (Migliorini *et al.*, 2015). In addition, a large proportion of first reports of *Phytophthora* species are associated with nurseries and urban areas also highlighting the importance of these activities in introducing and spreading these pathogens (Hulbert *et al.*, 2017).

Additionally, oomycetes may be particularly affected by agricultural activities. Monoculture crops, the introduction of naive host species, and irrigation may facilitate spread or increase abundance in crop pathogens which could result in community changes and more frequent epidemics. Landscape structure can affect invasion success in disturbed habitats because of increased resource or habitat availability (With, 2002). Large, spatially connected habitats may increase the potential for successful spread (With, 2002). Agricultural crops provide an ideal habitat for invasive species as there is an abundance of resources and opportunity for spread via extensive corridors of connected host. In addition, oomycetes benefit from agricultural systems because irrigation facilitates the spread of zoospores which require water for dispersal and germination. *Phytophthora* and *Pythium* species are frequently associated with irrigation water, and several species of both genera have been implicated in disease outbreaks associated with infected water (Hong & Moorman, 2005).

Little is known about the diversity, abundance or ecological roles in native habitats of *Phytophthora* species or of the other genera within the oomycetes. Recent studies suggest that a

diverse community of *Phytophthora* species are present in natural ecosystems (Hansen *et al.*, 2012a; Burgess *et al.*, 2017b). One challenge with studies of *Phytophthora* diversity in natural ecosystems is that we are lacking baseline data on species presence and abundance. Surveys that attempt to describe species diversity often discover previously unknown or undescribed species. Many of the species discovered do not cause noticeable disease or impact to the ecosystem and appear to occupy a saprophytic niche in their natural habitats (Hansen *et al.*, 2012a). Without knowledge or data on the native communities it is difficult to determine if these species are introduced, if they can cause potential damage, or if they are endemic and native but have not been described yet.

Few studies have been conducted using unbiased methods to compare diversity and distribution of oomycetes in environments with varying levels of anthropogenic influence. The objective of this study was to conduct a DNA meta-barcoding analysis to test the hypothesis that anthropogenically-impacted environments (i.e. urban centres, residential or agricultural lands) comprise a more diverse oomycete community and more non-native species than natural environments (i.e. forests) due to activities such as agriculture, exotic plant movement, and gardening that may facilitate the introduction of new species and/or new hosts to these environments. We also hypothesized that interface environments (i.e. natural forests or parks within or bordering urban centers) would be impacted by these activities, but to a lesser extent. In order to test these hypotheses, we designed our sampling to target three environment types: 1) urban areas in close proximity to nurseries, residential, agricultural areas and a park with extensive horticultural activities, 2) natural forests with little to no urban influence, and 3) urban/natural interface areas defined as natural environments/parks in urban settings. We used

natural environments with minimal urban influence to establish a baseline of the expected natural oomycete communities in southwest B.C. Urban areas were chosen around areas with activities that were expected to influence the oomycete community. Since many oomycete species are known or suspected to have been introduced through horticultural material and the nursery trade, and many others are associated with agriculture, it was hypothesized that signatures of anthropogenic influence would most likely be detected in areas geographically close to these activities. Interface areas provided a means to determine if migration was occurring. Since they were not directly associated with activities suspected to influence the community, migration would be needed to produce similar signatures of anthropogenic influence as those found in urban areas. We characterized and compared oomycete soil communities between the three site types using DNA meta-barcoding of the internal transcribed spacer one (ITS1). We chose to use soil since it likely reflects the most complete oomycete community encompassing both soil and foliar or canopy species due to fallen plant detritus, and dormant spores that remain protected over long periods of time.

Methods

Field sampling

Sampling was done in the south west coastal area of British Columbia, Canada in the vicinity of Greater Metropolitan Vancouver and southeast Vancouver Island near Victoria (Figure 2.1, Table 2.1). Four sites were sampled in each of the three environment types (urban, interface and natural): two in Vancouver and two in Victoria (Figure 2.2). These were further divided into the two dominant biogeoclimatic subzones in the area according to the biogeoclimatic ecosystem

classification system (Pojar *et al.*, 1987): the coastal western hemlock very dry maritime subzone (CWHxm), and the coastal Douglas-fir moist maritime subzone (CDFmm). In addition, a fifth urban site in a highly agricultural area was sampled in the coastal western hemlock dry maritime subzone (CWHdm) (Figure 2.1). Each of the 13 total sites was given a two letter site code; full site names, two letter codes and details are found Table 2.1.

Sampling was conducted once in the fall of 2012 and a second time in spring of 2013 to obtain a more complete community profile and cover possible differences between seasons. At each site (or geographic region), ten soil samples (approximate volume 250-500 mls) were collected. Loose debris such as moss, leaves and twigs were removed and the top layer of soil was sampled with a trowel up to an approximate depth of 20 cm and stored in labeled Ziploc bags and brought to the lab for processing. The trowel was cleaned with water and a clean paper towel between each sample. Soil samples were frozen at -20 ºC until DNA extraction and DNA meta-barcoding for community analysis (below).

DNA meta-barcoding of *Phytophthora* communities

Soil samples were mixed and one gram of soil was weighed from each sample and placed into one or two extraction tubes, depending on volume. DNA was extracted from soil samples using the FastDNA Spin Kit for Soil (MP Biomedicals, USA) following the manufacturerôs protocols. For samples that were split into two tubes, DNA from the two extractions were combined into a single sample. Samples were quantified using an Invitrogen Qubit 2.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and diluted to 5 ng/ul. The ITS1 region between

the small subunit and the 5.8 subunit of the nuclear ribosomal RNA gene (rDNA) repeats was amplified using the primers ITS6 and ITS7 (Cooke *et al.,* 2000). The ITS6 primer was amended on the 5ô end with Titanium adapter A (Roche Switzerland) and a unique multiplex identifier tag (MID (a unique 10 base fragment; Roche Switzerland)) for each sample and the ITS7 primer was amended with the Titanium adapter B (Roche Switzerland) on the 5ô end. Each sample was amplified in three duplicate 25 µl PCR reactions containing 0.5 units of Invitrogen Platinum *Taq* DNA polymerase (ThermoFisher Scientific, Waltham, MA, USA), 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM Invitrogen dNTPs (ThermoFisher Scientific, Waltham, MA, USA), 1 µM of each primer, and 0.3 µg/µl bovine serum albumin (New England Biolabs, Ipswich, MA, USA) and 10 ng of DNA. Reactions were performed with an initial heating step at 94 °C for three minutes, followed by 35 cycles of a denaturation step at 94 °C for 30 seconds, an annealing step at 50 °C for 30 seconds, an elongation step at 72 °C for one minute. The reaction was finished with a 10 minute extension step. Amplicons were visualized on a 1% agarose gel stained with Ethidium Bromide. Samples were repeated if they failed and for samples that failed repeatedly, DNA concentrations of 7.5 and 5 ng/µl were tried.

Amplicons from the three duplicate reactions were combined and purified in two steps using the Agencourt Ampure XP magnetic PCR clean-up system (Beckman Coulter, Brea, CA, USA). Primer-dimers were eliminated using a 1.8:1 volume ratio of magnetic beads to amplicons and fragments smaller than 150-200 bp were eliminated with a 0.6:1 volume ratio. Samples were eluded in 26 µL EB buffer (Qiagen) and concentrations were measured using the Quant-iT Picogreen dsDNA assay kit (Invitrogen, Eugene, OR, USA). Samples were divided into five groups for the fall samples, with each group containing from 25 to 30 samples with a unique

multiplex identifier tag. The spring samples were divided into three groups each containing from 47 to 52 samples with a unique multiplex identifier tag. Amplicon libraries composed of samples in equimolar proportions were constructed and amplicon lengths were verified using an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA). Unidirectional sequencing was done on a Roche GS FLX+ system (Roche-454 Life Sciences, Branford, CT, USA) at McGill University and Genome Quebec Innovation Centre (Montreal, QC, Canada). Each library from the fall was sequenced on a quarter plate using the Titanium chemistry and each library from the spring was sequenced on half a plate using the FLX+ chemistry.

Data filtering and OTU assignment

Sequence data was analyzed using the QIIME software package (Caporaso *et al.*, 2010). Barcode sequences were used to assign reads to samples, and barcode and linker primer sequences were trimmed. Sequences were filtered for a minimum quality score of 25 using a 50bp sliding window, maximum homopolymer length of eight, maximum primer mismatch of two bases, and a maximum number of barcode errors of one. Data from the spring was processed and filtered as above; however due to erroneous additional bases found between the barcode and linker primer sequences, many sequences were filtered out. To fix this problem, prior to barcode assignment to samples, a custom Python script (Python Software Foundation. Python Language Reference, version 2.7. http://www.python.org) was developed to locate barcode sequences and linker primer sequences and eliminate any additional bases between the two. After filtering, the ITS1 sequence was extracted from all sequences (fall and spring) using the ITSx software which also screens for chimeric sequences (Bengtsson-Palme *et al.*, 2013). The resulting sequences were

filtered for a minimum length of 160bp. To choose a minimum sequence length, sequences of *Phytophthora* were downloaded from the *Phytophthora* database (Park *et al.*, 2013) and the ITS1 sequence was extracted using the ITSx software (Bengtsson-Palme *et al.*, 2013). The minimum lengths of the known *Phytophthora* sequences were used to choose a minimum sequence length threshold.

Open reference operational taxonomic unit (OTU) picking was done using the QIIME software (Caporaso *et al.*, 2010) with the UCLUST algorithm (Edgar, 2010) for sequence clustering at 97% and assignment of a representative sequence for each cluster. For taxonomic assignment, a database was created by downloading all Stramenopiles ITS sequences from NCBI. Additional ITS sequences of *Phytophthora* were downloaded from the *Phytophthora* database (Park *et al.*, 2013) and added to the NCBI sequences. The ITS1 sequence was extracted using the ITSx software (Bengtsson-Palme *et al.*, 2013). The resulting database was formatted to the specifications required in the QIIME software (Caporaso *et al.*, 2010). A corresponding taxonomy file was created which contained the hierarchical taxonomic information for each of the sequences in the database file. For use with the Ribosomal Database Project (RDP) classifier 2.2 (Wang *et al.*, 2007), the taxonomy file was created with six levels: 1) superphylum (=stramenopiles), 2) class, 3) order, 4) family, 5) genus and 6) species. Taxonomy was assigned to the OTUs with the RDP classifier 2.2 (Wang *et al.*, 2007) implemented in the QIIME package (Caporaso *et al.*, 2010), using the custom database as a reference and a confidence level cut-off of 0.5. Sequences with less than two copies (singletons), and samples with less than 100 total sequence reads were filtered out for downstream analysis. Sequences that could not be classified as oomycetes were eliminated for further analyses.

To extract the complete set of *Phytophthora* species as well as possible new species, OTU picking was done as above but with a similarity of 99% for the clustering step. Taxonomy was assigned to the representative OTUs (output file from QIIME where a single sequence represents a cluster of sequences in the full dataset) in two ways: 1) using the QIIME pipeline as above and 2) by searching each representative sequence from the OTU picking step against the NCBI nucleotide database using the BLASTn algorithm (Altschul *et al.*, 1990). Results from both methods, and the spring and fall datasets were combined and sequences identified as *Phytophthora* were used to create a final set of representative *Phytophthora* sequences. The final set of representative *Phytophthora* sequences were crossed with the full data set to extract all sequences clustered with the representative *Phytophthora* sequences using a custom Python script (Python Software Foundation. Python Language Reference, version 2.7. http://www.python.org). The full QIIME pipeline was run again on the full dataset from both spring and fall. Clustering thresholds of 98%, 98.5%, and 99% were tested to find a threshold that reduced the resulting number of clusters with the same taxonomic assignments while maximizing those with different assignments. A final value of 98.5% was used. Final taxonomy assignments were checked manually; BLAST results were checked for alignment coverage and percent similarity. OTUs assigned to a species complex with closely related ITS sequences or with ambiguous results were aligned with known representatives of the complex or closely related species downloaded from the *Phytophthora* database (Park *et al.*, 2013) or NCBI and searched for fixed nucleotide differences. OTUs were assigned to the best match based on fixed differences if possible; otherwise they were identified as a member of a species complex. OTUs with no close species matches were assigned to one of the ten previously defined *Phytophthora*

clades (as defined in (Cooke *et al.*, 2000) and Blair *et al.* (2008)) or just as unknown *Phytophthora*s. OTUs with less than two sequence reads in a sample were filtered out of that sample to retain a more conservative *Phytophthora* dataset.

For sequences with a best match to a *Phytophthora* species but low sequence similarity, sequence similarities were calculated between aligned ITS1 sequences of the unknown OTUs and the ITS1 of sequences from the *Phytophthora* database (Park *et al.*, 2013) and sequences of groups closely related to *Phytophthora* including *Halophytophthora*, *Salisapilia*, *Phytopythium* and *Pythium*. A multiple alignment was done using ClustalW (Thompson *et al.*, 1994) in the BioEdit Sequence Alignment Editor (Hall, 1999), and a similarity matrix was calculated using the Biological Structure Analysis (bio3d) package (Grant *et al.*, 2006) in the R statistical environment (R Core Team, 2017). Sequences were included as *Phytophthora* if the sequence similarities were within sequence similarities found between members of the *Phytophthora* genus, and if the sequence similarities were closer to members of the *Phytophthora* genus than members of other closely related groups. Read coverage of the unknown sequences was determined, and only sequences with a minimum coverage of five reads were included. Unknown sequences with over 99% identity to another unknown sequence were grouped together for read counts. In addition, the lengths of the ITS1 for the unknown sequences were compared to the length of the ITS1 for species found in the *Phytophthora* database (Park *et al.*, 2013). For the phylogenetic analysis the full ITS of the representative Phytophthora and other closely related species (above) were aligned with MAFFT (Katoh & Standley, 2013) and a maximum likelihood phylogeny was reconstructed using default settings in the Phylogeny.fr pipeline (Dereeper *et al.*, 2008) with PhyML 3.0 (Guindon *et al.*, 2010).

Data analysis

Sequence abundance for each genera and OTU was calculated using QIIME (Caporaso *et al.*, 2010). The number of OTUs and the number of samples each OTU was found in was calculated and compared for each major genus of the oomycetes for both fall and spring, and for each site type (urban, interface and natural). The number of *Phytophthora* positive samples from the manually curated *Phytophthora* dataset were calculated and compared between site types using a chi-square test. The number of OTUs as well as the number of *Phytophthora* and *Pythium* OTUs (97% sequence similarity) were compared between site types (urban, interface and natural) with a chi-square test. For comparisons of OTUs between site types, OTUs were filtered for a minimum of two reads per sample.

A canonical correspondence analysis (CCA) was done using the vegan package in R (Oksanen *et al.*, 2016) to determine if there was an impact from site type on the OTU profile.

Samples were rarefied to 300 sequences and 500 sequences and number of OTUs and Shannonôs diversity index were calculated for each site and site type using QIIME (Caporaso *et al.*, 2010). The difference between site type was compared using an analysis of variance (ANOVA).

Results

***Pythium* and *Phytophthora* represent the most dominant genera in Oomycete soil communities.** There were 291064 and 246641 ITS1 sequences after trimming and filtering the fall and spring datasets, respectively. A total of 266597 (91.6%) and 208128 (84.4%) of those sequences were identified as oomycetes by sequence homology in fall and spring, respectively. The percentage of sequences filtered at each step was similar in the datasets for the two sampling times (Appendix A, table S2.7). The number of Oomycete sequence reads ranged from 0 to 7273 per sample (averaging 2169 in fall and 1737 in spring; Tables 2.2 and 2.3). Within the oomycetes, *Pythium* was the genus represented by the highest number of sequences (79.2% in fall and 87.5% in spring; Figure 2.3A). *Apodachlya*, *Saprolegnia* and *Phytophthora* were less abundant, accounting for less than 5% each of the total Oomycete sequences (Figure 2.3A). However, even though *Pythium* sequences dominated the read counts, there was a higher OTU to read ratio for *Phytophthora* than for *Pythium*, and *Phytophthora* was represented by the second highest number of OTUs (Figure 2.3B). Out of 136 Oomycete OTUs in the fall and 112 in the spring sampling, 53 and 39 were *Pythium*, respectively, whereas 22 and 19 were *Phytophthora* OTUs. Even though *Phytophthora* sequences only accounted for 1.4% of the total reads in fall and 0.6% in the spring, they accounted for 16% of the total OTUs in the fall and 17% in the spring. Other oomycetes such as *Saprolegnia* and *Aphanomycetes* accounted for 10% or less of the OTUs for both sampling times (Figure 2.3B).

The Oomycete community was dominated by two *Pythium* species. The OTUs represented by the most sequences were identified as *Pythium attrantheridium* (55.5% of the Oomycete

sequences in fall and 69.6% in spring), followed by *Pythium monospermum* (17.1% of the Oomycete sequences in the fall and 6.3% of the sequences in the spring). These were also two of the most widespread species, found in 87% of the total samples (Figure 2.4). In contrast, most OTUs were not widespread (Figure 2.4); only 3% of the OTUs were found in more than 50% of the total samples. The majority of OTUs were rare with 78% of the OTUs found in less than 5% of the samples.

Anthropogenic activities influence Oomycete species composition and diversity. The site PR was significantly different in preliminary analyses (data not shown) and was removed from all analyses comparing site type for this chapter (data was retained for comparisons in chapter 3) to minimize any effects from having one extra urban site, or from one urban site in a biogeoclimatic zone (CWHdm) where there were no matching interface or natural sites. Site type (urban, interface or natural) had a significant influence on species composition in the canonical correspondence analysis ($F = 0.1906$, $p = 0.001$) (Figure 2.5). Urban sites were characterized by the highest number of genera (Figure 2.6), OTUs, OTUs that were exclusive to that site type (Figure 2.7) and low frequency OTUs (Figure 2.8). By contrast, natural sites had the fewest genera, OTUs, low frequency OTUs, and exclusive OTUs. Interface sites were closer to urban sites for these values. The difference in total number of OTUs between groups was significant (chi-square = 10.109, $p = 0.006$). There were more shared OTUs between urban and interface sites than between either and natural sites (Figure 2.7). This pattern was also observed within the two most abundant genera, *Pythium* and *Phytophthora*. There were 16, 20, and 27 *Phytophthora* OTUs in the natural, interface and urban groups, and 20, 35, and 45 *Pythium* OTUs in the natural, interface and urban groups respectively (Figures 2.8A, 2.8B). The difference in the

number of Pythium OTUs found in each site type was significant (chi-square = 6.969, p = 0.031).

Urban sites had the highest taxonomic diversity in both fall and spring. When samples were rarefied to the same number of sequences per sample (300), Shannonôs diversity and the number of OTUs were significantly higher ($p<0.05$) in urban than interface and natural sites (Figure 2.9; Tables 2.2, 2.3 and 2.4). When interface was removed from the comparison, urban sites were always significantly higher than natural sites (Table 2.4). Diversity was slightly lower in the spring than in the fall. Similar results were found with a sampling depth (rarefaction) of 500 (Tables 2.2 and 2.3). Shannonôs diversity and

***Phytophthora* are more frequent in urban than natural sites.** Forty percent of the samples were positive for *Phytophthora* after combining the manually curated fall and spring samples. A higher proportion of the urban samples were positive for *Phytophthora* (51%) compared to interface (41%) or natural samples (28%); however the difference was not significant. The most widespread *Phytophthora* OTUs were *P. syringae*, and *P.* taxon Kelmania found in 14 and 10 samples respectively. *Phytophthora syringae* was found only in interface and urban samples, whereas *P.* taxon Kelmania was found in all site types, however only once in natural samples. The most widespread OTUs in natural sites were *P. pseudotsugae*, *P. chlamydospora* and *P. cryptogea*, which were only found in four samples each. Almost all *Phytophthora* OTUs were rare. All except *P. syringae*, and *P.* taxon Kelmania were found in less than five percent of the samples; *P. syringae* and *P.* taxon Kelmania were found in 6% of the samples. Over half of the OTUs were found in 1% or less of the total samples.

New *Phytophthora* species are more frequent in urban areas. Three putative new *Phytophthora* species were found in urban or in urban/interface samples; none were found in natural sites. These were found in the manual curation of the *Phytophthora* dataset where there were eighteen OTUs whose representative sequences did not have a high homology match in the NCBI database (less than 95% identity; Table 2.5) but had a closest match to a *Phytophthora* species. The percent identity between the sequences for eight of these OTUs and *Phytophthora* taxa were within the range found within the genus (54.3% to 100% similarity in the ITS1) (Table 2.5). Based on sequence homology, five of these OTUs were clustered into two taxa, unknown sp 1 and unknown sp 2, and another single OTU made the third taxa, unknown sp 3 (Table 2.5). The remaining two were rejected due to low read coverage. The species with the closest homology to unknown sp 1 were two uncultured *Phytophthora* clones (99% and 100% identity). It shared 83% sequence similarity with several unrelated species including *Phytophthora clandestine*, *Phytophthora cactorum, Phytophthora asparagi* and *Peronospora sparsa*. The ITS 1 was not informative enough to determine if this sequence was a *Phytophthora* species: in the phylogenetic analysis it groups with *Peronospora* sequences, within two major clades of *Phytophthora* and is closest to clade five, (Figure 2.10). Unknown sp 1 was found in three sites (SV, MD and HM) in both the fall and the spring sampling. The species with the closest homology to Unknown sp 2 is *Phytophthora boehmeriae*, (only 71% sequence identity). In the phylogenetic analysis it groups with clade nine (Figure 2.10). It was found in three sites (PS, PR and SV) in the fall sampling. Unknown sp 3 is closest to *Phytophthora gallica* (79%), and groups with clade 10 in the phylogenetic analysis, (Figure 2.10). It was found in four sites in spring (CV, PS, MD and SV). Further work using the both ITS 1 and 2 would be required to confirm

that these species are *Phytophthora*, if they could represent new clades, or if they could represent new genera.

The remaining ten unknown OTUs were rejected as potential *Phytophthora* species because the percent identity between these sequences and known members of the *Phytophthora* genus was lower than what was found between members of the genus (Table 2.5). In addition, the length of the ITS1 in these OTUs was much longer than what was found in *Phytophthora* (*Phytophthora* average ITS1 length = 209, min = 114, max = 248).

Manual curation of the *Phytophthora* dataset resulted in a total of 34 OTUs versus 36 from the automated taxonomic identification. Sixty-four percent of the OTUs identified to the species level through the automated pipeline were identified to the same species with the manual identification (Appendix A, Table S2.8).

Oomycete communities fluctuate with season. There was a larger and more diverse Oomycete community in the fall than in the spring. Eighty-four percent of the total OTUs were recovered in the fall sampling whereas 69% were recovered in the spring sampling period (Table 2.6). Fifty-three percent of the OTUs were recovered in both sampling periods. There was a high turnover in OTUs between seasons suggested by 31% and 16% of species recovered only in the fall and spring, respectively. This high turnover was also observed in the *Phytophthora* community; out of a total of 37 OTUs, only 35% were recovered in both seasons, whereas 35% and 27% were recovered in fall and spring only, respectively. There was a reduction in total *Pythium* OTUs in

the spring; out of a total of 59 OTUs, 56% were found in both seasons, 34% were found in the fall and only 10% were exclusive to the spring sampling period (Table 2.6).

Discussion

The anthropocene (a newly defined epoch in which human impacts are as important as natural processes) is changing the way we view species range and distributions. Globalization is resulting in human-mediated dispersal that is causing a breakdown of the normal biogeographic barriers (Capinha *et al.*, 2015). This may result in a biotic homogenization across widespread similar environments which are socially and economically linked (Capinha *et al.*, 2015). The impact of anthropogenic activities on plant pathogen communities is still poorly understood. For some prominent species, global distributions and genetic structure follow patterns for human-mediated dispersal rather than natural dispersal (Prospero *et al.*, 2009; Goss *et al.*, 2011b; Dutech *et al.*, 2012; Grünwald *et al.*, 2012) (Brasier 2001); however anthropogenic impacts are less documented at a community level for microorganisms. We demonstrate that anthropogenic activities impact Oomycete communities by increasing overall species diversity, the number and general presence of species in the plant pathogenic genus *Phytophthora* and the overall abundance of individual taxa. Furthermore, urban/natural interface environments likely act as a bridge for invasion into natural environments. This could impact both the natural biota and natural ecosystem processes. *Phytophthora*s introduced from nursery trade or spread by human movement have caused landscape level outbreaks in California, Australia, New Zealand and Chile (Wills, 1993; Garkaklis *et al.*, 2004; Durán *et al.*, 2008; Grünwald *et al.*, 2012; Hansen *et al.*, 2012a).

One likely explanation for our results is the migration of species into anthropogenically impacted ecosystems. Anthropogenic activities can influence the diversity of microorganisms through several mechanisms. Global trade, total population and population density have all been associated with patterns of invasion (Desprez-Loustau *et al.*, 2010; Pautasso *et al.*, 2010, 2015; Santini *et al.*, 2013). Environmental disturbance, landscaping, forestry, and increased movement through dense road networks are typical activities associated with increased urbanization which may contribute to increased incidences of invasive species (Lake & Leishman, 2004; Alston & Richardson, 2006; Desprez-Loustau *et al.*, 2010). Increase in diversity and abundance of exotic plant species has been linked to disturbances, in particular those that are accompanied by an increase in nutrient rich water (Lake & Leishman, 2004). Patterns of invasion in plant pathogens likely mimic those of invasive plants. Increases in potential hosts, or stress to native plant hosts could directly influence pathogen populations in invaded environments. Invasive plant species may also cultivate different pathogen communities which was suggested as a cause for the different Oomycete communities found around invasive and native reed populations (Nelson & Karp, 2013). Furthermore, microorganisms and plant pathogens associated with invasive plants, ornamental plants, species used in reforestation or other forest management activities can be carried into new habitats. High frequencies and a high diversity of *Phytophthora* species have been found in the soil and roots of nursery plants (Prigigallo *et al.*, 2015; Jung *et al.*, 2016); furthermore, a high frequency of first reports of *Phytophthora* species are associated with nursery plants highlighting the risks associated with horticulture activities (Hulbert *et al.*, 2017). Higher rates of *Phytophthora* infestation in planted forests versus natural forests or seed generated forests, and a high frequency of Phytophthoras associated with planting stock

originating from nurseries all suggest that this is a very important pathway in the movement of plant pathogens (Jung *et al.*, 2016).

Global trade results in the accidental introduction of organisms that hitchhike on traded goods and many crop and forest pathogens are introduced through the plant trade (Brasier, 2008; Burgess et al. 2016) . Invasive forest pathogens are often first introduced into agricultural or urban areas and then spread to natural forests (Paap *et al.*, 2017). Our study clearly shows a pattern indicative of movement from urban to natural environments. There were ten times more shared species between urban and interface than between interface and natural environments. A similar pattern of decreasing diversity and frequency in *Phytophthora* populations going from urban to natural environments was found in Sweden (Redondo *et al.*, 2018). Urban/natural interface environments are subject to many influences from urban centres such as increased human access, higher levels of disturbance, soil enrichment and introductions of alien species through horticulture (Alston & Richardson, 2006). Similar to the pattern shown here with oomycetes, interface environments reflect a pattern of increased diversity in alien plant species which is more pronounced with increasing proximity to urban areas (Alston & Richardson, 2006). The proximity of urban/interface areas to large sources of propagules makes them a sink for invasive species. This increased reservoir of inoculum could act as a bridge for invasive species into natural environments. *Phytophthora ramorum* is an invasive pathogen that has been introduced through the nursery trade and subsequently migrated into adjacent natural forests where it causes sudden oak death in the United States, and sudden larch death in the UK (Ivors *et al.*, 2004; Prospero *et al.*, 2007; Mascheretti *et al.*, 2008; Brasier & Webber, 2010; Grünwald *et al.*, 2012). Similarly, *Phytophthora cinnamomi*, an important agricultural and horticultural

pathogen, escaped from nursery material and spread to natural systems causing Jarrah dieback and ecosystem wide damage in Australia (Weste & Marks, 1987; Wills, 1993; Garkaklis *et al.*, 2004; Hardham, 2005; Cahill *et al.*, 2008).

Much of our understanding of pathogenic species in the oomycetes is focused on *Phytophthora* because of their long and visible history as crop pathogens. However, other groups and species in the oomycetes can also cause diseases and become invasive. The Peronosporales appear to be common invaders suggested by their abundance in invasion records (Desprez-Loustau *et al.*, 2010). Our study suggests that anthropogenic activities likely facilitate the spread of many species in this important group of microorganisms, not only *Phytophthora* species. The trends observed in this study were not only present in the *Phytophthora* genus, but appeared to be present across the entire oomycete community.

Pythium was the dominant group of oomycetes in all habitat types examined in this study. Many *Pythium* species cause disease or damping off in young seedlings and can inhibit seed germination, but are only weakly pathogenic to mature plants, or can reduce growth without killing mature plants (Schroeder *et al.*, 2013). Negative feedback on plant growth due to pathogens is hypothesized to maintain species diversity in plant communities by limiting growth in certain species, and allowing the growth of other competitors (Mills & Bever, 1998; Laliberté *et al.*, 2015; Albornoz *et al.*, 2017). This feedback was demonstrated with *Pythium* species preferentially limiting growth of *Danthonia* and *Panicum* in a plant community without causing mass mortality or obvious disease symptoms (Mills & Bever, 1998). *Pythium attrantheridium*, described in 2004, has been associated with both agricultural and natural environments in North

America (Packer & Clay, 2000; Allain-Boulé *et al.*, 2004; Broders *et al.*, 2007; Reinhart *et al.*, 2010). *Pythium attrantheridium* has been suggested to play a role in shaping the distribution of trees in the ecosystem by inhibiting the establishment of younger trees near parent trees where it is associated with root systems (Packer & Clay, 2000; Allain-Boulé *et al.*, 2004). Packer & Clay, (2000) tested the Janzen-Connell hypothesis (Janzen, 1970; Connell, 1971) that predators, in this case *Pythium* soil pathogens, maintain tree diversity by influencing tree distribution. They showed that *Pythium* species caused high mortality in seedlings close to parent trees, but had no effect on seedling survival further away from trees (Packer & Clay, 2000). *Pythium monospermum* has a ubiquitous distribution but little is known about its pathogenicity (Van der Plaats-Niterink A J, 1981). It may be pathogenic or associated with disease on some agricultural species (Robertson, 1973; Van der Plaats-Niterink A J, 1981). The high abundance of *Pythium* observed in natural or forest ecosystems in this study suggests that they may contribute to stand structure and maintenance of biodiversity in these mature ecosystems. Other studies have suggested that oomycetes play an important role in driving succession in young ecosystems but that they decline with ecosystem age (Dickie *et al.*, 2017). In our study, the higher frequency of oomycetes in interface and urban sites likely reflects the disturbance from anthropogenic activities which increase diversity and abundance. These shifting patterns in Oomycete communities could interfere with natural ecosystem processes that these pathogens are a part of which may result in increases in disease or ecosystem declines.

Similarly, some *Phytophthora* species are likely natural inhabitants in healthy ecosystems. Recent studies have revealed many *Phytophthora* species in natural environments and have led to questions regarding their presence, role and potential impact (Hansen *et al.*, 2012a; Burgess *et*

al., 2017b). A similar role to *Pythium* in the maintenance of biodiversity has been suggested and shown for native *Phytophthora* species (Albornoz *et al.*, 2017). Overall, our data show that *Pythium* was the most abundant genus in Oomycete communities in all site types including natural environments, and that all other genera including *Phytophthora* are much less abundant. This is similar to what has been found in other studies. Low *Phytophthora* diversity was found in soil samples in Spain using pyrosequencing on environmental samples (Català *et al.*, 2015). Most studies of Oomycete communities also find a much lower abundance of *Phytophthora* compared to *Pythium* or other Oomycete genera (Arcate *et al.*, 2006; Nelson & Karp, 2013; Sapkota & Nicolaisen, 2015; Riit *et al.*, 2016; Esmaeili Taheri *et al.*, 2017). However, the reverse was true in the *Pythium* and *Phytophthora* populations around declining sweet chestnut forests where *Phytophthora* was known to be present (Vannini *et al.*, 2013). Although there was a relatively low abundance of *Phytophthora* sequences recovered from the soil samples in our study, we found a relatively high species diversity suggesting a small but diverse population of *Phytophthora* likely resides in natural healthy environments. It is possible that the natural environment has been invaded by exotic species; however it is unlikely that all of the species present are exotic. Regardless of their indigenous or exotic status, the reservoir of inoculum is important because under conducive environmental conditions, host presence, or an imbalance in the ecosystem, a single species can cause devastating damage. In declining oak stands in Spain, metabarcoding of *Phytophthora* populations revealed an overall low species diversity but high frequency and dominance of a few species possibly causing disease (Català *et al.*, 2017). Devastating pathogens such as *Phytophthora cinnamomi* can survive in soil for long periods in infected plant tissues or as chlamydospores or oospores and under the right conditions, disease develops (Hardham, 2005).

Culture-independent methods such as the DNA meta-barcoding used here are revolutionizing the way we conduct microbial community studies. Community DNA meta-barcoding was a useful and successful approach for exploring the Oomycete community in soils. A high percent of the sequence reads obtained in this study were identified as Oomycete or Stramenopiles; very few reads were identified from other groups. Sapkota and Nicolaisen (2015) tested the DNA meta-barcoding approach on samples with known amounts of *Pythium* and *Phytophthora* DNA, and although the number of sequences obtained did not match the actual abundance of DNA from the two groups, both groups were always detected, suggesting that the method is reliable and that the low frequency of *Phytophthora* in soils is biological.

The internal transcribed spacer region is a useful target for describing the taxa present in the environment. It can be used to identify species or complexes of closely related species as it is fast evolving giving a higher level of interspecific variation than more conserved regions of DNA. In addition, the vast number of barcoding studies done using the ITS has increased the number of sequences available for comparison in reference databases. This allowed us to identify most OTUs to the species level. Although there was good agreement between the manual and automated taxonomic identifications in the *Phytophthora* dataset, the differences between the two datasets highlights the importance of drawing conclusions made from species level taxonomic assignments. There are several *Phytophthora* species complexes that comprise species with greater than 99% sequence homology in the ITS. Manual identification using fixed differences allowed for a higher confidence in identifications; however using only the ITS1 region could still leave room for errors in this group of organisms. Further analysis using species

specific primers or culture-dependent methods would be needed to confirm results found using DNA meta-barcoding. Many regulatory agencies would require the presence of a culture to confirm positive findings of a quarantine organism. DNA meta-barcoding would be a useful tool for monitoring large geographic areas. More intensive sampling could be targeted to locations to confirm the presence of high risk pathogens or possible disease outbreaks.

The largest disadvantage of using the ITS is that the high divergence in this region makes it unsuitable for phylogenetic based diversity analyses. Differences in diversity between sites could be underestimated since only using species richness or counts of OTUs does not take into account how related or not the species are between the sites. Using primers targeted to specific genera could also be used to get a more in depth look at any particular genus. In this study and in those using more general primers, the diversity in less dominant genera could be underestimated, in particular for rare species.

The trends found in Oomycete communities were the same in the fall and spring sampling; however sampling in two seasons gave a higher coverage of the species present. There was a high turnover of OTUs between the two sampling seasons. This is likely due to both biological reasons and sampling methods. *Phytophthora* populations are known to fluctuate over seasons (Balci *et al.*, 2007; Reeser *et al.*, 2011). Sampling in multiple seasons and conditions gives a better look at the populations. In addition, when using DNA meta-barcoding, a very small amount of soil is used to make inferences about a much larger population. It is very likely that multiple samples would show different results. A larger sampling area and number of samples increases the confidence in the results found. Our study showed the same trends across seasons

and taxonomic groups, higher diversity in urban sites, more shared species in urban and interface sites, and similar relative abundance in different genera, giving us a high confidence in the conclusions drawn.

Overall, we have found a strong signature of anthropogenic activities on Oomycete communities. Further studies on important genera, in different ecosystems, and in healthy versus diseased forests will further help to understand how anthropogenic activities contribute to disease emergence, how anthropogenic activities influence this important group of organisms, and the role of these organisms in natural environments. Our study has provided important information of the composition and distribution of Oomycete communities in southwest British Columbia and contributes to our understanding of this important group of organisms worldwide.

Table 2.1: Details of sampling areas.

Site	Location	Region	Biogeoclimatic zone	Site type	Urban Site Type*	Collection areas	Dominant vegetation in sampled areas
BW	Bowen Island	Bowan Island	CWHxm	Natural		forested areas	Douglas-fir, alder, cedar, hemlock
CR	Chemainus River/ Cowichan Valley	Vancouver Island	CWHxm	Natural		forested areas	Alder, hemlock, cedar, Douglas-fir, salal, Abies fir
GT	Gowland Tod Park, Victoria	Vancouver Island	CDFmm	Natural		forested areas	Cedar, alder, Douglas-fir, ferns
SA	Saturna Island	Gulf Islands	CDFmm	Natural		forested areas	Cedar, hemlock, alder, Douglas-fir, big leaf maple, salal
CV	Campbell Valley Park, Langley	Metro Vancouver	CDFmm	Interface		forests/meadows/ replanted areas	Alder, birch, maple, blackberry, cedar
GL	Grant lake/ Cowichan Lake, Vancouver Island	Vancouver Island	CWHxm	Interface		forested areas bordering residential areas	Alder, cedar, hemlock, Abies fir, Douglas-fir, salal
HM	Cuthbert Holmes Park and Swan Lake ,Victoria	Vancouver Island	CDFmm	Interface		forests/meadows/beside ponds/beside trails	Willows, black berry, cedar, Douglas-fir, ivy, wild rose, other shrubs, oak
PS	Pacific Spirit Park, Vancouver	Metro Vancouver	CWHxm	Interface		forested areas adjacent to trails	Cedar, hemlock, alder, Douglas-fir

Site	Location	Region	Biogeoclimatic zone	Site type	Urban Site Type*	Collection areas	Dominant vegetation in sampled areas
PR	Pitt River area, Pitt Meadows	Metro Vancouver	CWHdm	Urban	Agricultural	areas adjacent to dykes, rivers, walking trails, agricultural fields	Blackberry, grass, blueberry farms, willows, cottonwood, alder, cedar
SV	Steveston area, Richmond	Metro Vancouver	CDFmm	Urban	Agricultural	areas adjacent to dykes, agricultural fields, roads, ditches	Alder, blackberry, poplar, grass
BP	Beacon Hill Park, Victoria	Vancouver Island	CDFmm	Urban	Park	flower beds, grassy meadow areas	Rhododendron, cedar, Douglas-fir, oak, cypress, many horticultural plants, shrubs
HV	Happy Valley, Victoria	Vancouver Island	CWHxm	Urban	Residential	public areas in residential neighborhoods: trails, sides of roads, forest patches	Big leaf maple, cedar, Douglas-fir, alder, hemlock
MD	Marine Drive Area, Burnaby	Metro Vancouver	CWHxm	Urban	Residential	public areas in residential neighborhoods: forest patches, trails, sides of roads edge of golf course and recreational fields	Alder, Big leaf maple, blackberry, cedar, shrubs

*Urban sites included areas influenced by different anthropogenic activities. The dominant influence based on a visual assessment of the area is listed.

Table 2.2: Descriptive statistics for fall samples including minimum and maximum number of sequences, Shannon's diversity index and number of OTUs for sequences classified as Oomycetes.

Site	Type	N	Min reads[c]	Max reads[c]	Depth 300*			Depth 500*		
					S[a]	H[a]	N[b]	S[a]	H[a]	N[b]
CR	Natural	9	29	3149	9	1.7	8	11	1.8	5
GT		10	662	5127	11	1.6	10	13	1.6	10
BW		11	29	2928	11	1.7	9	14	1.8	9
SA		9	1096	5882	14	1.8	9	16	1.8	9
GL	Interface	9	736	3502	9	1.1	9	11	1.2	9
CV		9	485	3897	13	1.4	9	14	1.2	8
PS		7	119	2609	21	2.5	6	29	2.9	5
HM		10	1264	7273	12	1.6	10	14	1.6	10
PR	Urban	9	461	1994	25	3.1	9	30	3.1	8
HV		10	466	4445	12	1.8	10	15	2.0	9
SV		10	935	3486	21	2.3	10	25	2.4	10
BP		10	602	5027	14	1.6	10	18	1.7	10
MD		10	341	2597	18	2.2	10	20	2.2	9
Natural		39	29	5882	11	1.7	36	14	1.7	33
Interface		35	119	7273	13	1.6	34	16	1.6	32
Urban (no PR)		40	341	5027	16	2.0	40	20	2.0	38

[a] S - number of species measured as the number of OTUs at 97 clustering, H - Shannonôs diversity index, N[b] - adjusted sample size with samples less than the sampling depth excluded, [c] - minimum and maximum number of sequences after removing unclassified sequences, *sampling depth - the number of sequences randomly sampled for diversity calculations.

Table 2.3: Descriptive statistics for spring samples including minimum and maximum number of sequences, Shannon's diversity index and number of OTUs for sequences classified as Oomycetes.

Site	Type	N	Min reads[c]	Max reads[c]	Depth 300*			Depth 500*		
					S[a]	H[a]	N[b]	S[a]	H[a]	N[b]
CR	Natural	9	0	2274	6	0.5	2	7	0.5	2
GT		10	91	2547	10	1.1	9	13	1.1	8
BW		11	126	2816	11	1.5	8	14	1.4	7
SA		7	680	5435	8	0.7	7	9	0.8	7
GL	Interface	6	0	2749	10	0.5	2	13	0.5	2
CV		9	193	1827	16	1.6	8	20	1.5	7
PS		8	217	1729	12	1.9	7	15	2.0	6
HM		10	666	2569	14	1.3	10	17	1.3	10
PR	Urban	10	20	1998	14	1.9	9	15	1.8	8
HV		10	1075	3042	12	1.1	10	14	1.1	10
SV		10	801	5560	17	1.7	10	20	1.7	10
BP		10	1512	3109	18	2.2	10	22	2.2	10
MD		10	1940	5371	20	2.0	10	24	2.0	10
Natural		37	0	5435	9	1.1	26	12	1.1	24
Interface		33	0	2749	14	1.5	27	17	1.4	25
Urban (no PR)		40	801	5560	16	1.7	40	20	1.7	40

[a] S - number of species measured as the number of OTUs at 97 clustering, H - Shannonôs diversity index, N[b] - adjusted sample size with samples less than the sampling depth excluded, [c] - minimum and maximum number of sequences after removing unclassified sequences, *sampling depth - the number of sequences randomly sampled for diversity calculations.

Table 2.4: ANOVA results for diversity comparisons on samples that were rarefied to 300 sequences per sample.

Group	Dependant variable	F-statistic	p-value
Fall: natural vs interface vs urban	Number of OTUs	5.730	0.004
Fall: natural vs urban	Number of OTUs	15.055	0.000
Fall: natural vs interface vs urban	Shannonôs Diversity	3.352	0.039
Fall; natural vs urban	Shannonôs Diversity	5.127	0.026
Spring: natural vs interface vs urban	Number of OTUs	14.547	0.000
Spring: natural vs urban	Number of OTUs	28.991	0.000
Spring: natural vs interface vs urban	Shannonôs Diversity	5.084	0.008
Spring: natural vs urban	Shannonôs Diversity	10.886	0.002

Table 2.5: OTUs sharing low sequence identity to *Phytophthora* taxa. Blast search results, read coverage, and range of percent identity to members of the *Phytopthora* genus.

Representative sequence	New ID*	Species with closest BLAST match	Length (bp)	Query coverage (%)	Identity (%)	# of reads	% Identity to other *Phytophthora*
OTU2029 [a]	Unknown sp 1	*Phytophthora clandestina*	212	99	83	61	55.0-87.7
OTU1617 [ac]	Unknown sp 1	*Phytophthora clandestina*	217	99	81	31	61.6-85.9
OTU4328 [ac]	Unknown sp 1	*Phytophthora clandestina*	215	99	82	206	61.6-86.3
OTU2987 [ab]	Unknown sp 2	*Phytophthora boehmeriae*	234	100	71	2	59.2-81.4
OTU578 [ab]	Unknown sp 2	*Phytophthora boehmeriae*	234	100	71	45	59.2-80.3
OTU3647		*Phytophthora colocasiae*	215	100	73	3	62.5-80.2
OTU1517 [a]	Unknown sp 3	*Phytophthora gallica*	214	100	79	40	56.5-78.8
OTU1914		*Phytophthora intercalaris*	212	100	76	2	59.1-77.1
OTU2190		*Phytophthora sp. REB326-69*	335	96	89	33	43.4-70.8
OTU4062		*Phytophthora cactorum*	251	41	93	7	50.3-67.9
OTU2785		*Phytophthora infestans*	324	95	89	20	45.2-67.0
OTU2513		*Phytophthora asparagi*	327	92	84	2	45.3-66.7
OTU2551		*Phytophthora megasperma/ P. asparagi*	323	98	87	2	45.6-66.4
OTU1670		*Phytophthora asparagi*	323	91	87	514	45.0-66.4
OTU3579		*Phytophthora cactorum*	214	99	81	4	51.1-85.8
OTU4280		*Phytophthora cactorum/ Peronospora sparsa*	208	99	80	2	50.6-84.1
OTU5636		*Phytophthora boehmeriae*	232	100	71	4	50.3-68.8
OTU6144		*Phytophthora quercina*	217	99	81	2	51.6-85.0

[a] Sequences included as potential new *Phytophthora* species; [b,c] OTUs combined as same OTU.

Table 2.6: Number of OTUs in each genus that were found in both fall and spring or that were exclusive to only one season.

Genus	Shared in both seasons	OTUs in fall only	OTUs in spring only
Total	89	53	27
Achlya	4	2	1
Albugo	1	0	0
Aphanomyces	5	1	1
Apodachlya	3	0	0
Brevilegnia	1	1	2
Dictyuchus	1	0	0
Globisporangium	1	0	1
Hyaloperonospora	0	1	0
Lagena	1	0	0
Lagenidium	2	1	0
Leptolegnia	1	0	1
Newbya	0	1	0
Other	7	1	1
Ovatisporangium	1	0	0
Paralagenidium	1	0	0
Peronospora	1	1	0
Phytophthora	13	13	10
Phytopythium	3	0	1
Plectospira	0	0	1
Protoachlya	0	0	1
Pythiogeton	2	1	1
Pythiopsis	1	3	0
Pythium	33	20	6
Saprolegnia	6	7	0
Thraustotheca	1	0	0

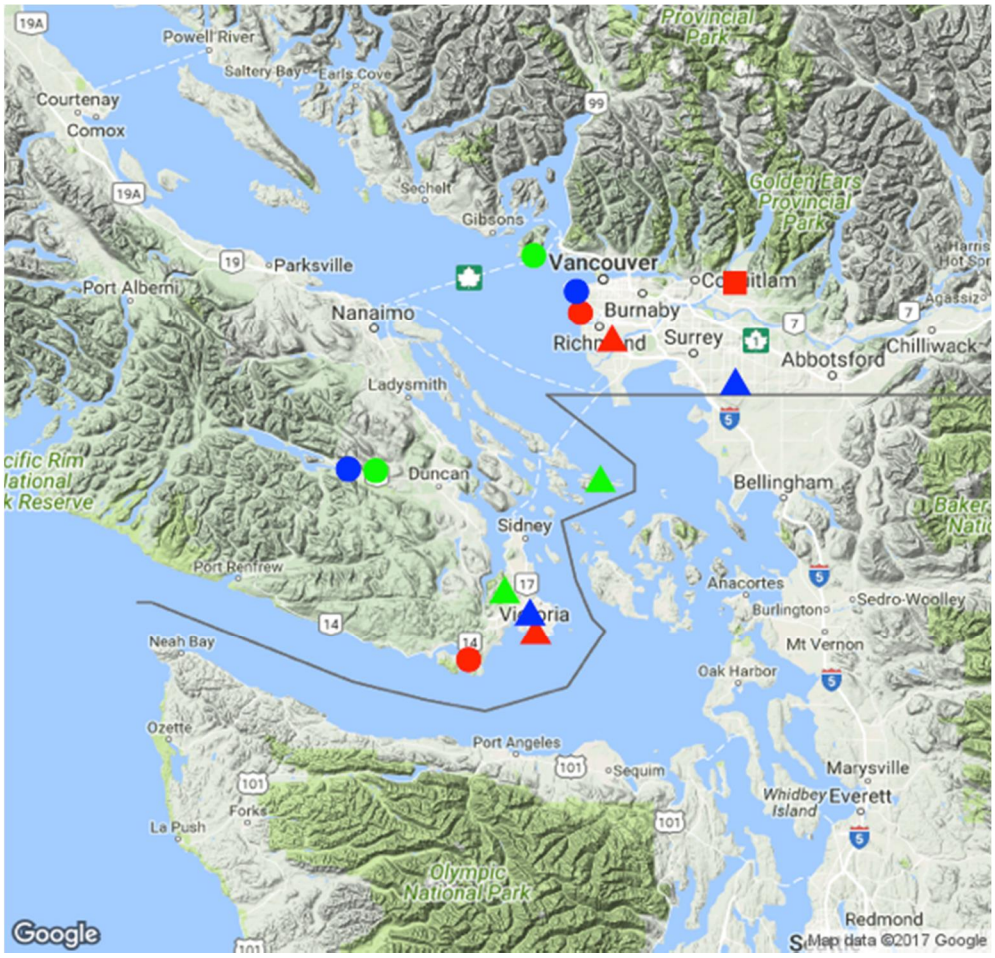


Figure 2.1: Map of sampling sites. Natural, interface and urban sites are represented by markers with green, blue and red colors respectively; biogeoclimatic zone are represented by circles (CWHxm), triangles (CDFmm) or square (CWHdm)

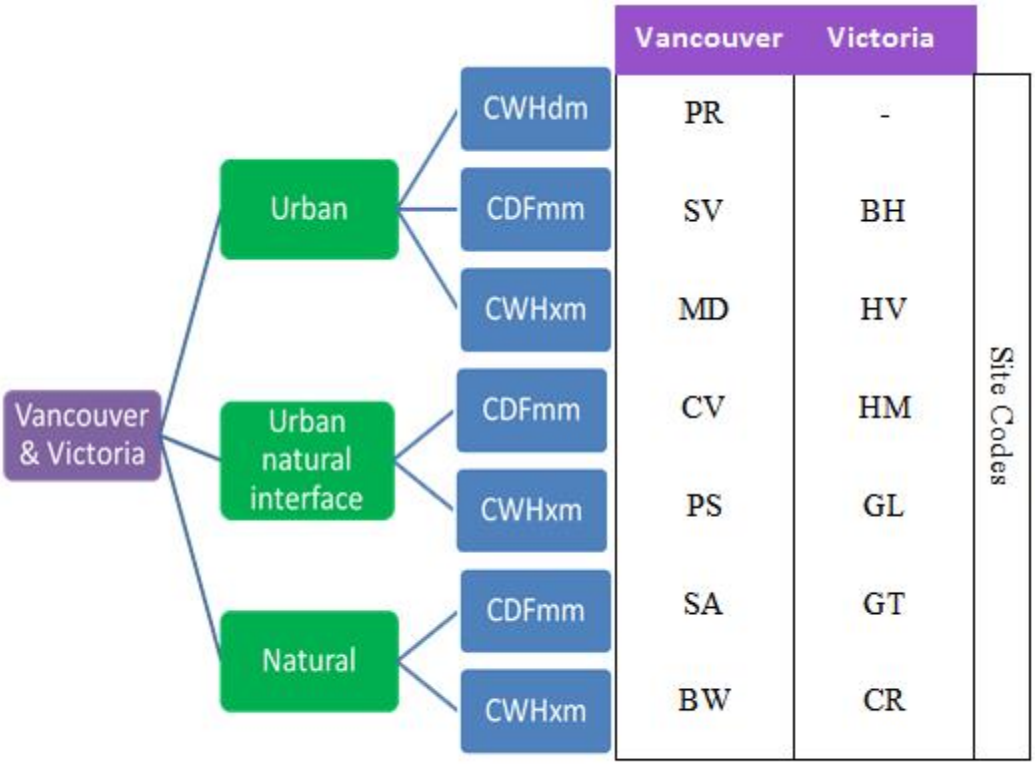


Figure 2.2: Sample design showing the hierarchical arrangement of sites within geographic area and site types

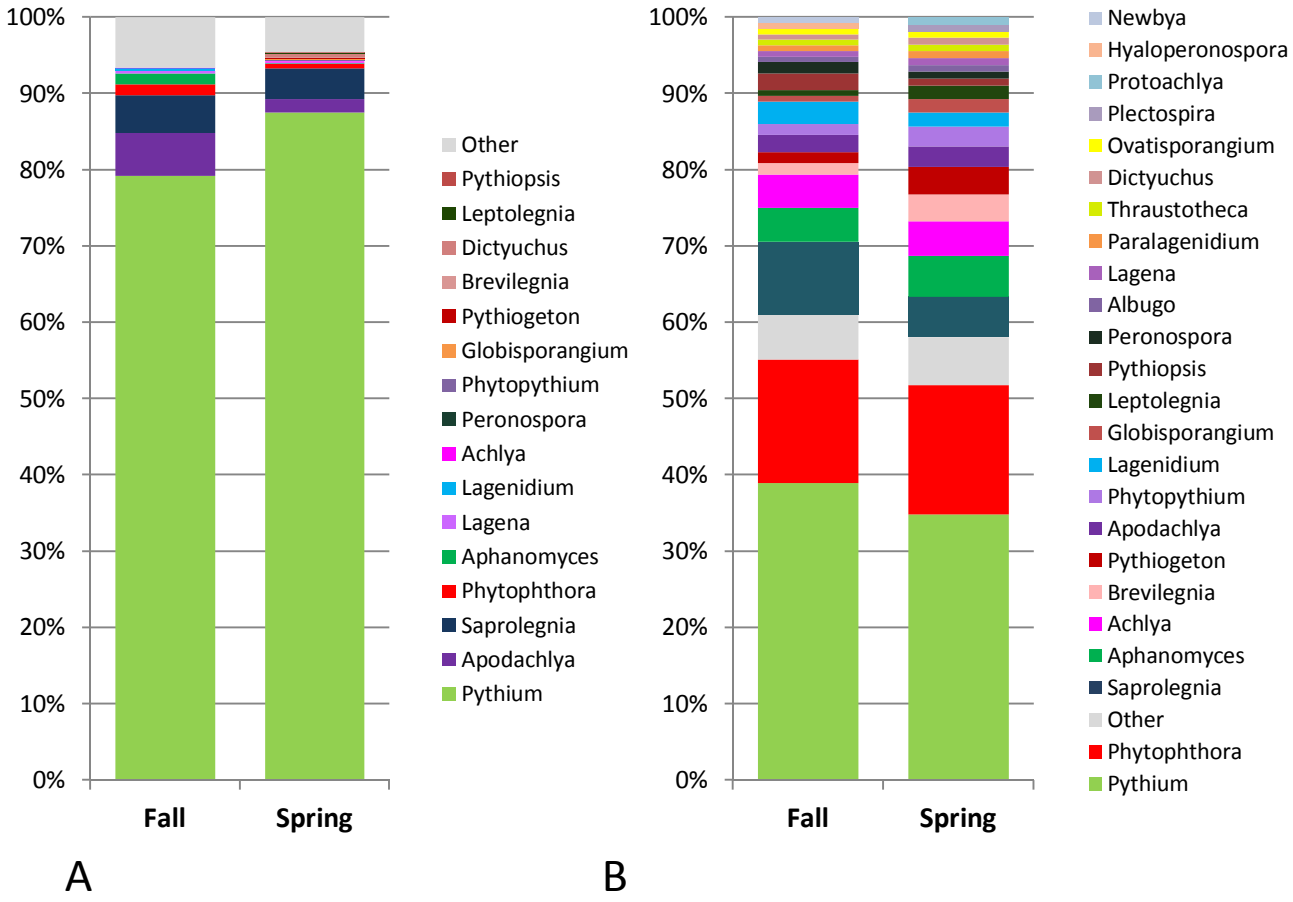


Figure 2.3: Oomycete community composition in southwest British Columbia for fall and spring collections. (A) Relative abundance of sequence reads by genera. (B) Relative abundance of OTUs by genera.

