Resolving the *Pythium ultimum* species complex

by

Quinn Eggertson

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

Carleton University
Ottawa, Ontario

Your file Votre référence
ISBN: 978-0-494-93569-9

Our file Notre référence
ISBN: 978-0-494-93569-9

ABSTRACT

The *Pythium ultimum* species complex is comprised of two morphological varieties: var. *ultimum* and var. *sporangiiferum*. Previous efforts to resolve whether the two varieties are genetically distinct have produced conflicting results. Using 29 isolates of *Pythium ultimum* - twelve isolates of var. *sporangiiferum* and seventeen isolates of var. *ultimum*, multi-gene genealogies were created from both newly identified highly variable genes and previously identified genes.

Observed conflict between species estimates given by nuclear and mitochondrial gene regions resulted in two separate studies. The first examined multiple nuclear genes to resolve the species complex, while the second tested the suitability of mitochondrial genes for phylogenetic analysis. Nuclear multi-gene genealogies found that the *Pythium ultimum* species complex is comprised of four genetically distinct species, while mitochondrial multi-gene genealogies indicate that the mitochondrial genome of *Pythium ultimum* is single copy, non-recombinant and uniparentally inherited.

ACKNOWLEDGEMENT

I would like to thank C. Robin Buell and John Hamilton from Michigan State University for providing me with SNP data and allowing me access to the *P. ultimum* genome website before it was released.

I would also like to express sincere gratitude towards André Lévesque for being a supportive and engaged supervisor, Tara Rintoul for her support, patience and always appreciated advice, Gregg Robideau for his general helpfulness and for kindly lending me his primers for OCM1, and Rafik Assabgui, Julie Chapados, and Kasia Dadej for their quick sequencing. I'd also like to acknowledge Myron Smith, who taught me the basics of both genetics and mycology and Nicolas Rodrigue whose willingness and patience in explaining the nuances of bioinformatics software was appreciated. Last but not least, I would like to thank my family and friends for their support, encouragement and for only looking slightly pained when I tried to explain my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

THESIS INTRODUCTION

In the past, species were often described based on the placement of a single isolate in the gene tree of a single locus. Due to technological and monetary limitations this was, to an extent, acceptable. With the rapid emergence of faster and cheaper sequencing technology, the ability to analyze large multi-locus data sets is now possible and the cost of full genome sequencing is rapidly declining. This trend is resulting in ever increasing numbers of published genomes, enabling not only the identification of hyper-variable genes (which are ideal for species level phylogenetic analysis) but also the ability to locate conserved regions within these genes for primer design (Townsend *et al.*, 2008). This in turn, allows for the generation of multi-locus data sets that can be used to describe and define species boundaries at a better resolution than ever before (Rintoul *et al.*, 2012).

The theoretical Phylogenetic Species Concept (PSC) defines a species as the smallest monophyletic group diagnosable by a shared combination of characters (Mishler *et al.*, 2000). When the PSC is implemented using genetic data, these characters are represented by differences in nucleotide or amino acid sequences. The operational Genealogical Concordance Phylogenetic Species Recognition (GCPSR) uses multi-gene genealogies to determine phylogenetic topology through the concordance of branches supporting the presence of separate species and conflicting branches indicating gene flow (Taylor *et al.*, 2000). When species identification within a genus is made difficult by nondescript morphological characteristics and variable reproductive behaviour, the GCPSR has the capacity to analyze taxonomic relationships at a higher resolution than other species recognition techniques (Moralejo *et al.*, 2008; Taylor *et al.*, 2000). Since

genetic changes can be detected long before morphological and behavioural changes, it has been argued that the GCPSR is one of the most sensitive recognition techniques to changes described by the Evolutionary Species Concept (Taylor *et al.*, 2000).

For genetic analysis to be effective, the genes selected for phylogenetic purposes must give an accurate and unbiased representation of the evolutionary history of the organism being analyzed. However, individual loci each have their own unique evolutionary history, which may not be representative of the evolution of the organism as a whole. Incidences of incomplete lineage sorting, introgression, selective sweeps and convergent evolution can mask the true evolutionary signal and give false species estimates. To overcome potential sampling error, it is recommended that multiple unlinked, neutral, single copy genes are analyzed (Taylor *et al.*, 2000; Wiens *et al.*, 2010).