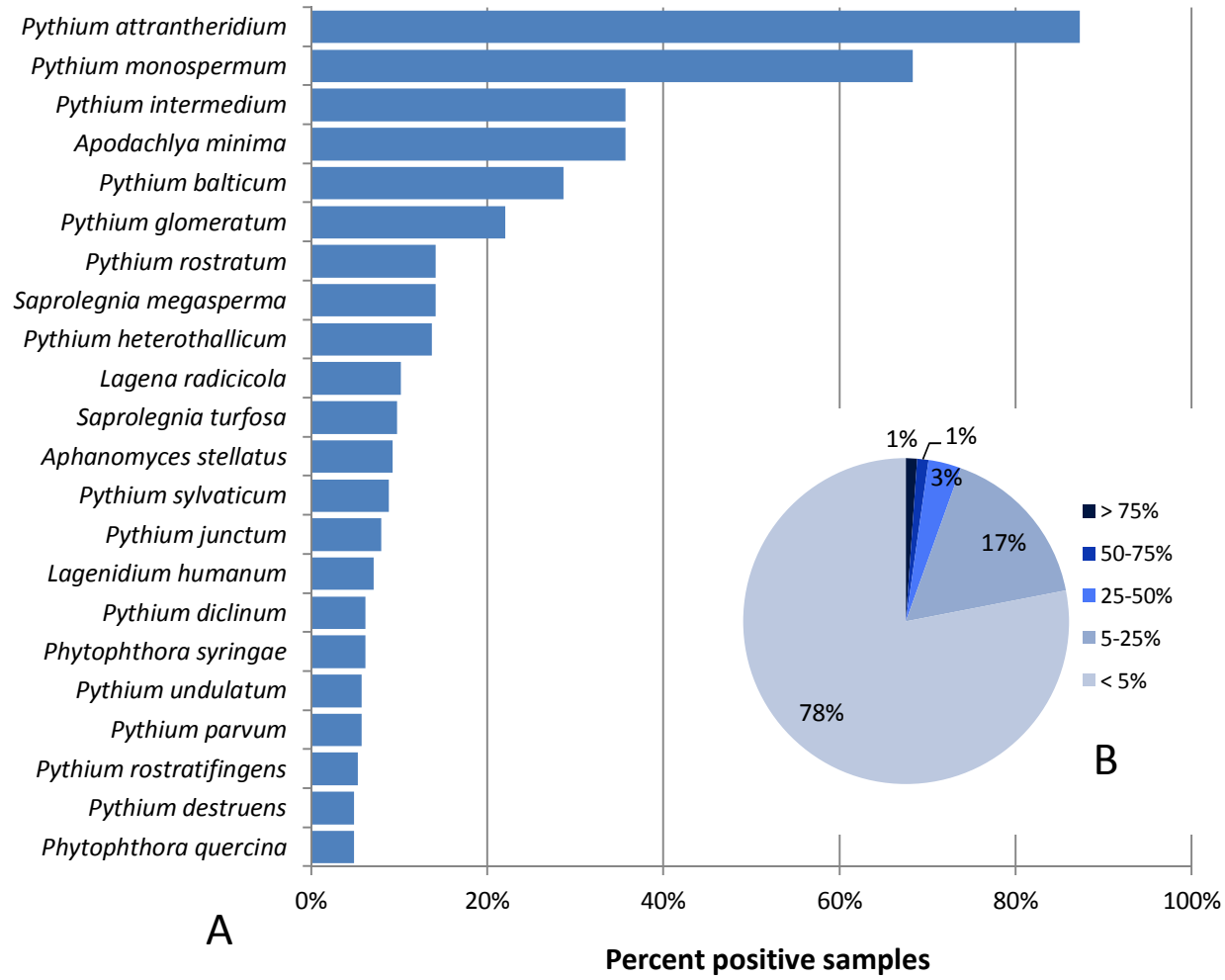


Figure 2.4: Distribution of OTUs across samples. (A) Percentage of samples each higher frequency OTU was found in (higher frequency OTUs were those observed in at least 5% of the total samples; only OTUs assigned to the species level are shown). (B) Percentage of the OTUs that were found across each percentage category of the total samples.

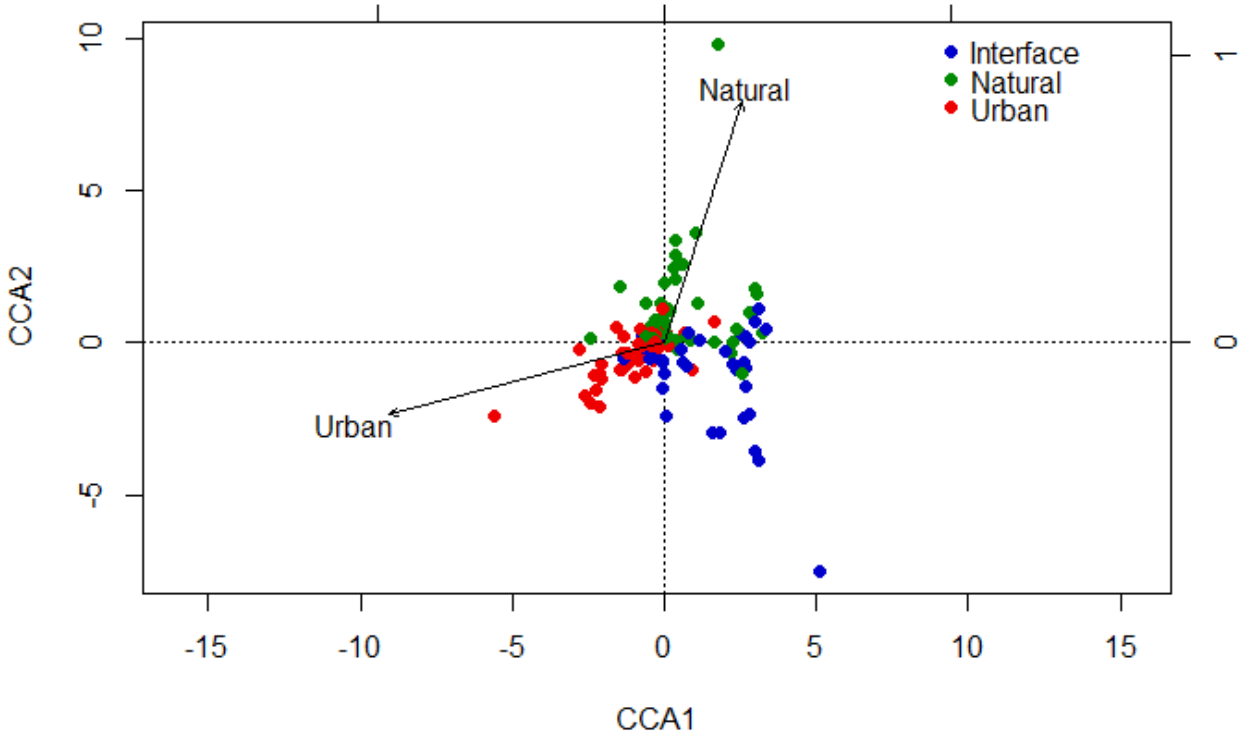


Figure 2.5: Influence of site type on species profile illustrated in a canonical correspondence analysis (CCA). Species profile is based on the relative abundance of species in the combined fall and spring samples.

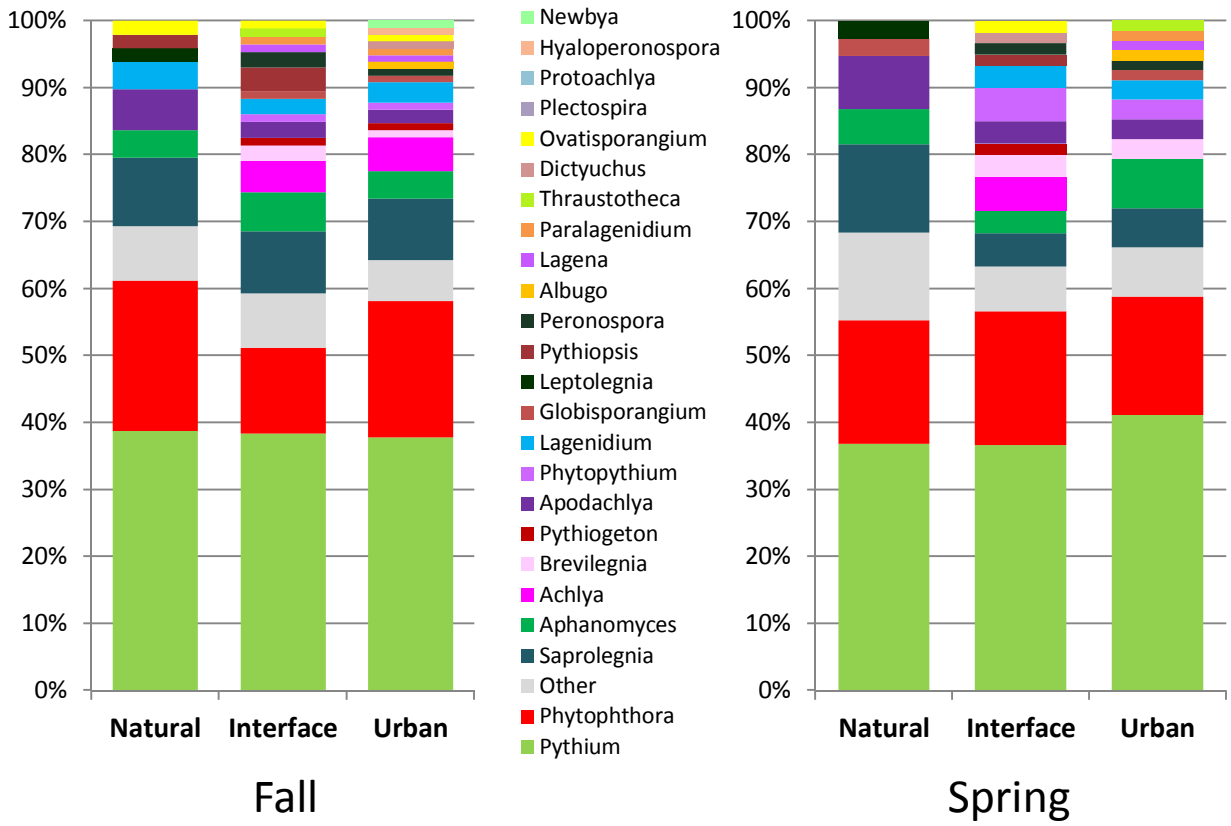


Figure 2.6: Relative abundance of OTUs by genera in each site type for the fall and spring Oomycete collections.

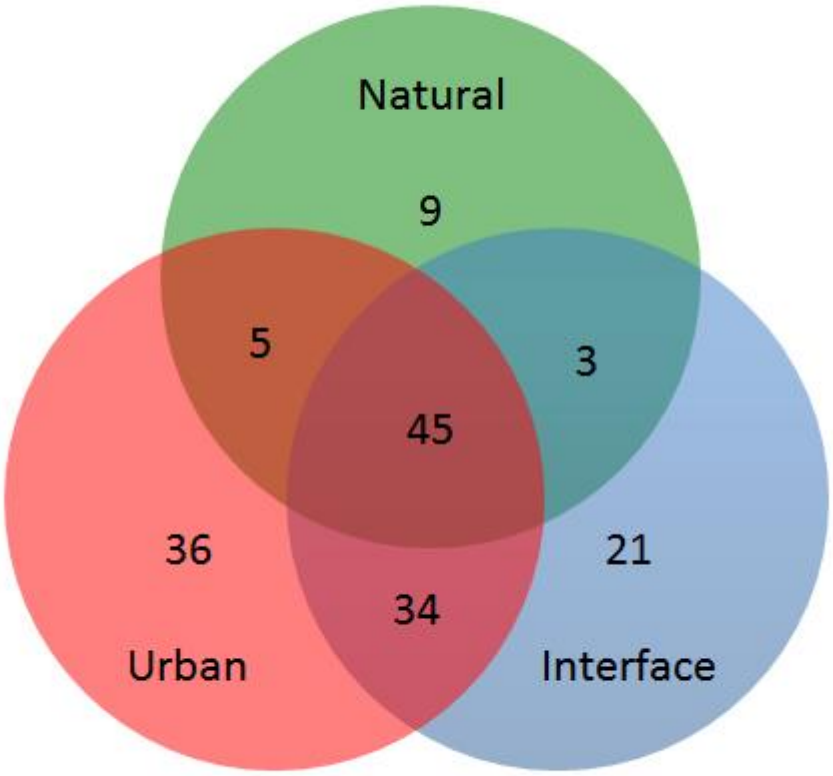
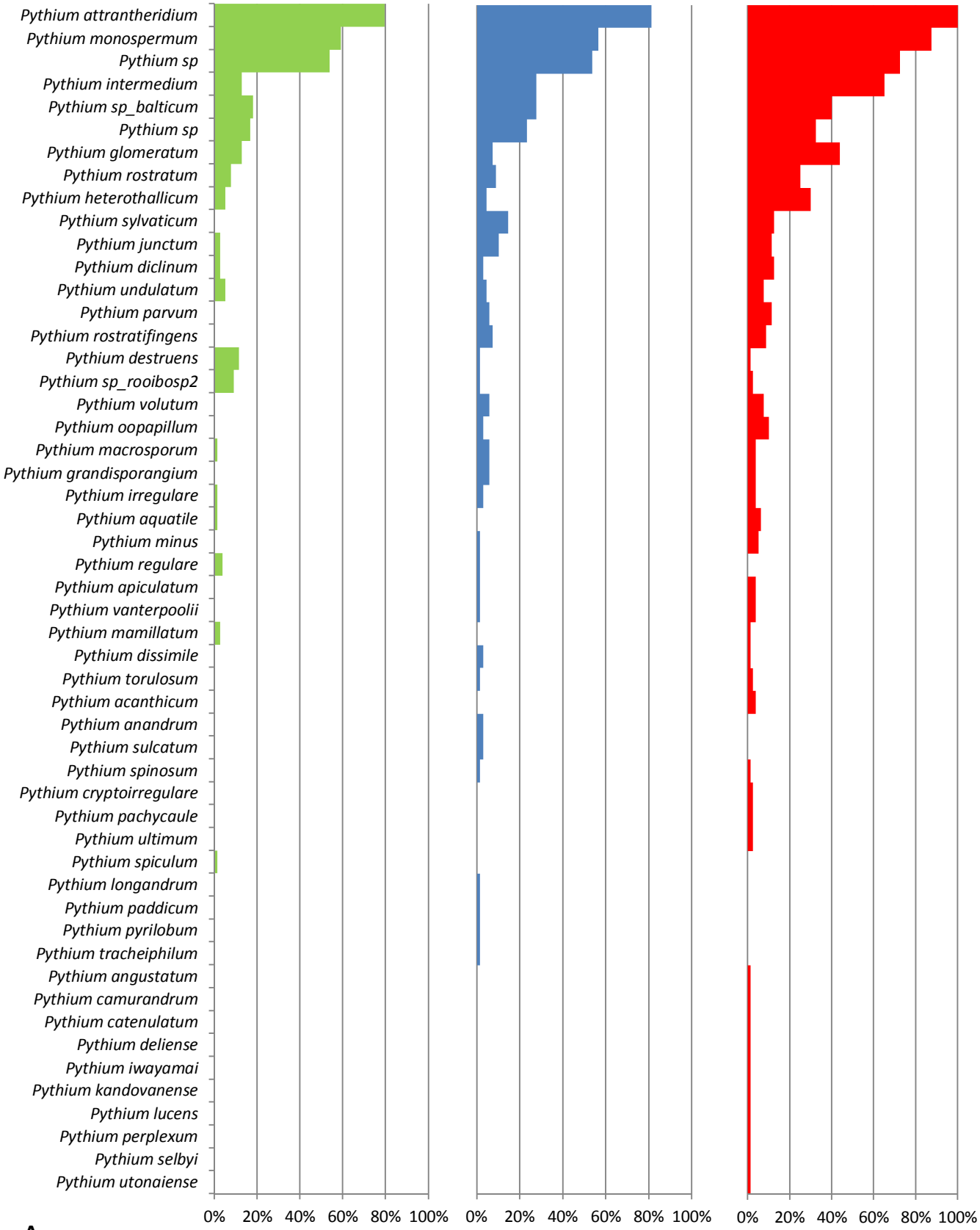
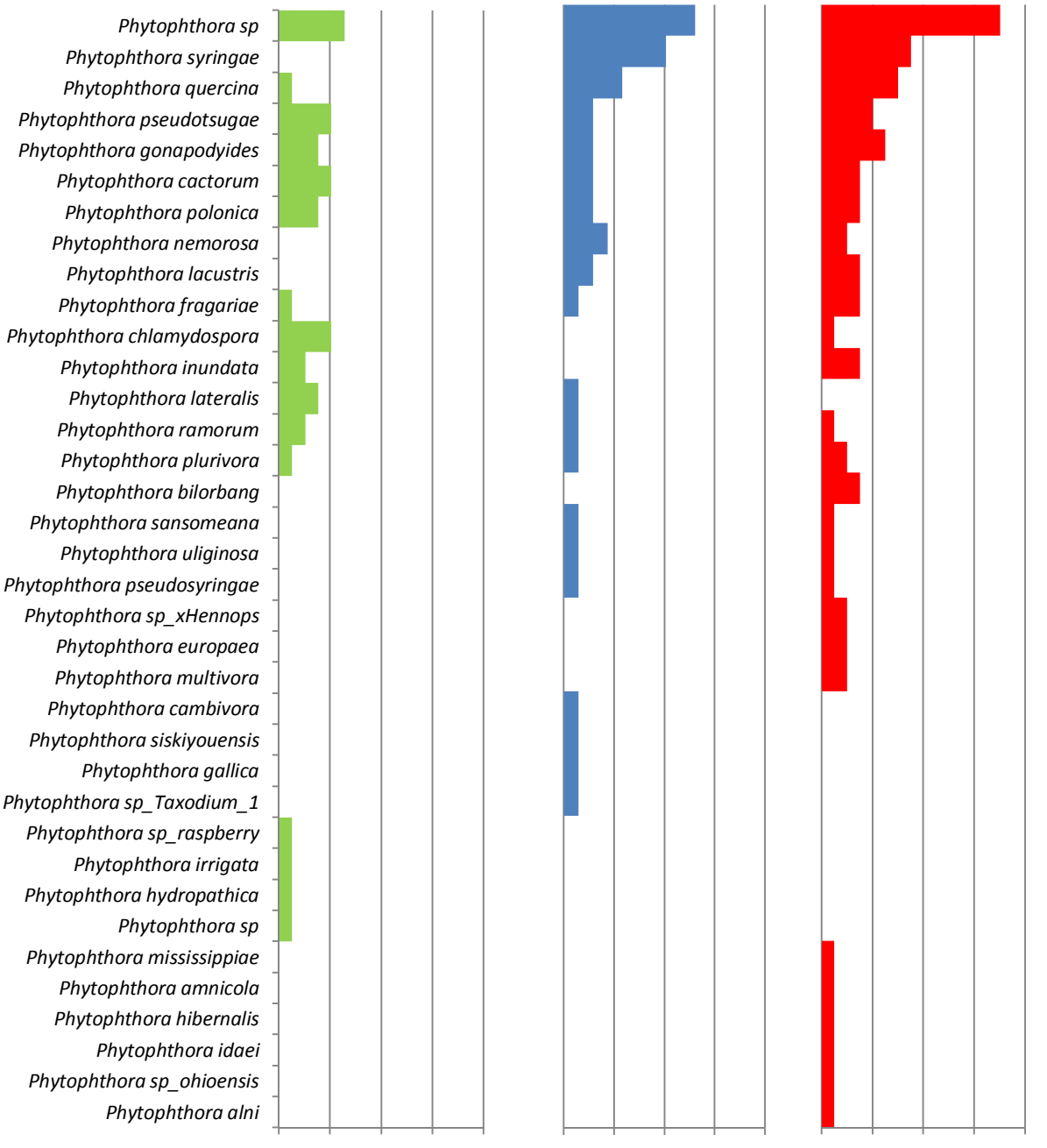


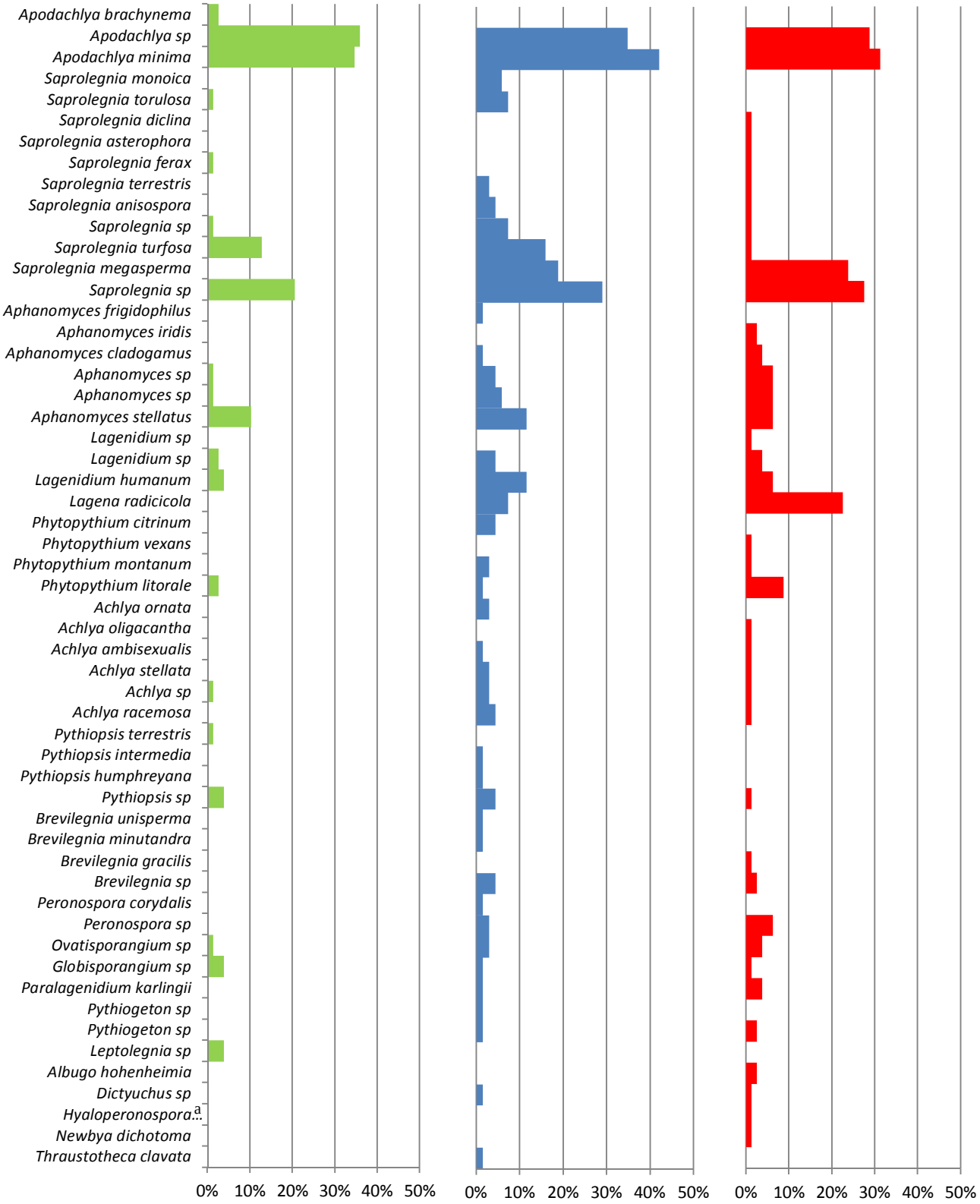
Figure 2.7: Number of shared and private oomycete OTUs between site types (fall and spring sampling combined).


Pythium attrantheridium
Pythium monospermum
Pythium sp
Pythium intermedium
Pythium sp_balticum
Pythium sp
Pythium glomeratum
Pythium rostratum
Pythium heterothallicum
Pythium sylvaticum
Pythium junctum
Pythium diclinum
Pythium undulatum
Pythium parvum
Pythium rostratifingens
Pythium destruens
Pythium sp_rooibosp2
Pythium volutum
Pythium oopapillum
Pythium macrosporum
Pythium grandisporangium
Pythium irregulare
Pythium aquatile
Pythium minus
Pythium regulare
Pythium apiculatum
Pythium vanterpoolii
Pythium mamillatum
Pythium dissimile
Pythium torulosum
Pythium acanthicum
Pythium anandrum
Pythium sulcatum
Pythium spinosum
Pythium cryptoirregulare
Pythium pachycaule
Pythium ultimum
Pythium spiculum
Pythium longandrum
Pythium paddicum
Pythium pyrilobum
Pythium tracheiphilum
Pythium angustatum
Pythium camurandrum
Pythium catenulatum
Pythium deliense
Pythium iwayamai
Pythium kandovanense
Pythium lucens
Pythium perplexum
Pythium selbyi
Pythium utonaiense
0% 20% 40% 60% 80% 100%
0% 20% 40% 60% 80% 100%
0% 20% 40% 60% 80% 100%

A


Phytophthora sp
Phytophthora syringae
Phytophthora quercina
Phytophthora pseudotsugae
Phytophthora gonapodyides
Phytophthora cactorum
Phytophthora polonica
Phytophthora nemorosa
Phytophthora lacustris
Phytophthora fragariae
Phytophthora chlamydospora
Phytophthora inundata
Phytophthora lateralis
Phytophthora ramorum
Phytophthora plurivora
Phytophthora bilorbang
Phytophthora sansomeana
Phytophthora uliginosa
Phytophthora pseudosyringae
Phytophthora sp_xHennops
Phytophthora europaea
Phytophthora multivora
Phytophthora cambivora
Phytophthora siskiyouensis
Phytophthora gallica
Phytophthora sp_Taxodium_1
Phytophthora sp_raspberry
Phytophthora irrigata
Phytophthora hydropathica
Phytophthora sp
Phytophthora mississippiae
Phytophthora amnicola
Phytophthora hibernalis
Phytophthora idaei
Phytophthora sp_ohioensis
Phytophthora alni
0%
5%
10%
15%
20%
0%
5%
10%
15%
20%
0%
5%
10%
15%
20%

B


Apodachlya brachynema
Apodachlya sp
Apodachlya minima
Saprolegnia monoica
Saprolegnia torulosa
Saprolegnia diclina
Saprolegnia asterophora
Saprolegnia ferax
Saprolegnia terrestris
Saprolegnia anisospora
Saprolegnia sp
Saprolegnia turfosa
Saprolegnia megasperma
Saprolegnia sp
Aphanomyces frigidophilus
Aphanomyces iridis
Aphanomyces cladogamus
Aphanomyces sp
Aphanomyces sp
Aphanomyces stellatus
Lagenidium sp
Lagenidium sp
Lagenidium humanum
Lagena radicicola
Phytopythium citrinum
Phytopythium vexans
Phytopythium montanum
Phytopythium litorale
Achlya ornata
Achlya oligacantha
Achlya ambisexualis
Achlya stellata
Achlya sp
Achlya racemosa
Pythiopsis terrestris
Pythiopsis intermedia
Pythiopsis humphreyana
Pythiopsis sp
Brevilegnia unisperma
Brevilegnia minutandra
Brevilegnia gracilis
Brevilegnia sp
Peronospora corydalis
Peronospora sp
Ovatisporangium sp
Globisporangium sp
Paralagenidium karlingii
Pythiogeton sp
Pythiogeton sp
Leptolegnia sp
Albugo hohenheimia
Dictyuchus sp
Hyaloperonospora...
Newbya dichotoma
Thraustotheca clavata
0% 10% 20% 30% 40% 50%
0% 10% 20% 30% 40% 50%
0% 10% 20% 30% 40% 50%

C

Figure 2.8: Number of samples each OTU is found in each site type. *Pythium* OTUs (A), *Phytophthora* OTUs (B) or other OTUs (C). Samples from natural shown in green, interface in blue and urban in red. OTUs were based on a sequence clustering threshold of 97% and minimum count of 2 sequences in a sample. [a] *Hyaloperonospora nasturtii-aquatici*.

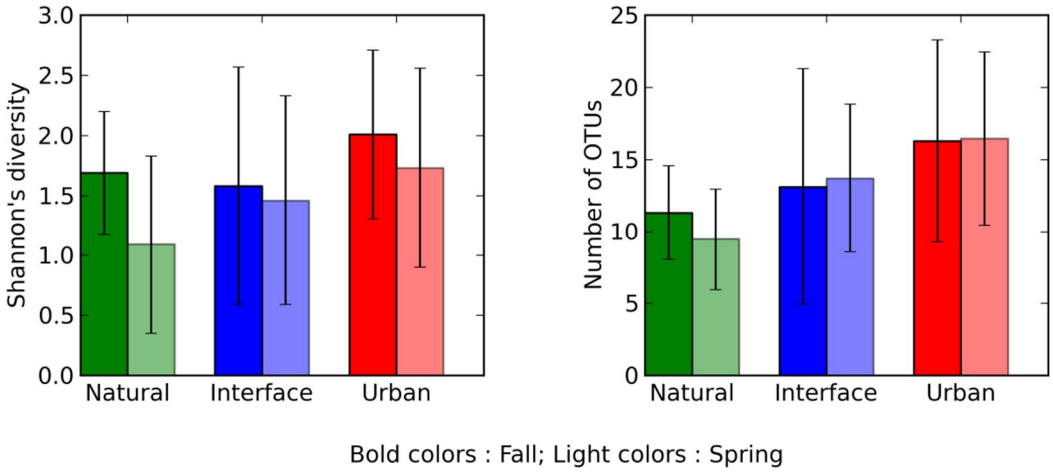


Figure 2.9: Diversity by site type. Samples were rarefied to a depth of 300 sequences. Samples with less than 300 sequences were eliminated. Number of OTUs is right and Shannon's diversity index is on the left. Error bars represent standard deviation. ANOVA results for comparisons are shown in Table 2.4.

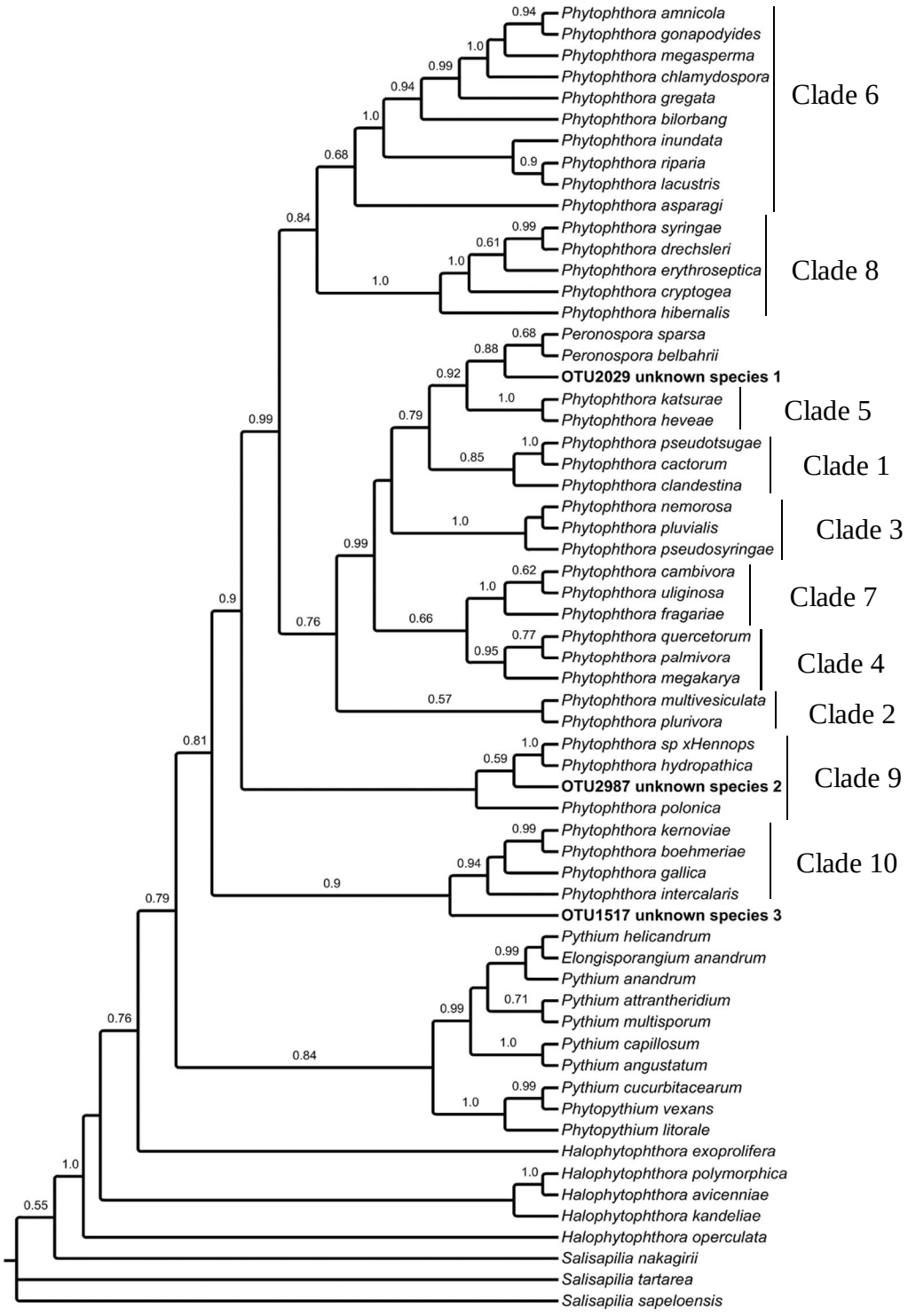


Figure 2.10: Phylogenetic analysis with unknown putative *Phytophthora* species. Phylogenetic reconstruction was done using the ITS 1 of representative *Phytophthora* species and close relatives.

Chapter 3: Anthropogenic activities introduce new species and increase the frequency of *Phytophthora* in urban habitats compared to native habitats

Introduction

Emerging plant and forest diseases threaten natural ecosystems, food supply, and the economies that depend on them. The most important drivers of emergent plant diseases are anthropogenic activities, in particular introductions of alien pathogens, climate change and agricultural practices (Anderson *et al.*, 2017). Alien invasive species cause major environmental damage and economic losses (Pimentel *et al.*, 2000). Emergent diseases of crops and trees are often caused by alien pathogens that benefit from attacking susceptible host species and from the lack of competitors. In new habitats, host species have not co-evolved with the pathogens and therefore have not evolved resistance (Anderson *et al.*, 2017). Alien species can be introduced to new habitats through several mechanisms, both natural and human aided; however human aided introductions have far exceeded natural introductions since human migration and commerce began (Mack *et al.*, 2000). Global trade has increased the risk of introductions by transporting pathogens on commodities such as living plant material, fruits or vegetables, wood products such as logs and lumber or on wood packaging materials.

In recent years invasive pathogenic *Phytophthora* species have impacted natural ecosystems, agriculture, forestry and horticulture industries and interrupted trade. The *Phytophthora* genus is often considered one of the most important and damaging groups of plant pathogens in history.

In the past few decades, an increasing number of emergent species have caused landscape level disturbances in natural and forest ecosystems (Hansen, 2015). Although the exact origin of those *Phytopthora* species is often not known, they are believed to have been introduced through international plant trade and the planting of infected nursery stock (Brasier & Jung, 2006; Jung *et al.*, 2016)(Jung et al., 2016)(Jung et al., 2016)(Jung et al., 2016). *Phytophthora ramorum*, *Phytophthora lateralis* and *Phytophthora cinnamomi* are examples of pathogens that were spread between geographic regions via trade and subsequently escaped into natural forests causing extensive damage to native hosts (Hansen *et al.*, 2000; Brasier & Jung, 2006; Brasier, 2008; Cahill *et al.*, 2008; Hansen, 2008, 2015; Brasier & Webber, 2010; Grünwald *et al.*, 2012; Burgess *et al.*, 2017a). Economic impacts include the costs of enacting quarantine regulations, monitoring pathogen spread, and eradicating disease outbreaks. Reduction of income is due to the loss of marketable species or restrictions imposed on traded commodities. Port Orford cedar (*Chamaecyparis lawsoniana* (A. Murr .) Parl.) is a highly valuable wood and ornamental species; however *Phytophthora lateralis* has caused the near complete shutdown of both the nursery and commodity trade for this tree (Hansen *et al.*, 2000). *Phytophthora ramorum* has had wide ranging economic impacts (Grünwald *et al.*, 2012). In some nurseries, the costs associated with controlling *P. ramorum* have resulted in bankruptcy (Grünwald *et al.*, 2012). The cost of government eradication efforts for *P. ramorum* in Curry County, Oregon between 2001 and 2008 was approximately 4.3 million US dollars (Kanaskie *et al.*, 2010). Furthermore, climate change may exacerbate the problems associated with many invasive species; *P. cinnamomi* is predicted to expand into new regions as the climate becomes more favorable (Burgess *et al.*, 2017a).

Non-indigenous pathogens can have serious wide ranging effects on ecosystems (Loo, 2009). Loss of a dominant host species can result in disequilibrium because of changes in microclimate, hydrology, and nutrient fluxes or the loss of food sources and habitat for wildlife (Ellison *et al.*, 2005; Loo, 2009). Reduction of Port-Orford cedar in Oregon and California by *P. lateralis* could result in significant ecological changes including changes in soil calcium content, loss of habitat structure, as well as loss of stream structure and stability along riparian zones (Hansen *et al.*, 2000; Loo, 2009). Death of tanoak (*Lithocarpus densiflorus*) in Oregon and California forests due to *P. ramorum* increases fire risk, may affect animals dependant on acorns as a food source, and could cause extended periods of chaparral vegetation in some areas (Hansen, 2008; Metz *et al.*, 2011).

Currently, we have very little knowledge of what native or invasive *Phytophthora* species are present in BC forests, or whether anthropogenic activities could influence *Phytophthora* community composition and distribution. Some common pathogenic species such as *P. cactorum P. cambivora*, *P. citricola*, and *P. megasperma* were reported in the irrigation water from orchards in the Okanagan and Similkameen valleys in Southern British Columbia in the 1960s (McIntosh, 1966); however these reports would have been based on morphological data and may not reflect updated species designations, and many of these species are now considered complexes of closely related sister species. *Phytophthora ramorum* has been found in nurseries in British Columbia (Bilodeau *et al.*, 2007), and *P. lateralis* has been found in urban areas of Victoria (Gorman, 1993). It is unknown if these pathogens have escaped into natural ecosystems in British Columbia.

Phytophthora monitoring is often conducted using different techniques, from different niches (water, soil, or foliage) over different seasons. The choice of method, niche, or sampling period may result in different community profiles and raises question on which methods provide the most accurate or inclusive data. One of the most frequently utilized survey methods is stream baiting, using Rhododendron leaves in mesh traps submersed in streams, followed by isolation of necrotic bait leaf tissue on selective media, however other baits are commonly utilized including pear tissue (Hulvey *et al.*, 2010, Reeser *et al.*, 2011, Hüberli *et al.*, 2013, Shrestha *et al.*, 2013, Brazee *et al.*, 2016). Additional survey methods are now becoming more frequently used including baiting in the lab where water and soil collected in the field are baited in bottles or other containers (soil samples are flooded with water), or filtration methods where water is filtered and *Phytophthora* are isolated from the filters (Balci *et al.*, 2007; Hwang *et al.*, 2008; Reeser *et al.*, 2011; Oh *et al.*, 2013; Hansen *et al.*, 2017). DNA meta-barcoding on environmental samples is also starting to be used, however this technique is relatively new in *Phytophthora* studies (Vannini *et al.* 2013; Sapkota and Nicolaisen 2015; Prigigallo *et al.*, 2016; Riit *et al.*, 2016; Esmaeili Taheri *et al.*, 2017). Methods employed, baits used in trapping methods, as well as substrate or niche, and seasonality all provide somewhat differing results (Balci *et al.* 2007; Reeser *et al.*, 2011; Hansen *et al.*, 2012a; Hüberli *et al.*, 2013; Oh *et al.* 2013).

In Chapter 2, we used culture-independent approaches to generate oomycete community profiles and found higher diversity in urban than in natural sites. We also found more *Phytophthora*-positive samples and more *Phytophthora* species in urban than in natural sites. Our objective in this work was to use culture-dependant methods to bait, isolate and barcode *Phytophthora* species to obtain a more detailed analysis of species present in both water and soil. We again

targeted sampling to the same environments with varying anthropogenic impact: urban areas (around agricultural or residential areas and an urban park), interface areas (natural areas within or bordering urban areas), and natural areas (forests with minimal urban influence). As in Chapter 2, we tested the hypothesis that anthropogenic activities influence *Phytophthora* communities and increase the frequency of putative non-indigenous species. We hypothesized that urban environments contain higher *Phytophthora* diversity and more introduced species compared to natural environments. In addition, we investigated differences in *Phytophthora* communities under two different urban influences: agricultural or residential. We hypothesized that different activities would result in different communities. We surveyed the *Phytophthora* species present across the three site types (urban, interface and natural) using a baiting in the lab approach where soil and water samples were collected in the field, baited in the lab for *Phytophthora* using rhododendron leaves and pear tissue. *Phytophthora* species were isolated from necrotic bait tissue and barcoded using the internal transcribed spacers 1 and 2 (ITS1 and ITS2). Both soil and water were sampled over two seasons (fall and spring) to obtain a more inclusive sampling of the community; two bait types were also used for the same reason. We hypothesized that fall and spring communities would harbor different species due to climate and moisture differences. In addition, to further test the hypothesis that different urban influences (agricultural or residential) result in different communities, and to compare survey methods, we surveyed streams in the urban sites using an in stream baiting method with rhododendron leaves in mesh traps. We hypothesized that different methods and different substrates would result in different species profiles.

Methods

Direct baiting and isolation of *Phytophthora* from streams

In the fall of 2011, four *Phytophthora* traps were set in small to moderate waterways in each of the five urban sites (described in Chapter 2, Figures 2.1 and 2.2, Table 2.1). Traps were constructed following the design from Hwang *et al.* (2008). Mesh screen was wrapped around a wooden dowel and stapled to make four pockets for leaf baits. Total trap size was approximately 20 x 20 cm and each pocket was four to five centimeters wide. *Rhododendron* leaves were used for baits, with two entire leaves and two leaves wounded with small cuts. Traps were tied to a wooden stake which was anchored into the waterway (Figure 3.1). Traps were left for ten weeks from the end of August to the beginning of November 2011. Foliage baits were collected approximately every 2 weeks and traps were reset with fresh bait leaves. Collected leaves were wrapped in moistened towel and brought back to the lab for immediate processing.

Bait leaves were rinsed in sterile water and pieces of leaf tissue (approximately 5mm x 5mm) were aseptically cut from the green edge of symptomatic zones and partially embedded in *Phytophthora* selective PARPH-CMA media (Kannwischer and Mitchell) containing 17 g/L cornmeal agar (Sigma-Aldrich, St. Louis, MO, USA), 10 mg pimaricin preparation (Sigma-Aldrich, St. Louis, MO, USA), 250 mg ampicillin sodium salt (Sigma-Aldrich, St. Louis, MO, USA), 10 mg Rifamycin SV sodium salt (Sigma-Aldrich, St. Louis, MO, USA), 100 mg pentachloronitrobenzene 95% (Sigma-Aldrich, St. Louis, MO, USA), and 50 mg hymexazol (Alfa Aesar, Haverhill MA, USA) in 1.0 liter deionized water. For each bait leaf, five pieces of leaf tissue were embedded on a single plate. Petri plates were incubated at room temperature for

two to four days and between one to five individual isolates were selected based on differing morphology, growth rates and morphology resembling *Phytophthora* (white, and mycelial growth form). Each selected isolate was sub-cultured onto a new individual plate of PARPH-CMA and incubated at room temperature.

Isolation and identification of *Phytophthora* from soil and water samples

Water and soil samples were collected from all 13 sites described in Chapter 2 in the fall of 2012 and in the spring of 2013 (Figures 2.1 and 2.2, Table 2.1). Ten soil samples were collected as described in Chapter 2, and ten water samples (one litre each) were collected in each stream or waterbody adjacent to the soil sampling location, placed in Ziplock bags and brought back to the lab for processing. Samples were chosen to cover the largest number of independent waterways available at each site. A total of 260 water samples and 260 soil samples were used for lab baiting and culture-based *Phytophthora* identification. Water samples used for lab baiting methods (below) were processed within 24 hours of collection, and soil samples used for lab baiting methods (below) were processed within nine days of collection.

Water samples were baited with pieces of *Rhododendron* leaves and pear placed in Ziploc bags. Soil samples were poured into 1 litre rectangular plastic containers, flooded with sterile water in an approximately 1:2 ratio of soil to water, covered with a layer of cheesecloth and baited with *Rhododendron* leaves and pear (Balci *et al.*, 2007). Soil and water samples were incubated at room temperature for five to seven days after which time sections of the leaves and pear used for baiting (approximately 5mm x 5mm) were aseptically cut from the edge of symptomatic zones

and partially embedded in *Phytophthora* selective PARPH-CMA media. Subcultures of individual isolates were made as described above for the *Phytophthora* trapping methods to obtain pure cultures of each isolate.

Fresh subcultures of each individual isolate that yielded *Phytophthora*-like morphology (white, and mycelial growth form) on PARPH-CMA media in fall 2011, fall 2012 and spring 2013, were transferred onto 2% malt extract agar (MEA) and cultures were incubated at room temperature. Mycelium was scraped from the cultures and placed in a sterile 1.5 ml Eppendorf tube. The tissues were ruptured by placing tubes in liquid nitrogen and then in a Tissuelyser (Qiagen, Germany) for 3 minutes at maximum speed. Sterile water (1.2 ml) was added to the samples followed by a 5 minute boiling step at (95.9 °C). Samples were then diluted 1:100 and 1 µl was used in PCR. The internal transcribed spacer region (ITS) of the nuclear ribosomal RNA genes (rDNA) was amplified in PCR using the primers ITS6 (Cooke *et al.*, 2000) and ITS4 (Bruns *et al.*, 1990). PCR reactions were performed in 25 µl volumes and contained 0.5 units of Invitrogen Platinum *Taq* DNA polymerase (ThermoFisher Scientific, Waltham, MA, USA), 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM Invitrogen dNTPs (ThermoFisher Scientific, Waltham, MA, USA), and 1 µM of each primer. Reactions were performed with an initial heating step at 94 °C for three minutes, followed by 35 cycles of a denaturation step at 94 °C for 30 seconds, an annealing step at 55 °C for 30 seconds, an elongation step at 72 °C for one minute. The reaction was finished with a 10 minute extension step. PCR products were visualized on a 1% agarose gel stained with Ethidium Bromide. PCR products were sent to the Plateforme de Séquençage et de Génotypage des Genomes, Centre de Recherche du CHUL (Quebec, Canada) for sequencing. Samples from 2011 were sequenced in both directions using both primers (ITS6 and ITS4), whereas samples

from 2012 and 2013 were sequenced in only one direction using the ITS4 primer. PCR reactions that failed were repeated using no dilution, and then again with a 1:10 dilution. Samples that failed 3 times were eliminated from the experiment.

For 2011 samples, ITS sequences were assembled in Geneious version (6.0.5) (http://www.geneious.com, Kearse *et al.*, 2012) using the default settings. All assembled and unassembled sequences were compared with sequence data from Genbank (US National Center for Biotechnology Information) using the Basic Local Alignment Search Tool (BLAST, Altschul *et al.*, 1990). Species identification was based on sequence homology (a minimum of 99% sequence similarity) and phylogenetic analysis (as described below). Initially, one phylogenetic tree was built with all isolates from this study. Isolates that grouped tightly together with a strong homology with only one species in Genbank were assigned to that species designation. For sequences with a close homology to more than one species from a species complex, or with discordant results between the homology search and phylogenetic results, representative sequences from each of the species in the complex or closely related species within the *Phytophthora* phylogenetic clade (as defined in Cooke *et al.* (2000) and Blair *et al.* (2008)) were downloaded and aligned with the query sequence using ClustalW (Thompson *et al.*, 1994) in the BioEdit Sequence Alignment Editor (Hall, 1999). Fixed nucleotide differences were used to assign species where possible. If no fixed differences separated the species, the sequence was assigned a species complex designation. A species was considered as a new taxon if there were no species with a close sequence homology and if the sequence did not group with others in the phylogenetic analysis with sequences of the closest available species. Sequences with many ambiguities were considered undetermined.

For phylogenetic analysis to determine putative species assignments, assemblies (ITS1 and ITS4 from 2011) and sequences matching *Phytophthora* based on homology were trimmed in Geneious 6.0.5 using the default settings. Sequences from 2011 that could not be assembled were aligned where possible. Assembled, aligned or single sequences were visually inspected for ambiguities and edited using the International Union of Pure and Applied Chemistry (IUPAC) nucleotide ambiguity codes (Cornish-Bowden, 1985). Sequences were aligned with representative sequences of closely related species obtained from Genbank, including where possible, sequences from publically available culture collections or from studies describing new species. Sequences were aligned using ClustalW (Thompson *et al.*, 1994) in the BioEdit Sequence Alignment Editor (Hall, 1999). Maximum likelihood phylogenies were reconstructed using default settings in the Phylogeny.fr pipeline (Dereeper *et al.*, 2008) with PhyML 3.0 (Guindon *et al.*, 2010), the approximate likelihood ratio test (aLRT, Anisimova *et al.* (2006)) for testing branch support, and TreeDyn for tree visualization (Chevenet *et al.*, 2006).

Resulting species assignments were screened to eliminate duplicate species from a sample, leaving only one representative per species for each sample in order to eliminate the possibility of over sampling the same individual.

Sequences from the culture based methods from all three years were combined and a variable file and a taxonomy file (using the taxonomy assigned from phylogenetic analysis and homology searches described above) were created matching the format for mapping files and taxonomy files in QIIME (Caporaso *et al.*, 2010). QIIME was used to cluster the sequences using a 98.75%

similarity cutoff, assign taxonomy, pick representative sequences and to create an OTU file as described in Chapter 2. Sequences were aligned in QIIME using MUSCLE (Edgar, 2004) with default settings. A maximum likelihood phylogeny was reconstructed as above using the Phylogeny.fr pipeline (Dereeper *et al.*, 2008). OTU tables from the culture based collection and the DNA meta-barcoding (Chapter 2) were exported as matrices and combined in Excel for downstream analyses. Prior to combining the DNA meta-barcoding, data was filtered for a minimum of 3 sequences per OTU, per sample, and data was converted to a presence (1)-absence (0) matrix.

Data analysis

The number of samples that yielded *Phytophthora* (herein called *Phythophthora*-positives) was determined for each site and each site type (urban, natural, and interface) and compared with a Chi-square test. The number of species, unique species, and unknown species was calculated for each site type and the number of shared species was calculated between site types.

Species over- and under-representation in each site type was tested by comparing, for each species, the number of individuals found in each site type to the total number of individuals. The degree of over- or under-representation of individuals for a species at a site type relative to the total number of individuals at the three sites was assessed by applying chi-square independence test. For each species, the approach consisted in placing the data in a table of n columns (representing the three different site types i.e. urban, natural, and interface) and m rows for the

different species. The A_{ij} value represent the expected number of individuals for site *i* and species *j* :

$$A_{ij} = \frac{\sum A_{in} \sum A_{mi}}{\sum A_{nm}}$$

And then the chi-square value calculated as follow:

$$\sum \frac{(A_{ij} - \bar{A}_{ij})^2}{\bar{A}_{ij}}$$

Shannonôs diversity index was determined for each site and averaged for each site type. Weighted UniFrac (Lozupone & Knight, 2005) implemented in QIIME (Caporaso *et al.*, 2010) was used to compare the phylogenetic Beta diversity between sites (using only culture data with full ITS1 and ITS2 sequences). UniFrac values were averaged for sites of the same site type to compare site types.

To look for possible effects from different anthropogenic activities, four of the urban sites were sub-divided into agricultural or residential (Chapter 2, Table 2.1). Agricultural sites were near large-scale nursery or agricultural operations while residential sites were near residential housing. The number and frequency of *Phytophthora* species, alpha diversity (Shannonôs diversity index) and beta diversity (weighted UniFrac) were calculated as above but using agricultural versus residential site types. A canonical correspondence analysis using the vegan package (Oksanen *et al.*, 2016) in the R statistical environment (R Core Team, 2017) was conducted to investigate the effect of urban site type and sample type (water or soil) on species

profiles. In addition, the package ggtree (Yu *et al.*, 2017) in the R statistical environment (R Core Team, 2017) was used to compare species abundance profiles for the two site types.

Putative hybrid species were defined as those with multiple heterozygous SNPs in the ITS sequence. Sequences with only a few heterozygous SNPs were not considered. The relationships of putative hybrid species to other similar species were investigated by comparing heterozygous SNPs in the putative hybrids to the alleles present in the other taxa to try to determine the closest matching species. Where possible, the ITS sequences were phased using a putative parent species as a template to create one ITS haplotype, and the other was created by subtracting the one putative parent to leave the alternate set of alleles.

Results

Anthropogenic activity increases the frequency and diversity of *Phytophthora*. There was a gradient of increasing *Phytophthora* abundance in samples from low anthropogenic influence (natural 23%), moderate influence (interface 33%) to high anthropogenic influence (urban 41% PR excluded, 39% PR included) (Figure 3.2). Since PR represents the only urban site in the biogeoclimatic zone CWHdm with no corresponding natural or interface sites, comparisons were made both with and without this site to account for possible increases in diversity or changes to the species profiles that could be due to sampling artefact. The percentage of *Phytophthora*-positive samples was higher in urban than in natural and interface sites (chi-square = 17.542, $p < 0.0001$, PR excluded). This gradient was consistent between all methods used: baiting water and soil in the lab and with culture-independent DNA meta-barcoding on soil from Chapter 2 (Figure

3.2). There were a total of 38 *Phytophthora* species found in fall and spring using a combination of baiting in the lab and DNA meta-barcoding. The highest number of species was observed in urban sites, followed by interface and natural sites (Figures 3.3 and 3.4). Urban sites also had more unique species (found in only one site type) and more unknown species. Only 12 species (32% of the total) were found in all three site types (Figures 3.3 and 3.4). Urban and interface sites shared more species (26%) than either urban and natural (5%) or interface and natural sites (11%) (Figure 3.3). Six species were unique to urban sites whereas only two species were unique to natural interface sites (Figures 3.3 and 3.4). We found a total of six unknown or previously undescribed taxa using baiting in the lab and DNA meta-barcoding (three additional unknowns, one of which was a putative hybrid, were found using stream baiting in urban sites; described later) (Figure 3.4). All five unknown taxa were found in urban sites, four in interface sites and none were found in natural sites. Three of those unknown taxa were found exclusively through DNA meta-barcoding (unknown species 1, 2 and 3, Figure 3.4; these are described in chapter 2). Two of the unknown species were putative hybrids (described later) that placed phylogenetically within clade nine. One was found multiple times in urban and interface sites (P. Unknown_Clade9_sp.4 (putative hybrid)) and the other one were found only once in an urban sites (P. Unknown_Clade9_sp.5 (putative hybrid)). There was a third putative hybrid, *P.* x *polonica*, which was found in all site types (Figure 3.4; described later).

Only two species were significantly over or under represented in a particular site type. *Phytophthora pseudotsugae* was found six times, exclusively in natural sites, and *P. cambivora* was found seven times, six times in interface sites and once in a natural site (Table 3.1). Species from nine of the ten clades described by Cooke *et al.* (2000) and Blair *et al.* (2008) were found in

this study; only clade five was not represented. Clades six, eight and nine were most commonly represented (Figure 3.4). The most common species overall were *P. gonopodyides*, *P. lacustris*, and *P. chlamydospora* (clade 6), followed by *P. plurivora* (clade 2) and *P. syringae* (clade 8) (Figure 3.4). *Phytophthora plurivora* was found in all site types, but less frequently in natural samples and *P. syringae* was only found in urban and interface sites.

Beta diversity (weighted UniFrac) was highest among natural sites and lowest among interface sites (Figure 3.5). Beta diversity was also low between interface and urban sites (Figure 3.5). PR had the highest UniFrac distances (Figures 3.5). The UniFrac distances between site types were significant, both with and without PR (ANOVA with PR, UniFrac values log transformed: $F = 3.045$, $p = 0.015$; ANOVA without PR, no transformation: $F = 3.030$, $p = 0.017$). Natural sites had the lowest and urban sites had the highest alpha diversity (Shannonôs diversity index) (Figure 3.6). SV had the highest Shannonôs diversity. There was no significant difference in Shannonôs diversity index among site types (Figure 3.6).

Different anthropogenic activities drive differences in *Phytophthora* communities. *Phytophthora* communities differed in agricultural and residential sites. Species composition was significantly associated with urban site type ($F = 3.787$, $p = 0.001$) and substrate sampled (water or soil; $F = 5.648$, $p = 0.001$) in a canonical correspondence analysis, where the differences in soil and water communities were observed within both residential and agricultural sites (Figure 3.7). Agricultural sites had slightly more species (29 *versus* 26 in residential), and higher overall abundance of *Phytophthora* species than residential sites (Figure 3.8). Interestingly, taxa within clade nine were significantly more frequent in agricultural sites (five taxa/OTUs found 55 times)

than in residential sites (three taxa/OTUs found three times; Figure 3.8; Table 3.1). Two clade 9 putative hybrids, *Phytophthora* x *polonica* and *P.* unknown Clade 9 sp.4, as well as *Phytophthora irrigata* were significantly over-represented in agricultural samples (Table 3.1). Clade 2, specifically *Phytophthora plurivora* were significantly over-represented in residential samples as well as *Phytophthora gonopodyides* (clade 6; Table 3.1).

Beta diversity (weighted UniFrac) was highest amongst agricultural sites and between agricultural sites *versus* all other site types, largely due to PR (Figure 3.5). Both agricultural sites had high Shannonôs diversity values, as did one of the residential sites (Figure 3.6).

Several unknown, undescribed and putative hybrid species were found. We obtained cultures of five *Phytophthora* spp. that belong to unknown taxa and one additional isolate that was a putative hybrid of *P. polonica*. Three of the unknown/undescribed isolates, also from clade 9, had several heterozygous SNPs in the ITS sequences suggesting they could be hybrids. *Phytophthora* unknown Clade 9 sp. 4 had high ITS identity (96 - 99%) to several other undescribed species from Oregon, West Virginia, South Africa and Ecuador (Table 3.2). In a comparison to a hybrid isolate from South Africa, *P.* sp. xHennops, our isolate is a close match possessing either a heterozygous SNP with nucleotides of both *P. hydropathica* and *P.* sp. xHennops, or matches *P.* sp. xHennops where *P.* sp. xHennops and *P. hydropathica* differ. The isolate in our study was found multiple times, in particular at the PR site. Identity between sequences of our isolates ranged from 98.6 to 100% to each other; however some differences were from ambiguous bases or difficulties with base calling due to background noise in the sequence traces, so we could not determine if these were isolates of the same species or multiple

hybridization events. Two other isolates, which may also be hybrids, *Phytophthora* unknown Clade 9 sp.5 and unknown Clade 9 sp. 6 did not share a close identity (above 99%) to any other species; the closest described species were *P. parsiana* or *P. hydropathica* (Table 3.2). These isolates were both only found once; unknown Clade 9 sp. 6 was only found through stream baiting.

The putative hybrid, *P.* x *polonica* appeared to have fixed heterozygous SNPs shared between the isolates (not confirmed). There were no matches observed in NCBI, but isolates were approximately 98% similar to *P. polonica*. For every heterozygous SNP, one of the bases always matched *P. polonica*, suggesting that this taxon is a hybrid between an unknown species and *P. polonica*. When we reconstructed or phased the two putative versions of the ITS sequences (one with the *P. polonica* alleles and one with the alternate alleles), the closest match to the alternate ITS sequence was with an undescribed species isolated from a stream in Maryland in the United States (Balci and Goss, 2013, unpublished data; GenBank Accession Number KC479206); however there were still at least five nucleotide differences between the alternate allele sequence obtained in our study and this sequence.

The two other unknown species were from clade six, both found only once from stream baiting in urban sites. Both isolates matched other undescribed species in GenBank (Table 3.2).

Different *Phytophthora* detection methods show different community profiles. A total of 42 *Phytophthora* species were found over the three years of sampling with all methods: stream baiting, baiting in the lab, and DNA meta-barcoding (Figure 3.9). Twelve species (mostly from

clade 6 and 9) were found with all three methods (Figure 3.9 and 3.10). Six species were shared between lab baiting and DNA meta-barcoding methods, three between lab baiting and stream baiting, and only one between stream baiting and DNA meta-barcoding (Figure 3.9 and 3.10). DNA meta-barcoding recovered the most species; however all were found in low frequency. Clades 6 and 9 were the most frequent clades encountered using stream baiting and baiting in the lab, whereas clade 8 was the most frequent using DNA meta-barcoding (Figure 3.9).

Phytophthora were found in 32% of samples in all years using all three methods. Baiting in the lab on soil yielded the lowest percentage of *Phytophthora* positive samples (17%) whereas stream baiting yielded the highest percentage of positive samples (64%; Figure 3.11). However DNA meta-barcoding on soil yielded the highest number of *Phytophthora* species (34; Figure 3.11). Baiting water in the lab yielded the least number of species (16); 13 of those were found in the urban sites. Stream baiting in the same urban waterways yielded 20 different species (Figure 3.11).

The spring/summer sampling period yielded over double the percentage of *Phytophthora* positive samples compared to the fall sampling period (69% compared to 31%). In addition there were 32 species recovered in the spring/summer sampling period compared to 26 from the fall sampling (data not shown).

Discussion

Anthropogenic disturbances and specifically alien invasive species are one of the most important drivers of emergent plant diseases (Anderson *et al.*, 2017). Globalization and international trade are responsible for the introduction and emergence of many alien invasive plant pathogens (Desprez-Loustau *et al.*, 2010). Environmental change, or changes in the pathogen can also lead to emergence (Stukenbrock & McDonald, 2008). We have found a strong signature of anthropogenic influences on *Phytophthora* populations in urban areas. Our data suggests a higher *Phytophthora* species diversity and frequency in urban and urban/natural interface areas likely resulting from both introductions of alien species and increases in natural populations due to environmental disturbance. Secondly, we find two putative hybrid *Phytophthora* species showing possible patterns of early emergence in agriculturally influenced environments. These results suggest anthropogenic influences not only shape *Phytophthora* communities, but likely also influence the frequency of disease emergence.

Phytophthora species are especially prone to introductions since they can easily be transported in soil of asymptomatic plants and be overlooked during inspections (Desprez-Loustau *et al.*, 2010; Migliorini *et al.*, 2015). Surveys of nursery plants have uncovered a relatively high frequency of *Phytophthora* species associated with soil (Dart & Chastagner, 2007; Parke *et al.*, 2014; Jung *et al.*, 2016; Prigigallo *et al.*, 2016). Higher species diversity in urban habitats could be explained by the influx of non-indigenous *Phytophthora*s brought through both horticultural and agricultural activities. This pathway of introduction was strongly suggested by the results of a large-scale survey of *Phytophthora* diversity in nurseries producing stock for forests,

horticulture, ornamental, amenity and landscape plantings (source populations) as well as areas where stock would have been planted (sink populatlions) including forests, riparian zones, landscapes, ornamental and horticultural plantations (Jung *et al.*, 2016). Similar infestation rates were found in nurseries producing planting material as in the areas where planting stock would have been destined. These rates of infestation were much higher than what had been found in mature forests (Jung *et al.*, 2016).

Movement of *Phytophthora* from urban to natural environments is likely facilitated through interface environments. There was a gradient in the impact of the anthropogenic activities where the measures of abundance and diversity in interface sites were always intermediate between those in urban and natural sites. We also found the highest number of shared species between urban and interface sites. This indicates that the most anthropogenically affected sites are a repository of *Phytophthora* species diversity but that the interface sites could act as a bridge toward natural environments supporting previous suggestions that urban trees act as bridge-head for invasive species establishment and spread to natural environments (Paap *et al.*, 2017). Under conducive conditions, a single *Phytophthora* species can cause an epidemic, a pattern that has been observed for some of the most damaging *Phytophthora* species, including two species that we recovered with DNA meta-barcoding. *Phytophthora ramorum* is believed to have been introduced through nursery material in California and Oregon and subsequently spread to adjacent forests where it caused extensive mortality of coast live oak (*Quercus agrifolia* Nee.) and tanoak (Rizzo *et al.*, 2002; Mascheretti *et al.*, 2008). It is now found in natural forest ecosystems extending from California through Oregon and Washington (Rizzo & Garbelotto, 2003; Chastagner *et al.*, 2009; Strenge *et al.*, 2016). In our study, *P. ramorum* DNA was found

using DNA meta-barcoding in urban, interface and natural sites; however we could not confirm results using a species specific quantitative PCR assay (data not shown), possibly because of the lower sensitivity in the PCR detection method versus DNA meta-barcoding, PCR inhibitors in soil, or a very low concentration of *P. ramorum* DNA in the sample. Nevertheless, these results suggest that this pathogen has escaped the nursery environment in BC, or that it is spreading from other sources in the Pacific Northwest. In spite of extensive monitoring and eradication from nurseries in BC, there are no previous reports of the disease occurring outside of nurseries in BC. Likewise, *Phytophthora lateralis* may be spreading from urban into natural sites. Although only found in natural and interface sites in our study, it is considered introduced in North America (Hansen *et al.*, 2000) and has previously been found in urban areas of BC (Hamelin, unpublished). This pathogen causes mortality of *Chamaecyparis lawsoniana* in the natural range of the tree in Oregon (Hansen *et al.*, 2000). *C. lawsoniana* does not occur naturally in BC, so our finding of *P. lateralis* in natural and interface sites in BC suggests that it must survive in natural environments on another host. *Phytophthora lateralis* has been found in the UK on western red cedar, and its host range extends to other members of the Cupressaceae family, including yellow cedar (*Cupressus nootkatensis*), Pacific yew and juniper, all of which occur naturally in BC (Hansen *et al.*, 2000; Schlenzig *et al.*, 2015, 2017). The possibility of *P. ramorum* and *P. lateralis* having spread to natural environments should be of concern as it could cause damage to natural ecosystems and to international trade.

A lack of knowledge on the biodiversity in a community or natural geographic range of a species can make it challenging to discern introduced species from native ones (Desprez-Loustau *et al.*, 2007). Recent survey studies are recovering many new species, however it is difficult to know if

these species are native or introduced (Linzer *et al.*, 2009). In addition to information on distributions, comparisons of the species composition and frequencies in natural *versus* urban environments may help to distinguish introduced from native species. We found several species only in urban areas that could represent introductions. For example, *P. drechsleri* was detected in only one urban sample in an agriculturally intense region. This pathogen causes disease on a variety of ornamental plants. In the eastern United States molecular evidence, specifically a single shared clone, suggests it was spread through several nursery facilities via movement of infected plants from a shared distributor (Lamour *et al.*, 2003). *Phytophthora hibernalis* was detected in one sample from a region that contains several horticultural nurseries and ornamental plant stores. This species causes citrus brown rot but was recently found infecting Rhododendron plants in nurseries in California and Oregon (Blomquist *et al.*, 2005; Yakabe *et al.*, 2009; Reeser *et al.*, 2011; Sims *et al.*, 2015). *Phytophthora multivora* and *P. gregata* were all limited to urban sites in our study and are thought to originate from Western Australia (Jung & Burgess, 2009). *Phytophthora multivora* is mostly limited to nurseries, plantations and agricultural crops outside of Western Australia where it is widespread in natural ecosystems (Jung & Burgess, 2009). *Phytophthora gregata* also has a limited occurrence outside of Australia. This species was associated with dying plants in alpine heathland vegetation in Australia (Jung *et al.*, 2011). *Phytophthora gregata* is from clade 6, which generally comprises weak or opportunistic pathogens; however it could still cause disease in other geographic ranges or on non-native hosts. Other species may be more difficult to determine natural *versus* introduced ranges. Several species were found in low frequency in our study, and were absent in natural sites; however they could still be endemic suggested by their occurrence in natural areas in other studies focused on

the Pacific Northwest. For example, *Phytophthora riparia* was found frequently in Oregon and Alaska but was not associated with disease (Reeser *et al.*, 2011; Hansen *et al.*, 2017).

Distribution alone may not be enough to determine the origin of a pathogen. *Phytophthora pseudosyringae* and *P. nemorosa* have been found frequently in Oregon and in California and, similar to *P. ramorum*, are associated with Tanoak, Bay Laurel and Coast live oak (Hansen *et al.*, 2017). *Phytophthora nemorosa* has only been found rarely in nurseries and not outside of Oregon and California, while *P. pseudosyringae* has been also been found in Alaska and extensively in Europe (Murphy & Rizzo, 2006; Wickland *et al.*, 2008; Hansen *et al.*, 2017). Their geographic range, lower pathogenicity on local hosts, and variation in the internal transcribed spacer and mitochondrial cytochrome oxidase spacers could suggest the pathogens are indigenous in Oregon and California (Hansen *et al.*, 2017). However low levels of genetic diversity and absence on herbarium specimens could also suggest they are introduced (Monahan *et al.*, 2008; Linzer *et al.*, 2009; Kozanitas *et al.*, 2017). If *P. pseudosyringae* is introduced, the extensive distribution along the Pacific Northwest may be concerning as this pattern could suggest emergent behavior that could be due to either a particularly invasive species or that spread may be facilitated by another factor such as climate change.

Agricultural practices create ecosystems which are often conducive to pathogen spread. Densely planted and genetically uniform host populations, irrigation practices and tillage may all support pathogen spread, and the homogenous landscape may protect pathogen communities from environmental fluctuations that reduce population sizes (Stukenbrock & McDonald, 2008). These agro-ecosystems may also support evolutionary events such as hybridization, host jumps,

host-tracking or horizontal gene transfers leading to the emergence of new pathogens (Stukenbrock & McDonald, 2008). The results of our study suggest that *Phytophthora* communities are evolving in response to agricultural practices. Higher pathogen diversity in urban areas in our study may not only result from introductions, but also from an increase in the frequency of rare endemic species. We detected some pathogens that are likely present due to a close relationship with their agricultural hosts, specifically a species from the *Phytophthora rubi/ P. fragariae* complex. Both *P. rubi* and *P. fragariae* are prevalent in the agricultural industry in the Pacific Northwest and *P. rubi* is possibly endemic (Gigot *et al.*, 2013; Stewart *et al.*, 2014). Furthermore, sites with agricultural activities had very different communities among each other compared to sites in residential areas or parks. Frequency appeared to increase for different species or different clades depending on the dominant urban activity. This was particularly evident for clade nine in agricultural areas and for two of the putative hybrid species, unknown Clade 9 sp.4 and *P.* x *polonica*, which were locally abundant in agricultural sites (each in a different agricultural site), but were less abundant (or absent) in other sites. This pattern not only suggests rapid local spread possibly resulting from agricultural activities, but may also suggest that these putative hybrids have emerged in an anthropogenic ecosystem.

Hybridization can provide an evolutionary fast-track. It can lead to changes in pathogen properties and in some cases, the ability to infect new hosts leading the the emergence of a new pathogen (Stukenbrock & McDonald, 2008). Some invasive tree pathogens, such as dutch elm disease (*Ophiostoma* spp.) and poplar leaf rusts (*Melampsora* spp.), have been shown to have hybridized, providing the hybrids with abundant new genetic material that selection can act upon. Hybridization between *Ophiostoma ulmi* and *Ophiostoma novo-ulmi* resulted in the

transfer of genetic material to *O. novo-ulmi* including a mating type allele which was missing in the introduced population, leading to sexual reproduction, increased diversity and increased fitness in *O. novo-ulmi* populations (Brasier, 2001). *Melampsora* x*columbiana,* a hybrid of the poplar leaf rusts *Melampsora medusae* and *Melampsora occidentalis*, is not only pathogenic on the parental host species but has also emerged on hybrid poplar which is resistant to both parental rust species (Newcombe *et al.*, 2000, 2001). *Phytophthora alni*, responsible for alder decline, emerged through multiple hybridization events leading to a more aggressive pathogen of alder than either of its parental species (Brasier *et al.*, 2004; Ioos *et al.*, 2006). We found evidence for at least three putative hybrids; one of which was limited to urban sites, another to urban and interface sites, and the third was widespread. For *Phytophthora* unknown Clade 9 sp.4 (locally abundant in one of the agricultural sites), we could not determine if there was a single hybridization event, or multiple hybridization events creating a hybrid swarm similar to what has been observed in the *Phytophthora alni* complex (Brasier *et al.*, 2004; Ioos *et al.*, 2006). Although it was difficult to determine the exact relationships of our unknown Clade 9 sp.4 isolates to several similar taxa reported in other studies without further genetic characterization, it appears as though they could either be the same species, or the result of a separate hybridization event between the same or similar parental species. The closely related hybrid, *P.* sp. xHennops from South Africa, was characterized by cloning the ITS gene region to reveal putative parents, *P. hydropathica*, and an unknown clade 9 species (Oh *et al.*, 2013). We can use the distribution of this and other similar taxa in NCBI to attempt to hypothesize the origin and invasion route of this new species. In addition to the South African *P.* sp. xHennops, our isolate also matches isolates found in the rhizosphere of red alder in Oregon (Sims *et al*. 2015), from Rhododendron in a nursery in Oregon (Knaus *et al.*, 2015), streams in West Virginia and Ohio

(Mansfield et al. 2008, unpublished data; GenBank accession numbers EU644712, EU644713, EU644715 and EU644718), and from the Cuyabeno Reserve in the Amazonian region of Ecuador (Coffey, 2012). This distribution suggests that this species could be from the southern hemisphere originating in the Amazon Forest, and that isolates found in South Africa, the US and Canada were introduced on imported plant material. The frequency of hybrids could indicate that *P. hydropathica* and the unknown parent, or parents, are capable of readily hybridizing when in contact, or that the hybrid is relatively fit, does well in the introduced environments and is spreading. The emergent pattern in this species could support the second scenario. Given the emergent pattern, further monitoring and pathogenicity testing would be warranted in this species.

The other species that was locally abundant in the other agricultural site was the putative hybrid *Phytophthora* x *polonica*, also from clade 9. The parent, *P. polonica*, could be European given its current known distribution and weak pathogenicity to alder in Poland (Belbahri *et al.*, 2006; Sárándi-Kovács *et al.*, 2016); however there are no reports of the hybrid there. Our isolates could be a clone spreading after a single hybridization event suggested by fixed heterozygous SNPs in all isolates. *Phytophthora polonica* could have been introduced to BC and hybridized with a local species; however, neither *P. polonica* nor any other closely matching isolates were recovered in our study. A hybridization event could have occurred with a local endemic and rare species that our survey did not uncover; however, the more widespread distribution of the hybrid in all site types and lack of either parental species makes this scenario less likely. In the Netherlands; *Phytophthora cactorum* hybridized with *Phytophthora hedraiandra* after *P. hedraiandra* was introduced (Man in ά Veld *et al.*, 2007). The hybrids appear to be spreading

affecting multiple new hosts, possibly aided by anthropogenic activities. Interestingly, *P. hedraiandra* has only been detected once in the Netherlands suggesting it cannot successfully establish, or that it exists in very low frequency (Man in ά Veld *et al.*, 2007). A similar situation could have occurred with the *P.* x *polonica* hybrid wherever it originates from. Hybridization appears to be common in *Phytophthora*, often resulting from closely related species being brought together under conducive conditions through anthropogenic activities (Brasier, 2001; Burgess, 2015); but may also occur in natural ecosystems as observed in Australia (Burgess, 2015). Furthermore, hybridization appears to be common amongst species from clades six and nine, both frequently found in waterways (Hüberli *et al.*, 2013; Nagel *et al.*, 2013; Oh *et al.*, 2013; Burgess, 2015; Stamler *et al.*, 2016). Although not all hybridization events result in viable species, the emergent nature of some hybrids resulting from a new host range or increased pathogenicity is concerning, in particular if these events are on the rise with globalization and increased international trade.

Overall, similar species were found in our study compared to other studies from the Pacific Northwest (Reeser *et al.*, 2011; Sims *et al.*, 2015). Eighteen different species were found in Oregon and Alaska, all of which, except ñnew species threeò, were found in our study (Reeser *et al.*, 2011). The most common species found in our study, *P. gonapodyides*, *P. lacustris* and *P. chlamydospora* (clade 6) were also the most common species in Oregon and/or Alaska streams (Reeser *et al.*, 2011) and amongst the most common species found in riparian Alder stands in Oregon (Sims *et al.*, 2015). These species were found in all site types in our study and could represent endemic species or species with a long history of establishment. It has been suggested that they are resident in streams, colonizing fresh plant litter and infecting riparian plants under

flooded conditions (Brasier *et al.*, 2003; Jung *et al.*, 2011). Many clade six species are ubiquitous, often not found associated with disease, or are opportunistic, causing disease on stressed plants or in overly wet conditions (Brasier *et al.*, 2003; Jung *et al.*, 2011; Hansen *et al.*, 2012b, 2017; Stamler *et al.*, 2016).

The three survey methods we used provide a complementary picture of *Phytophthora* communities and highlight that results are somewhat dependent on the method used. Our study suggests that soil harbors higher species diversity independent of the sampling method. The culture independent DNA meta-barcoding method (Chapter 2) yielded the highest number of species overall, as is expected. However some species, such as *P. cambivora* did not appear in our DNA meta-barcoding sampling but were readily found with our baiting approach. DNA DNA meta-barcoding yielded double the number of *Phytophthora* positive soil samples and nearly double the number of species. Similar results have been found in comparisons of baiting versus direct detection in soil. Comparisons of Peronomycete diversity from DNA extracted from hempseed baits versus DNA extracted directly from soil reveal higher diversity in soil suggesting not all species can be baited, that differences such as the vegetative state or spore types present in the soil influence detection of a species or that bait choice limits the species detected (Arcate *et al*. 2006). A higher number of *Phytophthora* species in soil were also found using DNA meta-barcoding compared to baiting by Vannini *et al.* (2013).

Similar to our study, Hansen *et al.* (2017) baited soil in the lab and also found a low frequency of *Phytophthora* species. Overall, *Phytophthora* was much less frequent in isolations from soils compared to canopy drip and plant foliage (Hansen *et al.*, 2017). However, Balci *et al.* (2007)

found a slightly higher percentage of *Phytophthora* positive soil samples in Oak forests in the USA, and Oh *et al.* (2013) found a much higher percentage of *Phytophthora* positive soil samples in South Africa, Interestingly, even though we had a lower percentage of *Phytophthora* positive samples, we found a higher number of species considering only the lab baiting results compared to both of these studies. This could be due to differences amongst sites, or land use patterns in the study areas and suggests that even with the low number of positives, we likely have a relatively good sampling of the diversity present.

Stream baiting recovered more species from the same locations than baiting in the lab did. The disadvantage to the stream baiting method is the labor involved with leaving traps in the water and returning to sites to collect and replace the bait leaves. Stream baiting may also underestimate the true species diversity. Stream baiting was shown underestimate diversity compared to water filtration (Hwang *et al.*, 2008; Brazee *et al.*, 2016). In addition, culture dependent methods likely favor fast growing species or species better suited to lab conditions. The dominance of clade six species in streams is likely a result of the ease of recovering *Phytophthora* species from waterways, and also due to their overall abundance in aquatic environments. Additionally, the baits used may limit the species found and using only a single bait type likely results in underestimating the diversity in stream habitats (Hüberli *et al.*, 2013).

Phytophthora are more frequent during the spring and early summer than in the fall, likely due to more available moisture and cooler temperatures. This was apparent in both the frequency of *Phytophthora* as well as the higher number of species found. Similar results were found by Balci *et al.* (2007) who also isolated twice as many *Phytophthora* in the spring compared to the fall.

Results from our study and the studies of others suggest that the most comprehensive picture of *Phytophthora* diversity is obtained by including multiple sampling substrates and environments, by using multiple methods, as well as by sampling over different times of the year to capture seasonal fluctuations.

One of the biggest challenges to surveying the *Phytophthora* community for invasive pathogens is how to determine if a species is introduced or native. Many species are largely unknown or new, and it can often be difficult to determine if a particular species is native or not. Studies such as this one not only uncover evolutionary patterns and mechanisms of pathogen emergence, but they serve to characterize *Phytophthora* communities and geographic ranges of *Phytophthora* pathogens, both natural and invasive. In addition, our data provides valuable information on putative invasive species in BC and will help to direct research and management efforts to high risk species.

Table 3.1: Species that were significantly over or under represented by site type (urban, interface or natural) and by urban site type (residential or agricultural) based on a chi-square analysis.

Comparison: Urban, Interface and Natural						
Species	**Urban**	**Interface**	**Natural**	**X^2 Urban**	**X^2 Interface**	**X^2 Natural**
P. pseudotsugae	0	0	6	2.907	1.921	19.868***
P. cambivora	0	6	1	3.391	6.307*	0.099

Comparison: Residential and Agricultural				
Species	**Agricultural**	**Residential**	**X^2 Agricultural**	**X^2 Residential**
P. gonapodyides	28	56	7.905**	10.210**
P. plurivora	2	18	7.628**	9.852**
P. x *polonica*	20	1	5.631*	7.273**
P. unknown clade 9 sp. 4 (putative hybrid)	17	1	4.631*	5.982*
P. irrigata	12	0	4.054*	5.236*
Clade 9	55	3	15.224 ***	19.665 ***
Clade 2	4	21	7.226**	9.334**

* Significance where * is significant at a probability lower than 0.05, ** is significant at a probability lower than 0.01, and *** is significant at a probability lower than 0.001.

Table 3.2: Unknown or putative new taxa and the origin of their closest hits in a homology search on Genbank.

Putative new taxa	Closest blast hit	% ID	Accesion number	Author	Origin of closest hit
P. unknown clade 9 sp. 4 (putative hybrid)	*Phytophthora parsiana*	99.4%	KJ666742.1	Sims,L.L., Sutton,W., Reeser,P.W. & Hansen,E.M.	Oregon *Alnus rubra* rhizosphere
	Phytophthora sp. Marion-2012	99.2%	KM501456.1	Knaus,B.J., Fieland,V.J., Graham,K.A. & Grunwald,N.J.	Rhododendron leaf in Oregon nursery
	Phytophthora sp. MC5	99.0%	EU644718.1	Mansfield,M.A., Balci,Y., Geiser,D. & Kang,S.	stream-baiting in West Virginia and Ohio
	Phytophthora sp. BF4	98.9%	EU644715.1	Mansfield,M.A., Balci,Y., Geiser,D. & Kang,S.	stream-baiting in West Virginia and Ohio
	Phytophthora sp. TF16	98.9%	EU644713.1	Mansfield,M.A., Balci,Y., Geiser,D. & Kang,S.	stream-baiting in West Virginia and Ohio
	Phytophthora sp. MC53	98.9%	EU644712.1	Mansfield,M.A., Balci,Y., Geiser,D. & Kang,S.	stream-baiting in West Virginia and Ohio
	Phytophthora sp. xHennops	98.9%	KC855182.1	Oh,E., Wingfield,M.J. & Burgess,T.I.	Pietermaritzburg, Kwa Zulu Natal, South Africa; forest stream
	Phytophthora sp. 1 EEO-2011 5	98.9%	GU799642.1	Oh,E., Gryzenhout,M., Wingfield,B.D. & Wingfield,M.J.	stream water in South Africa
	Phytophthora sp. 1 EEO-2011	98.9%	GU799640.1	Oh,E., Gryzenhout,M., Wingfield,B.D. & Wingfield,M.J.	stream water in South Africa
	Phytophthora sp. P16858	98.9%	GU594789.1	Coffey,M.D., Brar,A.K., Xu,E. & Zhang,Y.H.	Ecuador: Amazonian region, Cuyabeno Reserve
	Phytophthora sp. P16855	98.9%	GU594787.1	Coffey,M.D., Brar,A.K., Xu,E. & Zhang,Y.H.	Ecuador: Amazonian region, Cuyabeno Reserve

Putative new taxa	Closest blast hit	% ID	Accesion number	Author	Origin of closest hit
é Continued from above	*Phytophthora* sp. 1 EEO-2011	98.9%	GU799639.1	Oh,E., Gryzenhout,M., Wingfield,B.D. & Wingfield,M.J.	stream water in South Africa
	Phytophthora sp. Maryland 5	98.7%	KC479206.1	Balci,Y. & Goss,E.	Maryland stream
	Phytophthora hydropathica	96.8%	MF959534.1	Shrestha,S.K., Zhou,Y. & Lamour,K.	Tennessee, USA, stream water
P. unknown clade 9 sp. 5 (putative hybrid)	*Phytophthora* sp. x Kunnunara-like	97.5%	KU682602.1	Jung,T., Chang,T.-T., Bakonyi,J., Seress,D., Perez-Sierra,A., Xiao,X., Hong,C., Scanu,B., Fu,C.-H., Hsueh,K.-L., Maia,C., Abad-Campos,P., Leon,M. & Horta Jung,M.	Taiwan, river raft
	Phytophthora sp. Hennops river	97.3%	HQ292655.1	Oh,E., Gryzenhout,M., Wingfield,B.D. & Wingfield,M.J.	South Africa: Gauteng river water
	Phytophthora sp. Kunnunara	97.3%	EF437222.1	Cunnington,J.	Australia; water
	Phytophthora parsiana	96.7%	AY659741.1	Mostowfizadeh-Ghalamfarsa,R., Cooke,D.E. & Banihashemi,Z.	
P. unknown clade 9 sp. 6	*Phytophthora* sp. Umtamvuna	97.5%	KC855181.1	Oh,E., Wingfield,M.J. & Burgess,T.I.	Umtamvuna, Kwa Zulu Natal, South Africa; forest stream
	Phytophthora sp. 1 EEO-2011	97.1%	GU799641.1	Oh,E., Gryzenhout,M., Wingfield,B.D. & Wingfield,M.J.	stream water in South Africa
	Phytophthora hydropathica	97.0%	KC734446.1	Vydrzel,H., Prospero,S. & Schoebel,C.N.	Switzerland

Putative new taxa	Closest blast hit	% ID	Accesion number	Author	Origin of closest hit
P. unknown clade 6 sp. 7	*Phytophthora* sp. forestsoil-like	99.7%	KU682574.1	Jung,T., Chang,T.-T., Bakonyi,J., Seress,D., Perez-Sierra,A., Xiao,X., Hong,C., Scanu,B., Fu,C.-H., Hsueh,K.-L., Maia,C., Abad-Campos,P., Leon,M. & Horta Jung,M.	Taiwan, river raft
	Phytophthora sp. sylvatica	98.9%	EF522140.1	Belbahri,L., Lefort,F., Calmin,G. & Oszako,T.	Poland: Siestrzen; Alnus glutinosa rhizosphere
P. unknown clade 6 sp. 8	*Phytophthora* sp. 1 KS-2013	100%	KF286861.1	Milenkowic,I., Sikora,K., Borys,M., Oszako,T., Keca,N. & Nowakowska,J.A.	Poland: Fraxinus excelsior soil
	Phytophthora sp. 2 FFL-2008	100%	EU594600.1	Lakatos,F. & Szabo,I.	Hungary
	Phytophthora sp. IBL/2011/5/1/1	99.7%	JX276039.1	Nowakowska,J., Milenkovic,I., Oszako,T., Borys,M., Sikora,K., Belbahri,L. & Kubiak,K.	Polish and Serbian forests
	Phytophthora sp. 92-207	99.7%	EU106590.1	Lefort,F., Belbahri,L., Calmin,G., Bragante,F., Bolay,A., Gerber,B. & Viret,O.	Switzerland; Ilex aquifolia
	Phytophthora sp. AG43	99.7%	EU000143.1	Lefort,F., Belbahri,L., Calmin,G., Bragante,F., Bolay,A. & Viret,O.	Switzerland: Collonges-Bellerive; soil of infected *Ilex aquifolium*
	Phytophthora sp. P1050	99.7%	AF541905.1	Brasier,C.M., Cooke,D.E., Duncan,J.M. & Hansen,E.M.	Sweden from root of infected raspberry

Putative new taxa	Closest blast hit	% ID	Accesion number	Author	Origin of closest hit
P. x *polonica*	*Phytophthora polonica*	98.3%	many P. polonica	many matches	
	Phytophthora polonica	97.9%	GU258841.1	Coffey,M.D., Brar,A.K., Xu,E., Sarhan,E.A. &Cunningham,I.M.	World Phytophthora collection
Putative parental allele 1[ab]	*Phytophthora polonica*	99.6%	KX618506.1	Oszako,T., Tkaczyk,M., Nowakowska,J.A., Borys,M. & Tereba,A.	Plant tissue, *Alnus glutinosa* Gaertn., Poland
	Phytophthora polonica	99.6%	KF358225.1	Yang,X. & Hong,C.	Plant nursery in Mississippi
Putative parental allele 2[b]	*Phytophthora* sp. Maryland 8	98.9%	KC479194.1	Balci,Y. & Goss,E.	Stream from Maryland, USA

[a]matches many other *P. polonica*; [b]putative parental alleles of *P.* x *polonica*.

Figure 3.1: Stream baiting using mesh traps with rhododendron leaves.

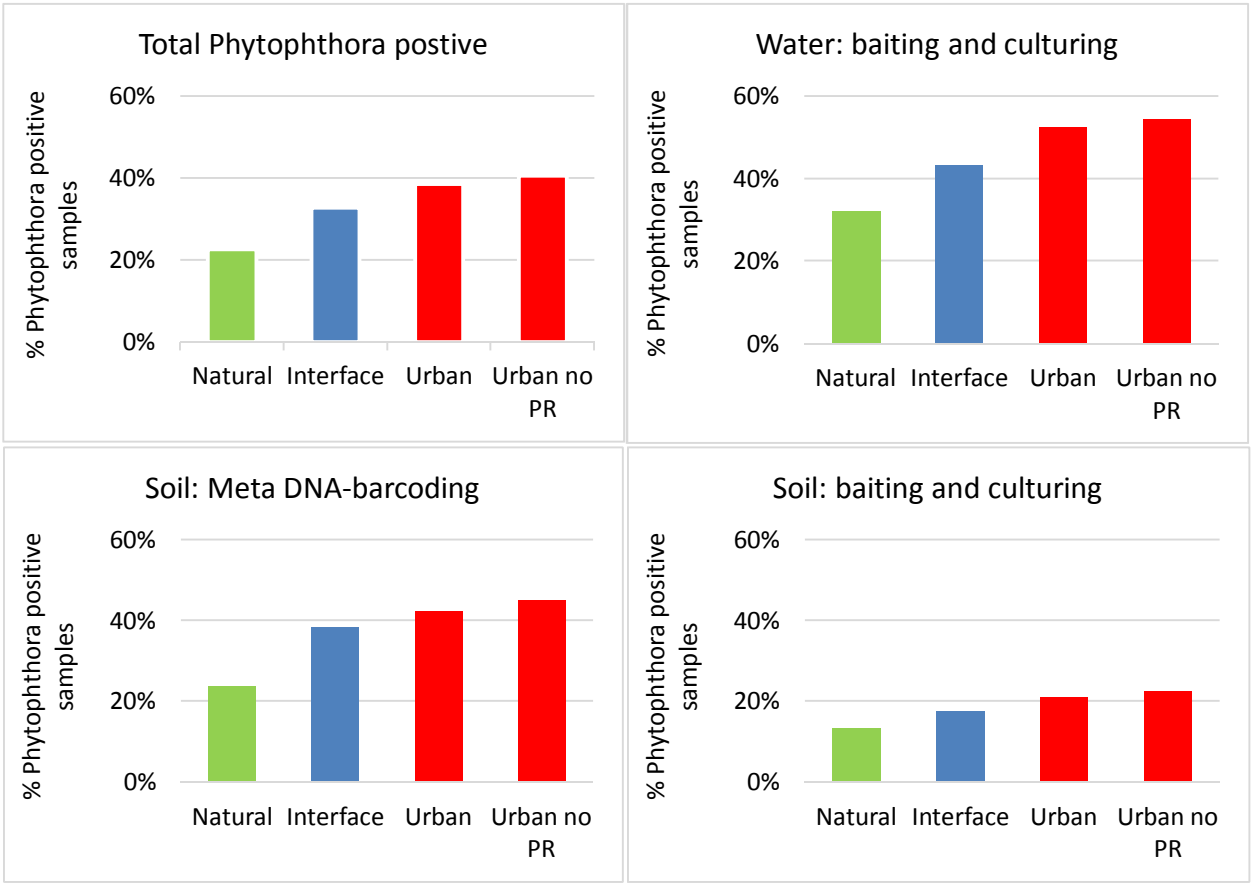


Figure 3.2: Percentage of *Phytophthora* positive samples by site type: Natural versus Interface or Urban; total of all methods (top left), and by each method where baiting and culturing was done in the lab and Meta DNA-barcoding was done on DNA extracts from soil samples.

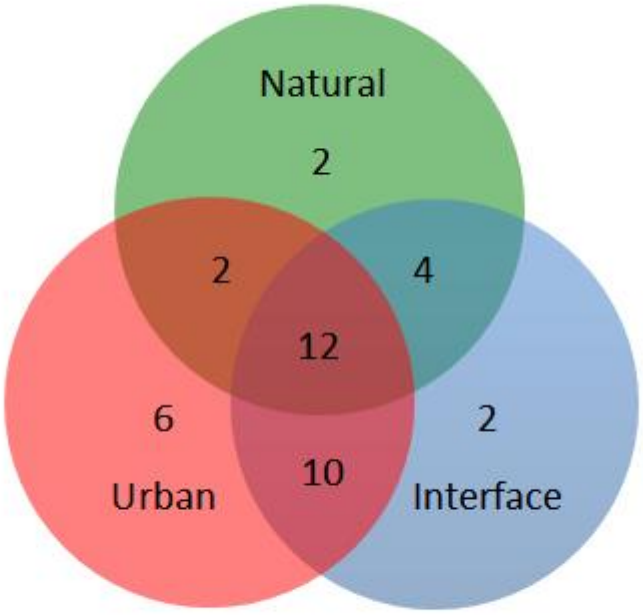


Figure 3.3: Total unique and shared species between the three site types, urban, interface and natural. Sampling occurred over both the fall and spring sampling periods, including both baiting in the lab and Meta DNA-barcoding on soil. Total samples for each group were 239 for urban, 229 for interface, and 234 for natural. PR excluded from counts.

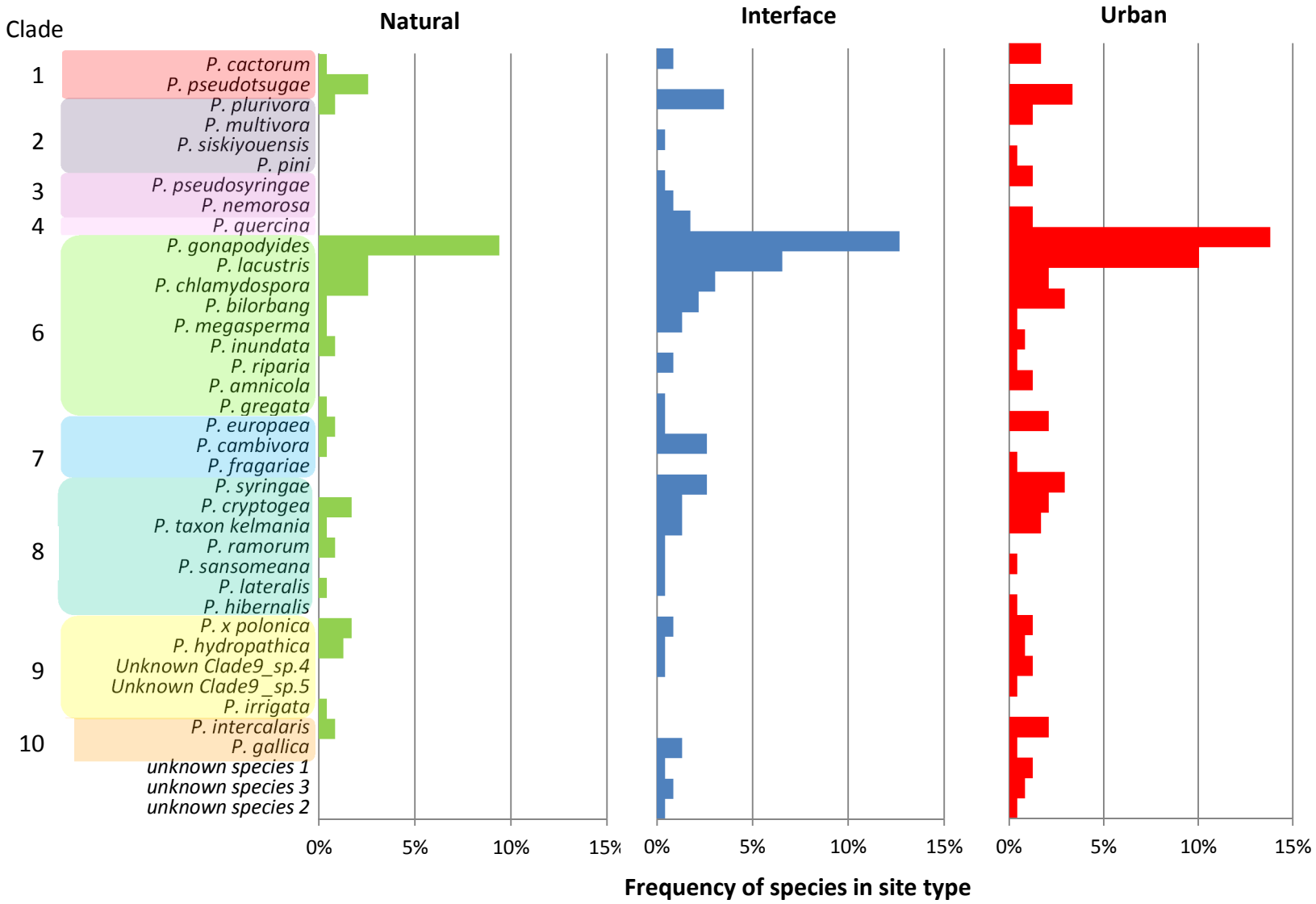


Figure 3.4: Frequency of each *Phytophthora* species by site type in fall and spring using baiting in the lab for water and soil and meta DNA-barcoding on soil (stream baiting not shown). Clade nomenclature: Cooke *et al.* (2000) and Blair *et al.* (2008).

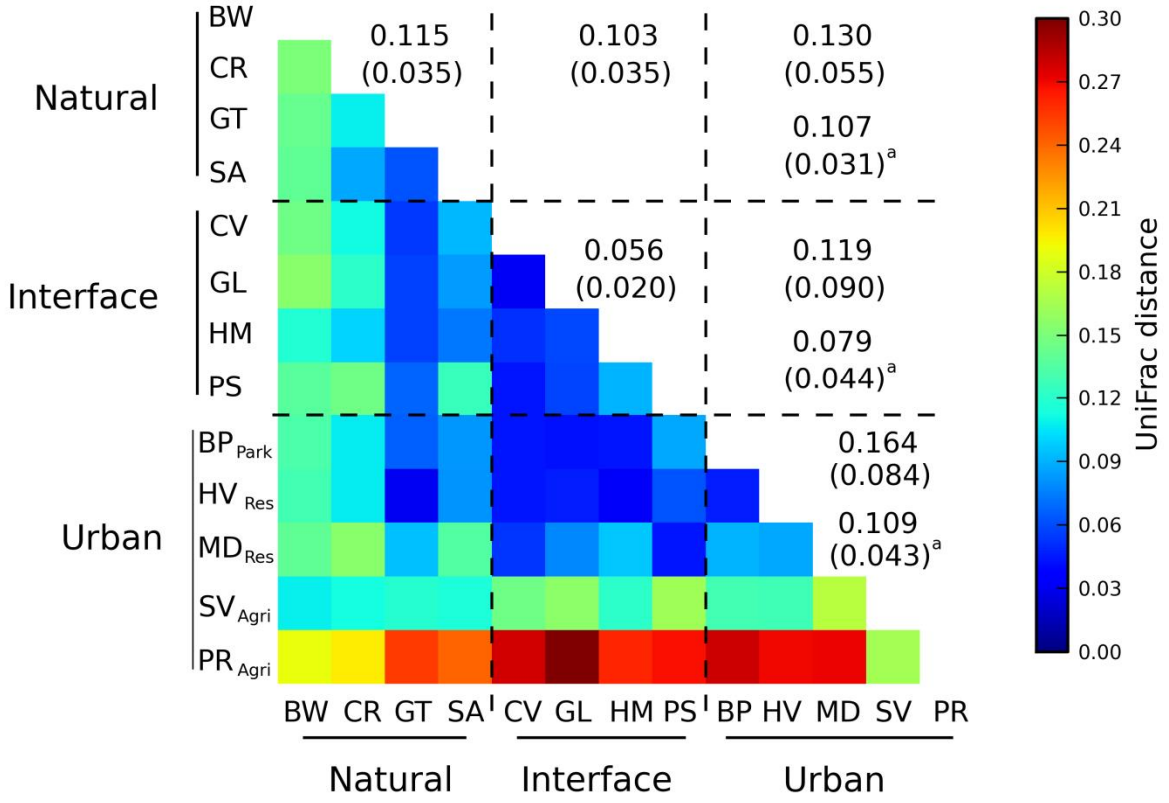


Figure 3.5: Weighted UniFrac distance between sites. Lower portion of the diagonal represents comparisons between sites within site types (natural, interface or urban). Values above the diagonal are the average of the sites with the standard deviation in brackets. [a] is the average and standard deviation without PR in the calculation. Urban sites are designated with Park, Res for sites within residential areas, and Agri for sites within agricultural areas.

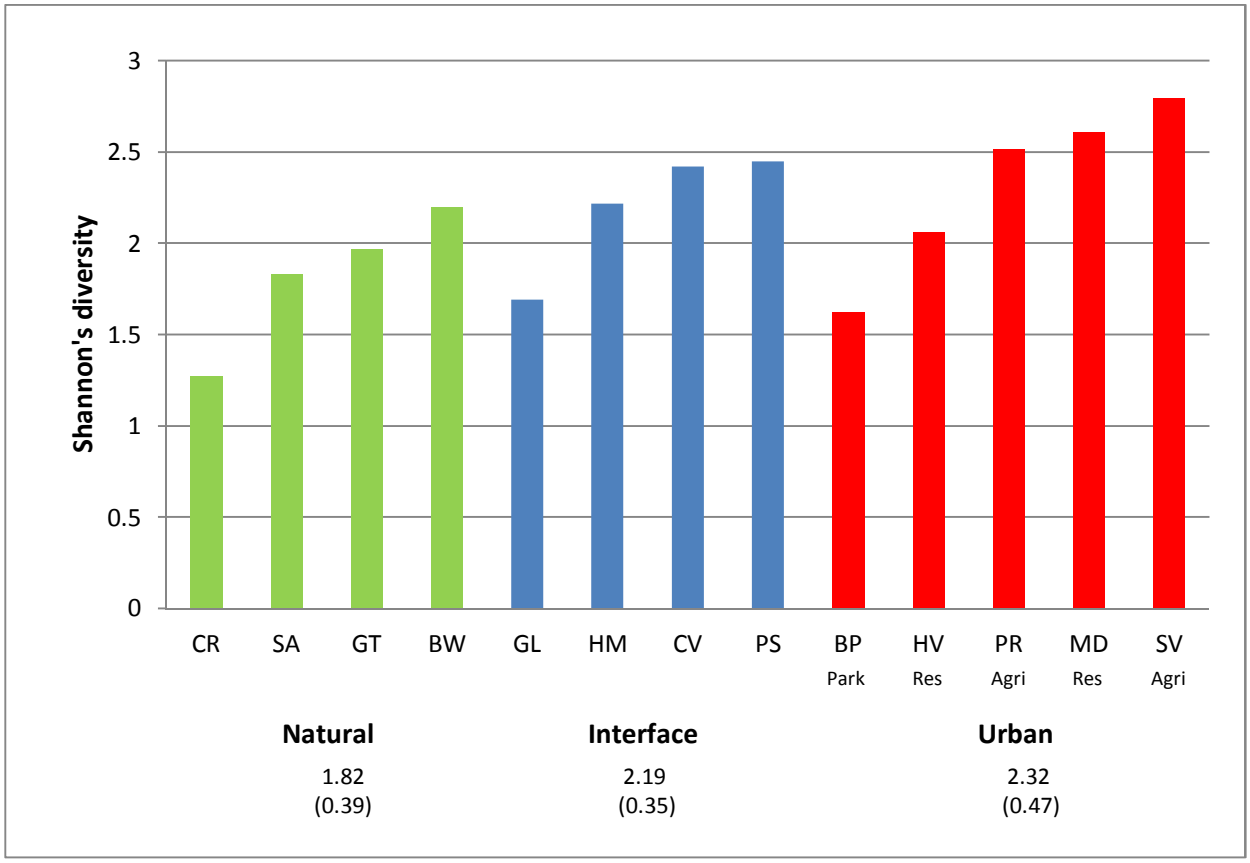


Figure 3.6: Shannon's diversity of sites. Values below site type are average Shannon's diversity for each site type with standard deviation in brackets. Urban sites are designated with Park, Res for sites within residential areas, and Agri for sites within agricultural areas.

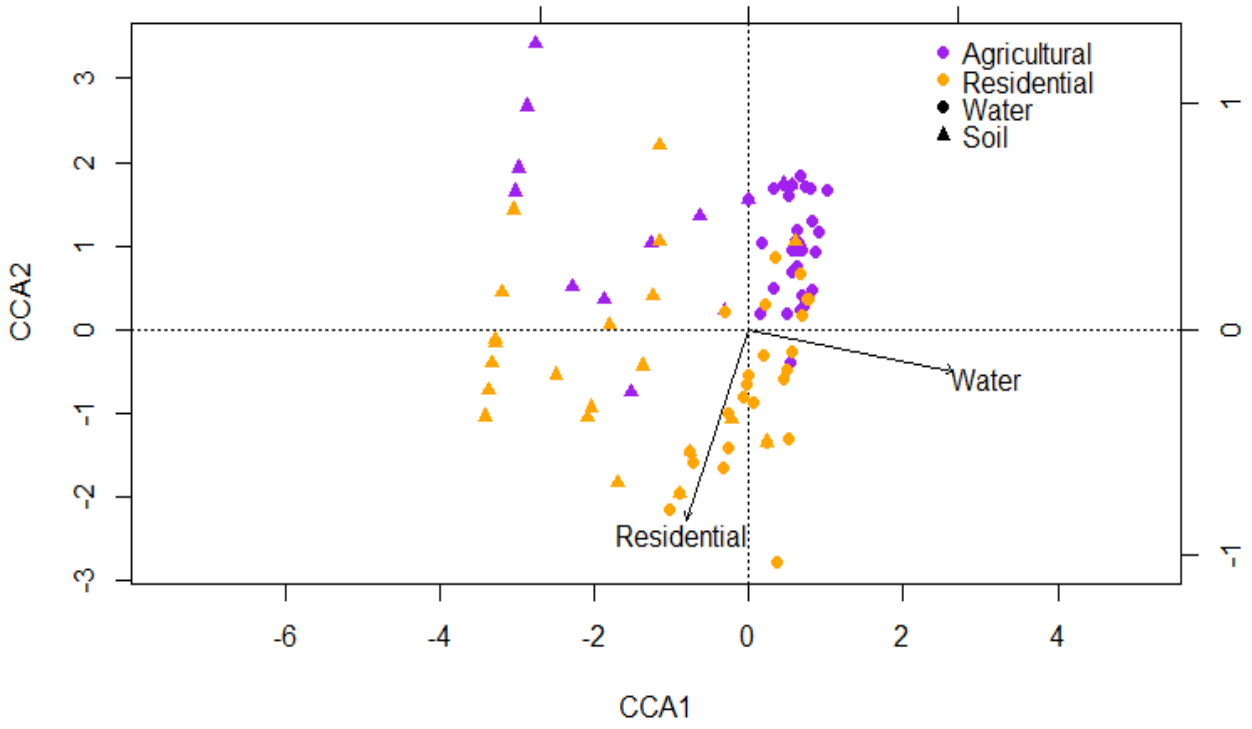


Figure 3.7: The relationship between urban site type (residential or agricultural), sample substrate (soil or water) and species presence and absence patterns. Results are based on a Canonical Correspondence Analysis (CCA). Data includes samples from soil and water baited in the lab and meta DNA-barcoding on soil only. Both site type and substrate were significantly associated with species composition.