When scientists started using nucleotides as a source of informative characters for the PSC, mitochondrial genes (mtDNA) were often the marker of choice. This was because they were generally assumed to be maternally inherited, non-recombinant, selectively neutral, with a small effective population size and relatively high mutation rate (Bossu *et al.*, 2009; Galtier *et al.*, 2009). These characteristics were considered more sensitive to evolutionary signal between closely related taxa and were argued to result in more informative characters and faster lineage sorting relative to nuclear genes.

Due to its unique pattern of inheritance, the mitochondrial genome functionally acts as a single locus. With the acknowledged difference between a gene tree and species tree (along with increasing reports of incongruence between mitochondrial and nuclear species estimates) there is currently a debate over whether mtDNA is appropriate for

phylogenetic and taxonomic studies, especially when used as a solitary means for species delimitation (Bossu *et al.*, 2009; Fort *et al.*, 1984; Galtier *et al.*, 2009; Rubinoff *et al.*, 2006; Rubinoff *et al.*, 2005). The basic assumptions of a uniparentally inherited, single copy, non recombinant mitochondrial genome is pivotal to this debate, yet rarely tested. Using mitochondrial multi-gene genealogies to examine these assumptions is unconventional, yet could provide species specific information on mitochondrial inheritance patterns without having to do breeding experiments.

Bayesian statistical analysis of phylogenetic data is based on the likelihood function (or the probability of observing given data, under a known probabilistic model, as a function of model parameters) and incorporates a user specified model of nucleotide substitution (Archibald *et al.*, 2003; Lemey *et al.*, 2009). Bayesian analysis is able to efficiently sample tree space by implementing the Markov Chain Monte Carlo (MCMC) simulation and is distinguished by its ability to incorporate prior knowledge about the data set (priors) into the analysis (Archibald *et al.*, 2003). The support values generated from this analysis show the probability of the tree given the model, priors and data (Archibald *et al.*, 2003). Several programs have recently been developed to examine multi-gene data sets using Bayesian analysis. The recent modification to the program Bayesian Evolutionary Analysis of Sampling Trees (BEAST, Drummond *et al.*, 2007), Species Tree Ancestral Reconstruction –BEAST (*BEAST, Heled *et al.*, 2010), allows for the co-estimation of multiple gene trees as well as the species tree in which they are embedded. This method generates a species tree by following the stochastic coalescent process back in time from the present (Heled *et al.*, 2010).

Pythium ultimum is a ubiquitous plant pathogen found in soil and is comprised of two morphological varieties. The taxonomic status of these varieties has yet to be clarified and has proven difficult to resolve. Multi-gene genealogies are becoming increasingly popular for resolving phylogenetic uncertainty in fungal taxonomy and have yet to be applied to the *P. ultimum* species complex. The publication of the annotated *P. ultimum* genome, as well as the unpublished genomic scaffold data for the neotype and ex-type strains of the morphological varieties, have enabled the identification of hyper-variable gene regions appropriate for phylogenetic analysis.

Early in this study, an observed conflict between species estimates given by nuclear and mitochondrial gene regions gave rise to the second chapter of this thesis, which specifically addresses the suitability of mitochondrial genes for multi-gene phylogenetic analysis. The first chapter was consequently revised to focus on using multiple nuclear gene regions to resolve taxonomic uncertainty in the *Pythium ultimum* species complex.

Chapter 1: Resolving the *Pythium ultimum* species complex using nuclear multi-gene genealogies under the GCPSR

INTRODUCTION

Pythium ultimum is one of the most common plant pathogens found in soil and is also one of the most pathogenic *Pythium* spp. on crops (Lévesque *et al.*, 2010; Plaats-Niterink, 1981). With its global distribution and wide host range, it annually contributes to considerable economic loss caused by damping-off and root rot of seedlings, in both horticultural and agricultural industries (Plaats-Niterink, 1981). Two morphological varieties distinguished by their ability to produce zoospores have been identified in this species. *Pythium ultimum* var. *ultimum* is the more common and pathogenic variety, which produces oogonia but rarely sporangia and zoospores, whereas var. *sporangiiferum* is distinctly less common but readily produces zoospores at room temperature and has a slightly higher maximum cardinal temperature (Drechsler, 1960; Plaats-Niterink, 1981).

Accurate taxonomic classification is required when the pathogenicity varies within a species complex, particularly when different disease control strategies must be implemented depending on the species present. It has been pointed out that the presence or loss of sporangia and zoospore production may be epidemiologically significant as it has implications for dispersal potential, infection processes and pathogen persistence (Jeger *et al.*, 2008).

Previous efforts to resolve whether the two morphological varieties of this species are genetically distinct are extensive and have included the following techniques: protein analysis (Abdelzaher *et al.*, 1995; Martin, 1995), isozyme analysis (Abdelzaher *et al.*, 1995; Barr *et al.*, 1996), random amplified polymorphic DNA (RAPD) (Francis *et al.*,

1994), restriction fragment length polymorphisms (RFLPs) (Francis *et al.*, 1994), sequence characterized amplified regions (SCARs) (Francis *et al.*, 1994), and phenetic analysis (Gherbawy *et al.*, 2005; Lévesque *et al.*, 2004). All of these studies arrived at uncertain or conflicting conclusions regarding the taxonomic distinction of the morphological varieties. This study focuses on two papers that are key to the current taxonomic status of this species complex. Barr *et al.* (1996) examined isozyme variation, morphology and optimal growth temperatures for 97 isolates of *Pythium ultimum* in an effort to develop alternative techniques to reliably identify the species. Isolates in their study were checked for their ability to produce zoospores and subsequently classified as either var. *sporangiiferum* or *ultimum* based on this characteristic. Isolates were further sorted into ten multilocus genotypes as defined by banding patterns of nine isozyme loci; eight of the ten genotypes identified by Barr are incorporated in our analysis (21 of the 29 isolates). Francis *et al.* (1994) used RFLPs, RAPD and SCARs to evaluate three morphological varieties within the *P. ultimum* species complex that differed in their reproductive strategy (var. *ultimum*, var. *sporangiiferum* and group HS). They determined that these morphological varieties were not genetically distinct and therefore should be classified as a single species. In our study, we include the key isolate (Pu33 –ATCC 58811) that Francis *et al.* (1994) used to disqualify the validity and genetic distinction of var. *sporangiiferum* from var. *ultimum*.

The theoretical PSC and the operational GCPSR are considered to be a superior framework for determining speciation in fungi when compared to those based on morphology or reproductive behaviour (Taylor *et al.*, 2000). Even though *Pythium* and other oomycetes are not true fungi (genetically they are more similar to heterokont blue

green algae and diatoms) the classification of different species within this genus is challenging for the same reasons, which are - nondescript morphological characteristics and variable reproductive behaviour. These challenges result in the GCPSR having the capacity to analyze the taxonomic relationships at a higher resolution than other species recognition techniques (Moralejo *et al.*, 2008; Taylor *et al.*, 2000).

The Bayesian analysis program *BEAST has already been successfully used to resolve species limits in other oomycetes and filamentous fungi, including *Phytophthora* (Blair *et al.*, 2012), *Penicillium* (Henk *et al.*, 2011), *Xanthoparmelia* (Leavitt *et al.*, 2011) and *Fusarium* (Sarver *et al.*, 2011). Of the above papers, use of the GCPSR for species delimitation is explicitly mentioned by Henk *et al.* (2011) and Sarver *et al.* (2011).

The main objectives of this study were to utilize the newly available genomes of *P. ultimum* var. *ultimum* and one isolate of var. *sporangiiferum* to identify hyper-variable genes for phylogenetic analysis and design primers for these identified gene regions. Then, using both newly identified and previously identified loci, on a comprehensive set of isolates for the species, create a multi-gene genealogy to examine the *Pythium ultimum* species complex under the GCPSR framework with the newly available phylogenetic software *BEAST.