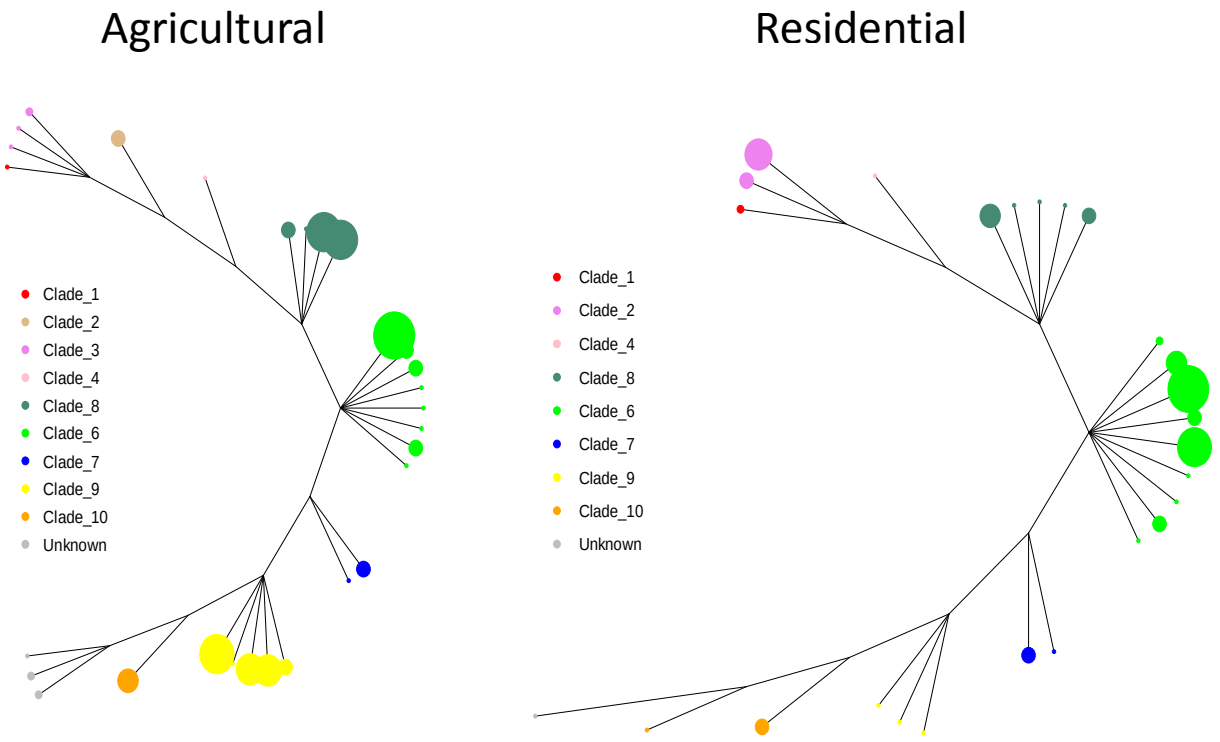


Figure 3.8: Representation of the abundance of *Phytophthora* clades (circle diameter) in urban sites that are in residential or agricultural areas. The dendrograms represent clades as defined using sequences in *Phytophthora*db.org

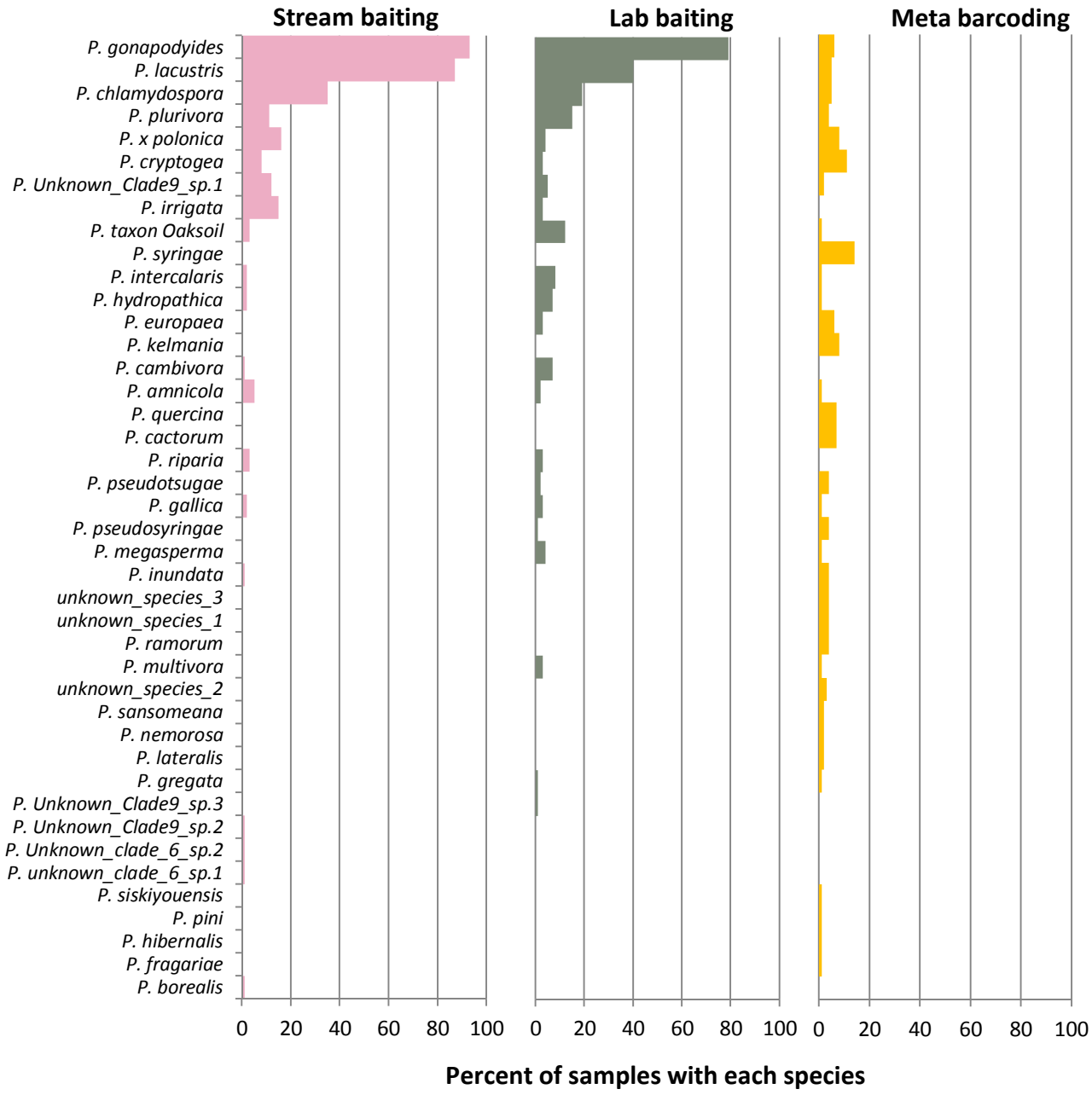


Figure 3.9: *Phytophthora* species present in south west BC and their frequency based on collection method. Stream baiting was done in fall (2011) in urban sites only, lab baiting was done on water and soil collected in fall (2012) and spring (2013) and Meta DNA-barcoding was done on soil collected in fall (2012) and spring (2013).

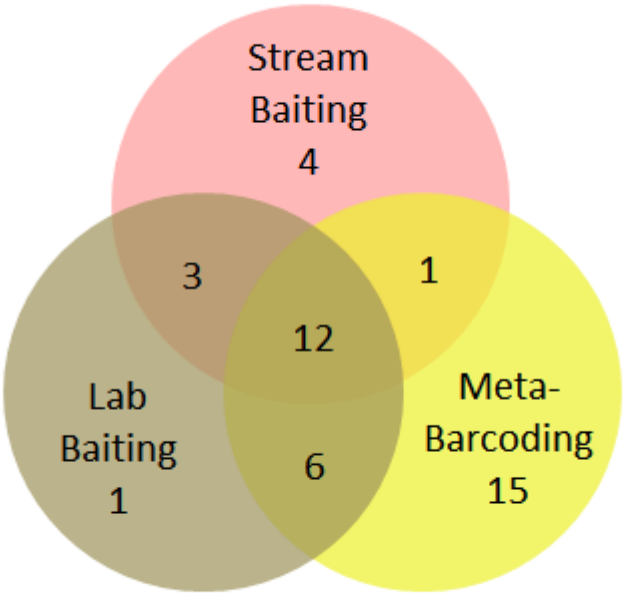


Figure 3.10: Unique and shared species by method: 1) Stream baiting, conducted only in urban sites in 2011, 2) Baiting soil and water samples in the lab, conducted on all samples in 2012 and 2013 and 3) DNA meta-barcoding, conducted on soil samples in 2012 and 2013. There were a total of 42 species found with all methods combined.

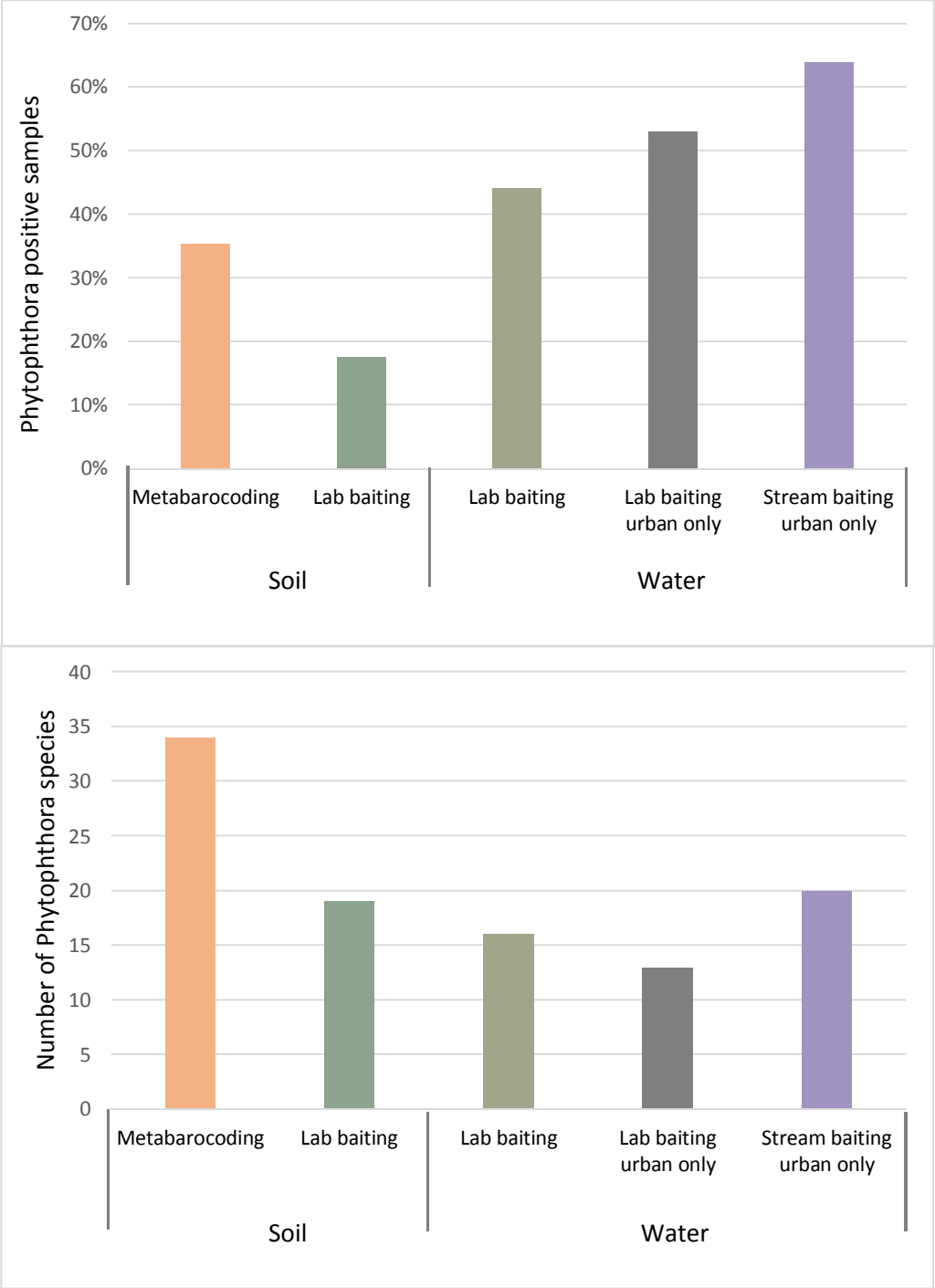


Figure 3.11: Comparison of *Phytophthora* positive samples and number of species by method and by substrate. DNA meta-barcoding was done on soil only in 2012 and 2012, baiting in the lab was done on soil and water in 2012 and 2013, and stream baiting was done in five urban streams in 2011.

Chapter 4: Mitotic recombination and a two-speed genome drive evolution in asexual lineages of the sudden oak death pathogen *Phytophthora ramorum*

Introduction

The previous two chapters of this thesis suggest that anthopogenic activities are an important pathway of introduction for Oomycetes and *Phytophthora* species. However, it is unknown how many introduced species will become established and invasive. Furthermore, it is unknown if particular characteristics of *Phytophthora* species make them better invaders. In order to understand success in invasive pathogens and to predict risk of invasion it is important to understand the evolution of known successful invasive alien pathogens.

Invasive alien tree pathogens are responsible for some of the most devastating forest disease epidemics. Successful alien tree pathogens can spread over thousands of km within a few decades invading new habitats and attacking new hosts under new environmental conditions. There are numerous examples of invasive tree pathogens that became successfully established and created ecosystem-wide changes by nearly eliminating their hosts (Loo, 2009). But there is a genetic paradox to many of these invasions: pathogen populations often undergo genetic bottlenecks following founder events caused by the introduction of a limited number of propagules (Pérez *et al.*, 2006, Gladieux *et al.*, 2015). This can result in a loss of genetic diversity, often accompanied by the loss of sexual reproduction. In heterothallic species (that require two mating types to complete a successful sexual cycle) founder events reduce the

likelihood of sexual reproduction (Fry *et al.*, 1992; Garbelotto, 2008; Brasier, 2012). Sometimes the absence of one mating type completely eliminates sexual reproduction, resulting in the proliferation of clonal lineages. The ability to reproduce sexually is considered an important life trait that can impact a pathogenôs ability to overcome host resistance (McDonald & Linde, 2002), adapt to new environments (North *et al.*, 2011), and contribute to invasiveness (Philibert *et al.*, 2011). Sexual reproduction can generate genotypic diversity, facilitate rapid dispersal of beneficial mutations, produce novel combinations of beneficial mutations and purge deleterious ones (Heitman, 2006; Heitman *et al.*, 2013; Calo *et al.*, 2013). Sexual reproduction can substantially speed up the rate of adaptation in diploid species, because sexual populations can produce the fittest homozygous genotype through segregation and mating of heterozygous individuals (Mandegar & Otto, 2007).

Despite the reduction in genetic diversity and frequent loss of sexual reproduction following an introduction, alien plant pathogens are often successful invaders and retain the ability to adapt, overcome host resistance and successfully establish in new habitats. One explanation proposed for this ability to adapt in invasive species is the generation of novel genotypes without sexual recombination. Mitotic recombination (MR) is one mechanism that can uncover previously hidden genetic variation without sexual genetic segregation and could help explain the genetic paradox of invasions. In an asexually reproducing diploid lineage, two rare mutational events must occur before a beneficial mutation can become fixed: one producing a heterozygous carrier and the second, at the same site, converting a heterozygous to a homozygous carrier. In simulation studies, MR increased the spread of beneficial mutations in asexual populations, allowing adaptation to proceed as fast as in sexual populations (Mandegar & Otto, 2007).

A second mechanism that could help explain the genetic paradox of invasion is the presence of a two-speed genome in invasive alien pathogens (Raffaele & Kamoun, 2012; Dong *et al.*, 2015). Genes encoding virulence effectors that modulate host plant processes are often associated with genome regions enriched in repetitive sequences and transposable elements, which can serve as a cradle for adaptive evolution (Dong *et al.*, 2015). These genome regions are often gene-sparse, but rich in effector genes in contrast to regions that are gene-dense and rich in core orthologous genes (Haas *et al.*, 2009; Raffaele *et al.*, 2010b,a). These genome regions have been found to evolve faster than the gene dense regions and may have facilitated host jumps and rapid adaptation in pathogens (Haas *et al.*, 2009; Raffaele *et al.*, 2010b,a).

Phytophthora (Oomycetes, Pythiaceae) is a genus of that comprises some of the most destructive invasive plant pathogens affecting crops and forests. *Phytophthora infestans* (Mont.) de Bary caused one of the worst agricultural disease outbreaks in human history and has been responsible for a famine that resulted in large-scale human migration. *Phytophthora ramorum* Werres, De Cock & Man in 't Veld is an invasive pathogen with a broad host range that causes sudden oak death, sudden larch death and Ramorum blight on horticultural plants and trees (Rizzo *et al.*, 2002; Brasier & Webber, 2010). Four lineages (EU1, EU2, NA1 and NA2; named after the continent they were first observed in) with low genotypic diversity and apparent lack of sexual reproduction have spread in the USA, Europe and Canada (Prospero *et al.*, 2009; Grünwald *et al.*, 2009, 2012; Vercauteren *et al.*, 2010; Goss *et al.*, 2011b). Since *P. ramorum* is diploid and heterozygosity is observed within lineages, it is suspected that the ancestral source population

reproduced sexually (Goss *et al.*, 2009). However, the absence of recombination in the extant lineages raises questions about adaptation and diversification in *P. ramorum*.

Mitotic recombination and two-speed genomes have been reported in *Phytophthora*s (Haas *et al.*, 2009; Raffaele *et al.*, 2010b,a; Lamour *et al.*, 2012; Kasuga *et al.*, 2016). However, the extent to which these factors play a role in genome evolution and adaptation in *P. ramorum* lineages is not known. In order to determine how genome features and mechanisms drive evolution and adaptation in asexual populations, we sequenced the genomes of 107 *P. ramorum* samples representing the four known lineages and known geographical range of the pathogen. We found evidence that extensive MR is reshaping the genomic profiles of the asexual lineages generating large-scale non-synonymous changes that could impact the phenotypic outcome. Mitotic recombination can lead to the expression of recessive or new alleles which are subject to selection. We show evidence for a selective sweep of a genomic region rich in putative pathogenicity genes that had undergone an MR event. Mitotic recombination events are associated with transposons and low gene density suggesting that they originate in fast evolving genomic regions. We also observed non-core and lineage specific regions of the genome which were enriched in genes associated with host-pathogen interactions and transposable elements. These non-core regions are characterized by rapid gene loss, a distinct nucleotide composition, and a high repeat and transposon content also suggesting faster evolving genomic regions. A two speed genome appears to drive lineage divergence and population evolution in *P. ramorum*.

Methods

Genome resequencing

A total of 107 individuals (Table S4.1) were selected for genome resequencing using the Illumina sequencing technology (*Appendix B, DNA extraction, genome sequencing, SNP extraction and filtering*). We selected representatives of the four lineages: 38 individuals of NA1, 17 individuals of NA2, 46 individuals of EU1 and 6 individuals of EU2. The individuals selected covered the chronological range (1995-2012) and the geographical range (Canada, the United States, and Europe) of the epidemics. Illumina reads were mapped onto the *P. ramorum* reference genome (version 1.1; DOE Joint Genome Institute, Walnut Creek USA) (Tyler *et al.*, 2006) and a SNP set was obtained for the full set of individuals (*Appendix B, DNA extraction, genome sequencing, SNP extraction and filtering*). The illumina reads of two representative individuals of each of the four clonal lineages were assembled for inter-lineage comparisons (*Appendix B,* De novo *genome assemblies*).

Phylogenetic analysis

A neighbor-joining phylogenetic analysis based on Euclidean distances was done with all individuals using the full SNP set to determine inter and intra-lineage diversity patterns in *P. ramorum* (*Appendix B, Phylogenetic analysis*). In addition, a set of genes under neutrality determined with Tajimaגs D values (Tajima, 1989) was used to reconstruct a maximum likelihood phylogeny to observe evolutionary patterns in the coding genome (*Appendix B, Phylogenetic analysis*). Divergence time between the lineages was estimated using BEAST v1.7.5 (Drummond *et al.*, 2012) (*Appendix B, Divergence time between lineages*).

Runs of homozygosity

We searched for runs of homozygosity (ROH) in each individual and within lineages. Using VCFtools (Danecek *et al.*, 2011), the multi-sample VCF file was separated into independent files for each individual. Using a custom python script (Python 2.6), the VCF files were searched for stretches of homozygous sites uninterrupted by heterozygotes. The length of these stretches was compared between individuals within a lineage to identify those with extensive ROH relative to the population and to identify scaffolds with extensive ROH relative to other scaffolds. For individuals that had extensive ROH, VCFtools (Danecek *et al.*, 2011) was used to extract observed number of heterozygotes and homozygotes for each SNP site for individuals with and without ROH for comparison. The percent of sites homozygous for either allele and the percent of heterozygous sites were calculated for individuals with ROH and representative individuals without ROH for each lineage. Individuals were also checked for genome-level copy number variation including aneuploidy (*Appendix B, Genome-level copy number variation*).

Effects on protein content from conversion to homozygosity

To determine the potential effects of conversion to homozygosity, the genes on the scaffolds with the ROH pattern were reconstructed for each strand of each individual (*Appendix B, Phylogenetic analysis*). Using a custom Python script, the genes were translated to proteins and the protein translations were compared for individuals with and without ROH to determine the number of amino acid changes between alleles for each group. Gene Ontology (GO) analysis was done on proteins found in the ROH regions using the model-based gene set analysis

approach (MGSA) in R (R Core Team, 2017), which analyzes all inferred GO terms at once by embedding them in a Bayesian network and provides summarized views of core biological processes (Bauer *et al.*, 2011).

Identification of a pathogenicity gene hot spot

A long run of homozygosity that occurred in scaffold seven in all NA1 and eight EU1 individuals was further investigated to identify protein candidates potentially involved in host-pathogen interactions. The protein set was extracted from the reference genome (version 1.1) (Tyler *et al.*, 2006) and re-annotated using Blast2GO (Conesa *et al.*, 2005). Putative effectors (van Damme *et al.*, 2012) and other proteins potentially involved in pathogenicity (carbohydrate active enzymes, some transporters, peptidases), and transposons (or transposon-like elements) were counted and mapped along the scaffold using GenoPlotR (Guy *et al.*, 2010) in R (R Core Team, 2017). MGSA (Bauer *et al.*, 2011) was used to identify protein sets that were enriched in scaffold seven and a chi-square was used to test if putative effectors were located on scaffold seven more often than expected.

Mitotic recombination breakpoints

To determine possible causes or mechanisms for mitotic recombination, we searched for scaffolds with putative recombination breakpoints. Mitotic recombination breakpoints were identified as regions where a switch from heterozygosity to homozygosity occurred. The region around each recombination breakpoint was scanned for gene content using annotations from the *P. ramorum* reference genome (version 1.1) (Tyler *et al.*, 2006). Intergenic distances (5ô to 5ô

and 5ô to 3ô) were calculated on the scaffolds with potential breakpoints and graphed on a heatmap using a custom Python script (Saunders *et al.*, 2014). For comparison, ten data sets, each with nine 15 kbp genomic regions with no potential breakpoints, no ROH, and no stretches of ambiguous nucleotides were randomly generated. Gene content and intergenic distances were determined for the 15 kbp regions. An analysis of variance (ANOVA) was used to compare the observed and random data sets using the package RCommander (Fox, 2005) in R (R Core Team, 2017).

Testing for phenotype differences in individuals with ROH

To test for phenotypic differences resulting from ROH, we compared the fitness of eight EU1 individuals with ROH in scaffold seven, to eight EU1 individuals without ROH (Table S4.2). We assessed the pathogenôs ability to cause lesions on woody stems of two hosts: *Pseudotsuga menziesii* (Mirb.) Franco (Douglas-fir) and *Larix kaempferi* (Lamb.) Carr. (Japanese larch). We also tested the ability to infect detached *Rhododendron* sp. leaves, and to grow on sapwood agar (*Appendix B, Effects of ROH on phenotype*).

Core and non-core genomes of *Phytophthora ramorum* lineages

To identify core and lineage-specific genomic regions, we used the mpileup function of SAMtools (Li *et al.*, 2009) to assess coverage depth over 1 kb sliding windows of all 107 re-sequenced individuals over two *de novo* genome assemblies of each of the four lineages (*Appendix B,* De novo *genome assemblies*). A window was considered as non-core when a contiguous region representing more than 90% of the window size was missing in 50% of the

individuals of at least one of the four *P. ramorum* lineages (thresholds of 75% and 99% were also tested in order to obtain a measure of variability in windows coverage [data not shown]). Lineage-specific, or non-core genomic regions shared between only two or three lineages were identified by reciprocal mapping to the *de novo* assemblies of each lineage. Groups of homologous non-core windows were identified with reciprocal BLASTn searches (E-value < 1e-20) within and between lineages followed by graph clustering using a TCL implementation of the Deep-First Search algorithm (identity cutoff = 40%; overlap cutoff = 100nt) (Kozik *et al.*, 2003; Joly *et al.*, 2010).

A neighbor-joining tree was reconstructed from a distance matrix of similarity based on non-core genome regions (as in *Appendix B, Divergence time between lineages: Phylogenetic analysis of one-to- one orthologs*). The extent of gene gain or loss was determined by attributing costs to gain and loss events and minimizing total cost (Maximum parsimony criterion, (Mirkin *et al.*, 2003)). GC content, transposon and repeat content were compared for the non-core and core genome.

Lineage-specific proteins and evolution of gene family size

Genome assemblies obtained for the four *P. ramorum* lineages and *P. lateralis* (CBS_168.42) (Feau *et al.*, 2016) were annotated using AUGUSTUS ver. 2.7 (Keller *et al.*, 2011) trained with models from *P. sojae, P. infestans* and *P. ramorum* available on the Joint Genome Institute (JGI) server (Grigoriev *et al.*, 2012). The protein models obtained were combined with those of five *Phytophthora* genomes and submitted to a clustering process with OrthoMCL (Li *et al.*, 2003)

(*Appendix B, OrthoMCL analysis and evolution of gene family size*). The resulting OrthoMCL clusters were filtered to minimize the confounding effect of truncated proteins and gene mis-predictions and CAFE (Computational Analysis of gene Family Evolution)(De Bie *et al.*, 2006) was used to infer gene family expansions/contractions (*Appendix B, OrthoMCL analysis and evolution of gene family size*). Genes encoding secreted proteins were predicted using the pipeline described in (Dhillon *et al.*, 2015). Putative effectors were identified using EffectorP (Sperschneider *et al.*, 2016).

Adaptation in the effector repertoire of Phytophthora ramorum lineages

RxLR and Crinkler effector genes were extracted from the *de novo* genome assemblies of each of the four lineages (*Appendix B, Genes encoding effectors*). Effector groups were clustered with OrthoMCL (Li *et al.*, 2003), aligned and dN/dS was estimated with Codeml implemented in PAMLV4.0 (Yang & Nielsen, 1998) as described in *Appendix B* (*Genes encoding effectors*). Observed dN/dS were compared to dN/dS calculated on a random set of genes and the set of sequence homologs of the 458 proteins of the core eukaryotic CEGMA dataset (Parra *et al.*, 2007) from the *P. ramorum* reference genome (*Appendix B, Genes encoding effectors*).

Results

Mitotic recombination generates diversity and novel phenotypes

Extensive genotypic diversity in clonal *P. ramorum* lineages. The four lineages of *P. ramorum* are well resolved and divergent based on a neighbor-joining tree (Fig. 4.1) using 485 327 bi-

allelic SNPs (*Appendix B, DNA extraction, genome sequencing, SNP extraction and filtering*; Table S4.3) and a maximum likelihood phylogenetic analysis (*Appendix B, Phylogenetic analysis;* Fig. S4.5). We estimate that these lineages diverged approximately between 0.75 million years ago (MYA) for the split between EU1 and NA1 and 1.3 MYA for the split between EU2 and the other lineages (*Appendix B*, *Divergence time between lineages*). We observed an unexpected amount of intra-lineage diversity; several individuals were divergent from the main group including a cluster of eight individuals in EU1 (Fig. 4.1; Fig S4.5). These individuals were characterized by a high number of loci that were homozygous in one or more individuals but heterozygous in the rest of the members of the lineage (Table S4.2 and S4.3). These homozygous sites were clustered in runs of homozygosity (ROH) ranging in size from 60 to 339 Kb and were most abundant within the EU1 lineage (30% of the individuals) (*Appendix B, Effects of ROH on genotype*). The ROHs were not randomly distributed throughout the genome; rather entire scaffolds or large portions of scaffolds were entirely homozygous in some individuals, interrupted by only a few heterozygous loci (Fig. S4.6). In addition, four individuals also exhibited plasticity in chromosome numbers, potentially contributing to the observed diversity (*Appendix B, Genome-level copy number variation,* Table S4.4 and Fig. S4.7). One NA1 individual with CCNV was amongst the more diverged individuals in the neighbor-joining tree (Fig. 4.1).

Protein content differences within clonal *P.ramorum* lineages. The ROHs observed have the potential to alter the protein content within clonal lineages if individuals with ROH lost one of two haplotypes with non-synonymous differences. We observed this pattern in 52% of the 5 172 genes located within scaffolds that contained ROHs (*Appendix B, Effects of ROH on genotype;*

Table S4.5). In 21% of those genes, there were five or more non-synonymous heterozygous nucleotide positions in individuals without ROH compared to those with ROH and up to 35 amino acid changes were observed in some genes (Table S4.5). Such large scale differences have the potential to generate substantial phenotypic changes in individuals with ROH.

ROH can uncover recessive mutations and be subject to selection. Scaffold seven had the longest ROH, covering approximately 650 Kb. All 38 NA1 individuals have this ROH (Fig. 4.2A) suggesting a selective sweep. Interestingly eight (out of 38) EU1 individuals possess the alternate haplotype suggesting that a similar selection event could be occurring in that lineage. This ROH was absent in all NA2 and EU2 individuals. The secretome differed on scaffold seven compared to the rest of the genome: the highest percentage of secreted proteins occurred on scaffold seven (24.6%) when compared to the 25 largest scaffolds (Fig. S4.8; chi-square = 45.127, $p < 0.001$). Fifteen percent of the genes on that scaffold were annotated as putative effectors, 5% of which encode necrosis inducing proteins or NPP1s; this represents the highest proportion of NPP1s on any scaffold, and is significantly higher than expected across the genome (chi-square = 79.261, $p < 0.001$). The NPP1 encoding genes occur in a cluster within the homozygous region, as do most of the putative effectors. Scaffold seven also comprises several genes that act in some plant pathogens as virulence factors such as carbohydrate active enzymes, peptidases and sugar transporters (*Appendix B, Effects of ROH on genotype*, Fig. S4.8, Table S4.6).

Transposable elements and low gene density associated with mitotic recombination. Mitotic recombination breakpoints were identified as locations where there was a transition from

heterozygosity to homozygosity in the regions affected by ROH. Transposable elements (TEs) were found in close proximity to mitotic recombination breakpoints in seven out of nine instances, which was four times more frequent compared with ten randomly generated datasets (Fig. 4.3, Table S4.7). The mitotic recombination breakpoints occurred in regions with a lower gene density than the rest of the genome. In scaffolds with mitotic recombination breakpoints, the intergenic distances measured as the 3ô flanking intergenic region (FIR) and the 5ô FIR were significantly larger than the intergenic distances in random sets of scaffolds (Table S4.7). Five percent of the genes had a 5ô FIR greater than 1×10^4, and seven percent of the genes had a 3ô FIR greater than 1×10^4 (Fig. 4.3), compared to two to four percent for the 5ô FIR and three to six percent for the 3ô FIR in the random data sets (Fig. 4.3, Table S4.7).

Phenotypic characterization of individuals affected by ROH. Lesion size on larch and Douglas-fir caused by EU1 individuals with ROH on scaffold seven were larger than lesions caused by EU1 individuals without ROH. Individuals with ROH produced longer lesions on larch log sections (average 57.4 mm ± versus 38.0 mm ±; Fig. S4.9 and S4.10; Kruskal-Wallis chi-square test= 4.412; p = 0.036) and slightly longer lesions on Douglas-fir log sections four weeks post-inoculation than those produced by non-ROH individuals (average 56.6 mm ± versus 54.7 mm; p = 0.753). The trends were reversed on both host species eight weeks post inoculation (*Appendix B, Effects of ROH on phenotype*).

The core and non-core genomes differ markedly among *Phytophthora ramorum* lineages

Reduced non-core genome in EU2. The four lineages of *P. ramorum* share a core genome representing 98.2% to 99.2% of the full content, leaving a non-core ranging from 0.34 Mbp (EU2) to 0.90 Mbp (NA2) (Fig. 4.4A). The non-core genome shared between EU2 and the other lineages was smaller (0.007 Mbp) than the non-core genome shared among the other lineages (1.7 Mbp; Fig. 4.4B). This corresponds to 0.16% to 0.37% of the EU2 non-core genome that was not found in the genomes of the three other lineages and 1.24% to 1.68% of the EU1, NA1 and NA2 non-core genomes that is missing in the EU2 genome (Fig. 4.4C). The neighbor-joining tree based on lineage similarity of non-core regions was topologically incongruent with the *P. ramorum* lineage phylogeny reconstructed from the core gene set (Fig. 4.4C and S4.11). This strong discordance between the two phylogenies suggests differences in the rate and extent of genome fluctuation (regions gain and losses) among lineages in the non-core genome, or a two-speed genome. We estimated that most of the fluctuations were the result of region losses in EU2 (68.6%), whereas they resulted from gains in the other lineages (Fig. 4.4C).

Evidence for a two-speed genome. The nucleotide and gene composition of the non-core genome in the four *P. ramorum* lineages were clearly distinct from the rest of the genome. The non-core region had a significantly lower G+C content (49.4% for non-core *vs.* 54.0% for core averaged over the four lineages; paired t-test value = 25.79, $P < 0.001$), a higher repeat content (46.3% for non-core vs. 15.5% for core averaged over the four lineages) coupled with a significant enrichment in TE-like gene sequences ($G2 = 19.3$, $P < 0.001$), four to seven time less

genes per Mb of DNA, and a higher number of unannotated genes (Table S4.8). Genes in the non-core genome were shorter and had a lower G+C content than those from the core genome (average = 54.9% *vs.* 57.9%; t = 5.14, $P = 0.01$) and genes found in repetitive regions (57.3%; t = 3.87, $P < 0.05$) (Table S4.8). A large number of genes showed strong bias in codon usage (from 82% in NA2 to 100% in EU2), reflecting a clear preference towards codons ending with T or A nucleotides (Fig. S4.12). In addition, one third to half of the genes found in the non-core regions (35% in EU1 and EU2, 41% in NA1 and 50% in NA2) share paralogs with genes in the core genome and show high mutation loads making them likely degenerated copies (*Appendix B, Core and non-core genomes*; Fig. S4.13).

Among the protein-coding genes identified in non-core genome regions (17 in EU2 to 44 in NA2), we did not find significant enrichment in genes encoding secreted proteins ($G2 = 0.15$, $p = 0.98$) or effectors. However, we found presence/absence patterns for effector-like proteins that are part of the RxLR effector family (5 RxLR gene models; Table S4.9) and for genes with functions related to degradation of the plant cell wall (proteases and ester hydrolases; Table S4.9). Remarkably, none of the 17 non-core protein models predicted in EU2 encoded effector-like proteins (Table S4.9).

The non-core proteome of *P. ramorum* is enriched in plant cell wall modification enzymes. We identified a core proteome of 10,197 clusters representing 44,295 shared proteins between the four *P. ramorum* lineages, that have likely been conserved since the lineages split (Table S4.10). About 5.3% (541 clusters) of this core proteome is not shared with any other *Phytophthora* species and represents *P. ramorum* unique gene families (Table S4.10). We

identified only slight protein family expansion within each *P. ramorum* lineage followed by rapid divergence, resulting in proportions of 2.5 to 3.2% lineage-specific proteome (i.e. not shared with any other *Phytophthora* species or between *P. ramorum*-lineages) (Fig. S4.11). GO terms associated with peptidase activities (lineages EU1, EU2 and NA1) and pectin and glucan modification (EU1, NA1 and NA2), were enriched in the unique protein models suggesting that rapidly diverging proteins unique to each of the four *P. ramorum* lineages are enriched in functions related to plant cell wall modification and/or degradation (Table S4.11).

Evidence of host driven adaptation in *Phytophthora ramorum* lineages

***Phytophthora ramorum* effectors under positive selection.** Host driven adaptation between the lineages was suggested by positive selection in 20% of the RxLR and 38% of the Crinkler effector genes (CRN) compared to only 1% of the CEGMA gene set and only 3% in a random set of genes in the *P. ramorum* genomes (Fig. S4.14). CEGMA (72%) and random genes (70%) were under strong negative selection (dN/dS < 0.3). The distribution of dN/dS ratios was significantly different between the effector and the CEGMA gene sets (F-value = 376.93, $p < 0.001$) as well as the between the effector and the random gene sets (F-value = 298.82, $p < 0.001$). We identified one CRN family that likely diversified within the evolutionary history of the *P. ramorum* lineages through duplication, recombination and episodes of accelerated nucleotidic evolution (*Appendix B, Genes encoding effectors*; Fig. S4.15).

Discussion

Phytophthora ramorum provides a classic example of a successful emerging pathogen with a small effective population size, limited to asexual reproduction. The clonal lineages are without apparent meiotic recombination either due to the uneven distribution of the mating types or the presence of reproductive barriers (Brasier & Kirk, 2001; Ivors *et al.*, 2006; Goss *et al.*, 2009). We highlight two mechanisms that can generate diversity and adaptation in these asexual lineages. First, we observed a large amount of variation within lineages that could impact the pathogenôs ability to adapt and strong evidence that mitotic recombination is at the origin of this variation, potentially affecting hundreds of genes simultaneously. We found compelling evidence that a sweep in one long ROH has affected a genome region enriched in putative pathogenicity-related genes. Remarkably, this appears to have occurred in two independent events in NA1 and EU1. Secondly, we observed accelerated evolution in lineage specific regions affecting genes putatively involved in host-pathogen interactions. These lineage-specific regions have a distinct nucleotidic composition, are repeat-rich, gene-poor and comprise a large number of unannotated proteins suggesting rapid evolution. It is likely that this two-speed genome is contributing to phenotypic differences observed among the clonal lineages, as suggested for other *Phytophthora* pathosystems (Haas *et al.*, 2009; Raffaele *et al.*, 2010a; Dong *et al.*, 2015).

Mitotic recombination appears to be a dominant mechanism driving evolution in the asexual *P. ramorum* lineages. This process is expected to cause a loss of heterozygosity in all loci distal from the breakpoint creating runs of homozygosity (Goodwin, 1997; Tischfield, 1997; Llorente *et al.*, 2008). ROHs resulting from mitotic recombination are much longer than meiotic

recombination, and not uniformly distributed throughout the genome, the pattern observed in this study (Lee *et al.*, 2009; Magwene *et al.*, 2011). Mitotic recombination has been suggested in *P. ramorum* (Ivors *et al.*, 2006; Goss *et al.*, 2009; Kasuga *et al.*, 2016) and in other *Phytophthora* species (Chamnanpunt *et al.*, 2001; Dobrowolski *et al.*, 2003; Lamour *et al.*, 2012; Hu *et al.*, 2013); however the extent of the effect shown here is much larger than previously reported. LOH was suggested as a contributing factor in generating diversity and adaptation in the broad-host-range pathogen *Phytophthora capsici* (Lamour *et al.*, 2012). LOH was found in progeny from crosses as well and in clonal individuals affecting up to 36.5% of the genome; the LOH affected more than 11 000 sites and was associated with changes in pathogenicity and in mating type. However only 17% of those sites were in coding regions and only 264 genes were predicted to be affected by non-synonymous changes (Lamour *et al.*, 2012), a far lower number than the 5172 genes with non-synonymous changes that were predicted in our study. In *P. sojae*, mitotic gene conversion occurred in a genomic region coding for an avirulence/virulence gene potentially leading to virulence differences in new pathogen races (Chamnanpunt *et al.*, 2001). In *P. ramorum*, mating type switches observed in EU1 individuals could be explained by mitotic gene conversion (Vercauteren *et al.*, 2011). Mitotic recombination is also observed in fungi and has been associated with increases in diversity or with adaptation to environmental changes. For example, mitotic recombination resulting in loss of heterozygosity was the cause of high genotypic variation in the clonal species *Batrachochytrium dendrobatidis* responsible for the amphibians chytridiomycosis epidemic (James *et al.*, 2009). In *Candida albicans* LOH rates increase proportionally to exposure to oxidative stress, heat stress, and antifungal drugs (Forche *et al.*, 2011). In *Saccharomyces cerevesiae* a trade-off between outcrossing and mitotic recombination seems to play a role in shaping genome architecture in response to nutrient stress

(Magwene *et al.*, 2011). The variation generated by mitotic recombination in *P. ramorum* could be a response to host or environmental stresses and could influence adaptation to novel environments. The high number of genes with nonsynonymous changes in regions affected by ROH in *P. ramorum* could be the result of phenotypes that were selected for on different hosts, and are retained in the population.

One of the proposed advantages of mitotic recombination is that it facilitates the rapid evolution of asexual eukaryotic microorganisms facing changing environments (Forche *et al.*, 2011; Magwene *et al.*, 2011). Conversion of genomic regions to homozygosity may result in the expression of new or recessive alleles, and can improve the efficiency of selection on both beneficial and deleterious alleles (Mandegar & Otto, 2007). This can be particularly important for a pathogen that is reproducing clonally as new mutations or recessive alleles can remain masked by dominant alleles. Mandegar & Otto (2007) theorized that mitotic recombination can drastically increase the rate of adaptation in asexual populations by reducing the time between the emergence of beneficial mutations in a heterozygote and the appearance of a mutant homozygote; under some scenarios adaptation can occur at the same rate as in sexual populations, or even faster if the rate of mitotic recombination is very high. In lab experiments with diploid isogenic individuals of *Aspergillus nidulans*, increased rates of evolution were found in individuals that built up recessive mutations in diploid genomes, which were later expressed after individuals reverted to a haploid state (Schoustra *et al.*, 2007). A similar result was found with an individual which remained diploid, but had likely undergone a loss of heterozygosity in part of the genome, likely due to mitotic crossing over or non-disjunction.

We identified a genome region that appears to be a mitotic recombination hotspot in *P. ramorum*. ROH on scaffold seven affected all members of the NA1 lineage. The same region was affected in eight individuals in EU1 lineage, but the opposite haplotype was present. Two separate ROH events affecting the same genomic region suggests that genes on scaffold seven are likely under selection and that resulting phenotypes could have a selective advantage. To support this hypothesis, we showed that scaffold seven has a higher number of pathogenicity-related genes than expected and that LOH fixed non-synonymous mutations for a substantial proportion of these genes. Changes resulting from ROH in this genomic region could have direct effects in the interactions of the pathogen with its hosts. As *P. ramorum* is a generalist pathogen with over 100 host species, changes to the gene content or alleles could provide advantages in response to different environments. The slightly larger lesions caused by individuals with ROH in the early stages of infection on Japanese larch and Douglas-fir indicates faster colonization suggesting that these individuals could have a fitness advantage in overcoming host defences. There was no indication of loss of fitness associated with ROH on infected Rhododendron leaves or in artificial media. Further investigations into other host interactions or on sporulation efficiency could uncover other phenotypic differences associated with ROH.

Genotypic diversity and effective population sizes are important features for the successful establishment, persistence and adaptability of an invading population (Sakai *et al.*, 2001; Forsman, 2014; Gladieux *et al.*, 2015; Blackburn *et al.*, 2015). Rapid evolution in plant pathogens may be related to genome characteristics such as high transposon content and activity, genome size, or the physical arrangement of genes within the genome (Haas *et al.*, 2009; Raffaele *et al.*, 2010b,a; Judelson, 2012; Raffaele & Kamoun, 2012; Jiang & Tyler, 2012; Moller

& Stukenbrock, 2017). We hypothesize that the effective population size of *P. ramorum* has increased due to mitotic recombination triggered by TE activity and low gene density. The higher frequency of mitotic recombination breakpoints in regions with low gene density and physically associated with transposons suggests that these elements could be associated with a mechanism generating genotypic diversity in *P. ramorum* populations. In the asexual plant pathogen *Verticillium dahliae,* pathogenicity has been hypothesized to evolve through recombination between highly similar copies of retrotransposons which may generate diversity by facilitating mitotic crossover events and creating new beneficial genetic combinations (de Jonge *et al.*, 2013; Faino *et al.*, 2016). TEs have been suggested as the driving force behind adaptive evolution of *V. dahliae* (Faino *et al.*, 2016). TEs have been associated with large chromosomal rearrangements in the wheat pathogen *Zymoseptoria tritici* also suggesting their in role adaptation (Plissonneau & Stürchler, 2016; Moller & Stukenbrock, 2017).

Non-core genome regions in pathogenic micro-organisms generally contain effectors that may be involved in ecological adaptation (Hatta *et al.*, 2002; de Jonge *et al.*, 2013). These effectors are often adapted specifically to interact with host proteins leading to host or niche specialization and often result in speciation. As a consequence, the selective pressures acting on effectors can be especially efficient since having the wrong set of genes has a high fitness cost (Giraud *et al.*, 2010). Effectors that are no longer needed following a host jump are either eliminated or degenerated (Dong *et al.*, 2015). The positive selection and evolutionary history in effector genes in the *P. ramorum* lineages are further evidence of a two-speed genome where genes and regions of the genome involved in host-pathogen interactions are evolving faster than the rest of the genome. In addition, the *P. ramorum* EU2 lineage showed a striking reduction in gene content,

particularly in genes putatively involved in host-pathogen interactions. Although the non-core genomic regions were not significantly enriched in putative effectors, most of those found were missing in EU2 indicative of niche differentiation and speciation. Specialization of EU2 to Japanese larch (Larix *kaempferi*) (Brasier & Webber, 2010; Van Poucke *et al.*, 2012) and adaptive differences observed for this lineage could be related to the reduced non-core genome. Localization of race specific effectors in lineage-specific regions was suggested as a contributing factor in niche adaptation and virulence in *Verticillium dahliae* (de Jonge *et al.*, 2013). Lineage-specific regions were also not enriched in effectors compared to the non-core genome; however, plant-induced genes were upregulated in the lineage specific transcriptome (de Jonge *et al.*, 2013). Genes involved in pathogenicity are often found clustered together in less conserved, gene poor regions of the genome such as subtelomeric regions usually enriched in repeated elements or transposons (Haas *et al.*, 2009; Raffaele *et al.*, 2010a; Rouxel *et al.*, 2011; Bopp *et al.*, 2013; Dutheil *et al.*, 2016). In our study, not only were non-core regions enriched in genes associated with plant-pathogen interactions, they were also enriched in TE-like gene sequences, some of which were found in tandem with effector genes. This suggests that the non-core genome plays an important role in genome evolution and adaptation, and in shaping differences among the lineages. TEs were also associated with putative effectors in *Leptosphaeria maculans*, and lineage-specific effector genes in *Zymoseptoria tritici* (Rouxel *et al.*, 2011; Plissonneau & Stürchler, 2016). In *Magnaporthe oryzae,* transposons were linked to multiple translocations of an effector in different strains of the pathogen (Chuma *et al.*, 2011).

Several species of fungi possess genome defense mechanisms thought to limit the accumulation of transposable elements. The best known is repeat-induced point mutation or RIP. The RIP

process was described in Ascomycete fungi and consists of inactivating repeated sequences by introducing point mutations in CpG sites, resulting in mutational loads favoring GC to AT mutational changes (Galagan *et al.*, 2003; Galagan & Selker, 2004). RIP has never been demonstrated in *Phytophthora* species; however we found characteristics in the non-core genome that resembled a genome defense mechanism similar to RIP. The non-core genome was clearly distinct from the core genome in the *P. ramorum* lineages with a lower GC content, strong bias in codon usage favoring codons ending in AT, a high number of degenerated paralogs and enrichment of transposons or transposon like genes. Similarly, sequences exclusive to individuals found in *Z. tritici* also showed a mutational load likely resulting from the effect of a genome defense mechanism similar to RIP (Plissonneau & Stürchler, 2016). The existence of a genome defense mechanism that can deactivate transposable elements would restrict genome reshuffling during low-stress periods or could potentially lead to a more stable genome after TE activation during periods of stress or change.

Pathogen populations limited to asexual reproduction face a long term threat of extinction if they cannot adapt to changes in the environment or host. Mitotic recombination could promote the long term existence of diploid asexuals by increasing the rate of adaptation (Mandegar & Otto, 2007). Invading pathogens can carry some of the parental populationôs genetic diversity in the genome through heterozygosity. Mitotic recombination can release the diversity back into the population thereby increasing fitness as it will be able to respond to increased selection pressure. Several *Phytophthora* species are successful and adaptable invaders. Mitotic recombination may be an important contributing factor to the adaptability of these species.

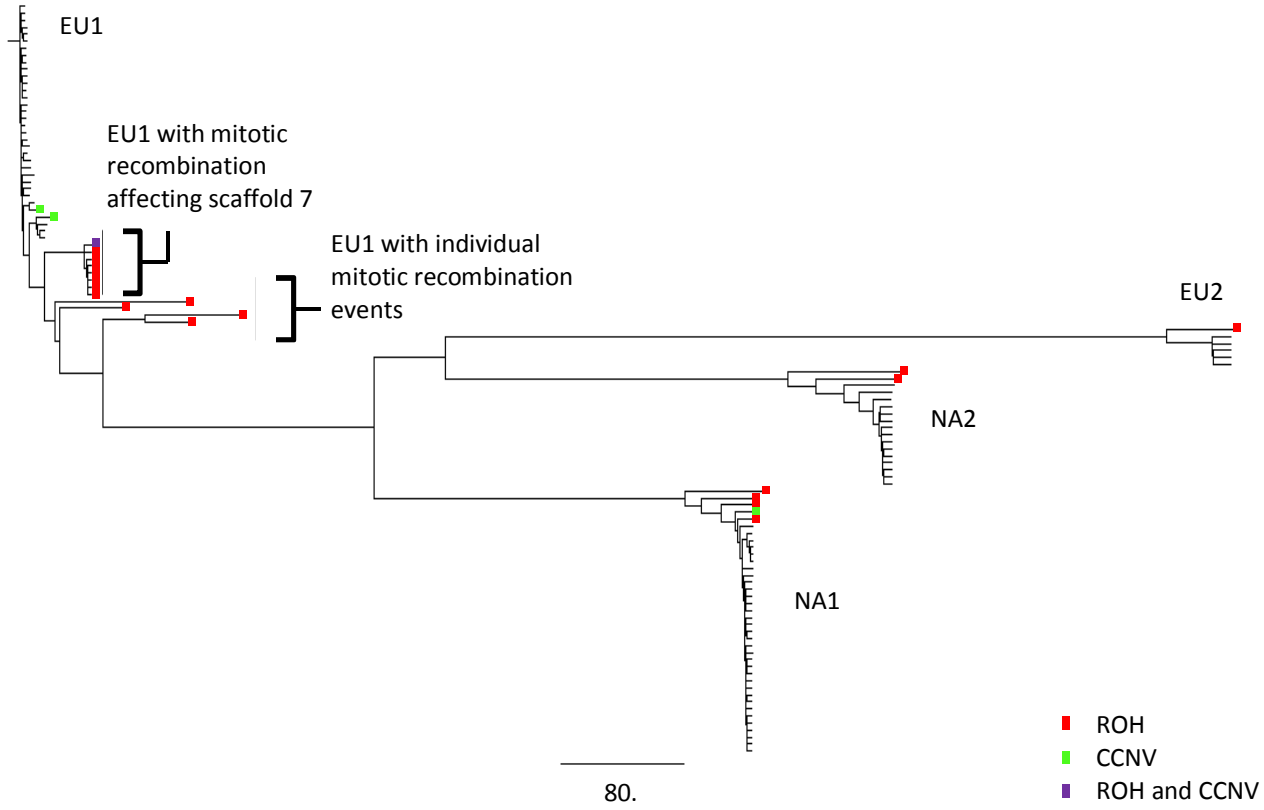


Figure 4.1: Evolutionary relationships in *Phytophthora ramorum*. Neighbor-joining tree constructed from Euclidean distances between individuals using 485 327 bi-allelic SNPs. EU1 individuals with mitotic recombination patterns shown.

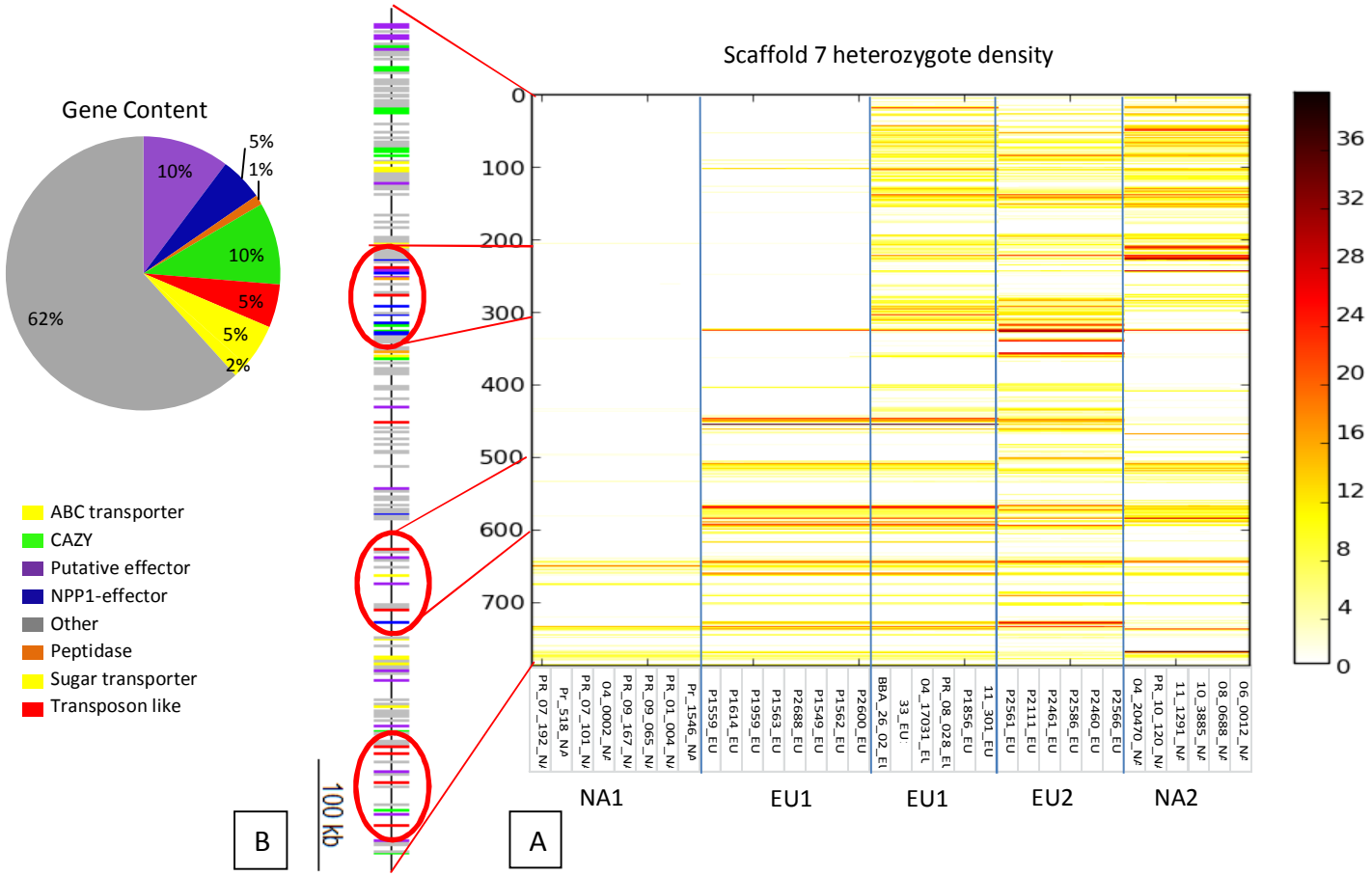


Figure 4.2: Comparison of heterozygote density, gene content and gene location in scaffold seven of *Phytophthora ramorum*. (A) Heterozygote density of representative isolates of each lineage. ROH pattern displayed for NA1 is observed in all 38 NA1 individuals. EU1 individuals possessed two distinct patterns on scaffold seven: an ROH pattern shared by eight individuals and a normal pattern in the remaining 38 individuals. (B) Gene map showing approximate location of genes along scaffold seven. (C) Piechart showing the percentage of each category of putative plant pathogenicity genes.