MATERIALS AND METHODS

In house BLAST, Identification and Selection of Isolates

The Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) was used for an in house database search to confirm and establish the identity of *Pythium ultimum* isolates in the AAFC culture collection. The database for BLAST was created by using a batch file to extract all available *Pythium* ITS (internal transcribed spacer of

nuclear ribosomal DNA) and COI (cytochrome c oxidase subunit 1, also abbreviated as Cox1) sequences and then importing them into Lasergene SeqMan (DNAStar) to create a single FASTA file. The ITS 1 region (ITS1) and COI sequences of the var. *ultimum* neotype (as designated by Plaats-Niterink, 1981) and the var. *sporangiiferum* ex-type strain (Drechsler, 1960) were used as the query. The ITS1 fragment was used for the query as not all isolates had the full ITS1-5.8S-ITS2 region sequenced.

The top 170 isolates selected using the BLAST search were compiled into a new FASTA file, aligned using the public domain alignment software MUSCLE (Edgar, 2004), and a preliminary phylogenetic tree was generated by conducting a neighbour joining analysis in PAUP (4.0b10).

By using previously identified isolates of var. *sporangiiferum* and var. *ultimum,* potential varietal groups in the ITS1 phylogeny were extrapolated and putative isolates of var. *sporangiiferum* were identified. All putatively identified (as determined by phylogenetic topology in ITS1 and COI NJ trees), previously identified (those that were consistent with the BLAST), and all obtainable published isolates of var. *sporangiiferum* were used. Therefore, not all strains classified as var. *sporangiiferum* in this study were observed to produced sporangia or zoospores.

Isolates of var. *ultimum* were selected to include: genetic variability as determined by phylogenetic topology in both COI and ITS1 gene phylogenies, a genetically identified isolate of var. *ultimum* that had been observed to produce zoospores – ATCC58811 (Francis *et al.*, 1994), 8 of the 10 genotypes described by Barr *et al.* (1996) and, to negate genetic effects caused by geographical isolation, var. *ultimum* strains were selected to have geographical overlap with putative isolates of var. *sporangiiferum.* In

total, 29 isolates were selected, seventeen of var. *ultimum* and twelve of var. *sporangiiferum* (Table 1).

Acquisition of Strains and Culture Conditions

Cultures obtained from the AAFC culture collection were grown on 2.5% V8 agar at room temperature. After sufficient colonization of the medium had occurred, a 1 mm^2 piece of agar was cut from the margins of the colony and transferred to a capped glass test tube filled half full with Difco potato dextrose broth.

DNA Extraction

DNA extraction was carried out using a protocol modified from Moller *et al.*(1992) as follows. Quarter sections of mycelial mats grown in potato dextrose broth were excised and added to 2mL screw cap tubes containing 500µL of TES Buffer (100mM Tris, pH 8.0, 10mMEDTA, 2% SDS), 50µL of proteinase K stock solution (1mg/mL) and 300 mg of zirconium oxide spheres (Fox Industries). A 6mm zirconium oxide sphere (Fox Industries) was added to tubes before samples were ground by being shaken vertically in a FastPrep® machine (BIO 101) for 45s at speed 4 and incubated at 65°C for 45 min. Samples were centrifuged at 15000 rcf for 4 min at 4°C. Supernatant was pipetted into a new 1.5mL snap-cap microcentrifuge tube along with 150µL of 5M NaCl and 70µL of 10% CTAB. Microtubes were vortexed for 30s and incubated for 10 min at 65°C. Following incubation, 700µL of chloroform: isoamyl-alcohol in a 24:1 ratio was added to the microtubes before being vortexed for a minute and placed on ice for 10 min. Tubes were centrifuged for 5 min at 15 000 rcf at 4°C before supernatant was transferred into a new 1.5mL snap-cap

Table 1: Isolates of *P. ultimum* separated into variety and genetic colour groups. Varieties were putatively identified based on the topology of ITS1 and COI phylogenetic trees, while the genetic colour groups were used to ease visual comparison of gene trees and were based on major clades observed in the *BEAST species tree.

Genetic Group	*Isolate ID*	*Location*	*Genotype* [a]	*Host*
var. *ultimum*				
bright yellow	BR144	Canada, Ontario	U1	*Nicotiana tabacum*
	BR816	South Africa, Transvaal Province[b]	U1	*Triticum*
	BR844	United States of America, Washington	U1	*Triticum*
	CBS291.31	n/a	-	*Dioscorea batatas*
	CBS296.37	United Kingdom	-	*Pisum sativum* root
	Lev2075	United States of America, Washington	-	*Malus*
	ATCC58811[c]	United States of America, Washington	-	*Triticum*
pale orange	CBS398.51[d]	Netherlands	U3	*Lepidium sativum*
	CBS729.94	Canada, British Columbia	-	*Triticum* growing on loam soil
	BR793	South Africa, Cape Province[b]	U3	*Citrus*
pale yellow	BR638	Canada, Alberta	U2	*Pisum sativum*
	BR858	United States of America, Colorado	U2	unknown
	BR583	Canada, Alberta	U2	*Carthamus*
	BR657	Canada, Alberta	U2	*Carthamus*
medium orange	BR825	United Kingdom	U4	*Viola*
	BR827	India	U4	*Psoralia*
dark orange	BR840	United States of America, California	U7	soil
var. *sporangiiferum*				
blue	CBS219.65[cef]	United States of America, Maryland	U6	*Chenopodium album,* rotting root
	CBS114.79[e]	Spain, Ibiza	U6	clay soil, under *Euphorbia canariensis* and *Artemisia canariensis*

	CBS111.65[e]	Lebanon, Beirut	U6	*Medicago sativa*
	BR776	South Africa, Transvaal Province[b]	U9	*Persea*
	BR781	South Africa, Cape Province[b]	U9	*Citrus*
	BR944	South Africa	-	*Chamaecytisus palmensis*
	DAOM240290	Canada, Ontario	-	soil
purple	BR783	South Africa, Transvaal Province[b]	U8	*Citrus*
	BR841	United States of America, Florida	U8	soil
	BR790	South Africa, Transvaal Province[b]	U8	*Citrus*
green	BR805[e]	South Africa, Transvaal Province[b]	U6	*Citrus*
	CBS122650	France	-	soil, identified at CBS as *Pythium glomeratum*

[a] Multi-loci genotype described by Barr *et al.*(1996)
[b] The provinces of Transvaal and Cape were subdivided in 1994 and are no longer recognized
[c] Zoospore production observed by Francis *et al.* (1994)
[d] Neotype for var. *ultimum* (Plaats-Niterink, 1981)
[e] Zoospore production observed by D. Barr (1996)
[f] Type strain of var. *sporangiiferum* (Drechsler, 1960)

microcentrifuge tube along with 225µL of 5M ammonium acetate. Tubes were inverted several times and placed on ice for 30 min before being centrifuged at 15 000 rcf for 10 min at 4°C. Supernatant was carefully poured into a new 2mL screw cap tube along with 700µL of isopropanol and chilled for 10 min before being centrifuged at 15 000 rcf for 10 min at 4°C to form a pellet. Isopropanol was poured off into a waste beaker and the inside of the tube was washed with 800µL of chilled 70% ethanol. Ethanol was pipetted out of the tubes before they were inverted over a folded Kimwipe in a laminar flow cabinet for 10 min to drain any ethanol residue. Pellets were resuspended with 50µL 0.1xTE containing 50µL 50µg/m of RNase A and incubated at 65°C for 10 min. Extracted DNA was stored at -20°C.