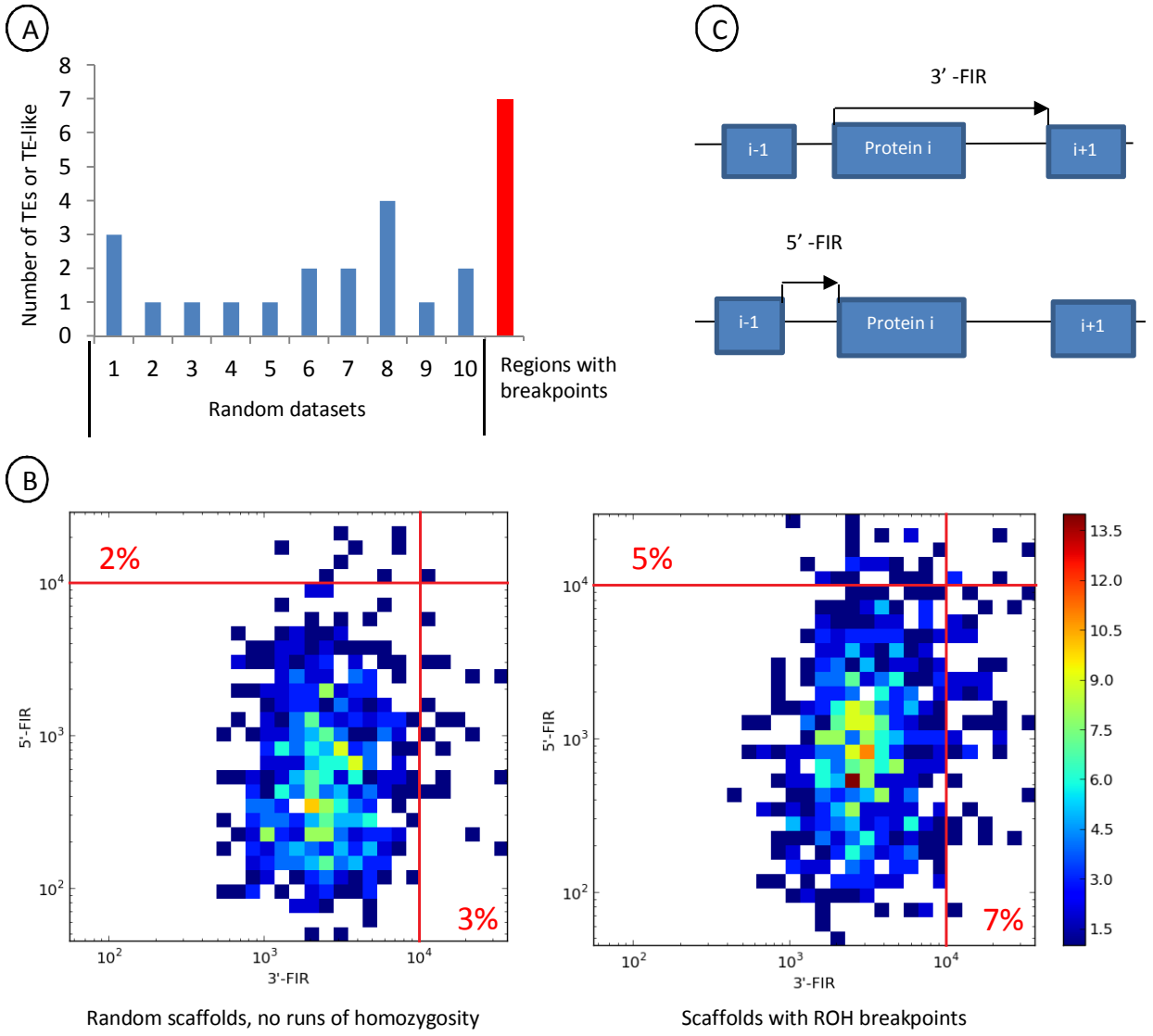


Figure 4.3: Genome characteristics around mitotic recombination breakpoints. (A) Transposable element content in the nine regions with ROH breakpoints in comparison to random regions of ten random datasets; (B) intergenic distances in nine scaffolds with ROH breakpoints compared to one of the random datasets with no recombination breakpoints; (C) intergenic distances represented by 5' and 3' distances between genes.

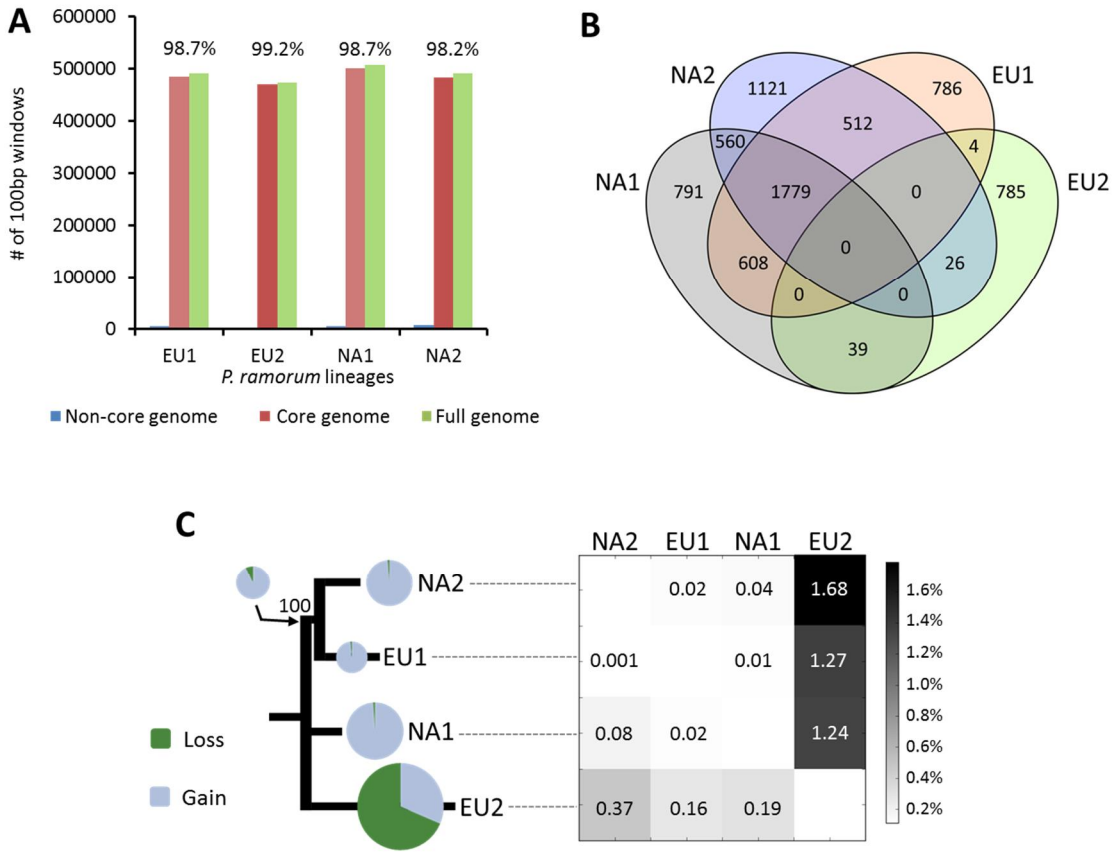


Figure 4.4: Core and non-core genomes predicted in the four *Phytophthora ramorum* lineages. (A) Proportion of core (present in all lineages) and non-core (missing in at least 50% of the individuals of a lineage) genomes; values above the bars are the proportion of core genome relatively to the size of the reference genome. (B) Number of non-core 100bp windows unique and shared between lineages. (C) Non-core regions shared between lineages. Neighbor Joining-tree based on genome similarity in terms of non-core regions (presence/absence of 100bp non-core region in the four lineages); pie-charts are proportional to the number of non-core regions that changed in each genome as predicted by CAFE, with expansions in blue and contractions in green. Proportions of lineage non-core genome (rows) that are found in the genomes of the other three lineages (columns); e.g. top right corner indicates that 1.68% of the non-core genome of NA2 was found in EU2 and bottom left corner shows that 0.37% of the non-core genome of EU2 was found in NA2.

Chapter 5: Conclusions

The objective of this work was to look at the influence of anthropogenic activities on oomycete communities, and on the pathogenic genus *Phytophthora*. We hypothesized that anthropogenic activities facilitate the introduction of alien species, the emergence of pathogens, and the post-introduction secondary migration of alien species to natural environments. Secondly, we hypothesized that season would affect diversity, as well as the habitat (water versus soil) and the method used to survey the community. For this reason, we chose to survey both soil and water in the spring and fall to obtain a more inclusive sampling of the community. Furthermore, we employed multiple approaches (DNA meta-barcoding soil, lab baiting soil and water, and stream baiting) to survey the population for the same reason. These additional approaches allowed us to compare seasonality in oomycete and *Phytophthora* communities as well as to compare the results obtained using multiple sampling methods. Lastly, we wanted to determine how invasive *Phytophthora* species overcome the invasion paradox after they are introduced into new environments. We hypothesized that *Phytophthora* species are highly adaptable due to a rapidly evolving genome, and that we would find genomic signatures of adaptation in an invasive population.

The first objective was addressed by looking on a broad scale at the oomycetes. There have been few studies characterizing oomycete communities in natural environments and none have previously been conducted in B.C. Furthermore, few studies have investigated *Phytophthora* communities in B.C. The impact of anthropogenic activities on these communities was an

important aspect of this study given the increasing number of emerging diseases believed to be linked to globalization. There are concerns regarding the potential of *Phytophthora* species being introduced or spread by human activities due to the recent emergence of a few devastating forest pathogens. In chapter 2, we compared natural oomycete communities to those in urban and interface environments to determine if those with varying levels of anthropogenic influence show a signature in the overall oomycete community. In chapter 3, we looked more in depth at the *Phytophthora* community, expanding on the methodologies used, to revisit that question, to assess the impact of different anthropogenic activities on species profiles, and to determine if anthropogenic activities increase the opportunity for inter species hybridizations.

Our results showed that anthropogenic activities influenced both urban oomycete as well as *Phytophthora* communities by increasing species diversity. This increase in oomycete and *Phytophthora* species diversity is likely caused in large part by the introduction of new species via the nursery and agricultural industry. Communities in urban sites were more diverse and several species were more abundant than in natural and interface sites, suggesting that anthropogenic activities also produce an environment conducive to population growth and spread. Furthermore, our results suggest that interface environments act as a bridge between urban and natural environments, and that there is a directional migration from urban sites to interface sites that could eventually lead to establishment of alien species in natural sites. The *Phytophthora* genus made up a relatively small portion of the oomycete community which was dominated by *Pythium*. However, the *Phytophthora* community was still quite diverse and the introduction of even just a single species that is not indigenous can have a devastating outcome.

In chapter 3 we were able to refine the analysis by focusing on the *Phytophthora* community, an important group because of their worldwide impact on crop and forest health. We used a culture-dependant approach coupled with DNA barcoding to allow for a deeper characterization of the taxa, and the discovery of putative hybrids. We observed again increased diversity in both urban and interface communities and more shared species between interface and urban sites. This convergence of our findings in chapters 2 and 3, using a completely different methodology, brings strong support to our initial hypothesis of the impact of anthropogenic activities on these communities. In addition, we identified several species in low frequency in urban environments that could represent recent introductions from other areas suggested by data on native ranges. Furthermore, we found two invasive pathogens, *P. ramorum* and *P. lateralis*, previously determined to be introduced in North America, in all site types suggesting that both introductions of new species and spread of invasive pathogens into natural environments is occurring.

Our sampling also allowed us to compare the impact of the different anthropogenic activities. Agricultural activities appear to create diverse *Phytophthora* communities that could be facilitating increased interspecific hybridization. Furthermore, agricultural activities may be supporting population growth of particular species which may lead to disease emergence. Two agricultural areas had an increased abundance of two separate putative hybrid species. At least one of these species could have spread through the global plant trade. Alternatively, one of the parental species could be spreading through global plant trade and hybridizing with local species in the locations it is introduced in. The second putative hybrid could also be introduced and emerging as suggested by the absence of any putative parental species.

This study has likely underestimated both the Oomycete and *Phytophthora* diversity present in British Columbia. Each method employed has limitations; however using multiple methods over multiple time frames has resulted in one of the most comprehensive surveys of *Phytophthora* communities. Continued work could include more sampling timeframes, sampling the canopy and foliage in addition to water and soil and by using DNA meta-barcoding on all sample types in addition to traditional culture based methods. Further work using DNA meta-barcoding could utilize both oomycete and *Phytophthora* primers. Genus specific *Phytophthora* primers would allow us to detect more rare species and provide a higher confidence in the results for low frequency species. Furthermore, more intensive sampling in different forest types would help to develop a better understanding of natural oomycete communities. Sampling both the oomycete diversity, as well as the above ground plant community would provide insight into ecosystem processes. How plant species might cultivate pathogen or microbial communities, and how microbial communities might influence forest diversity and stand structure.

In chapter 4 we looked more in depth at an invasive *Phytophthora* pathogen that has been introduced through the nursery trade. *Phytophthora ramorum* has emerged several times in the USA, Europe and Canada. Despite being restricted to clonal reproduction and a reduced gene pool, in each region where the pathogen has emerged, it has successfully invaded the new territory. We suspected that the unique genome features of *Phytophthora* pathogens enabled *P. ramorum* populations to rapidly evolve and overcome the invasion paradox. Our results strongly support the hypothesis that *P. ramorum* populations are diversifying, which is likely being driven by rapidly evolving, gene sparse and transposon rich genomic regions causing mitotic recombination. In addition, we also found evidence that genome structure is driving the

divergence between the lineages, possibly leading to observed niche differences, particularly in the EU2 lineage which showed a striking pattern of gene loss compared to the other three lineages. These findings suggest that the hypothesized two speed genome in *Phytophthora* species could also drive population evolution and species evolution.

This research is a significant contribution to our knowledge of oomycete and *Phytophthora* communities and how those communities are impacted by anthropogenic activities. Studies on oomycete communities are in their infancy and we are only beginning to understand community structure and possible roles of the different species in natural environments. A growing body of evidence is suggesting that native *Phytophthora* species are present in all habitats and many do not cause widespread disease. Many oomycetes including some *Phytophthora* likely play a role as saprotrophs. Plant pathogenic oomycetes likely contribute to succession patterns, maintaining diversity, or creating stand level structure in healthy forests. Our study will help in the global understanding of this group of microorganisms, and how we unintentionally facilitate the emergence of new forest diseases. Devastating forest diseases are often caused by alien invasive pathogens and understanding how these pathogens move into natural environments will allow us to improve our practices to reduce international spread, as well as to reduce or change activities that spread pathogens from urban to natural environments. This could include improved sanitation methods for exports in trade, improved inspection of material crossing borders including molecular testing for invasive pathogens, and pathogen tracking to uncover dispersal routes and stop the movement of invasive species. At a local scale, pathogen spread can be limited by reducing runoff and treating waste water in agricultural and nursery settings, as well as treating recycled water before reusing in nursery settings. In addition, public education programs

could highlight the effects of invasive species and encourage the public to properly dispose of garden waste and other plant material. Programs could teach the public not to plant alien species in natural environments.

Our study also provides new knowledge and a baseline to BC and Canada regarding the *Phytophthora* species present, including invasive species and species that are likely introduced. This is important information for forest and pest managers who may be able to use it for further monitoring of invasive species and possible next steps including eradication if necessary.

In addition, this work has provided valuable information on methods to monitor and track invasive pathogens or new emerging pathogens including meta-barcoding which can be used as a first step to monitor a wide range of areas and a large number of samples. Follow up monitoring can be done in locations where invasive pathogens are detected with more thorough methods to retrieve living cultures of pathogens to confirm their presence. The results of this study have already been shared with the Ministry of Forests, Land and Natural Resources who have used our findings to initiate a *Phytophthora* monitoring program including monitoring in areas where invasive alien species were detected.

This work has also shown that mechanisms important in the evolution of *Phytophthora* species are facilitating rapid adaptation in *P. ramorum* populations. This not only contributes to our scientific understanding of invasive pathogens, population evolution and *Phytophthora*, but it provides valuable information that may help in the management of the disease. One follow-up study already planned is to conduct a broader assessment of the impact of genomic structural

rearrangements on pathogenicity and other fitness traits. As this pathogen continues to invade and emerge in new places, a better understanding of how it adapts will support decisions and programs to stop spread. This may include the creation of better tracking and detection tools or eradication programs where feasible. Already a team of scientists funded by a Genome Canada grant are following this up by creating new genomic tools for monitoring and risk assessment of *Phytophthora* species. That work will be greatly facilitated by the work presented in this thesis.

References

Abascal F, Zardoya R, Posada D. **2005**. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104ï 5.

Aguayo J, Halkett F, Husson C, Nagy ZÁ, Szigethy A, Bakonyi J, Frey P, Marçais B. **2016**. Genetic diversity and origins of the homoploid-type hybrid *Phytophthora ×alni*. *Applied and Environmental Microbiology* **82**: 7142ï 7153.

Alamouti SM, Haridas S, Feau N, Robertson G, Bohlmann J, Breuil C. **2014**. Comparative genomics of the pine pathogens and beetle symbionts in the genus *Grosmannia*. *Molecular biology and evolution* **31**: 1454ï 1474.

Albornoz FE, Burgess TI, Lambers H, Etchells H. **2017**. Native soilborne pathogens equalize differences in competitive ability between plants of contrasting nutrient-acquisition strategies. *Journal of Ecology* **105**: 549ï 557.

Allain-Boulé N, Tweddell R, Mazzola M, Bélanger R, André Lévesque C. **2004**. *Pythium attrantheridium* sp. nov.: Taxonomy and comparison with related species. *Mycological Research* **108**: 795ï 805.

Alston KP, Richardson DM. **2006**. The roles of habitat features, disturbance, and distance from putative source populations in structuring alien plant invasions at the urban/wildland interface on the Cape Peninsula, South Africa. *Biological Conservation* **132**: 183ï 198.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. **1990**. Basic local alignment search tool. *Journal of molecular biology* **215**: 403ï 410.

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. **2017**. Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**: 535ï 544.

Anisimova M, Gascuel O, Sullivan J. **2006**. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology* **55**: 539ï 552.

Arcate JM, Karp MA, Nelson EB. **2006**. Diversity of Peronosporomycete (Oomycete) communities associated with the rhizosphere of different plant species. *Microbial Ecology* **51**: 36ï 50.

Atkinson RG. **1965**. *Phytophthora* species inciting root rot of *Chamaecyparis lawsoniana* and other ornamentals in coastal British Columbia. *Canadian Journal of Botany* **43**: 1471ï 1475.

Bailey TL, Williams N, Misleh C, Li WW. **2006**. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research* **34**: W369ï W373.

Balci Y, Balci S, Eggers J, Macdonald WL, Plant D, Sciences S, Gottschalk KW, Service UF. **2007**. *Phytophthora* spp . associated with forest soils in Eastern and North-Central U . S . oak ecosystems. *Plant Disease* **91**: 705ï 710.

Baldauf SL. **2003**. The deep roots of Eukaryotes. *Science* **300**: 1703ï 1706.

De Bary A. **1876**. Researches into the nature of the potato fungus *Phytophthora infestans*. *Journal of the Royal Agricultural Society of England* **212**: 239ï 269.

Bauer S, Robinson PN, Gagneur J. **2011**. Model-based gene set analysis for Bioconductor. *Bioinformatics* **27**: 1882ï 3.

Becq J, Churlaud C, Deschavanne P. **2010**. A benchmark of parametric methods for horizontal transfers detection. *PLoS ONE* **5**: e9989.

Belbahri L, Moralejo E, Calmin G, Oszako T, García JA, Descals E, Lefort F. **2006**. *Phytophthora polonica*, a new species isolated from declining *Alnus glutinosa* stands in Poland. *FEMS Microbiology Letters* **261**: 165ï 174.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. **2004**. Improved prediction of signal peptides: SignalP 3.0. *Journal of molecular biology* **340**: 783ï 95.

Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, De Wit P, Sánchez-García M, Ebersberger I, de Sousa F, *et al.* 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**: 914ï 919.

Bertier L, Leus L, D'hondt L, de Cock AWAM, Höfte M. **2013**. Host adaptation and speciation through hybridization and polyploidy in *Phytophthora*. *PLoS ONE* **8**: e85385.

De Bie T, Cristianini N, Demuth JP, Hahn MW. **2006**. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269ï 71.

Bilodeau GJ. **2008**. Détection et génomique de Phytophthora ramorum agent causal de la mort subite du chêne (lêncre des chênes rouges). *Thesis, Université Laval, Québec*.

Bilodeau GJ, Lévesque CA, Cock AWAM de, Duchaine C, Brière S, Uribe P, Martin FN, Hamelin RC. **2007**. Molecular detection of Phytophthora ramorum by real time-PCR using TaqMan, SYBR®Green and molecular beacons. *Phytopathology* **97**: 632ï 642.

Birch PRJ, Rehmany AP, Pritchard L, Kamoun S, Beynon JL. **2006**. Trafficking arms: oomycete effectors enter host plant cells. *Trends in Microbiology* **14**: 8ï 11.

Birky-Jr. CW. **1996**. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**: 427ï 437.

Blackburn TM, Lockwood JL, Cassey P. **2015**. The influence of numbers on invasion success. *Molecular Ecology* **24**: 1942ï 1953.

Blair JE, Coffey MD, Park SY, Geiser DM, Kang S. **2008**. A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. *Fungal Genetics and Biology* **45**: 266ï 277.

Blomquist C, Irving T, Osterbauer N, Reeser R. **2005**. *Phytophthora hibernalis*: A new pathogen on *Rhododendron* and evidence of cross amplification with two PCR detection assays for *Phytophthora ramorum*. *Plant Health Progress* **DOI**: 10.1094/PHP-2005-0728-1001-HN.

Bonants PJM, Hagenaar-de Weerdt M, Man in 't Veld WA, Baayen RP. **2000**. Molecular characterization of natural hybrids of *Phytophthora nicotianae* and *P. cactorum*. *Phytopathology* **90**: 867ï 874.

Bopp SER, Manary MJ, Bright AT, Johnston GL, Dharia N V, Luna FL, McCormack S, Plouffe D, McNamara CW, Walker JR, *et al.* 2013. Mitotic evolution of *Plasmodium falciparum* shows a stable core genome but recombination in antigen families. *PLoS Genetics* **9**: e1003293.

Brasier CM. **1992**. Evolutionary Biology of *Phytophthora*. *Annual Review of Phytopathology* **30**: 153ï 170.

Brasier CM. **1995**. Episodic selection as a force in fungal microevolution, with special reference to clonal speciation and hybrid introgression. *Canadian Journal of Botany* **73**: 1213ï 1221.

Brasier CM. **2001**. Rapid evolution of introduced plant pathogen via interspecific hybridization. *Bioscience* **51**: 123ï 133.

Brasier CM. **2008**. The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathology* **57**: 792ï 808.

Brasier CM. **2012**. Rapid evolution of introduced tree pathogens via episodic selection and horizontal gene transfer. Proceedings of the fourth international workshop on the genetics of host-parasite interactions in forestry: Disease and insect resistance in forest trees. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 133-142.

Brasier CM, Cooke DEL, Duncan JM. **1999**. Origin of a new *Phytophthora* pathogen through interspecific hybridization. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 5878ï 5883.

Brasier CM, Cooke DEL, Duncan JM, Hansen EM. **2003**. Multiple new phenotypic taxa from trees and riparian ecosystems in *Phytophthora gonapodyides-P. megasperma* ITS Clade 6, which tend to be high-temperature tolerant and either inbreeding or sterile. *Mycological Research* **107**: 277ï 290.

Brasier CM, Jung T. **2006**. Recent developments in *Phytophthora* diseases of trees and natural ecosystems in Europe. *Progress in Research on Phytophthora Diseases of Forest Trees. Proceedings, 3rd Int. IUFRO Working Party* **7**: 11ï 17.

Brasier CM, Kirk SA. **2001**. Comparative aggressiveness of standard and variant hybrid alder phytophthoras, *Phytophthora cambivora* and other *Phytophthora* species on bark of *Alnus*, *Quercus* and other woody hosts. *Plant Pathology* **50**: 218ï 229.

Brasier C, Kirk S. **2004**. Production of gametangia by *Phytophthora ramorum* in vitro. *Mycological Research* **108**: 823ï 827.

Brasier CM, Kirk S, Delcan J, Cooke DEL, Jung T, Man in't Veld WA. **2004**. Phytophthora alni sp. nov. and its variants: designation of emerging heteroploid hybrid pathogens spreading on Alnus trees. *Mycological research* **108**: 1172ï 1184.

Brasier CM, Rose J, Gibbs JN. **1995**. An unusual *Phytophthora* associated with widespread alder mortality in Britain. *Plant Pathology* **44**: 999ï 1007.

Brasier C, Webber J. **2010**. Sudden larch death. *Science* **329**: 824ï 825.

Brazee NJ, Wick RL, Hulvey JP. **2016**. *Phytophthora* species recovered from the Connecticut river valley in Massachusetts, USA. *Mycologia* **108**: 6ï 19.

Broad Institute (broadinstitute.org). **2014**. ôPhytophthora parasitica Assembly Dev initiativeô : https://olive.broadinstitute.org/projects/phytopht.

Broders KD, Lipps PE, Paul PA, Dorrance AE. **2007**. Characterization of *Pythium* spp. associated with corn and soybean seed and seedling disease in Ohio. *Plant Disease* **91**: 727ï 735.

Brown A V, Brasier CM. **2007**. Colonization of tree xylem by *Phytophthora ramorum*, *P. kernoviae* and other *Phytophthora* species. *Plant Pathology* **56**: 227ï 241.

Brown JKM, Hovmøller MS. **2002**. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**: 537ï 541.

Browning SR, Browning BL. **2007**. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American journal of human genetics* **81**: 1084ï 1097.

Bruns T, Lee S, Taylor J. **1990**. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In: M A Innis, D H Gelfand JJS and TJW, ed. PCR Protocols: A Guide to Methods and Applications. Academic press, London, 315ï 322.

Burgess TI. **2015**. Molecular characterization of natural hybrids formed between five related indigenous Clade 6 *Phytophthora* species. *PLoS ONE* **10**: e0134225.

Burgess TI, Scott JK, Mcdougall KL, Stukely MJC, Crane C, Dunstan WA, Brigg F. **2017a**. Current and projected global distribution of *Phytophthora cinnamomi*, one of the world ôs worst plant pathogens. *Global Change Biology*: 1661ï 1674.

Burgess TI, White D, McDougall KM, Garnas J, A.Dunstan W, Catala S, Carnegie AJ, Worboys S, Cahill D, Vettraino A-M, *et al.* 2017b. Distribution and diversity of *Phytophthora* across Australia. *Pacific Conservation Biology* **23**: 1ï 13.

Cahill DM, Rookes JE, Wilson BA, Gibson L, McDougall KL. **2008**. *Phytophthora cinnamomi* and Australiaôs biodiversity: Impacts, predictions and progress towards control. *Australian Journal of Botany* **56**: 279ï 310.

Callaghan S, Guest D. **2015**. Globalisation , the founder effect , hybrid *Phytophthora* species and rapid evolution ż new headaches for biosecurity. *Australasian Plant Pathology* **44**: 255ï 262.

Calo S, Billmyre RB, Heitman J. **2013**. Generators of phenotypic diversity in the evolution of pathogenic microorganisms. *PLoS Pathogens* **9**: e1003181.

Capinha C, Essl F, Seebens H, Moser D, Pereira HM. **2015**. The dispersal of alien species redefines biogeography in the Anthropocene. *Science* **348**: 1248ï 1251.

Caporaso G, Kuczynski J, Stombaugh J, Bittinger K, Bushman F, Costello E, Fierer N, Peña A, Goodrich J, Gordon J, *et al.* 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**: 335ï 336.

Català S, Pérez-sierra A, Abad-campos P. **2015**. The use of genus-specific amplicon pyrosequencing to assess *Phytophthora* species diversity using eDNA from soil and water in Northern Spain. *PLoS ONE* **10**: e0119311.

Català S, Pérez-Sierra A, Abad-campos P. **2017**. Metabarcoding and development of new real-time specific assays reveal *Phytophthora* species diversity in holm oak forests in eastern Spain. *Plant Pathology* **66**: 115ï 123.

Chamnanpunt J, Shan W, Tyler BM. **2001**. High frequency mitotic gene conversion in genetic hybrids of the oomycete *Phytophthora sojae*. *Proceedings of the National Academy of Sciences* **98**: 14530ï 14535.

Chastagner GA, Hansen EM, Riley KL, Sutton W. **2004**. Susceptibility of conifer shoots to infection by *Phytophthora ramorum*. *Phytopathology* **94**: S16.

Chastagner G, Oak S, Omdal D, Ramsey- A, Coats K, Valachovic Y, Lee C, Jeffers S, Elliott M. **2009**. Spread of *P . ramorum* from nurseries into waterways ð Implications for pathogen establishment in new areas. Fourth Sudden Oak Death Science Symposium. Santa Cruz, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, Pages 22-26.

Chastagner GA, Riley KL. **2010**. First report of *Phytophthora ramorum* infecting California red fir in California. *Plant Disease* **94**: 1170.

Chevenet F, Brun C, Bañuls A-L, Jacq B, Christen R. **2006**. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7**: 439.

Chuma I, Isobe C, Hotta Y, Ibaragi K, Futamata N, Kusaba M, Yoshida K, Terauchi R, Fujita Y, Nakayashiki H, *et al.* 2011. Multiple translocation of the AVR-Pita effector gene among chromosomes of the Rice Blast fungus Magnaporthe oryzae and related species. *PLoS Pathogens* **7**: e1002147.

Coffey MD. **2012**. The World Phytophthora Genetic Resource Collection (WPC). http://phytophthora.ucr.edu/.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. **2005**. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674ï 3676.

Connell JH. **1971**. On the role of natural enemies in preventing competitive exclusion in some marine animals and rain forest trees. Dynamics of population. PUDOC, 298ï 312.

Cooke DEL, Drenth A, Duncan JM, Wagels G, Brasier CM. **2000**. A molecular phylogeny of *Phytophthora* and related Oomycetes. *Fungal Genetics and Biology* **30**: 17ï 32.

Cornish-Bowden A. **1985**. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Research* **13**: 3021ï 3030.

van Damme M, Cano LM, Oliva R, Schornack S, Segretin ME, Kamoun S, Raffaele S. **2012**. Evolutionary and functional dynamics of Oomycetes effector genes. In: Martin F,, In: Kamoun S, eds. Effectors in plant-microbe interactions. Chichester, UK: John Wiley & Sons, Inc., 426.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156ï 2158.

Darriba D, Taboada GL, Doallo R, Posada D. **2012**. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.

Dart NL, Chastagner GA. **2007**. High Recovery Rate of Phytophthora from Containerized Nursery Stock Pots at a Retail Nursery Highlights Potential For Spreading Exotic Oomycetes.

Plant Health Progress.

Davidson JM, Garbelotto M, Koike ST, Rizzo DM. **2002**. First Report of *Phytophthora ramorum* on Douglas-Fir in California. *Plant Disease* **86**: 1274.

Davidson JM, Wickland AC, Patterson HA, Falk KR, Rizzo DM. **2005**. Transmission of Phytophthora ramorum in mixed-evergreen forest in California. *Phytopathology* **95**: 587ï 596.

Denman S, Kirk SA, Brasier CM, Webber JF. **2005**. In vitro leaf inoculation studies as an indication of tree foliage susceptibility to *Phytophthora ramorum* in the UK. *Plant Pathology* **54**: 512ï 521.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, *et al.* 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* **36**: W465ï W469.

Derrien T, Estellé J, Sola SM, Knowles DG, Raineri E, Guigó R, Ribeca P. **2012**. Fast computation and applications of genome mappability. *PLoS ONE* **7**: e30377.

Desprez-Loustau M-L, Courtecuisse R, Robin C, Husson C, Moreau P-A, Blancard D, Selosse M-A, Lung-Escarmant B, Piou D, Sache I. **2010**. Species diversity and drivers of spread of alien fungi (sensu lato) in Europe with a particular focus on France. *Biological Invasions* **12**: 157ï 172.

Desprez-Loustau M-L, Robin C, Buée M, Courtecuisse R, Garbaye J, Suffert F, Sache I, Rizzo DM. **2007**. The fungal dimension of biological invasions. *Trends in Ecology & Evolution* **22**: 472ï 480.

Dhillon B, Feau N, Aerts AL, Beauseigle S, Bernier L, Copeland A, Foster A, Gill N, Henrissat B, Herath P, *et al.* 2015. Horizontal gene transfer and gene dosage drives adaptation to wood colonization in a tree pathogen. *Proceedings of the National Academy of Sciences* **112**: 3451ï 3456.

Dickie IA, Wakelin A, Martinez-Garcia L, Richardson SJ, Makiola A, Tylianakis JM. **2017**. Oomycetes during 120,000 years of temperate rainforest ecosystem development. *bioRxiv* **DOI:** https://doi.org/10.1101/042341.

Dobrowolski MP, Tommerup IC, Shearer BL, O'Brien PA. **2003**. Three clonal lineages of Phytophthora cinnamomi in Australia revealed by microsatellites. *Phytopathology* **93**: 695ï 704.

Donahoo RS, Lamour KH. **2008**. Interspecific hybridization and apomixis between *Phytophthora capsici* and *Phytophthora tropicalis*. *Mycologia* **100**: 911ï 920.

Dong S, Raffaele S, Kamoun S. **2015**. The two-speed genomes of filamentous pathogens: waltz with plants. *Current Opinion in Genetics & Development* **35**: 57ï 65.

Drummond AJ, Suchard MA, Xie D, Rambaut A. **2012**. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969ï 1973.

Durán A, Gryzenhout M, Slippers B, Ahumada R, Rotella A, Flores F, Wingfield BD, Wingfield MJ. **2008**. *Phytophthora pinifolia* sp. nov. associated with a serious needle disease of *Pinus radiata* in Chile. *Plant Pathology* **57**: 715ï 727.

Dutech C, Barrès B, Bridier J, Robin C, Milgroom MG, Ravigné V. **2012**. The chestnut blight fungus world tour: Successive introduction events from diverse origins in an invasive plant

fungal pathogen. *Molecular Ecology* **21**: 3931ï 3946.

Dutheil JY, Mannhaupt G, Schweizer G, M.K. Sieber C, Münsterkötter M, Güldener U, Schirawski J, Kahmann R. **2016**. A tale of genome compartmentalization: The evolution of virulence clusters in smut fungi. *Genome Biology and Evolution* **8**: 681ï 704.

Eddy SR. **2008**. A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLoS Computational Biology* **4**: e1000069.

Edgar RC. **2004**. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**: 113.

Edgar RC. **2010**. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460ï 2461.

Elliott M, Sumampong G, Varga A, Shamoun SF, James D, Masri S, Grünwald NJ. **2011**. Phenotypic differences among three clonal lineages of *Phytophthora ramorum*. *Forest Pathology* **41**: 7ï 14.

Ellison AM, Bank MS, Clinton BD, Colburn EA, Elliott K, Ford CR, Foster DR, Kloeppel BD, Knoepp JD, Lovett GM, *et al.* 2005. Loss of foundation species: consequences for the structure and dynamics of forested ecosystems. *Frontiers in Ecology and the Environment* **3**: 479ï 486.

Esmaeili Taheri A, Chatterton S, Gossen BD, McLaren DL. **2017**. Metagenomic analysis of oomycete communities from the rhizosphere of field pea on the Canadian prairies. *Canadian Journal of Microbiology* **63**: 758ï 768.

Excoffier L, Laval G, Schneider S. **2005**. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47ï 50.

Faino L, Seidl MF, Shi-Kunne X, Pauper M, Van Den Berg GCM, Wittenberg AHJ, Thomma BPHJ. **2016**. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Research* **26**: 1091ï 1100.

Farrer RA, Henk DA, Garner TWJ, Balloux F, Woodhams DC, Fisher MC. **2013**. Chromosomal copy number variation, selection and uneven rates of recombination reveal cryptic genome diversity linked to pathogenicity. *PLoS Genetics* **9**: e1003703.

Fawke S, Doumane M, Schornack S. **2015**. Oomycete interactions with plants: Infection strategies and resistance principles. *Microbiology and Molecular Biology Reviews* **79**: 263ï 280.

Feau N, Taylor G, Dale AL, Dhillon B, Bilodeau GJ, Birol I, Jones SJM, Hamelin RC. **2016**. Genome sequences of six *Phytophthora* species threatening forest ecosystems. *Genomics Data* **10**: 85ï 88.

Forche A, Abbey D, Pisithkul T, Weinzierl MA, Ringstrom T, Bruck D, Petersen K, Berman J. **2011**. Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *mBio* **2**: e00129-11.

Forestry Commision Great Britian. **2012**. Phytophthora ramorum in larch trees ï update. : www.forestry.gov.uk/forestry/INFD-8EJKP4.

Forsman A. **2014**. Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of*

Sciences **111**: 302ï 307.

Fox J. **2005**. The R Commander: A basic statistics graphical user interface to R. *Journal of Statistical Software* **14**: 1ï 42.

Fry WE, Goodwin SB, Matuszak JM, Spielman LJ, Milgroom MG, Drenth A. **1992**. Population genetics and intercontinental migrations of *Phytophthora infestans*. *Annual Review of Phytopathology* **30**: 107ï 130.

Gaastra W, Lipman LJA, De Cock AWAM, Exel TK, Pegge RBG, Scheurwater J, Vilela R, Mendoza L. **2010**. *Pythium insidiosum*: An overview. *Veterinary Microbiology* **146**: 1ï 16.

Galagan JE, Calvo SE, Borkovich K a, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S, *et al.* 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**: 859ï 868.

Galagan JE, Selker EU. **2004**. RIP: The evolutionary cost of genome defense. *Trends in Genetics* **20**: 417ï 423.

Garbelotto M. **2008**. Molecular analysis to study invasions by forest pathogens¡ examples from Mediterranean ecosystems. *Phytopathologia Mediterranea* **47**: 183ï 203.

Garbelotto M, Ivors K, Hüberli D, Bonants P, Wagner A. **2006**. Potential for sexual reproduction of *Phytophthora ramorum* in Washington state nurseries. Proceedings of the Sudden Oak Death Second Science Symposium: The State of Our Knowledge. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, Pages 129-130.

García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A. **2012**. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**: 2678ï 2679.

Garkaklis MJ, Calver MC, Wilson BA, Hardy GESJ. **2004**. Habitat alteration caused by an introduced plant disease, Phytophthora cinnamomi: A potential threat to the conservation of Australian forest fauna. Conservation of Australiaôs Forest Fauna. Royal Zoological Society of New South Wales, 899ï 913.

Gigot J, Walters TW, Zasada IA. **2013**. Impact and occurrence of *Phytophthora rubi* and *Pratylenchus penetrans* in commercial red raspberry (*Rubus ideaus*) fields in northwestern Washington. *International Journal of Fruit Science* **13**: 357ï 372.

Giraud T, Gladieux P, Gavrilets S. **2010**. Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution* **25**: 387ï 395.

Gladieux P, Feurtey a., Hood ME, Snirc a., Clavel J, Dutech C, Roy M, Giraud T. **2015**. The population biology of fungal invasions. *Molecular Ecology* **24**: 1969ï 1986.

Gladieux P, Ropars J, Badouin H, Branca A, Aguileta G, de Vienne DM, Rodríguez de la Vega RC, Branco S, Giraud T. **2014**. Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Molecular ecology* **23**: 753ï 73.

Goodwin SB. **1997**. The population genetics of *Phytophthora*. *Phytopathology* **87**: 462ï 473.

Goodwin SB, Ben M'barek S, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, Van der Lee T a J, Grimwood J, Aerts A, *et al.* 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and

stealth pathogenesis. *PLoS genetics* **7**: e1002070.

Gorman M. **1993**. Tree and ornamental disease control in the city of Victoria. *B.C. Pest Monitor* **2**.

Goss EM, Carbone I, Grünwald NJ. **2009**. Ancient isolation and independent evolution of the three clonal lineages of the exotic sudden oak death pathogen *Phytophthora ramorum*. *Molecular Ecology* **18**: 1161ï 1174.

Goss EM, Cardenas ME, Myers K, Forbes G a, Fry WE, Restrepo S, Grünwald NJ. **2011a**. The plant pathogen *Phytophthora andina* emerged via hybridization of an unknown Phytophthora species and the Irish potato famine pathogen, *P. infestans*. *PloS ONE* **6**: e24543.

Goss EM, Larsen M, Vercauteren A, Werres S, Heungens K, Grünwald NJ. **2011b**. *Phytophthora ramorum* in Canada: Evidence for migration within North America and from Europe. *Phytopathology* **101**: 166ï 171.

Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. **2006**. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**: 2695ï 2696.

Greslebin AG, Hansen EM, Sutton W. **2007**. *Phytophthora austrocedrae* sp. nov., a new species associated with *Austrocedrus chilensis* mortality in Patagonia (Argentina). *Mycological Research* **111**: 308ï 316.

Greslebin AG, Hansen EM, Winton LM, Rajchenberg M. **2005**. *Phytophthora* species from declining *Austrocedrus chilensis* forests in Patagonia, Argentina. *Mycologia* **97**: 218ï 228.

Grigoriev I V, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm R a, *et al.* 2012. The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Research* **40**: D26-32.

Grünwald NJ, Garbelotto M, Goss EM, Heungens K, Prospero S. **2012**. Emergence of the sudden oak death pathogen *Phytophthora ramorum*. *Trends in Microbiology* **20**: 131ï 138.

Grünwald NJ, Goss EM, Ivors K, Garbelotto M, Martin FN, Prospero S, Hansen E, Bonants PJM, Hamelin RC, Chastagner G, *et al.* 2009. Standardizing the nomenclature for clonal Lineages of the Sudden Oak Death pathogen, *Phytophthora ramorum*. *Phytopathology* **99**: 792ï 795.

Grünwald NJ, Goss EM, Press CM. **2008**. *Phytophthora ramorum*: a pathogen with a remarkably wide host range causing sudden oak death on oaks and ramorum blight on woody ornamentals. *Molecular Plant Pathology* **9**: 729ï 740.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. **2010**. New algorithms and methods to astimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307ï 321.

Guy L, Roat Kultima J, Andersson SGE. **2010**. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**: 2334ï 2335.

Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, *et al.* 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**: 393ï 398.

Hall TA. **1999**. BioEdit: A user-friendly biological sequence alignment editor and analysis

analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95_98.

Hansen EM. **2008**. Alien forest pathogens: *Phytophthora* species are changing world forests. *Boreal Environ. Res.* **13**: 33ï 41.

Hansen EM. **2015**. *Phytophthora* species emerging as pathogens of forest trees. *Current Forestry Reports* **1**: 16ï 24.

Hansen EM, Goheen DJ, Jules ES, Ullian B. **2000**. Managing Port-Orford-Cedar and the introduced pathogen *Phytophthora lateralis*. *Plant Disease* **84**: 4ï 14.

Hansen EM, Kanaskie A, Prospero S, McWilliams M, Goheen EM, Osterbauer N, Reeser P, Sutton W. **2008**. Epidemiology of *Phytophthora ramorum* in Oregon tanoak forests. *Canadian Journal of Forest Research* **38**: 1133ï 1143.

Hansen EM, Parke JL, Sutton W. **2005**. Susceptibility of Oregon forest trees and shrubs to *Phytophthora ramorum*: A Comparison of Artificial inoculation and natural infection. *Plant Disease* **89**: 63ï 70.

Hansen EM, Reeser PW, Sutton W. **2012a**. *Phytophthora* beyond agriculture. *Annual Review of Phytopathology* **50**: 359ï 378.

Hansen EM, Reeser PW, Sutton W. **2012b**. *Phytophthora borealis* and *Phytophthora riparia*, new species in *Phytophthora* ITS Clade 6. *Mycologia* **104**: 1133ï 1142.

Hansen EM, Reeser PW, Sutton W. **2017**. Ecology and pathology of *Phytophthora* ITS clade 3 species in forests in western Oregon, USA. *Mycologia* **109**: 100ï 114.

Hansen EM, Rizzo D, Garbelotto M. **2006**. *Phytophthora* species from oak and tanoak forests in California and Oregon. Sudden Oak Death Second Science Symposium: The State of Our Knowledge. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 63-65.

Hardham AR. **2005**. *Phytophthora cinnamomi*. *Molecular Plant Pathology* **6**: 589ï 604.

Hardy GESJ, Colquhoun IJ, Shearer BL, Tommerup I. **2001**. The impact and control of *Phytophthora cinnamomi* in native and rehabilitated forest ecosystems in Western Australia. *Forest, Snow Landscape Research* **76**: 337ï 343.

Hatta R, Ito K, Hosaki Y, Tanaka T, Tanaka A, Yamamoto M, Akimitsu K, Tsuge T. **2002**. A conditionally dispensable chromosome controls host-specific pathogenicity in the fungal plant pathogen *Alternaria alternata*. *Genetics* **161**: 59ï 70.

Hegarty MJ. **2012**. Invasion of the hybrids. *Molecular Ecology* **21**: 4669ï 4671.

Heitman J. **2006**. Sexual reproduction and the evolution of microbial pathogens. *Current Biology* **16**: R711ï R725.

Heitman J, Sun S, James TY. **2013**. Evolution of fungal sexual reproduction. *Mycologia* **105**: 1ï 27.

Hong CX, Moorman GW. **2005**. Plant pathogens in irrigation water: challenges and opportunities. *Critical Reviews in Plant Sciences* **24**: 189ï 208.

Hu J, Diao Y, Zhou Y, Lin D, Bi Y, Pang Z, Trout Fryxell R, Liu X, Lamour K. **2013**. Loss of heterozygosity drives clonal diversity of *Phytophthora capsici* in China. *PLoS ONE* **8**: e82691.

Hubbs CL. 1955. Hybridization between fish species in nature. *Systematic Zoology* **4**: 1ï 20.

Hüberli D, Hardy GESJ, White D, Williams N, Burgess TI. 2013. Fishing for *Phytophthora* from western Australiaôs waterways: a distribution and diversity survey. *Australasian Plant Pathology* **42**: 251ï 260.

Huitema E, Bos JIB, Tian M, Win J, Waugh ME, Kamoun S. 2004. Linking sequence to phenotype in *Phytophthora*ï plant interactions. *Trends in Microbiology* **12**: 193ï 200.

Hulbert JM, Agne MC, Burgess TI, Roets F, Wingfield MJ. 2017. Urban environments provide opportunities for early detections of *Phytophthora* invasions. *Biological Invasions* **19**: 3629ï 3644.

Hulvey J, Gobena D, Finley L, Lamour K. 2010. Co-occurrence and genotypic distribution of *Phytophthora* species recovered from watersheds and plant nurseries of eastern Tennessee. *Mycologia* **102**: 1127ï 1133.

Hunt M, Newbold C, Berriman M, Otto TD. 2014. A comprehensive evaluation of assembly scaffolding tools. *Genome Biology* **15**: R42.

Husson C, Aguayo J, Revellin C, Frey P, Ioos R, Marçais B. 2015. Evidence for homoploid speciation in *Phytophthora alni* supports taxonomic reclassification in this species complex. *Fungal Genetics and Biology* **77**: 12ï 21.

Hwang J, Oak SW, Jeffers SN. 2008. Detecting *Phytophthora ramorum* and other species of *Phytophthora* in streams in natural ecosystems using baiting and filtration methods. Proceedings of the Sudden Oak Death Third Science Symposium. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, Pages 55-58.

Ioos R, Andrieux A, Marçais B, Frey P. 2006. Genetic characterization of the natural hybrid species *Phytophthora alni* as inferred from nuclear and mitochondrial DNA analyses. *Fungal genetics and biology* **43**: 511ï 29.

Ivors K, Garbelotto M, Vries IDE, Ruyter-Spira C, Hekkert BTE, Rosenzweig N, Bonants P. 2006. Microsatellite markers identify three lineages of *Phytophthora ramorum* in US nurseries, yet single lineages in US forest and European nursery populations. *Molecular Ecology* **15**: 1493ï 1505.

Ivors KL, Hayden KJ, Bonants PJM, Rizzo DM, Garbelotto M. 2004. AFLP and phylogenetic analyses of North American and European populations of *Phytophthora ramorum*. *Mycological Research* **108**: 378ï 392.

James TY, Litvintseva AP, Vilgalys R, Morgan JAT, Taylor JW, Fisher MC, Berger L, Weldon C, du Preez L, Longcore JE. 2009. Rapid global expansion of the fungal disease Chytridiomycosis into declining and healthy amphibian populations. *PLoS Pathogens* **5**: e1000458.

Janzen DH. 1970. Herbivores and the number of tree species in tropical forests. *The American Naturalist* **104**: 501ï 528.

Jiang RHY, Tripathy S, Govers F, Tyler BM. 2008. RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences* **105**: 4874ï 4879.

Jiang RHY, Tyler BM. **2012**. Mechanisms and evolution of virulence in Oomycetes. *Annual Review of Phytopathology* **50**: 295ï 318.

Jiang RHY, Tyler BM, Whisson SC, Hardham AR, Govers F. **2006**. Ancient origin of elicitin gene clusters in *Phytophthora* genomes. *Molecular Biology and Evolution* **23**: 338ï 351.

Joly DL, Feau N, Tanguay P, Hamelin RC. **2010**. Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (*Melampsora* spp.). *BMC Genomics* **11**: 422.

Jombart T. **2008**. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403ï 1405.

de Jonge R, Bolton MD, Kombrink A, van den Berg GCM, Yadeta K a, Thomma BPHJ. **2013**. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Research* **23**: 1271ï 1282.

Judelson HS. **2012**. Dynamics and innovations within oomycete genomes: Insights into biology, pathology, and evolution. *Eukaryotic Cell* **11**: 1304ï 1312.

Jung T, Burgess TI. **2009**. Re-evaluation of *Phytophthora citricola* isolates from multiple woody hosts in Europe and North America reveals a new species, *Phytophthora plurivora* sp. nov. *Persoonia* **22**: 95ï 110.

Jung T, Orlikowski L, Henricot B, Abad-Campos P, Aday AG, Aguín Casal O, Bakonyi J, Cacciola SO, Cech T, Chavarriaga D, *et al.* 2016. Widespread *Phytophthora* infestations in European nurseries put forest, semi-natural and horticultural ecosystems at high risk of Phytophthora diseases. *Forest Pathology* **46**: 134ï 163.

Jung T, Stukely MJC, Hardy Ges, White D, Paap T, Dunstan WA, Burgess TI. **2011**. Multiple new *Phytophthora* species from ITS Clade 6 associated with natural ecosystems in Australia: Evolutionary and ecological implications. *Persoonia : Molecular Phylogeny and Evolution of Fungi* **26**: 13ï 39.

Jurka J, Kapitonov V, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. **2005**. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and genome research* **110**: 462ï 467.

Kamoun S. **2006**. A catalogue of the effector secretome of plant pathogenic Oomycetes. *Annual Review of Phytopathology* **44**: 41ï 60.

Kamoun S, Furzer O, Jones JDG, Judelson HS, Ali GS, Dalio RJD, Roy SG, Schena L, Zambounis A, Panabières F, *et al.* 2015. The Top 10 oomycete pathogens in molecular plant pathology. *Molecular Plant Pathology* **16**: 413ï 434.

Kanaskie A, Goheen E, Osterbauer N, Mcwilliams M, Hansen E, Sutton W. **2007**. Eradication of Phytophthora ramorum from Oregon forests؛ Status after 6 years. Proceedings of the Sudden Oak Death Third Science Symposium. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 15-17.

Kanaskie A, Hansen E, Goheen EM, Mcwilliams M, Laine J, Savona S, Timeus H, Woosley B, Sutton W, Reeser P, *et al.* 2010. Detection and eradication of *Phytophthora ramorum* from Oregon Forests , 2001-2008. Proceedings of the Sudden Oak Death Fourth Science Symposium. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of

Agriculture, Pages 5-7.

Kasuga T, Bui M, Bernhardt E, Swiecki T, Aram K, Cano LM, Webber J, Brasier C, Press C, Grünwald NJ, *et al.* 2016. Host-induced aneuploidy and phenotypic diversification in the Sudden Oak Death pathogen *Phytophthora ramorum*. *BMC Genomics* **17**: 385.

Kasuga T, Kozanitas M, Bui M, Hüberli D, Rizzo DM, Garbelotto M. **2012**. Phenotypic diversification is associated with host-induced transposon derepression in the Sudden Oak Death pathogen *Phytophthora ramorum*. *PLoS ONE* **7**: e34728.

Katoh K, Standley DM. **2013**. MAFFT Multiple Sequence Alignment Software Version 7ؾ improvements in performance and usability. *Molecular Biology* **30**: 772ï 780.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, *et al.* 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647ï 1649.

Keller O, Kollmar M, Stanke M, Waack S. **2011**. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* **27**: 757ï 763.

Kliejunas JT. **2011**. Sudden oak death; ramorum blight. A risk assessment of climate change and the impact of forest diseases on forest ecosystems in the Western United States and Canada. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 18-24.

Knaus BJ, Fieland VJ, Graham KA, Grünwald NJ. **2015**. Diversity of foliar *Phytophthora* species on *Rhododendron* in Oregon nurseries. *Plant Disease* **99**: 1326ï 1332.

Kozanitas M, Osmundson TW, Linzer R, Garbelotto M. **2017**. Interspecific interactions between the Sudden Oak Death pathogen *Phytophthora ramorum* and two sympatric *Phytophthora* species in varying ecological conditions. *Fungal Ecology* **28**: 86ï 96.

Kozik A, Chan B, Michemore R. **2003**. BLAST parser, distance matrix file and protein sequence clustering. How to obtain a distance matrix file from BLAST search results and use it for sequence clustering analysisؾ http://cgpdb.ucdavis.edu/BlastParser/Blast_Parser.html.

Kroon LPNM, Brouwer H, de Cock AW a M, Govers F. **2012**. The genus *Phytophthora* anno 2012. *Phytopathology* **102**: 348ï 64.

Lake JC, Leishman MR. **2004**. Invasion success of exotic plants in natural ecosystems: The role of disturbance, plant attributes and freedom from herbivores. *Biological Conservation* **117**: 215ï 226.

Laliberté E, Lambers H, Burgess TI, Wright SJ. **2015**. Phosphorus limitation , soil-borne pathogens and the coexistence of plant species in hyperdiverse forests and shrublands. *New Phytologist* **206**: 507ï 521.