SNP Analysis and Gene Selection

Candidate genes were selected based on four criteria, a) that they had relatively high rate of single nucleotide polymorphisms (SNPs), b) they were at least 1000 bp in length, c) they were single-copy in the genome and d) they had orthologous gene regions in other oomycetes. The intra-varietal variation (SNP/kb) for every gene in the *P. ultimum* genome was determined by doing a BLAST search using the fully annotated genome data of BR144 (Lévesque *et al.*, 2010) against scaffold data of the genome for CBS 398.51 (Hamilton, unpublished). Genes that had high contig SNPs and were at least 1000 bp long were then checked for orthologs and duplicate copies using the Genome Solutions website (http://oomycetes.genomeprojectsolutions-databases.com, unpublished). This website has pre-generated trees of homologous genes found within all published oomycete and diatom genomes, and was searched using the BR144 gene model name. If orthologs were identified, FASTA files of the corresponding gene regions were

obtained from the relevant published genome website. Finally, the presence of high intra-varietal SNPs were visually confirmed for each candidate gene and FASTA files for BR144 and CBS 398.51 were obtained from the genome website. Through this process, four single-copy, highly variable genes over 1000 bp long with orthologs in other oomycetes were identified: G002286, G002278, G002263 and G001129 (Table 2). Additionally, three previously identified gene regions, for which primers were already available, were also included in the analysis: ITS, beta-tubulin (Btub) and tubular mastigoneme protein (OCM1).

Primer Design

Primers were developed using published and unpublished genomic data from three isolates of *Pythium ultimum* and five other *Pythium* species. Nineteen primers were designed for genes G002286, G002278, G002263 and G001129. Conserved regions in *P. ultimum* were identified by aligning genomic sequences from BR144 (Lévesque *et al.*, 2010), CBS 398.51 (Hamilton, unpublished) and CBS 219.65 (Hamilton, unpublished) in MegAlign (DNASTAR). Secondary alignments were also created that included orthologous genes in closely related groups, in hope of finding conserved gene regions in *Pythium* or oomycete taxa suitable for primer design. Primers were designed in Oligo 7 (Molecular Biology Insights, Inc.) either manually or by using the program's "Primer and Probe" search function. Forward and reverse primers were designed to coincide with conserved regions in *P. ultimum,* and when possible with other taxa, amplify a product between 900-1200 bp in length, and have an approximate melting temperature of 55°C. For each gene, 2-6 primers were developed and tested in pairs. The location of the primer on the gene is indicated in the primer name.

Table 2: Information on single-copy[a] nuclear genes used for phylogenetic analysis of *Pythium ultimum*. Annotated function (with the organism where the gene was first annotated enclosed in parentheses), gene length, number of introns and scaffold ID were acquired from the *Pythium ultimum* genome website[b], while intra-varietal variation was obtained through an in-house BLAST search of the annotated genome of BR144 (var. *ultimum*) and the scaffold data of CBS 398.51 (the neotype of var. *ultimum*).

Gene [c]		***Annotated Function***	***Intra-varietal Variation (SNP/kb)***	***Length (bp)***	***Introns***	***Scaffold***
G002286	T002289	gtf2h1: General transcription factor IIH subunit 1 (*Dictyostelium discoideum*)	19.2	1671	0	scf1117875582008
G002278	T002281	RPT1:26S protease regulatory subunit 7 (*Spinacia oleracea*)	19.7	1323	1	scf1117875582008
G002263	T002266	CGN: Cingulin (*Homo sapien*)	18.1	3426	7	scf1117875582008
G001129	T001129	stfR: Side tail fiber protein homolog from lambdoid prophage Rac (*Escherichia coli*, strain K12)	12.2	1310	5	scf1117875582023
G003503	T003513	OCM1- tubular mastigoneme protein	3.2	1815	0	scf1117875582006
G014589	T014620	Tubulin beta chain (*Phytophthora cinnamomi*)	1.0	1634	2	scf1117875582021

[a] ITS is a multi-copy nuclear ribosomal gene and not annotated in the genome
[b] http://pythium-anno.plantbiology.msu.edu/cgi-bin/gbrowse/pug1_v4/
[c] First column indicates the locus identifier and the second column indicates the gene model name