Lamour KH, Daughtrey ML, Benson DM, Hwang J, Hausbeck MK. **2003**. Etiology of *Phytophthora drechsleri* and *P. nicotianae* (=*P. parasitica*) diseases affecting floriculture crops. *Plant Disease* **87**: 854ï 858.

Lamour KH, Mudge J, Gobena D, Hurtado-gonzales OP, Kuo A, Miller NA, Rice BJ, Raffaele S, Liliana M, Bharti AK, *et al.* 2012. Genome sequencing and mapping reveal loss of

heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Molecular Plant-Microbe Interactions* **25**: 1350ï 1360.

Latijnhouwers M, de Wit PJGM, Govers F. **2003**. Oomycetes and fungi: Similar weaponry to attack plants. *Trends in Microbiology* **11**: 462ï 469.

Lee PS, Greenwell PW, Dominska M, Gawel M, Hamilton M, Petes TD. **2009**. A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. *PLoS Genetics* **5**: e1000410.

Li H, Durbin R. **2009**. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754ï 1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. **2009**. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078ï 9.

Li L, Stoeckert CJ, Roos DS. **2003**. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* **13**: 2178ï 2189.

Linzer RE, Rizzo DM, Cacciola SO, Garbelotto M. **2009**. AFLPs detect low genetic diversity for *Phytophthora nemorosa* and *P. pseudosyringae* in the US and Europe. *Mycological Research* **113**: 298ï 307.

Llorente B, Smith CE, Symington LS. **2008**. Break-induced replication: What is it and what is it for? *Cell Cycle* **7**: 859ï 864.

Loo JA. **2009**. Ecological impacts of non-indigenous invasive fungi as forest pathogens. *Biological Invasions* **11**: 81ï 96.

Lozupone C, Knight R. **2005**. UniFrac: A New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* **71**: 8228ï 8235.

Mack RN, Simberloff D, Mark Lonsdale W, Evans H, Clout M, Bazzaz FA. **2000**. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* **10**: 689ï 710.

Magwene PM, Kayıkçı Ö, Granek JA, Reininga JM, Scholl Z, Murray D. **2011**. Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* **108**: 1987ï 1992.

Maloney PE, Rizzo DM, Koike ST, Harnik TY, Garbelotto M. **2002**. First report of *Phytophthora ramorum* on oast redwood in California. *Plant Disease* **86**: 1274.

Man in ‘t Veld WA, de Cock AWAM, Summerbell RC. **2007**. Natural hybrids of resident and introduced *Phytophthora* species proliferating on multiple new hosts. *European Journal of Plant Pathology* **117**: 25ï 33.

Man in ’t Veld WA, Veenbaas-Rijks WJ, Ilieva E, de Cock AWAM, Bonants PJM, Pieters R. **1998**. Natural hybrids of *Phytophthora nicotianae* and *Phytophthora cactorum* demonstrated by isozyme analysis and random mmplified polymorphic DNA. *Phytopathology* **88**: 922ï 929.

Mandegar MA, Otto SP. **2007**. Mitotic recombination counteracts the benefits of genetic segregation. *Proceedings of the Royal Society B: Biological Sciences* **274**: 1301ï 1307.

Marano A V, Jesus AL, de Souza JI, Jerônimo GH, Gonçalves DR, Boro MC, Rocha SCO, Pires-Zottarelli CLA. **2016**. Ecological roles of saprotrophic Peronosporales (Oomycetes, Straminipila) in natural environments. *Fungal Ecology* **19**: 77ï 88.

Martin FN, Loper JE. **1999**. Soilborne plant diseases caused by Pythium spp.: Ecology, epidemiology, and prospects for biological control. *Critical Reviews in Plant Sciences* **18**: 111ï 181.

Mascheretti S, Croucher PJP, Vettraino A, Prospero S, Garbelotto M. **2008**. Reconstruction of the Sudden Oak Death epidemic in California through microsatellite analysis of the pathogen Phytophthora ramorum. *Molecular Ecology* **17**: 2755ï 2768.

Matari NH, Blair JE. **2014**. A multilocus timescale for oomycete evolution estimated under three distinct molecular clock models. *BMC Evolutionary Biology* **14**: 101.

McDonald BA, Linde C. **2002**. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**: 349ï 379.

McIntosh DL. **1966**. The occurrence of *Phytophthora* spp. in irrigation systems in British Columbia. *Canadian Journal of Botany* **44**: 1591ï 1596.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al.* 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**: 1297ï 1303.

Meentemeyer RK, Rank NE, Shoemaker DA, Oneal CB, Wickland AC, Frangioso KM, Rizzo DM. **2008**. Impact of sudden oak death on tree mortality in the Big Sur ecoregion of California. *Biological Invasions* **10**: 1243ï 1255.

Metz MR, Frangioso KM, Meentemeyer RK, Rizzo DM. **2011**. Interacting disturbances: wildfire severity affected by stage of forest disease invasion. *Ecological Applications* **21**: 313ï 320.

Migliorini D, Ghelardini L, Tondini E, Luchi N, Santini A, Richardson DM. **2015**. The potential of symptomless potted plants for carrying invasive soilborne plant pathogens. *Diversity and Distributions* **21**: 1218ï 1229.

Miller MA, Pfeiffer W, Schwartz T. **2010**. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. *2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA*: 1ï 8.

Mills KE, Bever JD. **1998**. Maintenance of diversity within plant communities: Soil pathogens as agents of negative feedback. *Ecology* **79**: 1595ï 1601.

Mirkin BG, Fenner TI, Galperin MY, Koonin E V. **2003**. Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evolutionary Biology* **3**: 2.

Möller EM, Bahnweg G, Sandermann H, Geiger HH. **1992**. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic acids research* **20**: 6115ï 6116.

Moller M, Stukenbrock EH. **2017**. Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology* **15**: 756ï 771.

Monahan WB, Tse J, Koenig WD, Garbelotto M. **2008**. Mycological Research News1. *Mycological Research* **112**: 757ï 758.

Murphy SK, Rizzo DM. **2006**. Incidence of Phytophthora ramorum, P . nemorosa and P . pseudosyringae in three coastal California forest communities. Proceedings of the sudden oak death second science symposium: the state of our knowledge. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 69-71.

Nagel JH, Gryzenhout M, Slippers B, Wingfield MJ, Hardy GESJ, Stukely MJC, Burgess TI. **2013**. Characterization of *Phytophthora* hybrids from ITS clade 6 associated with riparian ecosystems in South Africa and Australia. *Fungal Biology* **117**: 329ï 347.

Nelson EB, Karp MA. **2013**. Soil pathogen communities associated with native and non-native *Phragmites australis* populations in freshwater wetlands. *Ecology and Evolution* **3**: 5254ï 5267.

Newcombe G, Stirling B, Bradshaw HD. **2001**. Abundant pathogenic variation in the new hybrid rust *Melampsora ×columbiana* on hybrid poplar. *Phytopathology* **91**: 981ï 985.

Newcombe G, Stirling B, McDonald S, Bradshaw HD. **2000**. *Melampsora × columbiana*, a natural hybrid of *M. medusae* and *M. occidentalis*. *Mycological Research* **104**: 261ï 274.

Noga EJ. **1993**. Water mold infections of freshwater fish: Recent advances. *Annual Review of Fish Diseases* **3**: 291ï 304.

Normark BB, Judson OP, Moran NA. **2003**. Intraclonal genetic variation: ecological and evolutionary aspects. *Biological Journal of the Linnean Society* **79**: 69ï 84.

North A, Pennanen J, Ovaskainen O, Laine A-L. **2011**. Local adaptation in a changing world: the roles of gene-flow, mutation, and sexual reproduction. *Evolution* **65**: 79ï 89.

Nowell RW, Green S, Laue BE, Sharp PM. **2014**. The extent of genome flux and its role in the differentiation of bacterial lineages. *Genome biology and evolution* **6**: 1514ï 1529.

Oh E, Gryzenhout M, Wingfield BD, Wingfield MJ, Burgess TI. **2013**. Surveys of soil and water reveal a goldmine of *Phytophthora* diversity in South African natural ecosystems. *IMA Fungus* **4**: 123ï 131.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB, Simpson GL, Solymos P, Stevens, M.H.H, Wagner H. **2016**. Vegan: Community ecology package. R package version 2.3-5.

Oliva RF, Kroon LPNM, Chacón G, Flier WG, Ristaino JB, Forbes GA. **2010**. *Phytophthora andina* sp. nov., a newly identified heterothallic pathogen of solanaceous hosts in the Andean highlands. *Plant Pathology* **59**: 613ï 625.

Paap T, Burgess TI, Wingfield MJ. **2017**. Urban trees¸ bridge-heads for forest pest invasions and sentinels for early detection. *Biological Invasions* **19**: 3515ï 3526.

Packer A, Clay K. **2000**. Soil pathogens and spatial patterns of seedling mortality in a temperate tree. *Nature* **404**: 278ï 281.

Panabières F, Ponchet M, Allasia V, Cardin L, Ricci P. **1997**. Characterization of border

species among Pythiaceae: several *Pythium* isolates produce elicitins, typical proteins from *Phytophthora* spp. *Mycological Research* **101**: 1459ï 1468.

Paradis E. **2010**. pegas: an R package for population genetics with an integratedï modular approach. *Bioinformatics* **26**: 419ï 420.

Paradis E, Claude J, Strimmer K. **2004**. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289ï 290.

Park B, Martin F, Geiser DM, Kim H-S, Mansfield MA, Nikolaeva E, Park S-Y, Coffey MD, Russo J, Kim SH, *et al.* 2013. Phytophthora database 2.0: Update and future direction. *Phytopathology* **103**: 1204ï 1208.

Parke JL, Knaus BJ, Fieland VJ, Lewis C, Grünwald NJ. **2014**. *Phytophthora* community structure analyses in Oregon nurseries inform systems approaches to disease management. *Phytopathology* **104**: 1052ï 1062.

Parke JL, Lucas S. **2008**. Sudden oak death and ramorum blight. *The Plant Health Instructor* **DOI**: 10.1094/PHI-I-2008-0227-01.

Parra G, Bradnam K, Korf I. **2007**. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**: 1061ï 1067.

Parra G, Bradnam K, Ning Z, Keane T, Korf I. **2009**. Assessing the gene space in draft genomes. *Nucleic acids research* **37**: 289ï 97.

Pautasso M, Dehnen-Schmutz K, Holdenrieder O, Pietravalle S, Salama N, Jeger MJ, Lange E, Hehl-Lange S. **2010**. Plant health and global change - Some implications for landscape management. *Biological Reviews* **85**: 729ï 755.

Pautasso M, Schlegel M, Holdenrieder O. **2015**. Forest health in a changing world. *Microbial Ecology* **69**: 826ï 842.

Pemberton CL, Salmond GPC. **2004**. The Nep1-like proteinsð a growing family of microbial elicitors of plant necrosis. *Molecular Plant Pathology* **5**: 353ï 359.

Pérez JE, Nirchio M, Alfonsi C, Muñoz C. **2006**. The biology of invasions: The genetic adaptation paradox. *Biological Invasions* **8**: 1115ï 1121.

Philibert A, Desprez-Loustau M-L, Fabre B, Frey P, Halkett F, Husson C, Lung-Escarmant B, Marçais B, Robin C, Vacher C, *et al.* 2011. Predicting invasion success of forest pathogenic fungi from species traits. *Journal of Applied Ecology* **48**: 1381ï 1390.

Phillips AJ, Anderson VL, Robertson EJ, Secombes CJ, van West P. **2008**. New insights into animal pathogenic oomycetes. *Trends in Microbiology* **16**: 13ï 19.

Pimentel D, Lach L, Zuniga R, Morrison D. **2000**. Environmental and economic costs of nonindigenous species in the United States. *BioScience* **50**: 53ï 65.

Van der Plaats-Niterink A J. **1981**. Monograph of the genus *Pythium*. *Studies in Mycology* **21**: 1ï 244.

Plissonneau C, Stürchler A. **2016**. The evolution of orphan regions in genomes of a fungal pathogen of wheat. *mBio* **7**: e01231-16.

Pojar J, Klinka K, Meidinger D V. **1987**. Biogeoclimatic ecosystem classification in British

Columbia. *Forest Ecology and Management* **22**: 119ï 154.

Van Poucke K, Franceschini S, Webber JF, Vercauteren A, Turner JA, McCracken AR, Heungens K, Brasier CM. 2012. Discovery of a fourth evolutionary lineage of *Phytophthora ramorum*: EU2. *Fungal Biology* **116**: 1178ï 1191.

Prigigallo MI, Abdelfattah A, Cacciola SO, Faedda R, Sanzani SM, Cooke DEL, Schena L. 2015. Metabarcoding analysis of *Phytophthora* diversity using genus-specific primers and 454 pyrosequencing. *Phytopathology* **106**: 305ï 313.

Prigigallo MI, Abdelfattah A, Cacciola SO, Faedda R, Sanzani SM, Cooke DEL, Schena L. 2016. Metabarcoding analysis of *Phytophthora* diversity using genus-specific primers and 454 pyrosequencing. *Phytopathology* **106**: 305ï 313.

Prospero S, Grünwald NJ, Winton LM, Hansen EM. 2009. Migration patterns of the emerging plant pathogen *Phytophthora ramorum* on the west coast of the United States of America. *Phytopathology* **99**: 739ï 749.

Prospero S, Hansen EM, Grünwald NJ, Winton LM. 2007. Population dynamics of the sudden oak death pathogen *Phytophthora ramorum* in Oregon from 2001 to 2004. *Molecular Ecology* **16**: 2958ï 2973.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, *et al.* 2017. PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* **81**: 559ï 575.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841ï 842.

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

R Core Team. 2017. R: A language and environment for statistical computing.

Raffaele S, Farrer R a, Cano LM, Studholme DJ, MacLean D, Thines M, Jiang RHY, Zody MC, Kunjeti SG, Donofrio NM, *et al.* 2010a. Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science* **330**: 1540ï 1543.

Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* **10**: 417ï 430.

Raffaele S, Win J, Cano LM, Kamoun S. 2010b. Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*. *BMC Genomics* **11**: 637.

Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution* **2**: vew007.

Ramsfield TD, Bentz BJ, Faccoli M, Jactel H, Brockerhoff EG. 2016. Forest health in a changing world: Effects of globalization and climate change on forest insect and pathogen impacts. *Forestry: An International Journal of Forest Research* **89**: 245ï 252.

Ravenhall M, Škunca N, Lassalle F, Dessimoz C. 2015. Inferring horizontal gene transfer.

PLOS Computational Biology **11**: e1004095.

Redondo MA, Boberg J, Stenlid J, Oliva J. **2018**. Functional traits associated with the establishment of introduced *Phytophthora* spp . in Swedish forests. *Journal of Applied Ecology*: 1ï 15.

Reeser P, Sutton W, Hansen E. **2010**. *Phytophthora* species in tanoak trees, canopy drip, soil, and streams in the sudden oak death epidemic area in southwestern Oregon. *The Fifth Meeting of the International Union of Forest Research Organizations (IUFRO), Working Party 7.02.09*.

Reeser PW, Sutton W, Hansen EM, Remigi P, Adams GC. **2011**. *Phytophthora* species in forest streams in Oregon and Alaska. *Mycologia* **103**: 22ï 35.

Reinhart KO, Tytgat T, Van der Putten WH, Clay K. **2010**. Virulence of soil-borne pathogens and invasion by *Prunus serotina*. *New Phytologist* **186**: 484ï 495.

Rice P, Longden I, Bleasby A. **2000**. EMBOSS: The european molecular biology open software suite. *Trends in Genetics* **16**: 276ï 277.

Rieseberg LH, Archer MA, Wayne RK. **1999**. Transgressive segregation, adaptation and speciation. **83**: 363ï 372.

Riit T, Tedersoo L, Drenkhan R, Runno-Paurson E, Kokko H, Anslan S. **2016**. Oomycete-specific ITS primers for identification and metabarcoding. *MycoKeys* **14**: 17ï 30.

Rizzo DM, Garbelotto M. **2003**. Sudden oak death: Endangering California and Oregon forest ecosystems. *Frontiers in Ecology and the Environment* **1**: 197ï 204.

Rizzo DM, Garbelotto M, Davidson JM, Slaughter GW, Koike ST. **2002**. *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease* **86**: 205ï 214.

Robertson GI. **1973**. Pathogenicity of *Pythium* spp. to seeds and seedling roots. *New Zealand Journal of Agricultural Research* **16**: 367ï 372.

Rollins L, Coats K, Elliott M, Chastagner G. **2016**. Comparison of five detection and quantification methods for *Phytophthora ramorum* in stream and irrigation water. *Plant Disease* **100**: 1202ï 1211.

Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, Dominguez V, Anthouard V, Bally P, Bourras S, *et al.* 2011. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature Communications* **2**: 202.

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, *et al.* 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* **32**: 305ï 332.

Salinas NR, Little DP. **2014**. 2matrix: A utility for indel coding and phylogenetic matrix concatenation. *Applications in Plant Sciences* **2**: apps.1300083.

Santini A, Ghelardini L, De Pace C, Desprez-Loustau ML, Capretti P, Chandelier A, Cech T, Chira D, Diamandis S, Gaitniekis T, *et al.* 2013. Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *New Phytologist* **197**: 238ï 250.

Sapkota R, Nicolaisen M. **2015**. An improved high throughput sequencing method for studying oomycete communities. *Journal of Microbiological Methods* **110**: 33ï 39.

Sárándi-Kovács J, Nagy L, Lakatos F, Sipos G. **2016**. Sudden *Phytophthora* dieback of wild cherry trees in Northwest Hungary. *Acta Silvatica et Lignaria Hungarica* **12**: 117.

Saunders DGO, Win J, Kamoun S, Raffaele S. **2014**. Two-dimensional data binning for the analysis of genome architecture in filamentous plant pathogens and other Eukaryotes - Plant-pathogen interactions: Methods and protocols. In: Birch P,, In: Jones JT,, In: Bos JIB, eds. Totowa, NJ: Humana Press, 29ï 51.

Savory F, Leonard G, Richards TA. **2000**. The role of horizontal gene transfer in the evolution of fungi. *Annual Review of Phytopathology* **38**: 325ï 363.

Schlenzig A, Campbell R, Chard J. **2015**. *Phytophthora* species infecting hardy ornamentals in nurseries and the managed environment in Scotland. *Journal of Phytopathology* **163**: 686ï 689.

Schlenzig A, Campbell RB, Roberts AMI. **2017**. The susceptibility of selected conifer foliage to infection with *Phytophthora lateralis*. *Forest Pathology* **47**: e12333ï n/a.

Schmieder R, Edwards R. **2011**. Quality control and preprocessing of metagenomic datasets. *Sciences-New York*: 3ï 5.

Schornack S, van Damme M, Bozkurt TO, Cano LM, Smoker M, Thines M, Gaulin E, Kamoun S, Huitema E. **2010**. Ancient class of translocated oomycete effectors targets the host nucleus. *Proceedings of the National Academy of Sciences* **107**: 17421ï 17426.

Schoustra SE, Debets AJM, Slakhorst M, Hoekstra RF. **2007**. Mitotic recombination accelerates adaptation in the fungus *Aspergillus nidulans*. *PLoS Genetics* **3**: e68.

Schroeder KL, Martin FN, de Cock AWAM, Lévesque CA, Spies CFJ, Okubara PA, Paulitz TC. **2013**. Molecular detection and quantification of Pythium species: Evolving taxonomy, new tools, and challenges. *Plant Disease* **97**: 4ï 20.

Scott P, Burgess T, Hardy G. **2013**. Globalization and *Phytophthora*. In: Lamour K, ed. Phytophthora: A Global Perspective. CABI, 226ï 232.

Seidl MF, Van den Ackerveken G, Govers F, Snel B. **2012**. Reconstruction of Oomycete genome evolution identifies differences in evolutionary trajectories leading to present-day large gene families. *Genome Biology and Evolution* **4**: 199ï 211.

Shishkoff N. **2007**. Persistence of *Phytophthora ramorum* in soil mix and roots of nursery ornamentals. *Plant Disease* **91**: 1245ï 1249.

Shrestha SK, Zhou Y, Lamour K. **2013**. Oomycetes baited from streams in Tennessee 2010ï 2012. *Mycologia* **105**: 1516ï 1523.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. **2015**. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210ï 3212.

Simpson JT, Wong K, Jackman SD, Simpson JT, Durbin R, Salzberg SL, Phillippy AM, Zimin A, Simpson JT, Wong K, *et al.* 2012. ABySS¡ A parallel assembler for short read sequence data structures assemblies and assembly algorithms. *Genome Research* **19**: 1117ï 1123.

Sims LL, Sutton W, Reeser P, Hansen EM. **2015**. The *Phytophthora* species assemblage and diversity in riparian alder ecosystems of western Oregon, USA. *Mycologia* **107**: 889ï 902.

Snel B, Bork P, Huynen MA. **1999**. Genome phylogeny based on gene content. *Nature Genetics* **21**: 108ï 110.

Soanes D, Richards T a. **2014**. Horizontal gene transfer in eukaryotic plant pathogens. *Annual review of phytopathology* **52**: 583ï 614.

Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM. **2016**. EffectorP: Predicting fungal effector proteins from secretomes using machine learning. *New Phytologist* **210**: 743ï 761.

Stamatakis A, Hoover P, Rougemont J. **2008**. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic biology* **57**: 758ï 771.

Stamler RA, Sanogo S, Goldberg NP, Randall JJ. **2016**. Identification of *Phytophthora* species in rivers and streams of the Southwestern United States. *Applied and Environmental Microbiology* **82**: 4696ï 4704.

Stewart JE, Kroese D, Tabima JF, Larsen MM, Fieland VJ, Press CM, Zasada IA, Grünwald NJ. **2014**. Pathogenicity, fungicide resistance, and genetic variability of *Phytophthora rubi* isolates from raspberry (*Rubus idaeus*) in the western United States. *Plant Disease* **98**: 1702ï 1708.

Stipes RJ, Campana RJ (Eds.). **1981**. Brasier Clive M,. Compendium of Elm diseases. St Paul, Minnesota: American Phytopathological Society, 76ï 79.

Stoletzki N, Eyre-walker A. **2011**. The positive correlation between dN/dS and dS in mammals is due to runs of adjacent substitutions. *Molecular Biology and Evolution* **28**: 1371ï 1380.

Strenge D, Elliott M, Chastagner G, Sclar C, Stern D. **2016**. Managing *Phytophthora ramorum* at the Bloedel Reserve. Proceedings of the Sudden Oak Death Sixth Science Symposium.43ï 44.

Stukenbrock EH. **2013**. Natural selection and genome plasticity in play: a genomic view on speciation in fungal plant pathogens. *New Phytologist*: 895ï 907.

Stukenbrock EH. **2016**. The role of hybridization in the evolution and emergence of new fungal plant pathogens. *Phytopathology* **106**: 104ï 112.

Stukenbrock EH, Bataillon T. **2012**. A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. *PLoS Pathogens* **8**: 1ï 4.

Stukenbrock EH, McDonald BA. **2008**. The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology* **46**: 75ï 100.

Sung JM, Jin KS, Park JS. **1983**. Identification and pathogenicity of *Pythium* spp. associated with seedling damping off of rice. *Korean Journal of Mycology* **11**: 27ï 32.

Sutton W, Hansen EM, Reeser P, Kanaskie A. **2008**. Comparing *Phytophthora ramorum* diagnostic protocols for the national sudden oak death stream monitoring program. Proceedings of the Sudden Oak Death Third Science Symposium. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Pages 461-466.

Tajima F. **1989**. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585ï 595.

Thompson JD, Higgins DG, Gibson TJ. **1994**. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673ï 4680.

Tischfield JA. **1997**. Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination. *American Journal of Human Genetics* **61**: 995ï 999.

Torto-Alalibo T, Tian M, Gajendran K, Waugh ME, van West P, Kamoun S. **2005**. Expressed sequence tags from the oomycete fish pathogen *Saprolegnia parasitica* reveal putative virulence factors. *BMC Microbiology* **5**: 46.

Torto T, Li S, Styer A, Huitema E, Testa A, Gow NA, van West P, Kamoun S. **2003**. EST mining and functional expression assays identify extracellular effector proteins from plant pathogen *Phytophthora*. *Genome Res* **13**: 1675ï 1685.

Tyler BM. **2002**. Molecular basis of recognition between *Phytophthora* pathogens and their hosts. *Annual Review of Phytopathology* **40**: 137ï 167.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RHY, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, *et al.* 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313**: 1261ï 1266.

Vannini A, Bruni N, Tomassini A, Franceschini S, Vettraino AM. **2013**. Pyrosequencing of environmental soil samples reveals biodiversity of the *Phytophthora* resident community in chestnut forests. *FEMS Microbiology Ecology* **85**: 433ï 442.

Venables WN, Ripley BD. **2002**. *Modern Applied Statistics with S*. New York, NY.: Springer.

Vercauteren A, De Dobbelaere I, Van Bockstaele E, Maes M, Heungens K. **2011**. Genotypic and phenotypic characterization of the European A2 isolates of *Phytophthora ramorum*. *European Journal of Plant Pathology* **129**: 621ï 635.

Vercauteren A, De Dobbelaere I, Grünwald NJ, Bonants P, Van Bockstaele E, Maes M, Huengens K. **2010**. Clonal expansion of the Belgian *Phytophthora ramorum* populations based on new microsatellite markers. *Molecular Ecology* **19**: 92ï 107.

Wang Q, Garrity GM, Tiedje JM, Cole JR. **2007**. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**: 5261ï 5267.

Wernersson R, Pedersen AG. **2003**. RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Research* **31**: 3537ï 3539.

Werres S, Marwitz R, Man In't veld WA, De Cock AWAM, Bonants PJM, De Weerdt M, Themann K, Ilieva E, Baayen RP. **2001**. *Phytophthora ramorum* sp. nov., a new pathogen on Rhododendron and Viburnum. *Mycological Research* **105**: 1155ï 1165.

Weste G, Marks GC. **1987**. The biology of *Phytophthora cinnamomi* in australasian forests. *Annual Review of Phytopathology* **25**: 207ï 229.

Wheeler TJ, Eddy SR. **2013**. nhmmer: DNA homology search with profile HMMs. *Bioinformatics* **29**: 2487ï 2489.

Wickland AC, Jensen CE, Rizzo DM. **2008**. Geographic distribution, disease symptoms and pathogenicity of *Phytophthora nemorosa* and *Phytophthora pseudosyringae* in California, USA. *Forest Pathology* **38**: 288ï 298.

Wills RT. **1993**. The ecological impact of *Phytophthora cinnamomi* in the Stirling Range National Park, Western Australia. *Australian Journal of Ecology* **18**: 145ï 159.

Win J, Morgan W, Bos J, Krasileva K V, Cano LM, Chaparro-Garcia A, Ammar R, Staskawicz BJ, Kamoun S. **2007**. Adaptive evolution has targeted the C-Terminal domain of the RXLR effectors of plant pathogenic Oomycetes. *The Plant Cell* **19**: 2349ï 2369.

Winton LM, Hansen EM. **2001**. Molecular diagnosis of *Phytophthora lateralis* in trees, water, and foliage baits using multiplex polymerase chain reaction. *Forest Pathology* **31**: 275ï 283.

With KA. **2002**. The landscape ecology of invasive spread. *Conservation Biology* **16**: 1192ï 1203.

Yakabe LE, Blomquist CL, Thomas SL, MacDonald JD. **2009**. Identification and frequency of *Phytophthora* species associated with foliar diseases in California ornamental nurseries. *Plant Disease* **93**: 883ï 890.

Yang Z, Nielsen R. **1998**. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution* **46**: 409ï 418.

Yoshida K, Schuenemann VJ, Cano LM, Pais M, Mishra B, Sharma R, Lanz C, Martin FN, Kamoun S, Krause J, *et al.* 2013. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *eLife* **2013**: 1ï 25.

Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. **2017**. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**: 28ï 36.

Zdobnov EM, Apweiler R. **2001**. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**: 847ï 848.

Appendices

Appendix A: Supplementary tables chapter 2

Table S2.7: Sequencing, filtering and taxonomy assignment details for the two sampling periods.

	Fall	Spring
Number of samples	124	122
Total number of 454 reads	620355	760186
Filtering		
sequences too short	158917	193267
Poor quality	74118	93666
Ambiguous sequences and homopolymers	874	2563
primer or barcode mismatch	52518	36927
Filtered from ITSx	42201	74939
Number of sequences after quality filtering	291727	247053
Singletons	663	412
Number of seqs minimum count 2	291064	246641
Descriptive data		
Range of sequences per sample	12 to 7273	15 to 5592
Number of OTUs (total)	1311	1692
number of OTUs not assigned	138	99
Stramenopile reads	285746	240541
Stramenopile OTUs	1173	1593
Oomycete reads	266609	208128
Oomycete reads (minimum reads 2)	92	84
Oomycete OTUs	609	1015
Range of Oomycete sequences per sample	12 to 7273	0 to 5560
Mean number of Oomycete sequences per sample	2150	1706
Range of Oomycete OTUs per sample	2 to 78	0 to 58
Mean number of Oomycete OTUs per sample	24	22
Number of Oomycete OTUs identified to Genus	427	710
Number of Oomycete OTUs identified to Species	319	468

Table S2.8: OTU species level taxonomic identifications using manual identification and automated identification with clustering at 97% similarity.

OTUs identified with both manual and automated methods	Addition OTUs from manual taxonomic identification	Addition OTUs from automated taxonomic identification
Phytophthora amnicola	*Phytophthora* Unknown_Clade9_large_group_parsiana_like	*Phytophthora* sp_xHennops
Phytophthora cactorum	*Phytophthora cryptogea*	*Phytophthora alni*
Phytophthora chlamydospora	*Phytophthora gregata*	*Phytophthora bilorbang*
Phytophthora europaea	*Phytophthora intercalaris*	*Phytophthora cambivora*
Phytophthora rubi/fragariae	*Phytophthora megasperma*	*Phytophthora idaei*
Phytophthora gallica	*Phytophthora pini/citricola*	*Phytophthora irrigata*
Phytophthora gonapodyides	*Phytophthora* taxon_kelmania	*Phytophthora mississippiae*
Phytophthora hibernalis	*Phytophthora* taxon_Oaksoil	*Phytophthora* Other
Phytophthora hydropathica	unknown sp_1	*Phytophthora* s_
Phytophthora inundata	unknown sp_2	*Phytophthora* sp_ohioensis
Phytophthora lacustris	unknown sp_3	*Phytophthora* sp_raspberry
Phytophthora lateralis		*Phytophthora* sp_Taxodium_1
Phytophthora multivora		*Phytophthora* uliginosa
Phytophthora nemorosa		
Phytophthora plurivora		
Phytophthora polonica		
Phytophthora pseudosyringae		
Phytophthora pseudotsugae		
Phytophthora quercina		
Phytophthora ramorum		
Phytophthora sansomeana		
Phytophthora siskiyouensis		
Phytophthora syringae		

Appendix B: Chapter 4, supplementary material

DNA extraction, genome sequencing, SNP extraction and filtering

Methods

DNA extraction. DNA was extracted from *P. ramorum* isolates cultured on potato dextrose agar (PDA) media at 20°C under a 12-hour photoperiod for 7 to 10 days. We used a chloroform method (Möller *et al.*, 1992) with the following modification: tissue was collected from mycelia grown on a cellophane membrane (GE Healthcare Bio-Sciences Corp., Piscataway Township, NJ) placed on top of the agar. The mycelia was ground and placed in a 2ml tube with 500µl of TES Buffer and 4ul of Proteinase K (20mg/ml) and incubated for 30 min at 55°C, with occasional gentle mixing. One hundred and forty microlitres NaCl (5M) and 65µl 10% CTAB was added and samples were incubated for 10 min at 65°C. Seven hundred microlitres chloroform:isoamyl alcohol (24:1) was added and samples were vortexed and centrifuged for 5 min at 13 000rpm. The supernatant was transferred to a fresh 1.5ml tube and 3µl RNase (10mg/ml) was added. Samples were incubated for 30 min at 37°C, centrifuged for 1 min at 13 000 rpm. Supernatant was transferred to a fresh 1.5ml tube and 150µl ammonium acetate (3M) and 500µl isopropanol was added. Samples were placed on ice for 30 min and centrifuged 5 min at 13 000 rpm, 4°C. The supernatant was aspirated off and the pellet washed with cold 70% EtOH (~250µl) and centrifuged for 5 min at 13 000 rpm, 4°C. The pellet was air dried for 10 min and suspended in 50µl TE buffer. DNA concentration was quantified with a Qubit Fluorometer (Life Technologies Inc., Grand Island, NY).

Genome Sequencing. Genome sequencing was done at Canada's Michael Smith Genome Sciences Centre or GSC (Vancouver Canada). Libraries were constructed using one of two methods (method for each isolate in Table S4.1). The majority of the libraries were constructed on a SPRI-TE robot (Beckman-Coulter, USA) according to the manufacturerôs instructions (SPRIworks Fragment Library System I Kit, A84801). Briefly, 1ug of genomic DNA in a 60uL volume, and 96-well format, was fragmented by Covaris E210 sonication for 30 sec using a ñDuty cycleò of 20% and ñIntensityò of 5. Up to 10 paired-end genome sequencing libraries were prepared in parallel using the SPRI-TE 300-600bp size-selection program. Following completion of the SPRI-TE run the adapter ligated library templates were quantified using a Qubit fluorometer. Five ng of adapter ligated template was PCR amplified using Phusion DNA Polymerase (Thermo Fisher Scientific Inc. USA) and Illuminaôs PE indexed primer set, with cycle conditions: 98ęC for 30 sec followed by 10 cycles of 98ęC for 15 sec, 62ęC for 30 sec and 72ęC for 30 sec, and a final amplicon extension at 72ęC for 5min.

For the remaining samples, 2ug of genomic DNA in a 96-well format was fragmented by Covaris E210 sonication for 30 sec using a ñDuty cycleò of 20% and ñIntensityò of 5. The paired-end sequencing library was prepared following the BC Cancer Agencyôs Genome Sciences Centre 96-well Genomic ~350bp-450bp insert Illumina Library Construction protocol on a Biomek FX robot (Beckman-Coulter, USA). Briefly, the DNA was purified in a 96-well microtitre plate using Ampure XP SPRI beads (40-45uL beads per 60µL DNA), and was subject to end-repair, and phosphorylation by T4 DNA polymerase, Klenow DNA Polymerase, and T4 polynucleotide kinase respectively in a single reaction, followed by cleanup using Ampure XP SPRI beads and

3ô A-tailing by Klenow fragment (3ô to 5ô exo minus). After cleanup using Ampure XP SPRI beads, picogreen quantification was performed to determine the amount of Illumina PE adapters used in the next step of adapter ligation reaction. The adapter-ligated products were purified using Ampure XP SPRI beads, then PCR-amplified with Phusion DNA Polymerase (Thermo Fisher Scientific Inc. USA) using Illuminaôs PE indexed primer set, with cycle conditions: 98ęC for 30 sec followed by 6 cycles of 98ęC for 15 sec, 62ęC for 30 sec and 72ęC for 30 sec, and a final extension at 72ęC for 5 min.

For both library construction methods, PCR products were purified using Ampure XP SPRI beads, and quantified with Caliper LabChip GX for DNA samples using the High Sensitivity Assay (PerkinElmer, Inc. USA). PCR products of desired size range were purified using gel electrophoresis (8% PAGE or 1.5% Metaphor agarose gels in a custom built robot) and DNA quality was assessed and quantified using an Agilent DNA 1000 series II assay and Quant-iT dsDNA HS Assay Kit using Qubit fluorometer (Invitrogen), then diluted to 8nM. The final concentration was verified by Quant-iT dsDNA HS Assay prior to Illumina Sequencing. For sequencing, 100bp paired-end tagged (PET) reads were prepared. Twelve individuals were pooled per lane and sequencing was performed on the Illumina HiSeq 2000 (Illumina Inc). Reads were provided in BAM format for each individual.

Mapping and filtering reads. PRINSEQ v0.20.3 (Schmieder & Edwards, 2011) was used to filter the bam files for redundant reads, reads containing one or more ôNôs, and to trim low quality bases from the ends of the reads. Trimming was done by discarding the last ten nucleotides if the average quality fell below 20 (average quality was calculated on windows of

10bp). Trimmed reads were mapped onto the *P. ramorum* reference genome (Pr-102 isolate) (Tyler *et al.*, 2006) available through the Department of Energy-Joint Genome Institute (DOE-JGI, Walnut Creek USA) (version 1.1) with the Burrows-Wheeler Aligner (BWA) (Li & Durbin, 2009) using the default parameters. Mapping coverage, the percent of reads mapped and mapping quality statistics were calculated using Qualimap v.0.7.1 (García-Alcalde *et al.*, 2012).

Extracting SNPs and filtering. Variant sites were obtained in a mulitvcf format using the mpileup program of Samtools and the Bayesian variant calling models implemented in BCFtools (Li *et al.*, 2009). Single nucleotide polymorphisms (SNPs) were filtered using VCFtools (Danecek *et al.*, 2011) with the parameters determined from testing to minimize false positives and maximize true positives (*Appendix B, Testing SNP filtering parameters and false call rates*); a maximum mean depth of 90 and a minimum mean depth of 10, and for each SNP, a minimum quality of 30, a minimum depth of 10, a minimum mapping quality of 30, and a minimum distance to a gap of 10bp. For each genotype, we set a minimum depth of 4 reads, a maximum of 240, and a minimum genotype quality of 20. We then applied a filter for a minimum allele count of two (an allele needed to be found in at least two individuals in the data set) and for no missing data.

Testing SNP filtering parameters and false call rates. To set the final vcf filtering parameters, various values for each parameter were applied to a subset of genomes for 92 isolates using VCFtools (Danecek *et al.*, 2011). The resulting SNP datasets were screened for a set of true positive SNPs and false positive SNPs. For true positive SNPs, Sanger sequenced SNPs from previous studies (Bilodeau, 2008; Goss *et al.*, 2009) were used. For false positives, genes with

close paralogs in the JGI reference genome (NA1 lineage) were aligned to their paralogs, and differing sites between the two genes were used as heterozygous false positives. NA1 isolates were then screened for the false positives. For a maximum depth, we tested a mean maximum depth filter (the mean coverage of all individuals) set at just below the mean average mapping coverage of 92 genomes, just above the mean average coverage and approximately two times the mean average coverage. We then tested SNP filters for minimum coverage, minimum variant quality, minimum mapping quality, and distance of the SNP to a gap using VCF-annotate from the VCFtools package (Danecek *et al.*, 2011). Finally we tested genotype filters for minimum and maximum depth at an individual SNP, minimum genotype quality, and a site filter for minimum average coverage (using the mean coverage of all individuals). Mean minimum and maximum site filters allowed us to accept SNPs with low or high coverage in some individuals provided the mean values did not go below or above the set threshold. The number of SNPs filtered, the number of true positive SNPs found, and the number of false calls was compared for the different parameters tested to choose a combination of filters that maximized the number of true positives found and minimized the number of false calls.

To assess the rate of false SNPs generated during the sequencing process, one individual was sequenced twice and one individual three times in independent runs using the same conditions for library preparation and sequencing. These individuals were trimmed and mapped to the reference as above and SNPs extracted for each set of isolates relative to the JGI reference using the mpileup function in Samtools as above. The resulting BCF file was converted to a VCF and the total number of discordant SNPs between duplicate individuals in independent runs was obtained

for the raw data set. The VCF was filtered following the above parameters and the total number of discordant SNPs relative to the total number of SNPs was used to calculate a false call rate.

SNP counts and population diversity. To compare SNP content between lineages, for each SNP in the dataset, we first determined if a given SNP was fixed in all isolates within a lineage (but different between lineages), or if it was polymorphic in some of the individuals in a lineage. We then determined if the SNP was homozygous or heterozygous for each of the two categories (fixed and polymorphic).

Results and discussion

Genome Sequencing and mapping statistics. The *P. ramorum* reference genome (Pr-102 ; NA1 lineage) sequenced at the DOE-JGI was estimated at approximately 65Mbp with 15 743 genes (Tyler *et al.*, 2006). For the 107 isolates re-sequenced in this project, we achieved an average mapping coverage of 49.4X with a range of 16.7X to 80.9X. The average percent of the total sequenced reads that mapped to the DOE-JGI reference was 92.2% (ranged from 62.6% to 95.9%; however, all isolates except one were above 85%). Mapping of isolates from the different lineages to the NA1 reference genome showed similar ranges in the percent of reads mapped with the highest mapping occurring in the NA1 lineage. The EU2 isolates had the least variance in the percent of reads mapped, however the maximum percent of reads that mapped to the NA1 reference was lower (NA1: 62.6 to 95.9%; EU1: 85.8 to 94.9%; NA2: 87.2 to 94.8%; EU2: 92.7 to 92.9%).

SNP extraction and filtering. The total number of SNPs extracted from the initial set of 92 isolates used to test filter parameters was 912 622 (prior to filtering). No false positive validation SNPs were found (out of the set of false SNPs created from close paralogs) prior to or after applying any of the filters (Table S4.12). At all of the depth filtering parameters nearly twice the average coverage (mean depth filters 80, 90 and 100; mean depth 49.4X; Table S4.12), greater than 90% of the Sanger sequenced validation SNPs were found and only one unexpected, possibly false positive was found (TT expected and CA found). The only other discrepancies between the Sanger data and the Illumina data were cases where a SNP was called with one allele difference from the expected (TT expected, CT found) which could be due to either a false call or a new variant. The final set of filtering parameters used were a minimum average SNP coverage over all individuals of 10 reads, with a minimum of four reads covering each SNP for each individual; a maximum average SNP coverage over all individuals of 90 reads with no individual exceeding a coverage of 240X at a SNP; a minimum site quality of 30 with each individual reaching a minimum quality of 20; a minimum mapping quality of 30 around each SNP; and a minimum distance of 10 bases between each SNP and the closest gap. With this set of filtering parameters, 434 504 SNPs remained in the set of 92 test isolates (431 112 loci excluding singleton SNPs). Ninety-seven percent of the validation SNPs were found and 96.5% were the expected call. For the full data set of 107 isolates, the total number of variants including indels prior to filtering was 1 085 718. After filtering there was a total of 895 274 SNPs. Excluding singletons and missing data, there were a total of 485 327 bi-allelic sites.

The rate of false calls between independent runs of the same individuals ranged from 4.6% (two runs) to 8.3% (three runs). After applying the final filtering parameters, this rate dropped to 0.7%

to 1.8% (Table S4.13). The false call rate dropped to less than 0.2% when the SNP sites that were filtered out of the main data set for missing data and singletons were removed from the controls. From these results, the final filtering parameters were chosen (above) and the data set with a minimum allele count of two (no singleton SNPs) and no missing data were used in the population analyses.

SNP counts and lineage genetic diversity. The majority of the SNPs discovered by mapping the Illumina reads onto the reference NA1 genome were between the four clonal lineages rather than within, where the individuals within each clonal lineage were fixed for the alleles (either homozygous (72% to 80%) or heterozygous (20% to 30%)) (Table S4.3). The fixed heterozygous loci accounted for 81% to 92% of the total heterozygous loci in each lineage (% fixed heterozygous loci = 88% for EU1, 91% for NA1, 81% for NA2 and 92% for EU2). Genome wide SNP profiles of the isolates within lineages were very similar. The majority of the SNPs that were polymorphic (varied between the individuals within a lineage) were heterozygous, and the highest number was found in the NA2 lineage (Table S4.3). Nucleotidic diversity was calculated for each lineage and was highest in the EU1 lineage and lowest in the EU2 lineage (Pi = 0.00116 for EU1, 0.00108 for NA1, 0.00108 for NA2 and 0.00077 for EU2).

De novo genome assemblies

Methods

Two representative isolates of each of the four clonal lineages was selected for *de novo* genome assembly. For each isolate, Illumina sequencing reads were filtered by trimming the barcode

sequence using PRINSEQ v0.20.3 (Schmieder & Edwards, 2011) and assembled into contigs and pre-scaffolded using ABySS (Simpson *et al.*, 2012) with k-values ranging from 32 to 96. Final scaffolding was completed with SSPACE v.3.0 (Hunt *et al.*, 2014) and the best assembly was then selected based on genome size (~50.0Mb) and contiguity (best N50 and length of longest scaffold). Completeness of the *de novo*-genome assemblies was assessed using BUSCO (Benchmarking Universal Single-Copy Orthologs) (Simão *et al.*, 2015). The *de novo*-assemblies were then repeat masked with RepeatMasker (A.F.A. Smit, R. Hubley & P. Green RepeatMasker at http://repeatmasker.org) using 2,101 *Phytopthora* spp. repeats available in Repbase (Jurka *et al.*, 2005) and annotated using AUGUSTUS ver. 2.7 (Keller *et al.*, 2011) trained with models from *P. sojae*, *P. infestans* and *P. ramorum* available on the JGI server (Grigoriev *et al.*, 2012). Predicted protein models were functionally annotated with Blast2GO (Conesa *et al.*, 2005) following homolog searches using BLASTp (Altschul *et al.*, 1990) against NR (E-value cut-off of 1e-05) and a protein domain search using Interproscan (Zdobnov & Apweiler, 2001).

Results and discussion

Genome assembly statistics were in the range of those obtained for other *Phytophthora* and Oomycete *de novo* assemblies generated from short Illumina re-sequencing reads; completeness reached 82% of the BUSCO eukaryotic set, a value similar to those obtained for other Oomycete *de novo* assemblies and slightly higher than the 81.6% value obtained for the JGI Pr-102 genome assembly (Tyler *et al.*, 2006; Feau *et al.*, 2016) indicating these Illumina *de novo* assemblies can be considered complete(Table S4.14).

Phylogenetic analysis

Methods

Neighbor joining phylogenetic reconstruction. To reconstruct the phylogenetic relationships of the isolates using genome wide SNP patterns, the VCF was converted to a Plink format using VCFtools (Danecek *et al.*, 2011) and Plink v1.07 (Purcell *et al.*, 2017) was used to recode the data for use in the package Adegenet (Jombart, 2008). Euclidean distances were calculated using the stats-package version 3.4.0 in R (R Core Team, 2017) and used to construct a neighbor-joining tree using the APE package (Paradis *et al.*, 2004). Arlequin 3.0 (Excoffier *et al.*, 2005) was used to calculate pairwise F_{ST} values between the lineages.

Maximum Likelihood phylogenetic reconstruction. To assess evolutionary dynamics on *P. ramorum*, we reconstructed the phylogenetic relationships of all the isolates using a selection of genes under total neutrality. We extracted the complete set of alternate gene calls for the 107 isolates. Data from homozygous isolates were used to phase the heterozygous isolates using Beagle4 (Browning & Browning, 2007). Alternate genome sequences were generated using the FastaAlternateReferenceMaker script from the GATK suite (McKenna *et al.*, 2010), using the VCF file as the source of variants. We used BEDTools (Quinlan & Hall, 2010) to extract the genes from their genomic coordinates. Two Fasta formatted files per gene, one for each strand, were obtained for each of the isolates. Each alternate gene model was translated to its amino acid sequence. Each gene model and amino acid translation was aligned using RevTrans (Wernersson & Pedersen, 2003). RevTrans aligns the amino acid sequences and uses this alignment as a scaffold to construct the gene model alignment. After alignment, Tajimaôs D was measured using

the package pegas (Paradis, 2010). We selected the gene models with a Tajimaôs D value with a significance greater than 0.05 (genes under neutrality).

For each gene, models of nucleotide substitution were calculated in jModelTest 2 (Darriba *et al.*, 2012) for 500 random neutral genes across the genome. All genes were concatenated and partitioned using 2matrix (Salinas & Little, 2014). The partitioned matrix was ran in RAxML 8.1.24 (Stamatakis *et al.*, 2008) to reconstruct the phylogenetic tree using Maximum likelihood and node supports using non-parametric bootstrap resampling using the CIPRES Science gateway (Miller *et al.*, 2010). One thousand bootstrap repetitions were used and the bipartition tree (best tree with bootstrap support values) was obtained. The tree was midpoint rooted to determine the most divergent clade inside *P. ramorum.*

Results and discussion

Maximum likelihood phylogenetic reconstruction. In the phylogenetic reconstruction, a total of 5588 genes were found under neutrality, 2050 genes under positive selection and 1984 genes were non-segregating. The phylogenetic reconstruction using a subset of 500 neutral genes yielded a topology similar to the neighbor-joining tree (Fig. S4.1), where all clonal lineages were clustered and highly supported monophyletic clades. In the case of NA1, two subclades can be observed, but lack bootstrap support of this splitting event. In the phylogenetic analysis, the two strands of phased haplotypes are clustered as expected for an asexual species (the Meselson effect; (Birky-Jr., 1996)), except for a few individuals, which were the same individuals forming clusters or diverging in the neighbor-joining analysis. These were isolates with runs of

homozygosity (ROH; *Appendix B, Effects of ROH on genotype*) which possessed two identical or near identical alleles for each gene.

Divergence time between lineages

Methods

Phylogenetic analysis of one-to-one orthologs. Sequences from the 4929 one-to-one ortholog clusters obtained for seven *Phytophthora* species, including *P. lateralis* and the four lineages of *P. ramorum* (*Appendix B, OrthoMCL analysis and evolution of gene family size*) were aligned with MAFFT (Katoh & Standley, 2013) and filtered for sequence alignments > 100 amino-acids with no gaps. The resulting 492 alignments were then submitted to the PAML package of the program CODEML, to estimate the ratio ω = dN/dS and test for positive selection by comparing the ñsite-specificò models M1a/ M2a and M7/M8 (Katoh & Standley, 2013) as described in (Alamouti *et al.*, 2014). This resulted in 49 polymorphic alignments (Ó4 SNPs between the *P. ramorum* lineages) with no evidence of positive selection that were retained. Each alignment was submitted to ProtTest v. 2.4 (Abascal *et al.*, 2005) and then concatenated into different partitions according to the best protein evolution model fitted (DayHoff, WAG or JTT), resulting in a super-alignment of 13 879 amino-acids. A maximum likelihood tree search followed by a rapid bootstrap analysis with 1000 replicates was then performed under partitioned data mode with RAxML 8.0.12 (Stamatakis *et al.*, 2008).

Divergence time between *P. ramorum* lineages. We used the Bayesian clock method implemented in BEAST v1.7.5 (Drummond *et al.*, 2012) to estimate divergence times between

the four *P. ramorum* lineages under three distinct clock models with the alignment generated for the phylogenetic analysis. Analyses were run with the strict clock, random local clock and UCLD relaxed clock models used in (Matari & Blair, 2014). For each run, partitions were treated separately under WAG, JTT or DayHoff model depending on the protein evolution model fitted, and the RAxML tree obtained in the previous analysis was used as a user-specified starting tree; a Yule speciation process was assumed with a uniform distribution on the birthrate (0ï 100; initial value 0.01). The 3 clock models were set with prior parameters defined in (Matari & Blair, 2014). To calibrate the molecular clock analyses, normal distributions were specified for *Phytophthora* species of the phylogenetic clades seven and eight (*P. sojae*, *P. cinnamomi*, *P. lateralis* and *P. ramorum*; mean = 22.3 My, SD = 0.61), clades one and two (*P. infestans*, *P. capsicii* and *P. parasitica*; mean = 21.5 My, SD = 0.67) and the radiation of the *Phytophthora* genus (mean = 26.6 My, SD = 0.67), according to the results of the strict clock analysis of (Matari & Blair, 2014).

For each dataset, the three clock models were tested under three independent runs with 50 million generations each. LogCombiner v1.7.5. was used to combine log and tree files from the three independent runs, for each model. Tracer v1.6 (http://tree.bio.ed.ac.uk/software/) was used to evaluate convergence, estimate the appropriate burn-in for each run, and calculate Bayes factors for model comparisons. Trees were visualized in FigTree v1.4 .2. (http://tree.bio.ed.ac.uk/software/).

Results and discussion

Phylogenetic analysis. The RAxML phylogenetic analysis conducted on a 13 879 amino-acids alignment, resulted in a maximum likelihood tree with a likelihood value of ï ln 83715.663913 and a tree topology congruent with *Phytophthora* phylogenies reconstructed from smaller sequence datasets (Blair *et al.*, 2008; Matari & Blair, 2014). The tree topology is strongly supported with bootstrap values of 100% for six out of the eight nodes. The relationship among the four *P. ramorum* lineages was well supported, with values varying between 82% (node NA2 with NA1/EU1) and 92% (node NA1/EU1) (Fig. S4.7).

Divergence time estimates. The low coefficient of variation obtained from a root-to-tip analysis of the RaxML phylogenetic tree with TempEst V1.5 (Rambaut *et al.*, 2016) indicated a low degree of deviation of this dataset from the strict molecular clock hypothesis. However, posterior distributions of parameters were consistent across all three runs, and clock/ucld parameter estimates consistent across the three models (Table S4.15). Similarly, node divergence estimates were consistent across the three models with mean values obtained under one model generally falling within the 95% confidence intervals obtained under the other two models (Table S4.16). Estimated time points suggest a divergence between *P. lateralis* and *P. ramorum* at 6.38 My (SD, 5.91-6.84) and divergence between the lineages within *P. ramorum* at 1.31 My (SD, 1.14-1.48) (EU2 from the three other lineages), 1.06 (0.91-1.20) (NA2) and 750,000 (0.61-0.88) years (split between EU1 and NA1) (Fig. S4.7; Table S4.16). These estimates are in the same order of magnitude as those obtained from nuclear sequence data, with reproductive isolation between *P. lateralis* and *P. ramorum* estimated between 1.5 and 5.4 My, and a divergence between the NA1, NA2 and EU1 lineages ranging from 0.5 to 0.17 My (Goss *et al.*, 2009).

Genome-level copy number variation

Methods

Genome-level chromosomal copy number variation (CCNV) was detected by using two methods based on the relative number of mapped reads. First, we used the read depth approach implemented in Control-FREEC software for high resolution CCNV quantification. A mappability profile was generated from the *P. ramorum* reference genome (Tyler *et al.*, 2006) with the gem-do-index and gem-mappability programs of the GEM library (Derrien *et al.*, 2012). Copy-number was then estimated for each 100bp non-overlapping window of the reference genome by running Control-FREEC on the BWA-mapped reads (bam format) and the mappability profile files with the following parameters : window = 100, telocentromeric = 0, ploidy = 2 to 6n.

We also used the measure read-depth ratio of alleles at heterozygous sites as described in Farrer et al. (Farrer *et al.*, 2013) and Kasuga et al. (Kasuga *et al.*, 2016). For each re-sequenced genome, the SAMtools (Li *et al.*, 2009) mpileup and Bcftools programs were run on each scaffold to generate allele counts at heterozygous loci. Only heterozygous loci with read depth Ó 20 and mapping quality Ó 30 were retained; loci with read ratios of heterozygous alleles equal or larger than eight were also excluded. Output files from both methods were then visualized by the way of a custom Python program (available upon request), providing plots of CCNV across scaffold and average ratio of reads of heterozygous alleles in sliding non-overlapping windows of 10Kb across scaffold.