Gene Amplification with PCR

The primers used for PCR amplification and sequencing reactions for each gene are listed in Table 3. For each PCR reaction, 1μL of DNA diluted to 5-50ng/μL was added to 1.0μL of 10x Titanium Taq buffer (with 3.5mM $MgCl_2$), 0.5μL of 2mM dNTPs, 0.04μL of 20μM forward primer, 0.04μL of 20μM reverse primer, 0.1μL of 50x Titanium Taq polymerase and 7.32μL of sterilized HPLC water, for a final reaction volume of 10μL. Gene amplifications were carried out in a Mastercycler® ep gradient thermocycler (Eppendorf). Thermocycler profiles for the amplification of the seven gene regions include an initial denaturation at 95°C for 3 min, followed by 40 cycles of 95°C for 30s, annealing for 45s at a temperature specific for each primer pair (Table 3), and extension at 72°C for 1 min 30s (2 min for OCM1), with a final extension at 72°C for 8 min (10 min for OCM1). The primers used for each isolate are given in Appendix A.

Gel-Electrophoresis

Gel-electrophoresis was used to check negative controls for contamination and to ensure successful amplification of the target genes. Samples with distinct single bands were then processed for sequencing. Samples that did not produce a band were re-amplified using 2 μL of diluted DNA, or with different primers. The volume of PCR products with distinct multiple bands were increased to 20-25μL by either combining two reactions (for the same isolate and gene region) or with sterilized HPLC water. These samples were then purified using E-gel® CloneWell Agarose Gels (Invitrogen™) and visually re-assessed on agarose gels before being processed for sequencing.

Table 3: Primers used for PCR and sequencing reactions for isolates of *Pythium ultimum*. Annealing temperatures (Ta) are only provided for PCR primers.

Gene	*Primer name*	*Primer sequence*	*Ta*	*Reference*
ITS	UN-UP18S42	CGTAACAAGGTTTCCGTAGGTGAAC	68	Bakkeren, 2000
	UN-LO28S22	GTTTCTTTTCCTCCGCTTATTGATATG	68	Bakkeren, 2000
OCM1	OCM1_up130	ATGTGCAACTGCTAYAAGAACTAC	60	Robideau, pers. comm.
	OCM1_lo1842	CARTTGTCGTTVGTGTADCCCTTG	60	Robideau, pers. comm.
	OCM1_up631^{s}	TGCAARRTBGGYGTKGATCC		Robideau, pers. comm.
	OCM1_lo643^{s}	GGATCMACRCCVAYYTTGCA		Robideau, pers. comm.
Btub	Oom-BTub-Up415	CGCATCAACGTGTACTACAA	55	Bilodeau, 2007
	Oom-Btub-Lo1401^{p}	CGCTTGAACATCTCCTGG	55	Bilodeau, 2007
	Oom-Btub-lo1402^{s}	CGCTTGAACATCTCCTG		Levesque, unpublished
	OomPy-BTub-Up901^{s}	TACGATATCTGCTTCCG		Levesque, unpublished
	OomPy-BTub-Lo954^{s}	CACACACCAAGTGGTTC		Levesque, unpublished
G002286	PYU1_G002286_F_245	GACGCTCGTGCTKGAGTT	68	
	PYU1_G002286_R_1473_pul	GCGACCTTGGCTTTGAACGA	68	
	Pyt_G002286_F_207	GTCTCCCAAGGCRATGATCCG	65	
	Pyt_G002286_R_1552	TCGTCGTACTTGGAGCCCAT	65	
	Pyt_G002286_R_1451	CGCAGCAGRTCRTTCACGT	65	
G002278	PYU1_G002278_F_144	GCRVGTGGARGASGAYATCAA	62	
	PYU1_G002278_R_994^{p}	GTSCGCTGCACYTCRTTGTC	62	
	Oom_G002278_F_442	AAGTACGCSATCCAGATCCC	55	
	Oom_G002278_R_1352	TGATGACYTTGTTGACCGACT	55	
G002263	PYU1_G002263_F_1028	CARCTYRTYCGSTCGAACCAC	60	
	PYU1_G002263_R_2034	TKGTGRTACGCYTCRAAGTTCT	60	
	PYUvS_G002263_F_986	CCAATTTGCGCTCGAAGCTAC	60	
	PYU_G002263_R_2095	TGCTGCTTCTGGTTCTTGTCC	60	
	PYU1_G002263_F_1007^{s}	GAGTTTAAGGRCATCAAGCAG		
	PYU1_G002263_R_2029^{s}	SGCTTCAAAGTTCTTGTCG		