Results and discussion

Plasticity in chromosome numbers i.e. aneuploidy is a common feature of asexual micro-organism genomes (Normark *et al.*, 2003). Polyploidy resulted in major evolutionary implications for some modern asexual lineages of the potato late blight pathogen *Phytophthora infestans* (Yoshida *et al.*, 2013). Similarly, CCNV was recently reported in *P. ramorum* with some NA1 isolates showing atypical non wild-type colony morphology (n*wt*) isolated from oaks. Oaks were described as a dead-end host for *P. ramorum*, as sporulation of the pathogen from oak trunk cankers has never been reported. Based on their observations, Kasuga et al. (Kasuga *et al.*, 2012) hypothesized that oak defense mechanisms could induce chromosomal aberrations resulting in the *nwt* phenotype change in *P. ramorum*. Similarly, the same authors observed CCNV and the *nwt* phenotype in EU1 isolates infecting the rare host Lawson cypress (*Chamaecyparis lawsoniana*). We looked for CCNV and aneuploidy as a possible cause for the observed ROH in our isolates, as well as a potential driver of population diversity within the lineages. We observed chromosomal differences in a few NA1 and EU1 isolates, with inferred ploidy varying from 2n to 5n depending on the scaffold analyzed (Table S4.4; Fig. S4.3). CCNV was not responsible for the ROH pattern; although two of the isolates with CCNV had ROH, it was not observed on the same scaffolds as the ROH was. In addition, no signs of aneuploidy were observed in isolates with ROH. Similar to ROH, CCNV could affect phenotype, in particular by increasing the expression of genes with increased copies. It could also relax the selection pressure on the extra gene copy allowing for adaptation to occur.

Effects of ROH on genotype

Results and discussion

Runs of homozygosity and population evolution. ROH was observed in individuals in all four lineages producing homozygous scaffolds in single individuals which were heterozygous in the general population (Fig. S4.2); lengths ranged from 60 to 339Kb in EU1, 76 to 120Kb in EU2, 131 to 236Kb in NA1 and 90 to 99Kb in NA2. In one EU1 individual, 13.5% of the loci that were heterozygous in the rest of the EU1 population were homozygous in the affected individual (Table S4.2). In comparison, 0.2% to 0.3% of the total SNP loci which were heterozygous in the general EU1 population were homozygous in individuals without the runs of homozygosity (Table S4.2).

In individual isolates, there were between 53 and 1587 genes in the affected scaffolds, and between 24 to 893 proteins with amino acid differences between the alleles. ROH resulted in a loss of approximately 14 to 880 alleles (Table S4.5).

The ROH pattern was slightly different in the NA2 lineage. In contrast to the other lineages, some of the ROHs in the NA2 lineage were often interrupted by heterozygous loci, or were not completely contiguous over the full region of the scaffold affected. Some isolates had short ROH alternating with heterozygous regions on affected scaffolds and in some cases this resulted in a general loss of heterozygosity (LOH) with no long ROH patterns. In addition, the NA2 lineage had more ROHs shared between all of the isolates than the other lineages.

There were also less allele differences resulting from ROHs in the individuals and scaffolds with non-contiguous ROHs in NA2, which could suggest that these are older ROHs. In one individual, Pram_04_20470, scaffolds 12 and 40 had 86 and 41 proteins respectively with amino acid differences between the alleles; however only 21 and 3 proteins were affected by ROH. In scaffold five, which comprised several ROHs or a general loss of heterozygosity shared by all NA2 individuals, there were 26 proteins (out of a total of 144 in the affected region; Table S4.5) with amino acid differences, about half of what the other lineages had on scaffold five (NA1 = 79, EU1 = 73, EU2 = 56). An older sexual recombination event, possibly in the source population, followed by asexual propagation could produce the observed pattern of non-continuous ROHs and high fixed heterozygosity.

GO terms associated with protein binding, transferase activity, oxidoreductase activity, hydrolase activity and RNA-directed DNA polymerase activity were the most highly ranked terms in the gene ontology analysis for genes affected by ROH (Table S4.17).

Amino acid differences in scaffold 7 ROH. In the shared ROH region on scaffold seven, 63 alleles (36%) with amino acid differences were lost in the eight EU1 isolates with ROH (Table S4.5). In the region affected in NA1 between 73 and 85 alleles were lost (using the other three lineages for comparison). Fifteen proteins had five or more amino acid differences between alleles in the affected region of EU1, and 20 proteins had five or more amino acid differences between alleles in the affected region of NA1 (Table S4.5). In one protein, one allele had a premature stop codon. Isolates with the ROH in EU1 had the allele with the stop codon mutation so may no longer have a functioning copy of this gene.

Enrichment of putative plant pathogenicity genes on scaffold 7. GO terms associated with lyase activity and hydrolase activity were enriched on scaffold seven (Table S4.6). Hydrolysis of glycosyl compounds and bonds were both GO terms associated with the carbohydrate active enzymes. There was also a moderate enrichment in terms associated with the breakdown of pectin (aspartyl esterase activity and pectinesterase activity) (Table S4.6).

Effects of ROH on phenotype

Methods

Growth of isolates with ROH in culture. Isolates were grown on 1.5% MEA prior to the experiment (Table S4.2). Agar plugs 5 mm in diameter were cut from the margin of the culture and placed onto the center of a petri plate containing 18ml of larch sapwood agar or Douglas-fir sapwood agar (Stipes & Campana, 1981). Sapwood agar was prepared by peeling twigs less than 0.5 cm in diameter from freshly fallen trees and homogenizing the tissues to obtain a sapwood pulp which was incorporated into the agar prior to autoclaving. Three replicates were made of each isolate on each agar type. Cultures were incubated in the dark at 25°C and growth was measured four, seven, 10 and 14 days post-inoculation. The number of spores present was counted at four, seven and 11 days on Douglas-fir agar and at seven and 11 days for larch agar, by observing the bottom of the petri plate under a dissecting microscope at two diagonal locations beginning from the inoculation point for days four and seven, and at four points 5 mm from the plug in each 90 degree direction for day 11.

Growth of isolates with ROH on Rhododendron leaves. Seventeen leaves of *Rhododendron* sp. showing no signs of disease were collected from each of three different plants. Leaves were surface sterilized in 10% bleach for 30 sec, rinsed in a sterile water bath for one minute, and in a second sterile water bath for an additional minute. Leaves were air dried and then pricked with a sterile pin on the left, abaxial surface, leaving an equal distance between the wound and the leaf edge, and between the wound and the mid-rib. A 7 mm diameter plug of mycelium, taken from a 7 day old culture grown on 15% V8-agar (15mls V8 juice, 2 g $CaCO_3$, 15 g agar in 1 L water) was placed mycelium side down onto the wound of each leaf. Inoculated leaves were incubated in clean plastic bins lined with 200 g autoclaved vermiculite and 100ml sterile distilled water. Each bin was placed in a plastic bag, and incubated in a growth chamber at 20 °C in the dark. Area of each lesion was measured daily from day three to seven, and day 10 to 12, with a final measurement on day 17. Sterile water was added once on day five to moisten the vermiculite. Leaves were checked visually for any signs of sporulation.

Growth of isolates with ROH on Douglas-fir and larch logs. Three Douglas-fir and three Japanese Larch trees between 10 cm and 25 cm diameter were freshly felled from the Malcolm Knapp Research Forest in Maple Ridge, British Columbia. The logs were cut into 0.5 m to 0.7 m bolts and sealed with epoxy-resin (Intergard 740, International Paint, Houston, TX, USA). Four isolates were inoculated on each bolt by placing a mycelial plug under the bark at the depth of the cambium and sealing it with a cotton plug and aluminum foil secured with tape (Brasier & Kirk, 2001). The log bolts were incubated at room temperature in a plastic bag for eight weeks. The length and width at the longest and widest points of each lesion were measured four and eight

weeks post inoculation (wpi). Differences between the groups were compared with a Kruskal-Wallis test. Statistical significance was assumed at a probability value of less than 0.05.

Results and discussion

Growth of isolates with ROH in culture. There were no indications of any gain or loss of fitness in isolates with ROH when grown in culture. The isolates without ROH had a slightly larger diameter than the isolates with the ROH in both the larch and Douglas-fir agar; however the differences were not significant (Table S4.18). The differences were largely due to one isolate, Pram_P2599_EU1, in the group without ROH which grew faster than all other isolates on both media. When this isolate was removed, the groups had average diameters that were not statistically different at all time periods (data not shown). No differences were found in the number of spores produced by each group of cultures in each of the collection times (Table S4.18).

Growth of isolates with ROH on Rhododendron leaves. No signs of gain or loss in fitness were observed for isolates with ROH on detached Rhododendron leaves. There were no significant differences between lesion length in the individuals with and without ROH. The average of the group was slightly higher for the individuals without ROH for the first 12 days, and slightly higher for the individuals with ROH by 17 days post inoculation; however the range of lengths were very similar for each group throughout the 17 days of the experiment (Fig. S4.12).

Growth of isolates with and without ROH on larch and Douglas-fir log bolts. In Japanese larch bolts eight weeks post inoculation, several lesions had grown together and in some of the bolts the entire cambium had turned necrotic reducing the number of reps from different log bolts, as well as the overall sample size of each group (number of isolates with ROH = 5, number of isolates without ROH = 6). The average lesion length of the remaining isolates with no ROH was longer (272.7 mm ±) than isolates with ROH (235.7 mm ±); however the difference was not significant (Kruskal-Wallis chi-squared = 0.833, p = 0.3613) (Fig. S4.5).In Douglas-fir eight weeks post inoculation, the average lesion length was higher in the isolates without ROH (155.3 mm versus 140.4 mm), however the difference was not significant (Kruskal-Wallis chi-squared = 0.6562, p = 0.4179) (Fig. S4.5).

OrthoMCL analysis and evolution of gene family size

Methods

The protein models obtained for the *de novo* assemblies EU1 030002 (14 095 models), EU2 P2586 (14 028), NA1 PR09-175 (14 213), NA2 0438813 (14 186) and *P. lateralis* CBS_168.42 (17 463) (*Appendix B*, De novo *genome assemblies*) were combined with those from the sequenced genomes of *P. infestans* (18 140 models) (Haas *et al.*, 2009), *P. sojae* (26 584 models) (Tyler *et al.*, 2006), *P. capsicii* (19 805 models) (Lamour *et al.*, 2012), *P. parasitica* (27 942 models) (Broad Institute (broadinstitute.org), 2014) and *P. cinnamomi* var. *cinnamomi* (26 131 models) (Grigoriev *et al.*, 2012) and a tentative clustering of one-to-one orthologs was carried out using OrthoMCL (Li *et al.*, 2003) (BLASTp search with E-value cut-off of 1e-05, coverage of at least 50% of the query sequence and identity of at least 30%; OrthoMCL inflation value of I = 4)

and then automatically aligned with MAFFT ver. 7.123b (Katoh & Standley, 2013). The resulting list of 51,451 OrthoMCL clusters was then submitted to two filters to minimize the confounding effect of truncated proteins resulting from fragmented de-novo assemblies and gene mis-predictions (Nowell *et al.*, 2014). First, redundant OrthoMCL clusters were eliminated by searching the full 51 451 OrthoMCL clusters with 1- and 2-member OrthoMCL clusters; clusters with a significant hit (BLASTn cut-off of 80% amino acid identity over a 80% alignment) to another OrthoMCL cluster were considered spurious and eliminated. The process was repeated iteratively over increasing member size classes (3-members, 4-members, etc.) until reaching an error rate of <0.05% (Nowell *et al.*, 2014). Second, OrthoMCL clusters containing multiple protein fragments were corrected by eliminating protein sequences from the same genome with an average pairwise amino acid identity of Ó80%. We performed GO enrichment analysis of the unique protein models using MGSA (Bauer *et al.*, 2011).

To reconstruct a neighbor-joining (NJ) tree representing genome content similarity, we calculated a distance matrix based on the gene presence/absence matrix inferred from the OrthoMCL analysis in which the distances measured between pairs of taxa are inversely proportional to the number of gene they share (Snel *et al.*, 1999). The tree was reconstructed by using the fneighbor program of the PHYLIP Ver. 3.696 package, with 1000 bootstrap replicates sampling columns of the presence/ absence matrix. The maximum likelihood model of CAFE (Computational Analysis of gene Family Evolution) (Snel *et al.*, 1999) was then used to study gene family expansions/contractions while taking into account the one-to-one ortholog phylogeny reconstructed for *Phytophthora* spp. (*Appendix B, Divergence time between lineages - Phylogenetic analysis of one-to-one orthologs*).

Core and non-core genomes

Results and discussion

For each lineage, in addition to the bias in codon usage for the non-core genome, there was also a bias for the genes located on repetitive DNA, but in lower frequencies than in the non-core(e.g. 72.7% of the genes found in the repetitive genome of NA2 showed the bias; Fig. S4.8) (Table S4.8). Different compositional feature such as G+C content and codon usage biases are usually diagnostic of gene regions with different origin than the rest of the genome (Becq *et al.*, 2010; Soanes & Richards, 2014; Ravenhall *et al.*, 2015). For example, strong compositional differences between the dispensable chromosomes and the core genome of the wheat fungal pathogen *Zymoseptoria tritici* suggested an origin by ancient horizontal gene transfer (Goodwin *et al.*, 2011). In addition, genome analyses in Oomycetes have demonstrated that HGT has had major impacts on the evolution of these organisms, specifically in the plant pathogens of the *Phytophthora* genus (Savory *et al.*, 2000). To test the HGT-origin hypothesis we searched for the presence of foreign top blast hits for the genes of the non-core genomes in the four *P. ramorum* lineages. Proportions of non-oomycetes hits were low and not significantly different from those observed in the full proteome ($G2 = 0.35$, $P = 0.95$), indicating that most of the genes found in the non-core genome were likely inherited vertically from Oomycete ancestors, ruling out the possibility of a recent HGT at the origin of the non-core genome of the *P. ramorum* lineages. Rather, the characteristics of the non-core genome suggest that genes in this region may be affected by a degenerative process that likely results in the strong differences observed with those of the core regions.

Genes encoding effectors

Methods

RxLR effectors. RxLR protein annotation was performed on *de novo* assemblies and also on the *P. ramorum* reference genome (Pr-102 isolate) (Tyler *et al.*, 2006) available through the Department of Energy-Joint Genome Institute (DOE-JGI, Walnut Creek USA) (version 1.1) (*Appendix B,* De novo *genome assemblies*). The latter was performed as a means of evaluating optimal strategies for identifying and filtering candidate Avh proteins. Annotation of Avh effectors for each lineage was comprised of two major parts, with each part optimized to produce the most number of candidates, which were manually inspected at the end. The first part used a large database of Avh family genes predicted from genomes of 23 *Phytophthora* species, including the *P. ramorum* genome, to predict candidate Avh genes within each *de novo* assembly. In order to refine and expand this list of candidates, the predicted Avh gene set was used to recursively search the *de novo* assembly. In addition, the second step was repeated, with the addition of Avh proteins known from the *P. ramorum* genome. Results from each recursive step were then combined with duplicates removed and manually reviewed. Details of each step in this approach are provided below.

Part 1 ï *Predicting Avh candidates.* First, genomic sequences were translated to six-frame open-reading frame using Emboss, which translates regions between stop codons in the forward and reverse complement directions with the user-specified minimum length of 90 amino acid residues (Rice *et al.*, 2000). These protein sequences were trimmed to start with the m-signal peptide. A

database of previously identified Avh (avirulence homolog) proteins was used to search *de novo* assemblies using Hmmer version 3.1b (Wheeler & Eddy, 2013). Protein homology is determined using probabilistic models called profile hidden Markov models, which assign position specific scoring for substitutions, insertions, and deletions, and gain power from use of a database that contains multiple instances from a sequence family. The database of *Phytophthora* Avh proteins contained 9,779 proteins identified from genomes of 23 *Phytophthora* species (*P. cajani*, *P. europaea*, *P. foliorum*, *P. hibernalis*, *P. litchii*, *P. megakarya*, *P. melonis*, *P. parvispora*, *P. pistaciae*, *P. syringae*, *P. uliginosa*, *P. vignae*, *P. cinnamomi* var. *robiniae*, *P.* taxon *niederhauserii*, *P. pisi*, *P. cinnamomi*, *P. rubi*, *P. fragariae*, *P. palmivora*, *P. parasitica*, *P. sojae*, *P. ramorum*, *P. infestans*)(provided by Danyu Shen and Brett Tyler, Oregon State University). Candidate proteins with an HMM score less than 20 and those lacking both a RxLR and dEER motif were removed, following which, candidate proteins identified more than once by the search were reduced to single instance for each protein sequence. Presence / location of a signal peptide cleavage site was predicted for each protein sequence using Signal-P 3.0 (Bendtsen *et al.*, 2004), where only sequences with probability of a signal peptide greater than or equal to 0.8 were retained. To maximize the number of candidates recovered during recursive protein prediction, no other filtering of protein sequences was performed at this step.

Part 2 ï *Recursive search.* Avh protein candidates identified in part 1 were converted into a HMM database. The first step was to create multiple sequence alignment using MUSCLE version 3.8.31 (Edgar, 2004) that was used to build a HMM model of aligned amino acid sequences using Hmmer. Candidate proteins were trimmed to start with the m-signal peptide, where those with an HMM score less than 10 and those lacking both the RxLR and dEER motif

were removed. Candidate proteins identified more than once by the search were reduced to a single instance and only those candidates with a signal peptide score greater than or equal to 0.8 were retained. This recursive search was repeated using a HMM database comprised of Avh protein candidates identified in part 1 and 370 previously identified in the *P. ramorum* genome (Jiang *et al.*, 2008). Resultant candidate proteins identified in both recursive searches were combined, and those identified more than once were removed. To aid manual curation, common motifs among the combined candidate proteins were identified using MEME version 4.9.1 (Bailey *et al.*, 2006), with user specified setting that the motif have a minimum of four amino acids and maximum of eight. The top three motifs identified by MEME were typically the m-signal peptide region, RxLR and dEER, where strength of the m-signal peptide region was likely increased due to trimming of reads to this position.

Part 3 ï *Optimization and manual curation.* Effector annotation was performed on all *de novo* assemblies and also on the *P. ramorum* reference genome. Each candidate protein with an RxLR motif identified by MEME was subsequently categorized according to actual RxLR sequence. The 370 Avh proteins previously identified from the *P. ramorum* genome were similarly processed by MEME to generate a list of 312 that had a predicted RxLR motif, which was used as a baseline for optimizing filtering to avoid exclusion of true positives. It should be noted that including the additional 370 predicted *P. ramorum* Avh proteins in the recursive search step yielded an additional 19 RxLR candidates when applying this method to the *P. ramorum* reference genome, but yielded fewer candidates for some *de novo* assemblies. Combining results using both recursive search strategies resulted in the most number of candidates.

Crinkler protein effectors (CRN). These effectors were identified by using two complementary approaches similar to those developed in Haas et al. (Haas *et al.*, 2009) (Fig. S4.13). A total of 552 previously reported CRNs from *P. infestans* (196 effectors), *P. ramorum* (19), *P. sojae* (100) and *P. capsici* (237) (CRNdb) were aligned with MAFFT (Katoh & Standley, 2013). The recombination domain containing a LxKLAK motif in the first 60 aa of the alignment and the HVLVVVP motif were then used to set and train two HMM models with the hmmbuild and hmmcalibrate commands of HMMER v3 (Eddy, 2008). The whole proteomes predicted for the four *P. ramorum* lineages and *P. lateralis* (see *Core and non-core genomes*) were searched for these two models with HMMER v3 (Eddy, 2008) (cut-off value of $1e^{-0.05}$). Candidate CRN identified through this first approach were then aligned with other CRN of the CRNdb, before training a new HMM model based on full-length CRN sequences.

In the second approach, sequences of the CRNdb were searched against the genomes of *P. lateralis* and the four *P. ramorum* lineages using the TBLASTn algorithm (e-value cut-off of 1e-04). The coordinates of matches were captured and all matches overlapping genes found in the first approach were removed. For the other matches, the corresponding DNA sequence was retrieved and translated in ORFs by using the EMBOSS package getorf (minimum size cut-off of 100 nt and a maximum size cut-off of 6000 nt) (Rice *et al.*, 2000). Predicted ORFs were then submitted to an HMMsearch (-T 0) for the full-length CRN model developed in the previous approach. Additional editing was then carried out on the positive hits by checking the presence of both LxKLAK and HVLVVVP domains. This set of CRN was then merged with the one obtained with the first approach to generate a final non-overlapping set of CRN-like proteins.

Analysis of positive selection on *P. ramorum* RxLR and CRN effectors. Sequences of the RxLR and CRN protein datasets were individually clustered with OrthoMCL (Li *et al.*, 2003) using the following parameters: BLASTp search with E-value cut-off of 1e-05, coverage of at least 50% of the query sequence, identity of at least 50% and an OrthoMCL inflation value of I = 1.5. Protein clusters were filtered for truncated proteins as described in Appendix B, OrthoMCL analysis and evolution of gene family size. For each sequence cluster, tBLASTn was used to retrieve DNA sequence homologs in the P. ramorum de novo assemblies used for the RxLR search. To generate a set of proteins not expected to be under positive selection for comparison, sequence homologs of the 458 proteins of the core eukaryotic CEGMA dataset (Parra *et al.*, 2009) were also retrieved by using the same tBLASTn approach. In addition, a random set of 500 proteins with sequence homologs was extracted and used for comparison. Each sequence dataset was then aligned with MAFFT ver. 7.123b (Katoh & Standley, 2013). Only sequence alignments of high quality (less than 5% ambiguous data) and with at least 3 synonymous sites were retained to reduce statistical bias in the estimation of the dN/dS value (Stoletzki & Eyre-walker, 2011). This resulted 255, 51, 315 and 292 alignments for the RxLR, CRN, CEGMA and random datasets, respectively. For each alignment, the average dN/dS value was estimated by fitting the sequence alignment with the basic maximum likelihood model M0 of Codeml (CodonFreq option set to F3X4) implemented in PAMLV4.0 (Yang & Nielsen, 1998). For statistical comparisons, RxLR and CRN datasets were grouped into effectors and a Box-Cox transformation was done on the full dataset using the MASS package (Venables & Ripley, 2002) in R (R Core Team, 2017) and the transformed data sets were compared with an ANOVA.

Results and discussion

RxLR effectors. Applying this annotation pipeline to the *P. ramorum* reference genome identified 391 candidate proteins, where 331 contained a predicted RxLR motif and 391 contained the dEER motif. Among the 331 identified RxLR motifs, 291 were an exact match of 8 amino acids to the RxLR motif in the *P. ramorum* Avh database, and 309 matched one of 44 different RxLR patterns (17 were a perfect RxLR string, 27 were imperfect, such as RYLK, KLLR, RFSR, etc.). As mentioned by Jiang et al. (Jiang *et al.*, 2008), it is possible that some of the Avh proteins in the *P. ramorum* Avh database and those identified using the present approach are pseudogenes or inactive alleles. In order to categorize the diversity of RxLR patterns present within and between lineages, the *Phytophthora* Avh database of 9,779 proteins was analyzed by MEME to identify the RxLR motif. Perfect RxLR motifs are summarized in Fig. S4.14 and S4.15.

Crinkler protein effectors. HMM searches for Crinkler effectors (CRN) specific domains resulted in a total of 40 (EU2) to 50 (EU1) CRN-like proteins having both the LxKLAK and HVLVVVP domains (Table S4.19). These values are slightly below the number of CRN proteins found in the NA1 reference genome (61 models; (Haas *et al.*, 2009)), and are probably dependent on the quality of the assemblies obtained with these Illumina genomes. Based on a Chi-square test, significant enrichment was not observed in any of the *P. ramorum* lineages. A first OrthoMCL clustering analysis including these CRN protein models and those predicted for *P. lateralis* (this study; Table S4.19) and *P. sojae* (Tyler *et al.*, 2006), resulted in 59 clusters, among which 11 are unique to *P. lateralis* and/or *P.sojae* (i.e. all the other clusters share at least one protein with at least one of the *P. ramorum* lineage), 11 are unique to the group *P. ramorum*/*P.*

lateralis (i.e. not shared with *P. sojae*) and 22 are unique to *P. ramorum*. More than half of the clusters unique to *P. ramorum* (12) are balanced with one ortholog from each lineage, indicating that novelty in CRN proteins stayed conserved among lineages within *P. ramorum*.

A second clustering with OrthoMCL between Crinklers from the four lineages resulted in 59 multigene clusters and 12 singlets (Fig. S4.16). Half of the clusters (28) were found to be symmetrical with equal contribution (i.e. one Crinkler protein) from each lineage, indicating no preferential gene expansion from one lineage compared to another. The EU1 lineage showed the highest number of unique Crinklers (7), compared to the other lineages having only one (EU2) or two (NA1 and NA2) unique. These óuniqueô proteins may have not been assigned to any other cluster due to their higher divergence with Crinklers from the same or other lineage. Only one cluster showed an expansion in the EU1 (2 protein models) and NA2 (3 protein models) lineages. Genes of the cluster have only one single ortholog copy in *P. lateralis*, suggesting that the diversification happened after divergence between this species and *P. ramorum*. Branch specific estimates of synonymous and non-synonymous substitution (ω) along the phylogeny showed a significant difference in evolutionary rates between paralog copies of this Crinkler subfamily, indicating an increase of non-synonymous mutations for some members of this subfamily (Fig. S4.11A). Furthermore, a RDP analysis allowed the detection of 4 inter-lineages recombination breakpoints (Fig. S4.11B). This suggests that the history of this CRN effector subfamily was likely shaped by duplication events (gains) followed by recombination before the emergence of the different lineages.

Table S4.1: *Phytophthora ramorum* isolates sequenced in this study.

Isolate ID	Location	Year	Host	Lib. Prep.[1]	# of QC passed reads	# of mapped reads (%)
EU1						
P1549	SW England, UK	2003	*Fagus sylvatica*	1	32 572 330	29 933 640 (91.9%)
P1559	SW England, UK	2003	*Rhododendron* sp.	1	35 073 128	31 629 489 (90.18%)
P1562	NA	2003	*Quercus cerris*	1	32 555 546	29 801 130 (91.54%)
P1563	NA	2003	*Castanea sativa*	1	40 256 606	37 763 140 (93.81%)
P1614	NA	2004	*Nothofagus* sp.	1	33 518 942	30 151 372 (89.95%)
P1856	NA	2004	*Acer platanoides*	1	29 707 842	27 710 875 (93.28%)
P1898	NA	2005	*Fagus sylvatica*	1	26 664 156	24 511 964 (91.93%)
P1959	NA	2005	*Quercus cerris*	1	32 083 066	29 294 756 (91.31%)
P2098	NA	2007	*Cinnamomum camphora*	1	34 939 704	32 201 270 (92.16%)
P2673	UK	2009	*Larix kaempferi*	1	30 161 994	27 754 864 (92.02%)
P2677	UK	2009	*Larix kaempferi*	1	31 571 516	29 388 052 (93.08%)
P2687	UK	2009	*Larix kaempferi*	1	30 543 134	28 447 493 (93.14%)
P2688	UK	2009	*Larix kaempferi*	1	34 649 664	32 313 089 (93.26%)
P2707	UK	2009	*Larix kaempferi*	1	40 594 090	37 496 546 (92.37%)
P2738	UK	2009	*Larix kaempferi*	1	38 177 036	35 280 954 (92.41%)
03-0002	BC, Canada	2003	*Rhododendron* sp.	1	35 783 200	33 174 125 (92.71%)
03-0107	BC, Canada	2003	*Rhododendron* sp.	1	34 576 636	30 615 115 (88.54%)
04-17031	BC, Canada	2004	*Rhododendron* sp.	1	36 986 074	33 755 002 (91.26%)
07-13013	BC, Canada	2007	*Rhododendron* sp.	1	48 525 936	46 064 371 (94.93%)
08-2933	BC, Canada	2008	*Pieris japonica*	1	42 129 748	39 296 121 (93.27%)
10-5341b	BC, Canada	2010	*Rhododendron* sp.	1	32 166 628	29 176 869 (90.71%)

Isolate ID	Location	Year	Host	Lib. Prep.[1]	# of QC passed reads	# of mapped reads (%)
11-301	BC, Canada	2011	unknown	1	36 467 348	32 524 682 (89.19%)
6N0408-2	France	2006	*Rhododendron Cheer*	1	42 197 568	38 865 912 (92.10%)
10-2278a	BC, Canada	2002	*Rhododendron Germania*	1	59 386 418	55 967 132 (94.24%)
10-2807a	BC, Canada	2003	*Viburnum tinus*	1	30 177 970	28 388 606 (94.07%)
10-3884	BC, Canada	2003	*Rhododendron Scyphocalix*	1	37 484 444	34 424 704 (91.84%)
10-3885	BC, Canada	2002	*Viburnum tinus*	1	45 128 834	41 500 703 (91.96%)
10-43896	BC, Canada	2004	*Rhododendron* sp.	1	41 733 280	39 476 814 (94.59%)
10-4389a	BC, Canada	2005	*Rhododendron* sp.	1	38 796 700	36 498 575 (94.08%)
BBA 26/02	Belgium	2002	*Viburnum bodnantense*	1	34 521 624	32 479 615 (94.08%)
BBA 14/98-a	Germany	1998	*Rhododendron catawbiense*	1	35 420 988	33 246 850 (93.86%)
BBA 9/95	Germany	1995	*Rhododendron catawbiense*	1	44 229 884	40 781 496 (92.20%)
CSL 1620	UK	NA	*Viburnum* sp.	1	31 039 742	27 959 446 (90.08%)
P2599	England	2011	*Larix kaempferi*	2	35 011 528	32 853 974 (93.84%)
P2600	England	2011	*Larix kaempferi*	2	33 617 038	31 660 288 (94.18%)
P2633	England	2009	*Larix kaempferi*	2	34 166 566	32 015 694 (93.70%)
PD 98/8/6743	Netherlands	1998	*Rhododendron* sp.	1	51 419 912	47 277 383 (91.94%)
BBA 35/01-01	Germany	2001	*Rhododendron* sp.	1	38 819 524	35 225 656 (90.74%)
1369017-11	CA, USA	2006	*Osmanthus heterophyllus*	1	34 512 168	31 666 736 (91.76%)
BBA CSL 1604	UK	2002	*Rhododendron grandiflora*	1	47 359 894	44 890 851 (94.79%)
BBA CSL 1560	UK	2002	*Viburnum tinus*	1	37 675 962	33 675 842 (89.38%)
09-82-68	OR, USA	2009	*Pieris japonica*	1	37 013 252	34 206 433 (92.42%)
09-617-17	OR, USA	2009	*Rhododendron* sp.	1	41 073 932	38 554 871 (93.87%)
10-923-6-1j	OR, USA	2010	*Rhododendron* sp.	1	40 869 198	35 045 986 (85.75%)
BBA 16/99	Germany	1999	*Viburnum bodnantense*	1	43 800 238	41 055 238 (93.73%)

Isolate ID	Location	Year	Host	Lib. Prep.[1]	# of QC passed reads	# of mapped reads (%)
EU2						
P2111	North of Ireland	2007	*Quercus robur*	2	29 685 302	27 602 287 (92.98%)
P2460	North of Ireland	2010	*Larix kaempferi*	2	29 213 436	27 107 877 (92.79%)
P2461	North of Ireland	2010	*Larix kaempferi*	2	27 427 386	25 428 118 (92.71%)
P2561	Scotland	2011	*Larix kaempferi*	2	37 013 894	34 332 585 (92.76%)
P2566	North of Ireland	2011	*Rhododendron ponticum*	2	25 240 228	23 437 003 (92.86%)
P2586	Scotland	2011	*Larix kaempferi*	2	32 741 190	30 384 114 (92.80%)
NA1						
04-0002	BC, Canada	2004	*Camellia* sp.	1	45 995 292	42 621 657 (92.67%)
04-2231	BC, Canada	2004	*Camellia japonica*	1	29 964 586	28 093 396 (93.76%)
04-22572	BC, Canada	2004	*Rhododendron* sp.	1	60 933 030	57 884 344 (95.00%)
04-35009	BC, Canada	2004	*Salal* sp.	1	39 248 998	35 098 663 (89.43%)
10-6342a	BC, Canada	2010	*Rhododendron* sp.	1	42 709 422	40 676 077 (95.24%)
11-263	BC, Canada	2011	*Rhododendron* sp.	1	28 370 272	26 811 793 (94.51%)
1020.1	OR, USA	2001	*Lithocarpus densiflorus*	1	47 690 832	45 355 262 (95.10%)
2092	OR, USA	2002	NA	1	40 185 274	35 685 804 (88.80%)
06-81-2	OR, USA	2006	*Camellia japonica*	1	39 112 726	36 377 329 (93.01%)
33-WA-RD	WA, USA	2005	Soil sample	1	29 931 574	28 220 328 (94.28%)
52-WA-A	WA, USA	2007	NA	1	49 022 376	45 131 177 (92.06%)
1428631-1	CA, USA	2007	*Camellia japonica*	1	38 225 732	35 037 704 (91.66%)
1418886	CA, USA	2004	*Camellia* x óCoral Delightô	1	51 195 010	47 572 883 (92.92%)
GA-28C	GA, USA	2004	*Camellia sasanqua*	1	41 105 758	38 727 544 (94.21%)
WSDA-1107	WA, USA	2004	*Camellia* sp.	1	37 457 570	33 582 416 (89.65%)
WA 1107.1	WA, USA	2004	NA	1	31 237 818	27 197 217 (87.07%)

Isolate ID	Location	Year	Host	Lib. Prep.[1]	# of QC passed reads	# of mapped reads (%)
08-SNJ-GO/NC-5B	NC, USA	2008	*Rhododendron* sp.	1	44 628 632	41 812 161 (93.69%)
SPN-S1-C	SC, USA	2009	*Rhododendron catawbiense*	1	41 499 194	37 687 693 (90.82%)
Lakeland IV, G-6-b	MS, USA	2009	water	1	41 013 618	39 337 004 (95.91%)
SOD-GA-741W3 S1587	GA, USA	NA	NA	1	33 521 948	31 310 504 (93.40%)
09-SNJ-SC.09-0353a	SC, USA	2009	*Rhododendron catawbiense*	1	44 662 740	40 203 242 (90.02%)
263AL	GA, USA	2005	*Kalmia latifolia*	1	54 023 504	51 463 597 (95.26%)
304F	GA, USA	2005	*Camellia japonica*	1	43 207 376	40 890 039 (94.64%)
RWN-P20-AR	SC, USA	2010	Soil sample	1	33 751 716	31 483 350 (93.28%)
SPN5-P2-2d	SC, USA	2010	*Kalmia latifolia*	1	38 574 618	35 910 371 (93.09%)
SPN5-P2-3c	SC, USA	2010	*Kalmia latifolia*	1	46 473 738	43 289 205 (93.15%)
Pr 733	CA, USA	2010	Water sample	1	41 303 016	38 177 582 (92.43%)
9488	OR, USA	2006	*Lithocarpus densiflorus*	1	40 766 538	38 577 598 (94.63%)
11372	OR, USA	2011	*Lithocarpus densiflorus*	1	46 264 218	44 237 245 (95.62%)
1365048	CA, USA	2011	*Umbellularia californica*	1	51 671 892	49 402 199 (95.61%)
1496607	CA, USA	2011	*Camellia japonica*	1	52 540 946	49 908 631 (94.99%)
1560191-36	CA, USA	2011	Pieris sp.	1	51 864 454	49 032 929 (94.54%)
1621926-8	CA, USA	2011	*Laurus nobilis*	1	34 038 516	31 469 874 (92.45%)
1625769	CA, USA	2011	*Molinadendron sinaloense*	1	44 207 710	41 880 104 (94.73%)
PC#11-1366	CA, USA	2011	*Cinnamomum camphora*	1	34 162 180	31 046 716 (90.88%)
Pr 1546	CA, USA	2012	*Quercus agrifolia*	2	27 022 916	25 514 680 (94.42%)
Pr 420	CA, USA	2004	*Quercus agrifolia*	2	28 587 464	26 714 242 (93.45%)
Pr 518	CA, USA	2009	*Quercus agrifolia*	2	20 147 362	12 620 067 (62.64%)

Isolate ID	Location	Year	Host	Lib. Prep.[1]	# of QC passed reads	# of mapped reads (%)
NA2						
04-20470	BC, Canada	2004	*Rhododendron* sp.	1	38 656 650	34 610 518 (89.53%)
04-38813	BC, Canada	2004	*Viburnum tinus*	1	33 812 788	30 697 228 (90.79%)
05-16845	BC, Canada	2005	*Ardisia* sp.	1	51 967 210	48 406 956 (93.15%)
05-9611	BC, Canada	2005	*Viburnum* sp.	1	34 762 888	31 976 355 (91.98%)
06-0012	BC, Canada	2006	Soil sample	1	44 616 902	40 658 776 (91.13%)
06-4942	BC, Canada	2006	*Distyllium Myricoides*	1	33 730 186	31 071 533 (92.12%)
07-17204	BC, Canada	2007	*Gaultheria* sp.	1	43 194 974	39 904 438 (92.38%)
07-46063	BC, Canada	2007	*Rhododendron* sp.	1	25 742 572	23 772 653 (92.35%)
08-0688	BC, Canada	2008	*Rhododendron* sp.	1	53 952 202	50 236 079 (93.11%)
09-4390	BC, Canada	2009	*Camellia* sp.	1	42 129 748	39 296 121 (93.27%)
09-4415	BC, Canada	2009	*Rhododendron* sp.	1	42 129 748	39 296 121 (93.27%)
10-3885	BC, Canada	2010	*Leucothoe fontanesia*	1	44 798 706	41 455 703 (92.54%)
10-4389a	BC, Canada	2010	*Rhododendron* sp.	1	45 198 056	42 833 947 (94.77%)
11-1291	BC, Canada	2011	NA	1	37 164 490	32 387 192 (87.15%)
wsda 3765	WA, USA	2004	*Rhododendron* sp.	1	42 866 926	37 913 912 (88.45%)
MR31	WA, USA	2004	*Rhododendron* sp.	1	61 592 742	54 726 800 (88.85%)
1470614-10	CA, USA	NA	*Rhododendron* sp.	1	38 106 050	35 534 364 (93.25%)
1586251-1	CA, USA	2010	*Camellia japonica*	1	94 614 546	84 555 561 (89.37%)

[1]Library preparation and sequencing protocol : 1, SPRI-TE 300-600bp library HiSeq 100 base PET; 2, plate-based large gap, HiSeq 100 base PET

Table S4.2: Number and percent of homozygous and heterozygous SNPs in *Phytophthora ramorum* individuals with ROH and representatives of the general population without ROH.

			Number of SNPs			% of SNPs				
	ROH	Total loci	[0/0]	[0/1]	[1/1]	[0/0]	[0/1]	[1/1]	HoL[1]	Notes*
EU1										
03_0002	Yes	159 218	10 804	137 744	10 670	6.8	86.5	6.7	13.5	
PR_99_002	Yes	159 218	6193	146 051	6974	3.9	91.7	4.4	8.3	
P1856	Yes	159 218	7357	145 926	5935	4.6	91.7	3.7	8.3	Tested phenotype
07_13013	Yes	159 261	1758	154 576	2927	1.1	97.1	2927	1.8	
P1559	Yes - scaf 7	158 852	647	156 403	1802	0.41	98.5	1.1	1.5	Tested phenotype
P1614	Yes - scaf 7	158 852	651	156 403	1798	0.41	98.5	1.1	1.5	Tested phenotype
P2600	Yes - scaf 7	158 852	657	156 389	1806	0.41	98.5	1.1	1.6	Tested phenotype
P1549	Yes - scaf 7									Tested phenotype
P1562	Yes - scaf 7									Tested phenotype
P1563	Yes - scaf 7									Tested phenotype
P1959	Yes - scaf 7									Tested phenotype
P2688	Yes - scaf 7									Tested phenotype
P2673	No	158 852	85	158 761	6	0.05	99.9	0.0	0.1	Tested phenotype
P1898	No									Tested phenotype
P2098	No									Tested phenotype
P2599	No									Tested phenotype
P2633	No									Tested phenotype
P2677	No									Tested phenotype
P2707	No									Tested phenotype
P2738	No	158 852	68	158 780	4	0.04	99.9	0.0	0.0	

			Number of SNPs			% of SNPs				
	ROH	Total loci	[0/0]	[0/1]	[1/1]	[0/0]	[0/1]	[1/1]	HoL[1]	Notes*
PR_08_034	No	158 852	90	158 756	6	0.06	99.9	0.0	0.1	
10_5341b	No	159 218	361	158 852	5	0.2	99.8	0.0	0.2	
04_17031	No	159 218	204	158 907	107	0.1	99.8	0.1	0.2	
EU2										
P2586	Yes	99 266	3868	91 962	3436	3.9	92.6	3.5	7.4	
P2561	No	99 266	299	98 906	61	0.3	99.6	0.1	0.4	
P2460	No	99 266	316	98 881	69	0.3	99.6	0.1	0.4	
P2566	No	99 266	293	98 921	52	0.3	99.7	0.1	0.3	
NA1										
PR_07_191	Yes	146 840	825	145 271	744	0.56	98.9	0.5	1.1	
PR_01_004	Yes	146 840	3709	139 788	3343	2.53	95.2	2.3	4.8	
Pr_420	Yes	146 840	2034	142 903	1903	1.39	97.3	1.3	2.7	
PR_09_167	No	146 840	72	146 749	19	0.05	99.9	0.0	0.1	
PR_11_018	No	146 840	97	146 700	43	0.07	99.9	0.0	0.1	
PR_11_011	No	146 840	154	146 552	134	0.10	99.8	0.1	0.2	
NA2										
10_3885	Yes	147 086	6857	133 049	7180	4.7	90.5	4.9	9.5	
04_20470	Yes	147 086	6336	136 572	4178	4.3	92.9	2.8	7.1	
05_16845	No	147 086	1770	143 441	1875	1.2	97.5	1.3	2.5	
06_0012	No	147 086	3356	141 029	2701	2.3	95.9	1.8	4.1	
04_38813	No	147 086	605	144 828	1653	0.4	98.5	1.1	1.5	
09_4415	No	147 086	576	144 865	1645	0.4	98.5	1.1	1.5	

[1]HoL, total homozygous loci; *Tested phenotype - isolates used to inoculate Douglas-fir, larch and Rhododendron and for growth studies on sapwood agar

Table S4.3: Classification of 485 327 SNPs as polymorphic or fixed within *Phytophthora ramorum* lineages and as heterozygous or homozygous. Numbers exclude singletons and missing data.

Lineage	Fixed within lineage (variable between lineage)		Polymorphic within lineage		
	Homozygous	Heterozygous	Homozygous[1]	Heterozygous[2]	Both[3]
NA1	338 354	133 089	16	13 836	32
NA2	337 857	118 845	5	28 574	46
EU1	325 883	129 762	13	16 872	12 797
EU2	386 023	91 125	4	8168	7

[1]SNPs that are homozygous and different between individuals (ex AA in some TT in others); [2]SNPs that are homozygous in some and heterozygous in others (ex AA in some and AT in others); [3]SNPs with all three states (e.g. AA, AT, and TT).

Table S4.4: *Phytophthora ramorum* individuals with Chromosomal Copy Number Variation (CCNV).

Lineage	Isolate	Origin (Host)	Inferred Ploidy[1]	CCNV
EU1	04-17031	Canada (*Rhododendron* sp.)	2n	Scaffold 5: 2n and 3n
EU1	P1563	UK (*Castanea sativa*)	3n	Scaffold 12: 2n and 3n Scaffold 19: 2n
EU1	PR-09-106	USA (*Pieris japonica*)	4n	Scaffolds 1, 2, 5, 7 to 10: 5n
NA1	PR-05-055	USA (Forest)	3n	Scaffold_1, 6 and 7: 4n Scaffold 4 : 2n

[1]As infered by the Control-FREEC software.

Table S4.5: Number of *Phytophthora ramorum* proteins with amino acid differences between allele retained and allele lost through conversion to homozygosity after mitotic recombination.

Lineage	Isolate	Number of genes			
		In scaffolds with ROH	With amino acid difference between alleles	With allele loss due to ROH	Genes with ≥5 differences between alleles
EU1	03_0002	1587	893	880	202
EU1	07_13013	349	186	175	46
EU1	P1856	879	516	507	128
EU1	PR_99_002	875	514	501	127
EU2	P2586	905	444	387	41
NA1[1]	PR_01_004	641	328	298	47
NA1[1]	PR_07_191	204	108	55	9
NA1[1]	PR_11_011	280	150	14	2
NA1[1]	Pr_420	220	147	146	36
NA2	04_20470	650	303	195	33
NA2	10_3885	1202	638	615	104
NA1	scaffold_7[2]	175	84	73	20
NA1	scaffold_100[2]	53	24	24	1
EU1	scaffold_7[3]	175	94	63	15
EU1	scaffold_100[3]	53	25	24	2
NA2	Scaffold_5[2]	144	56	26	2
All[4]	All	5172[4]	2960[4]	2698[4]	555[4]

[1]Excludes scaffold 7 and 100 which were shared by all NA1 individuals; [2]Shared by all individuals in the lineage; [3]Shared by 8 individuals in the EU1 lineage; [4]Total across all four lineages and all affected individuals where the same gene affected in two or more individuals is only counted once.

Table S4.6: Proteins enriched in scaffold seven as determined by MGSA analysis (posterior probability estimate ≥ 0.25).

GO term	# in Genome	# in scaffold 7	Estimate	Std error	Biological function
GO:0016849	16	6	0.95	0.003	phosphorus-oxygen lyase activity
GO:0004553	255	14	0.55	0.010	hydrolase activity, hydrolyzing O-glycosyl compounds
GO:0016798	265	14	0.29	0.007	hydrolase activity, acting on glycosyl bonds
GO:0045330	11	2	0.24	0.004	aspartyl esterase activity
GO:0030599	13	2	0.20	0.004	pectinesterase activity
GO:0008474	19	2	0.14	0.004	palmitoyl-(protein) hydrolase activity
GO:0098599	19	2	0.14	0.004	palmitoyl hydrolase activity
GO:0030248	20	4	0.12	0.012	cellulose binding
GO:0016790	22	2	0.11	0.003	thiolester hydrolase activity

Table S4.7: Comparison of intergenic distances on scaffolds with mitotic recombination breakpoints with ten sets of random scaffolds.

	Number of genes			ANOVA test	mean	Std error	F-value	p
	3'FIR > 10 000	5'FIR > 10 000	Both > 10 000					
Scaffolds with breakpoints	49 (7%)	38 (5%)	8 (1%)	3`FIR	4018	4149	-	-
				5'FIR	2425	4144	-	-
random set 1	26 (5%)	15 (3%)	1 (0%)	3`FIR	3596	3805	3.20	0.074
				5'FIR	1776	3226	8.43	0.004
random set 2	19 (3%)	13 (2%)	1 (0%)	3`FIR	3001	3420	23.8	0.000
				5'FIR	1340	2706	31.41	0.000
random set 3	18 (4%)	8 (2%)	0 (0%)	3`FIR	3497	3555	4.62	0.032
				5'FIR	1610	2634	13.03	0.000
random set 4	28 (5%)	23 (4%)	1 (0%)	3`FIR	3670	4612	2.1	0.148
				5'FIR	2012	4567	2.99	0.084
random set 5	30 (3%)	22 (2%)	2 (0%)	3`FIR	3192	3110	21.47	0.000
				5'FIR	1511	2800	28.54	0.000
random set 6	23 (4%)	13 (2%)	0 (0%)	3`FIR	3224	2945	16.2	0.000
				5'FIR	1581	2594	19.65	0.000
random set 7	22 (4%)	15 (3%)	0 (0%)	3`FIR	3323	3335	10.15	0.002
				5'FIR	1672	3330	11.93	0.001
random set 8	28 (6%)	18 (4%)	3 (1%)	3`FIR	3569	3652	3.61	0.057
				5'FIR	1799	3432	7.31	0.007
random set 9	27 (3%)	15 (2%)	2 (0%)	3`FIR	3365	3023	12.38	0.000
				5'FIR	1543	2685	24.43	0.000
random set 10	26 (4%)	18 (3%)	1 (0%)	3`FIR	3397	3285	9.942	0.002
				5'FIR	1619	2949	18.11	0.000

Table S4.8: Nucleotidic and gene content of the *de novo* assemblies obtained for the four *Phytophthora ramorum* lineages.

	Size (Mb)	% repetitive	%GC	# of genes (Genes/Mb^{-1})	Av. transcript size (bp)	# of gene with BLASTp hit	# of gene with Oomycete hit	# of gene with Blast2GO annotation	# of TE-like genes
Full genome									
EU1	49.1	16.06	53.98	14 095 (287)	1448	13 847 (98.2%)	13 754 (97.6%)	7976 (56.6%)	76 (0.54%)
EU2	47.6	15.70	54.02	14 028 (295)	1458	13 785 (98.3%)	13 684 (97.6%)	7912 (57.4%)	22 (0.16%)
NA1	50.7	14.32	53.96	14 213 (280)	1438	13 900 (97.8%)	13 811 (97.2%)	7758 (55.8%)	88 (0.61%)
NA2	48.5	16.10	53.94	14 186 (292)	1439	13 876 (97.8%)	13 744 (96.9%)	7871 (56.7%)	118 (0.83%)
Repeats									
EU1	7.88	100.0	52.88	500 (63)	1002	491 (98.2%)	472 (94.4%)	210 (42.0%)	18 (3.6%)
EU2	7.47	100.0	52.89	463 (62)	962	441 (95.2%)	412 (88.9%)	203 (43.8%)	80 (17.3%)
NA1	7.26	100.0	52.75	414 (57)	999	400 (96.6%)	389 (94.0%)	191 (46.1%)	15 (3.6%)
NA2	7.81	100.0	52.11	392 (50)	902	375 (95.7%)	357 (91.1%)	180 (45.9%)	27 (6.8%)
Non-core									
EU1	0.63	49.33	50.01	26 (41)	859	21 (80.7%)	20 (95.2%)	4 (15.4%)	3 (11.5%)
EU2	0.24	37.75	49.35	17 (71)	952	14 (82.4%)	12 (85.7%)	5 (29.4%)	6 (35.3%)
NA1	0.64	50.95	49.93	34 (53)	864	31 (91.2%)	30 (96.8%)	8 (23.5%)	8 (5.8%)
NA2	0.85	47.32	49.86	44 (52)	824	37 (84.1%)	34 (91.8%)	9 (20.5%)	7 (9.0%)

Table S4.9: Protein family content predicted in the non-core genome of the four *Phytophthora ramorum* lineages. Putative effector families according to (van Damme *et al.*, 2012) are bolded.

	EU1	EU2	NA1	NA2
Ester hydrolases	**-**	**-**	**1**	**-**
Peptidases	**1**	**-**	**2**	**4**
RxLR	**2**	**-**	**1**	**2**
ABC transporter	1	-	1	1
Kinases	**-**	**-**	**-**	1
Methyltransferases	3	1	1	2
Glycoside transferases	1	-	-	**-**
Helicases	-	2	1	**-**
Transposons	3	6	8	7
Other hypothetical proteins	10	5	17	22
Putative proteins without homologs	5	3	2	5
Total	26	17	34	44

Table S4.10: Shared and unique protein models found in *Phytophthora ramorum* lineages.

	# of clusters (# of protein sequences)	
	Before filtering	**After filtering[1]**
All	56 560 (192 586)	51 458 (177 030)
Core (shared among 10 *Phytophthora* protein sets)	6201 (73 502)	6200 (69 770)
One-to-one orthologs[2]	3934 (39 340)	4929 (49 340)
Lineage specific sets		
P. ramorum + *P. lateralis*	238 (1229)	238 (1206)
P. ramorum	541 (2262)	541 (2197)
EU1	616 (635ï 4.50%[3])	349 (353ï 2.47%)
NA1	713 (739ï 5.02%)	427 (439ï 3.09%)
EU2	717 (764ï 5.11%)	437 (452ï 3.12%)
NA2	738 (762ï 5.37%)	453 (460ï 3.19%)

[1]Filters applied to eliminate redundant OrthoMCL clusters and correct those having multiple protein fragments; [2]Only one copy of the gene in each of the taxa considered [3]The proportion (%) of unique protein is given for the four *P. ramorum* lineages

Table S4.11: MGSA analysis for enriched GO terms in set of proteins unique to the *Phytophthora ramorum* lineages. Five most represented terms shown.

Lineage	Biological functions	Estimate	SD
EU1	GO:0030570, pectate lyase activity	0.50	0.03
	GO:0016837, carbon-oxygen lyase activity, acting on polysaccharides	0.30	0.01
	GO:0005544, calcium-dependent phospholipid binding	0.50	0.03
	GO:0004523, ribonuclease H activity	0.40	0.02
	GO:0008239, dipeptidyl-peptidase activity	0.20	0.01
EU2	GO:0008474, palmitoyl-(protein) hydrolase activity	0.49	0.04
	GO:0070001, aspartic-type peptidase activity	0.44	0.02
	GO:0004190, aspartic-type endopeptidase activity	0.42	0.03
	GO:0003810, protein-glutamine gamma-glutamyltransferase activity	0.37	0.03
	GO:0016790, thiolester hydrolase activity	0.25	0.01
NA1	GO:0005200, structural constituent of cytoskeleton	0.88	0.01
	GO:0004035, alkaline phosphatase activity	0.72	0.02
	GO:0070006, metalloaminopeptidase activity	0.43	0.02
	GO:0030570, pectate lyase activity	0.33	0.02
	GO:0052862, glucan endo-1,4-beta-glucanase activity	0.27	0.02
NA2	GO:0005200, structural constituent of cytoskeleton,	0.78	0.02
	GO:0003864, 3-methyl-2-oxobutanoate hydroxymethyltransferase act.	0.50	0.01
	GO:0004448, isocitrate dehydrogenase activity,	0.36	0.01
	GO:0043168, anion binding	0.20	0.02
	GO:0030570, pectate lyase activity	0.11	0.01

Table S4.12: Number of SNPs in *Phytophthora ramorum* before and after filtering at various coverage and quality values, and the percentage of validation SNPs, expected SNPs and false SNPs found after each round of filtering.

Mean depth		SNP depth		SNP quality[1]			Gap[2]	Mac[3]	Total SNPs before filter[4]	Total SNPs after filter	% validation SNPs found after filter	% expected SNP	% False SNP
Min	Max	Min	Max	Q	MQ	GQ							
n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	912 622	N/A	97.2	96.1	0
	40								912 622	262 702	0	N/A	0
	60								912 622	596 132	57	98.5	0
	70								912 622	761 989	86	95.5	0
	80								912 622	839 974	97.2	96.1	0
	90								912 622	872 609	97.2	96.1	0
	100								912 622	887 393	97.2	96.1	0
	90			20	20		10		872 609	848 193	97.2	96.1	0
	90			30	20		10		872 609	839 102	97.2	96.1	0
	90			20	30		10		872 609	848 193	97.2	96.1	0
	90			30	30		10		872 609	839 102	97.2	96.1	0
10	90			30	30		10		839 102	811 527	97.2	96.1	0
10	90	5		30	30		10		811 527	629 579	97.2	96.1	0
10	90	4		30	30		10		811 527	647 405	97.2	96.1	0
10	90	3		30	30		10		811 527	665 856	97.2	96.1	0
10	90	2		30	30		10		811 527	685 939	97.2	96.1	0
10	90		150	30	30		10		811 527	545 725	96.9	96.1	0
10	90		200	30	30		10		811 527	743 609	97.2	96.1	0
10	90		240	30	30		10		811 527	801 298	97.2	96.1	0
10	90			30	30	30	10		811 527	398 814	96.9	96.5	0
10	90			30	30	20	10		811 527	441 034	96.9	96.5	0

Mean depth		SNP depth		SNP quality[1]			Gap[2]	Mac[3]	Total SNPs before filter[4]	Total SNPs after filter	% validation SNPs found after filter	% expected SNP	% False SNP
Min	Max	Min	Max	Q	MQ	GQ							
10	90			30	30	15	10		811 527	465 216	96.9	96.5	0
10	90	4		30	30	20	10		811 527	438 385	96.9	96.5	0
10	90	4		30	30	20	10	2	811 527	434 966	96.9	96.5	0
10	90	4	240	30	30	20	10		811 527	434 504	96.9	96.5	0
10	90	4	240	30	30	20	10	2	811 527	431 112	96.9	96.5	0
10	90	4	200	30	30	20	10		811 527	398 589	96.9	96.5	0
10	90	4	200	30	30	20	10	2	811 527	395 388	96.9	96.5	0
10	90	4		30	30	30	10		811 527	397 846	96.9	96.5	0
10	90	4		30	30	30	10	2	811 527	395 110	96.9	96.5	0

[1]Q, SNP quality value; MQ, Minimum Mapping quality at the SNP; GQ, minimum genotype quality; [2]Gap, number of base pairs between the SNP and the closest gap; [3]Mac, minimum allele count; [4]Indels removed.

Table S4.13: False call rate between sequencing runs of the same *Phytophthora ramorum* individual before and after applying SNP filters for coverage and quality.

	Before Filtering			After Filtering				No missing, no singletons[1]	
Individual	**Total sites**	**Discordant sites**	**False call rate**	**Total sites**	**Discordant sites**	**No of masked sites**	**False call rate**	**Discordant sites**	**False call rate**
Pram_06_4942_run 1 vs 2	476 646	39 428	8.3%	353 152	6434	36	1.8%	541	0.2%
Pram_06_4942_run 1 vs 3	481 739	38 261	7.9%	335 629	5819	23	1.7%	n/a	n/a
Pram_04_17031	418 014	19 052	4.6%	266 020	1819	15	0.7%	n/a	n/a

[1]Filters applied across the full dataset to remove SNPs with missing data in some individuals and SNPs found in only one individual.

Table S4.14: Assembly statistics and completeness for eight *de novo* assemblies generated for *Phytophthora ramorum*.

Lineage	Isolate name – Host	Total size (Mbp)	# of scaffolds	N50 (bp)	Length of longest scaffold (bp)	BUSCO coverage[1]
EU1	03-0002 ï *Rhododendron* sp.	49.1	4846	24 101	261 007	352 (82.1%)
	07-13013 ï *Rhododendron* sp.	48.2	5224	22 468	204 338	351 (81.8%)
EU2	P2586 ï *Larix kaempferi*	47.6	3876	31 609	286 734	353 (82.3%)
	P2111 ï *Quercus robur*	45.9	3802	31 672	298 197	353 (82.3%)
NA1	PR-09-175 ï *Camellia japonica*	50.7	5202	22 339	261 091	353 (82.3%)
	PR-11-009 ï *Umbellularia californica*	48.3	4942	22 480	261 530	353 (82.3%)
NA2	05-16845 ï *Ardisia* sp.	48.5	4983	23 557	261 946	353 (82.3%)
	04-38813 ï *Viburnum tinus*	49.3	4940	27 888	299 169	352 (82.1%)

[1]Assessed over 429 single-copy ortholog genes common to Eukaryotes (Simão et al. 2015; Bioinformatics, 31:3210).

Table S4.15: Mean posterior values for eight selected parameters estimated under the three molecular clock models.

	Strict clock	UCLD relaxed clock	Random local clock
Likelihood	-84 847.60	-84 795.49	-84 799.32
Posterior	-84 890.00	-84 834.71	-84 862.12
Yule.birthrate	0.0755	0.0756	0.0765
Clock.rate	0.0043	-	0.0044
ucld.mean	-	0.0043	-
ucld.stdev	-	0.18	-
Coefficient of variation	-	0.169	0.141
Rate change count	-	-	1.985

Table S4.16: Median divergence times and 95% confidence intervals for *Phytophthora* nodes estimated under the three molecular clock models. Divergence time estimates in My.

Nodes	Strict clock	UCLD relaxed clock	Random local clock
(EU1,NA1)	0.75 (0.61,0.88)	0.77 (0.54,1.00)	0.76 (0.62,0.90)
(NA2,(NA1,EU1))	1.06 (0.91,1.20)	1.10 (0.84,1.39)	1.07 (0.91,1.23)
(EU2,(NA2,(NA1,EU1)))	1.31 (1.14,1.48)	1.39 (0.65,2.04)	1.33 (0.91,1.73)
(*P. lateralis*, *P.ramorum*)	6.38 (5.91,6.84)	6.70 (5.07,8.34)	6.53 (6.00,7.04)
(*P. sojae*, *P. cinnamomi*)	11.81 (11.15,12.52)	12.3 (9.71,14.84)	12.78 (11.73,13.85)
(Clade 7, Clade 8)[1]	21.77 (20.95,22.62)	22.45 (20.20,24.78)	22.71 (21.65,23.81)
(*P. infestans*, *P. parasitica*)	14.65 (13.89,15.41)	13.93 (11.94,15.81)	12.22 (11.02,13.36)
(Clade 1, Clade 2)	23.09 (21.11,24.83)	22.12 (19.46,24.59)	22.02 (20.96,23.14)
(Clade 1,(Clade 2,(Clade 7,Clade 8))	25.64 (23.89,27.48)	25.8 (24.00,28.70)	25.65 (23.70,27.38)

[1]Clade members: Clade 1 = *P. parasitica* and *P. infestans*; Clade 2 = *P. capsici*; Clade 7 = *P. sojae* and *P. cinnamomi*; Clade 8 = *P. lateralis* and *P. ramorum.*

Table S4.17: Enriched gene ontology for *Phytophthora ramorum* proteins affected by ROH as determined by MGSA analysis (posterior probability estimate ≥ 0.25).

GO accession	inPopulation	inStudySet	Estimate	Std error	Biological function
GO:0005515	1758	326	0.50	0.05	protein binding
GO:0016746	176	50	0.49	0.05	transferase activity, transferring acyl groups
GO:0016491	791	140	0.42	0.04	oxidoreductase activity
GO:0003964	326	4	0.39	0.05	RNA-directed DNA polymerase activity
GO:0016835	85	3	0.33	0.03	carbon-oxygen lyase activity
GO:0004553	255	52	0.32	0.03	hydrolase activity, hydrolyzing O-glycosyl compounds
GO:0005198	146	5	0.26	0.03	structural molecule activity

Table S4.18: Growth and sporulation of isolates with and without ROH on sapwood agar.

		Day 4		Day 7		Day 10	
Media type	**Group**	**Mean growth (SD)**	**Mean sporulation (SD)**	**Mean growth (SD)**	**Mean sporulation (SD)**	**Mean growth (SD)**	**Mean sporulation (SD)**
Larch	ROH	22.4 (1.4)	n/a	37.0 (2.1)	5.6 (1.9)	52.9 (2.7)	2.3 (0.9)
Larch	no ROH	23.1 (1.4)	n/a	39.9 (6.5)	6.0 (3.4)	56.4 (9.2)	1.8 (0.8)
	F-value	0.849	n/a	1.455	0.093	1.070	1.615
	prob	0.372	n/a	0.248	0.765	0.318	0.224
Douglas-fir	ROH	25.5 (1.2)	0.7 (0.8) [1]	42.8 (2.8)	1.1 (0.6) [1]	58.5 (4.1)	4.8 (0.7) [2]
Douglas-fir	no ROH	25.9 (1.6)	0.6 (0.5) [1]	44.8 (6.2)	1.2 (0.8) [1]	63.3 (10.1)	4.8 (1.6) [2]
	F-value	0.256	0.255	0.670	0.004	1.517	0.022
	prob	0.621	0.622	0.427	0.953	0.238	0.884

[1]Counts done at two diagonal locations beginning from inoculation point; [2]Counts done on day 11 at four locations 0.5 cm from center of inoculation point.

Table S4.19: Number of *Phytophthora ramorum* and *P. lateralis* Crinkler effectors identified in the two approaches of the Crinkler identification pipeline.

Lineage (# of proteins)	1st approach			2nd approach	Final count (Incomplete proteins[2])
	LFLAK	HVLVVVP	LFLAK + HVLVVVP[1]		
NA1 (16 752)	21	15	14	28	42 (47)
NA2 (16 187)	27	16	15	32	47 (52)
EU1 (16 183)	32	25	18	32	50 (55)
EU2 (15 539)	21	14	12	28	40 (47)
P. lateralis (20 040)	12	9	3	43	46 (70)

[1]Both a LFLAK and a HVLVVVP domain were identified in the proteins; [2]Incomplete proteins are Crinkler like proteins with only a LFLAK or a HVLVVVP domain.

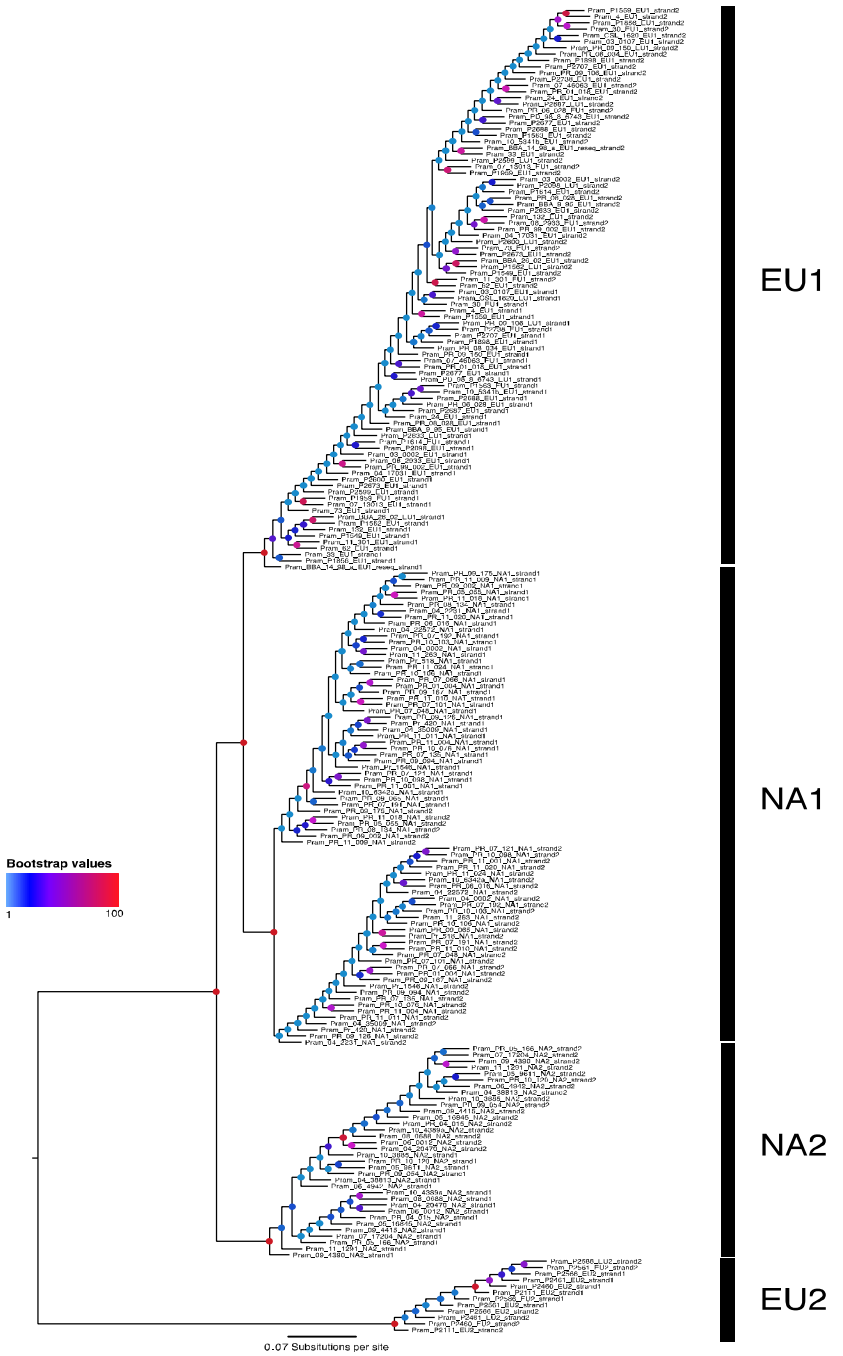


Figure S4.5: Phylogenetic reconstruction of *Phytophthora ramorum* individuals of four lineages. Maximum likelihood tree reconstructed between individuals using the sequence of 500 genes that are presumed to be neutral based on Tajima's D. The four lineages are located in highly supported clades, while most of the secondary clades have no support (Bootstrap values < 70)

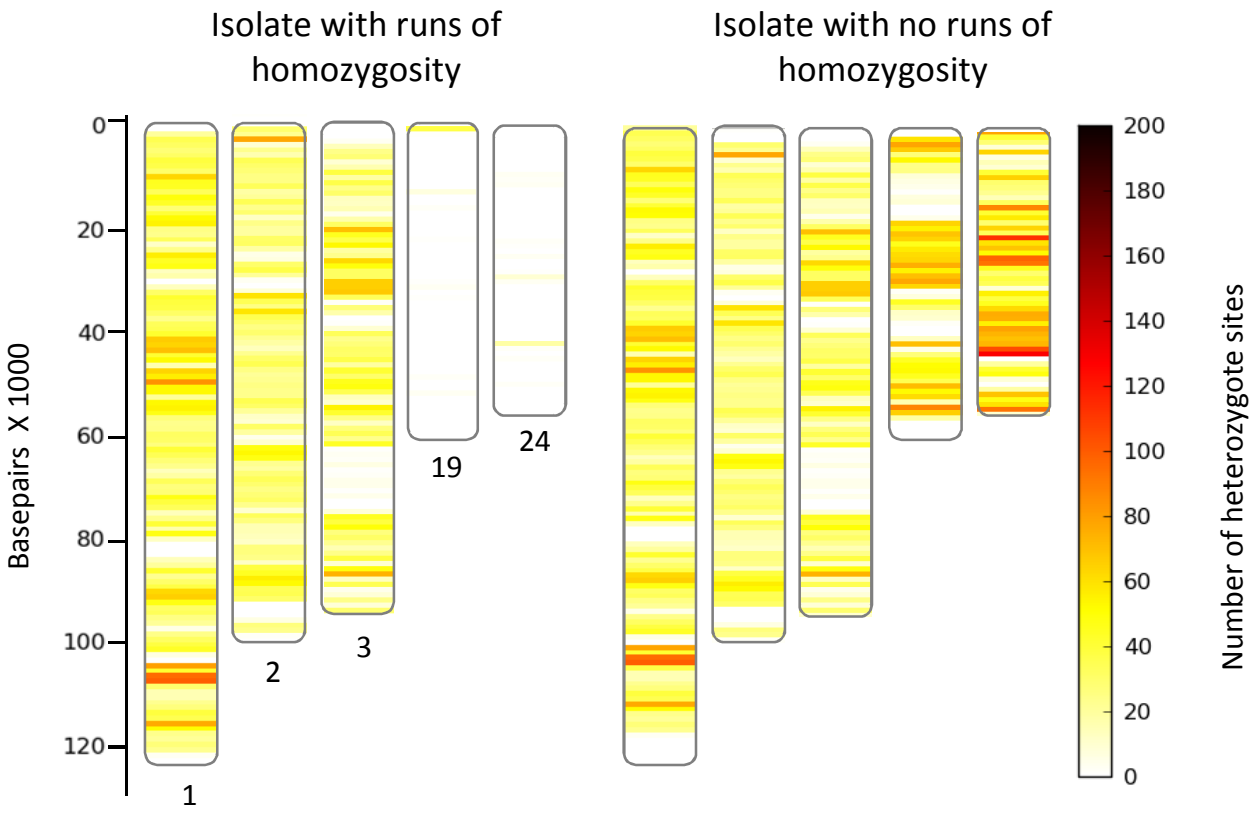


Figure S4.6: Runs of homozygosity caused by mitotic recombination. Heterozygote density on five scaffolds of two individuals of the EU1 lineage: one with long runs of homozygosity (ROH) caused by mitotic recombination (left) and one without ROH (right). The first three scaffolds of the isolate with the ROH pattern show no loss of heterozygosity, whereas heterozygotes are nearly absent in two affected scaffolds (19 and 24). In the individuals unaffected by MR (represented by the individual on the right) there is no loss of heterozygosity on the same two scaffolds.

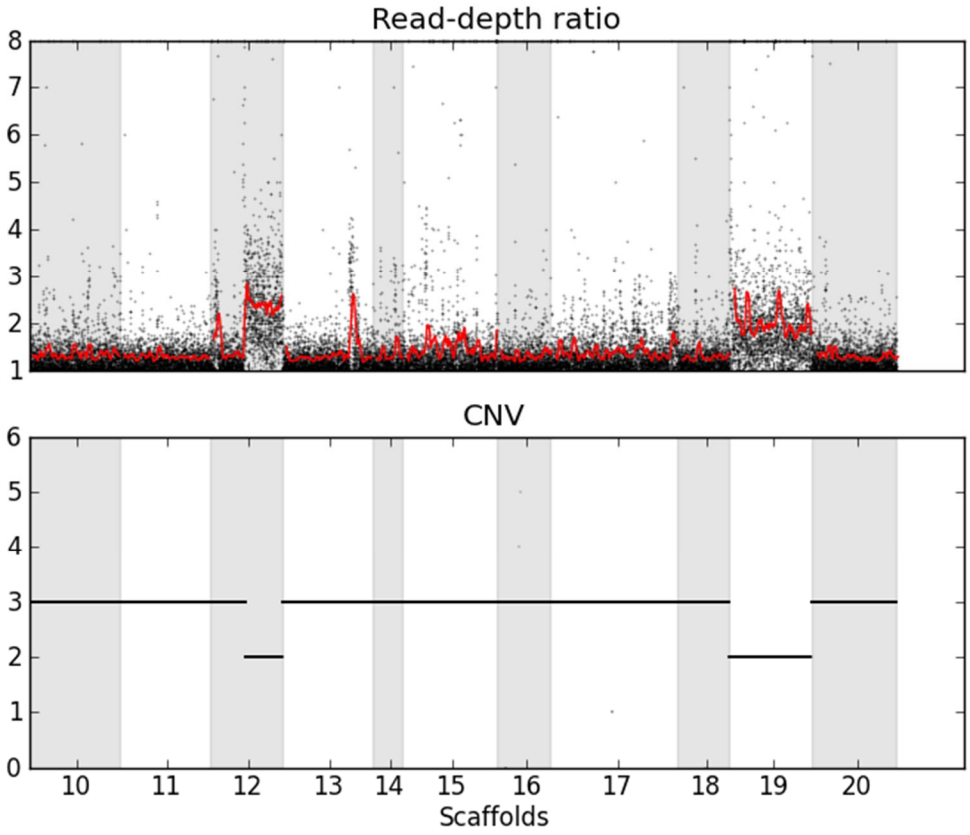


Figure S4.7: Read-depth analysis for copy number variation in the *Phytophthora ramorum* isolate P1563 (EU1) isolate.

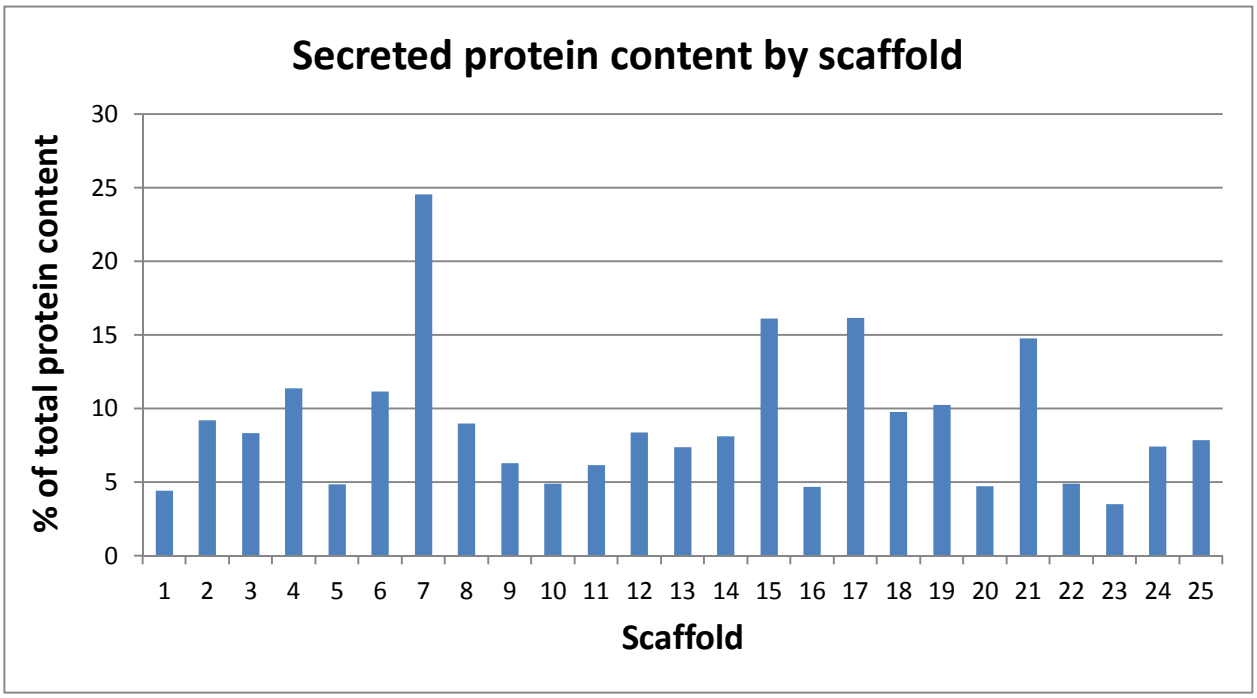


Figure S4.8: Comparison of predicted secreted protein content for the first 25 largest scaffolds in *P. ramorum*.

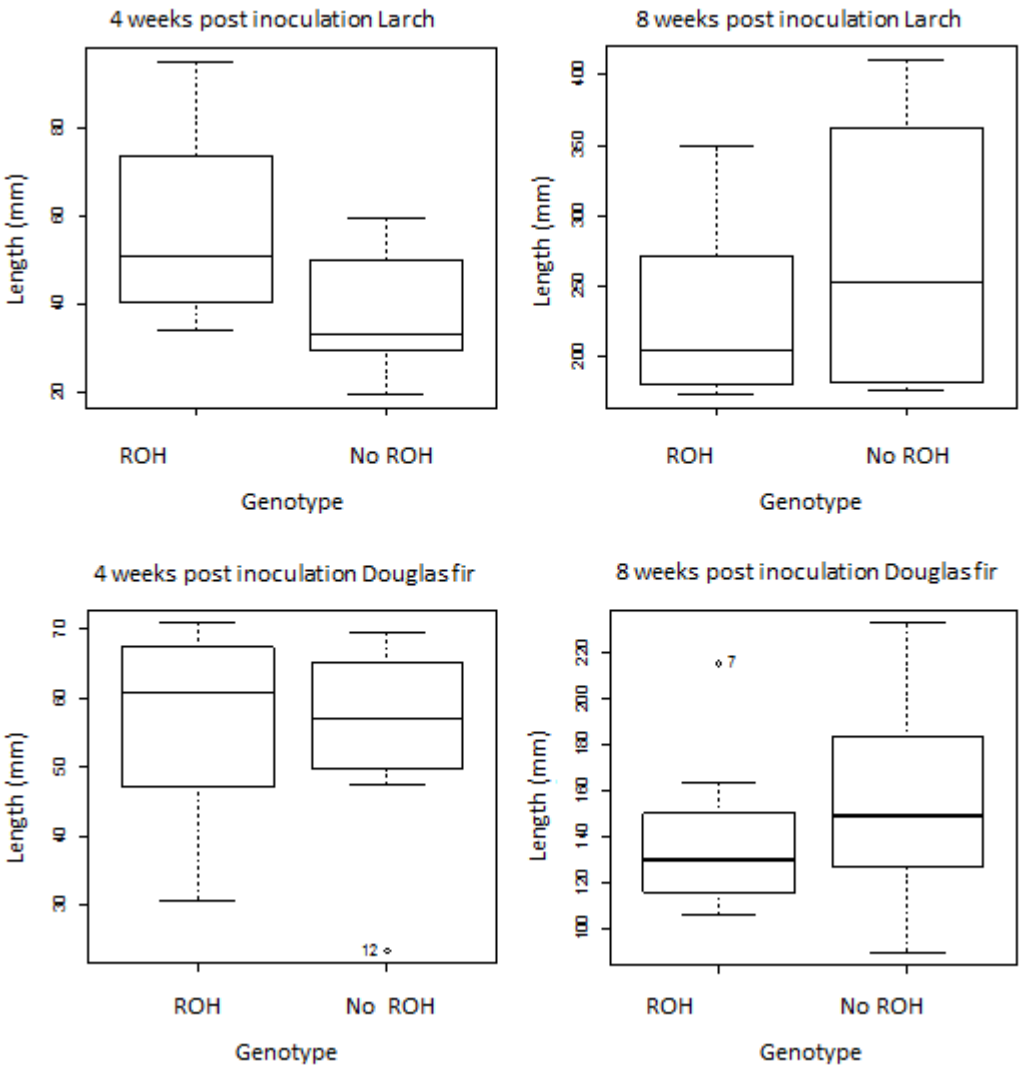


Figure S4.9: Lesion length on larch and Douglas-fir four and eight weeks post inoculation of individuals with ROH on scaffold seven compared to isolates without ROH.

P2677_EU1 NO ROH P1559_EU1 with ROH

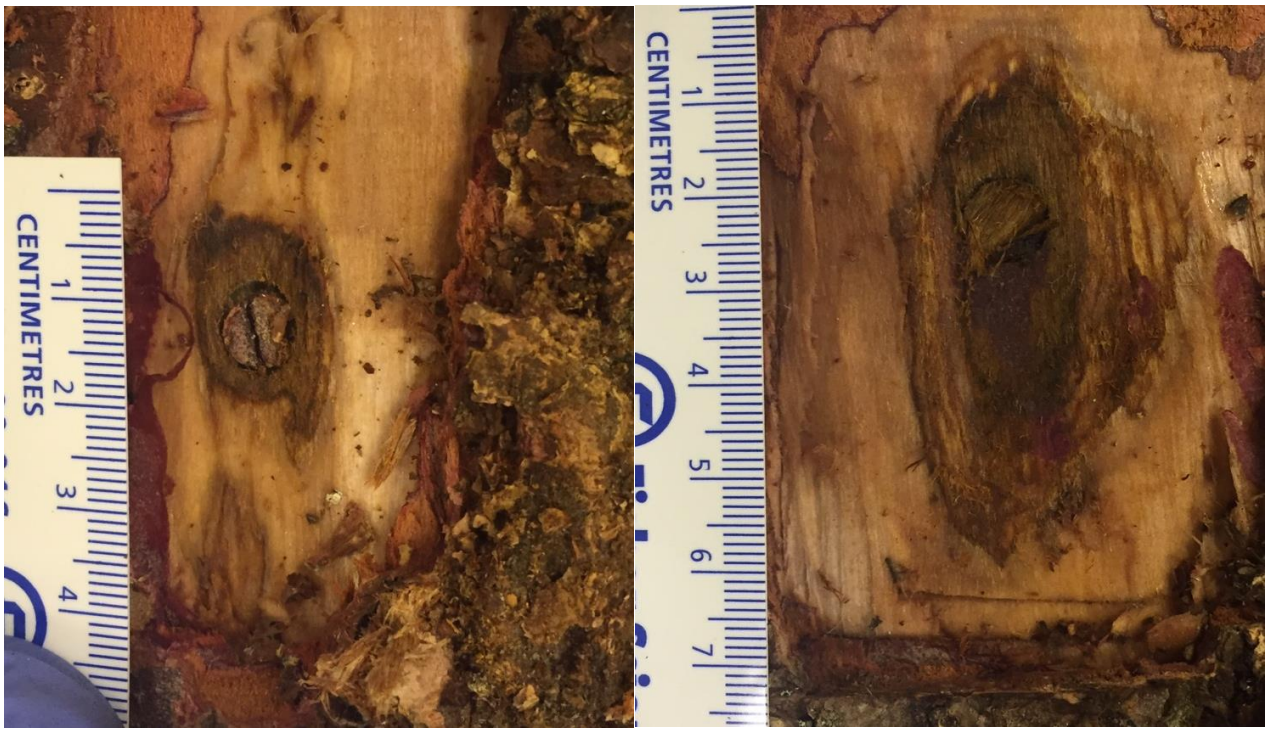


Figure S4.10: Lesions on larch four weeks four weeks post inoculation for two isolates, one with ROH and one without.

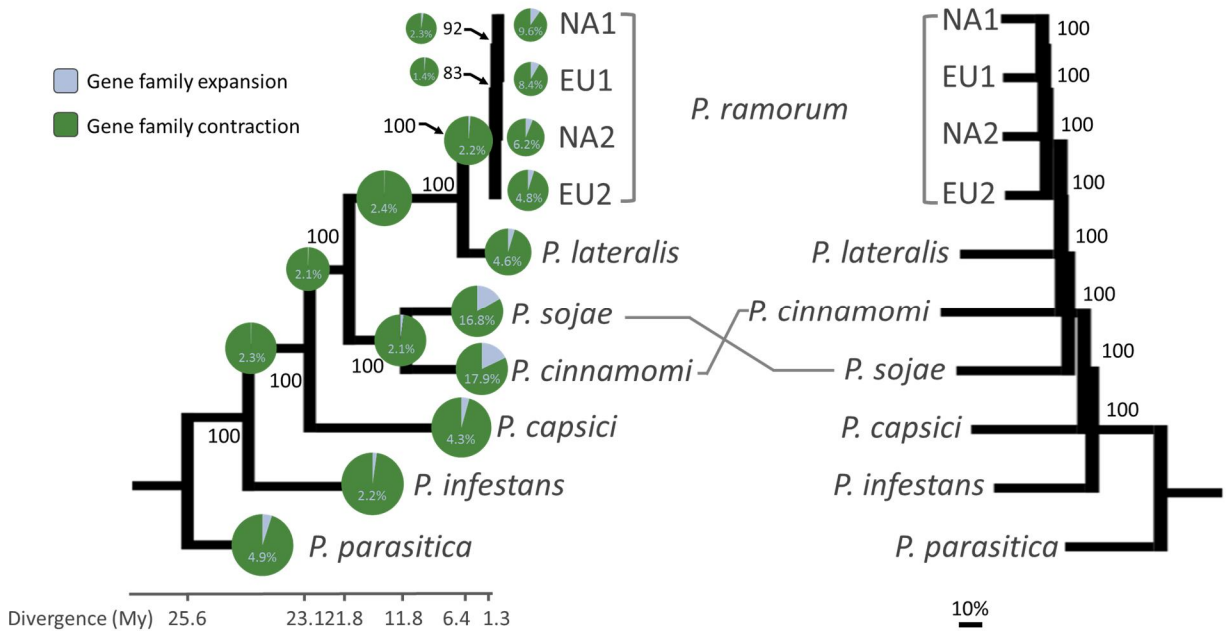


Figure S4.11: Gene family contraction and expansion in *Phytophthora* spp. Left: a CAFE analysis was conducted on the 51 458 OrthoMCL gene families and the maximum likelihood phylogenetic tree obtained for the four *P. ramorum* lineages, *P. lateralis* and five other *Phytophthora* species. The size of pie charts is proportional to the total number of gene families that changed; values in blue within the pie chart indicate the proportion of expansions; values in black at node are bootstrap support for the RAxML tree. Right: neighbor joining-tree reconstructed from a distance matrix based on genome similarity in terms of gene content (presence/absence of 51 458 OrthoMCL gene families). Values above nodes are statistical bootstrap support obtained from 100 resampling with replacement of gene families; scale bar indicates 1% genomic dissimilarity i.e. percentage difference in terms of number of shared genes.

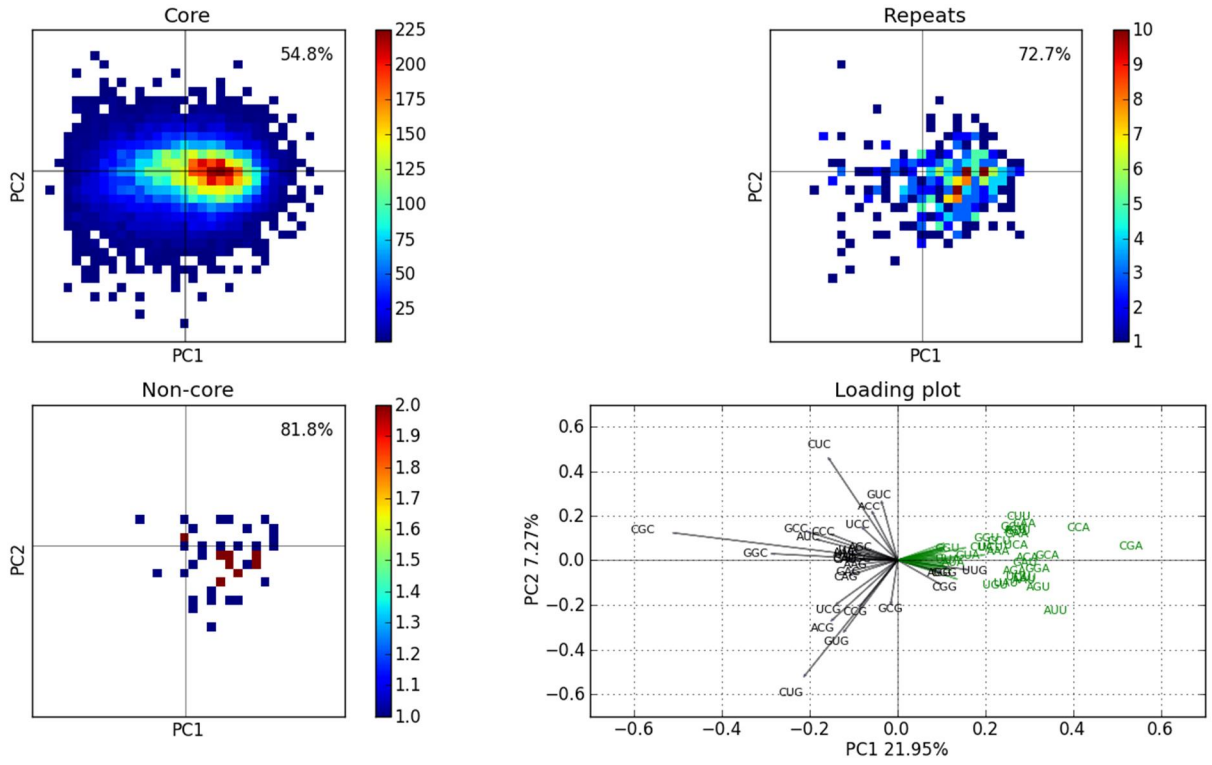


Figure S4.12: PCA of codon usage on protein coding genes of the *Phytophthora ramorum* NA2 isolate 05-16845 showing a strong bias toward usage of codons ending with 'A' or 'U'. Results are shown for genes included in the core (top left), repetitive (top right) and non-core (bottom left) genome; color scales on the right indicate protein density; percentage value on top right corner are the proportion of genes found on the right part of the PC1 axis. Bottom right graph represents the contribution of each codon variable to the different components; codons ending with 'G' or 'C' are labelled in black whereas those ending with 'A' or 'U' are in green.

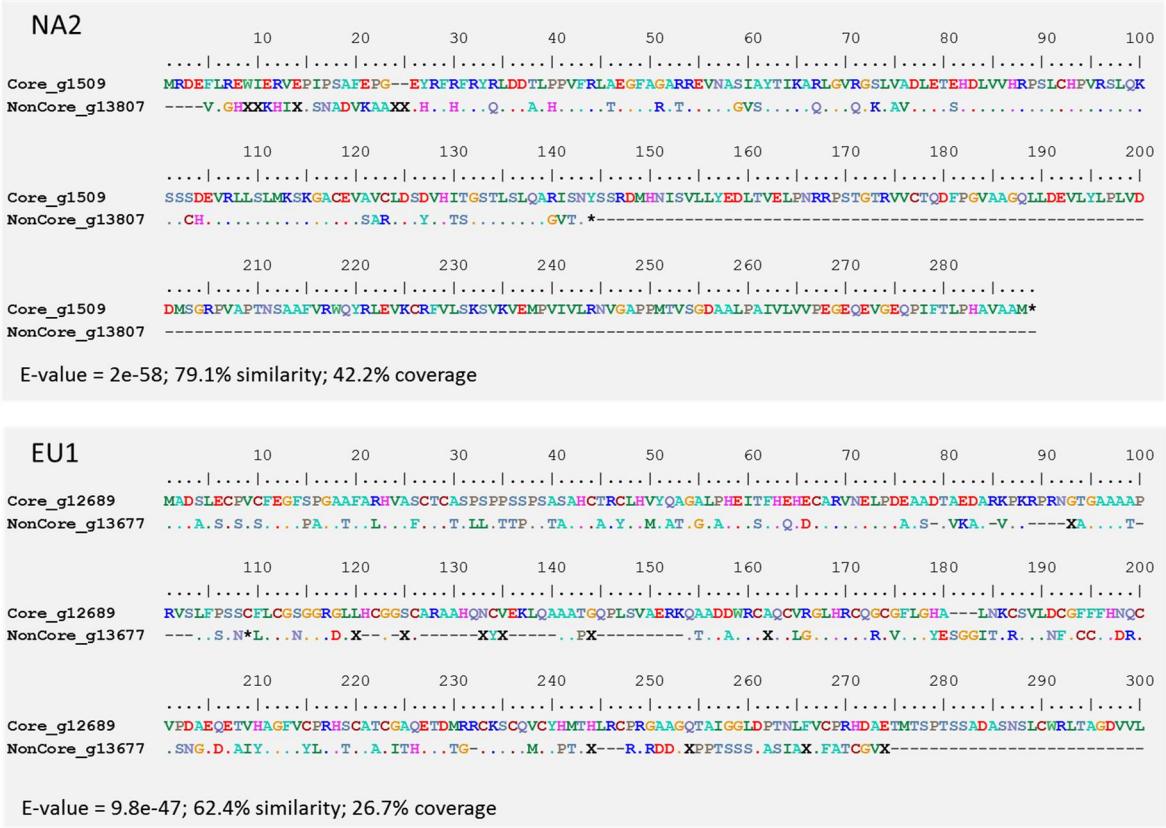


Figure S4.13: Example of degenerated protein sequence found in the non-core genomes of the NA2 and EU1 lineages. The non-core protein is aligned with its non-degenerated closest homolog found in the core genome; homologous amino-acids are represented with a dot. Waterman-Eggert similarity and coverage values are indicated under each alignment. An asterisk represents a stop codon.

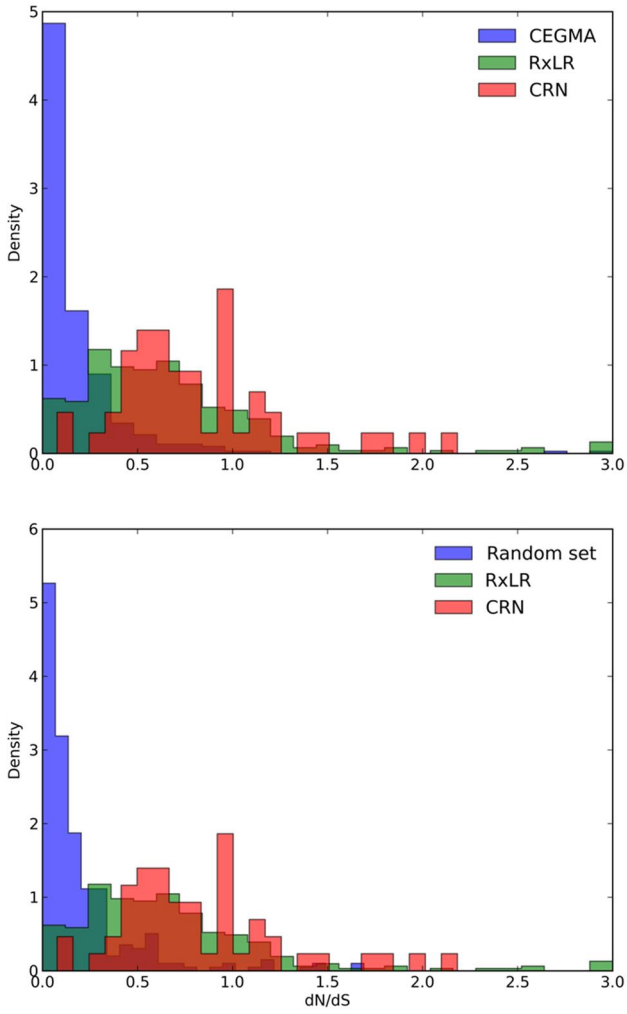


Figure S4.14: Distributions of dN/dS values obtained for the RxLR and the Crinkler gene sets compared with the CEGMA and random protein sets. RxLR gene set composed of 255 genes, Crinklers 51 genes, CEGMA set 315 genes, and random protein set 292 genes. Comparison with CEGMA set in top graph and random protein set on bottom graph.

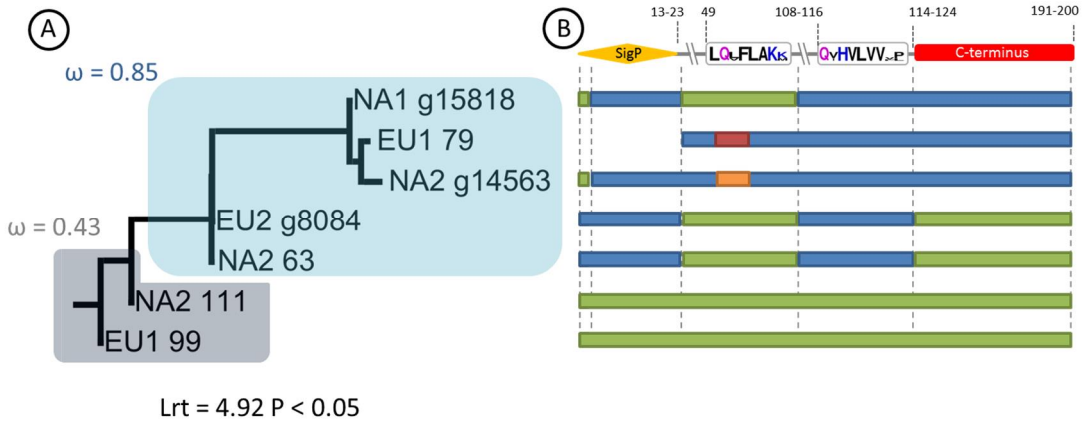


Figure S4.15: Phylogenetic relationships and evolutionary history of the CRN subfamily expanded in the *P. ramorum* EU1 and NA2 lineages. (A) CodeML branch test indicated a significantly increased dN/dS rate for five members of the CRN subfamily. (B) Recombination breakpoints observed among members of this CRN subfamily; segments of different colors represent genomic region between two breakpoints.

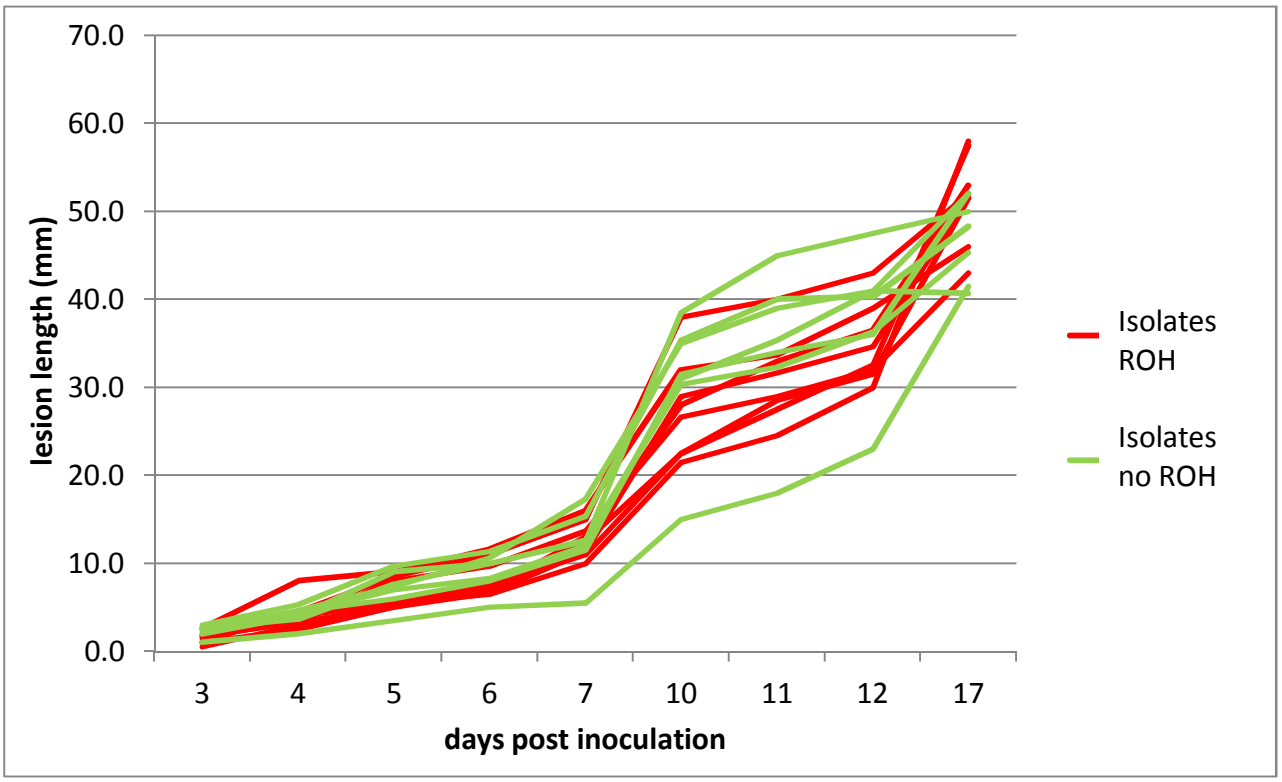


Figure S4.16: Lesion length on Rhododendron leaves of isolates with ROH on scaffolds seven and 100 compared to isolates with no ROH.

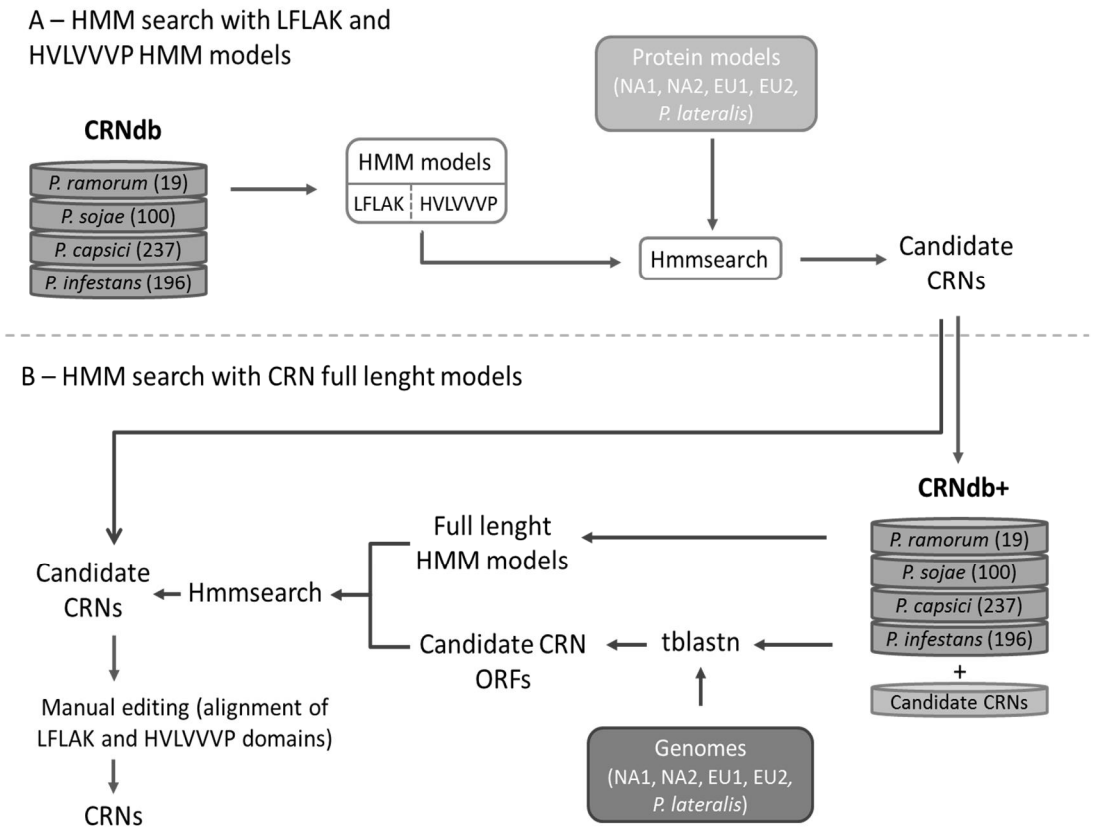


Figure S4.17: Complementary approaches used to identify *P. ramorum* and *P. lateralis* Crinkler effectors.

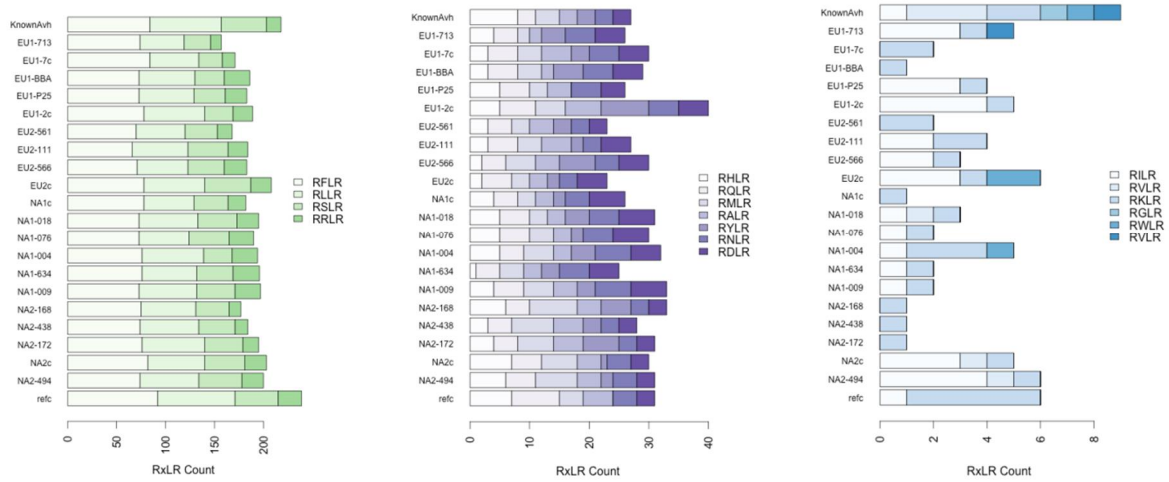


Figure S4.18: Count of each RxLR motif,where the top row shows the number present within the Avh database, the bottom row is the annotation pipeline applied to the reference genome, and all others between are the total obtained from each *de novo* assembly. The most abundant motifs (85.4%) were RFLR, RLLR, RSLR, and RRLR (Left), secondary in abundance (13.1%) were RHLR, RQLR, RMLR, RALR, RYLR, RNLR, and RDLR (Middle), and least abundant (1.55%) were RILR, RVLR, RKLR, RGLR, RWLR, and RVLR (Right).

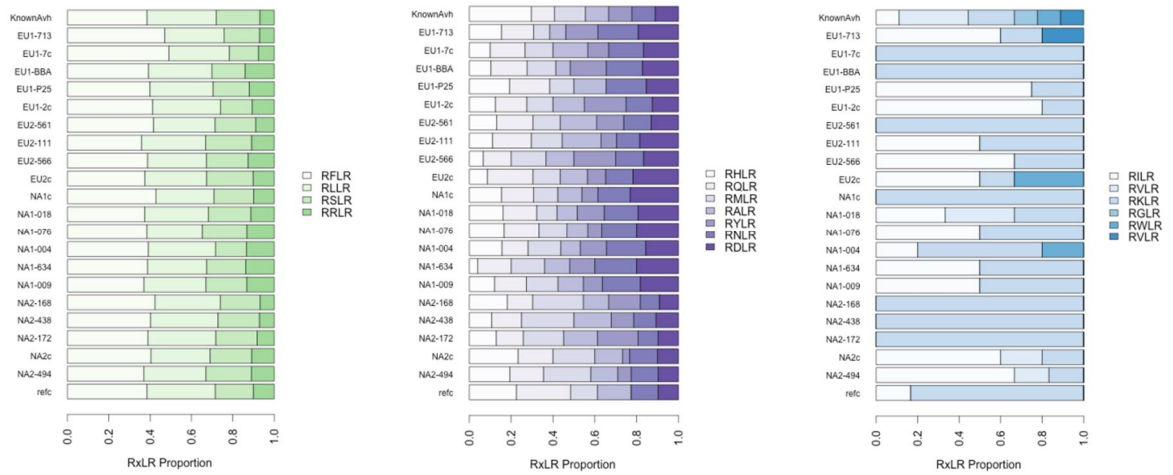


Figure S4.19: Proportion of each RxLR motif, where the top row shows the relative ratio of present in the Avh database, the bottom row is the annotation pipeline applied to the reference genome, and all others between are the relativized values for each *de novo* assembly. The most abundant motifs (85.4%) were RFLR, RLLR, RSLR, and RRLR (Left), secondary in abundance (13.1%) were RHLR, RQLR, RMLR, RALR, RYLR, RNLR, and RDLR (Middle), and least abundant (1.55%) were RILR, RVLR, RKLR, RGLR, RWLR, and RVLR (Right).

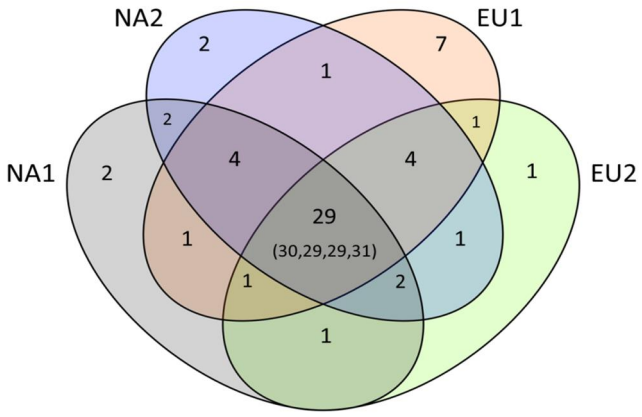


Figure S4.20: OrthoMCL clustering of Crinkler effectors found in the four *P. ramorum* lineages.