G001129	PYU1_G001129_F_391	CTGCTCGCGCCATTGACC	68
	PYU1_G001129_R_1445	TGCAATGACGGCTACGCAAA	68
	PYU1_G001129_F_107	TGGTCACGGGYCAGATCTCC	65
	PYU1_G001129_R_1386	AGCAGCACGTACARGAAGAAGA	65

[s] primer used only for sequencing
[p] primer used only for PCR

Sequencing

PCR products were amplified for sequencing using ABI BigDye™ Terminator v3.1 Cycling Sequencing Kits in a reaction volume of 10µL, with BigDye® Seq Mix diluted 1:8 with Seq buffer. Final concentrations of each reagent were 0.875x Sequencing buffer, 5% trehalose, 0.125x BigDye® Seq Mix, and 0.16 µM primer. Reaction volumes were brought to 10µL with sterile HPLC water and 1 µL of PCR product was added directly from the initial PCR amplification without doing a purification step (Robideau *et al.*, 2011). Samples that produced faint bands on agarose gels had 2 µL of PCR product added to the sequencing mixture, in order to increase amplicon concentration. Thermocycler profiles for the sequencing reactions of the seven gene regions had an initial denaturation at 95°C for 3 min, followed by 40 cycles at 95°C for 30s, annealing for 20s (40s for ITS and OCM1) at 55°C (58°C for ITS and 50°C for Btub) and extension for 60°C for 2 min (4 min for ITS and OCM1). An Applied Biosciences Prism® 3130*xl* Genetic Analyzer was used to generate DNA sequences from the sequencing amplification reactions.

Outgroups

Outgroup sequences for each locus, except for ITS, were selected following a BLAST search of preliminary genomic scaffold sequences of *P. aphanidermatum* (DAOM BR444), *P. iwayamai* (personal strain from Ned Tisserat, CSO), *P. irregulare* (CBS 250.28 = DAOM BR486 = LEV1481), *P. vexans* (CBS 119.80 = DAOM BR484), and *P. arrhenomanes* (CBS 324.62 = Lev1578) (unpublished). *Pythium heterothallicum* (CBS 450.67 and BR490) and *P. splendens* (CBS 462.48) sequences were generated

using primers and methods described above. Nucleotide BLAST was used for G002286, G002278, Btub, and OCM1, while a protein BLAST was needed for locating homologous gene sequences for G002263 and G001129. Gene sequences for ITS were taken from the DNA barcoding paper by Robideau *et al.* (2011), with accession numbers as follows: *P. irregulare* BR486 = CBS 25028 (HQ643596.1), *P. splendens* M0129 = CBS 46248 (HQ643795.1), *P. iwayamai* CBS 15664 (HQ643669.1), *P. heterothallicum* = CBS 45067 (HQ643553.1). Outgroups for final phylogenetic trees were *P. iwayamai*, *P. irregulare* and *P. splendens*.

Sequence Editing, Alignment and Phylogenetic Analysis

Ten phylogenetic trees were generated using sequencing data from seven nuclear genes (ITS, OCM1, Btub, G002286, G002278, G002263 and G001129) to create seven gene trees and three species trees.

Sequencing results were edited using Lasergene SeqMan software (DNAStar) and aligned using the online multiple alignment program MAFFT v. 6 using default settings (http://mafft.cbrc.jp/alignment/server/) (Katoh *et al.*, 2008). Aligned sequences were then visually evaluated in BioEdit Sequence Alignment Editor version 7.0.9, with sequences sorted by title and outgroups moved to the end. Alignments were converted from FASTA to nexus files with either MacClade 4.06 or SeaView version 4.3.2 (Gouy *et al.*, 2010). Phylogenetic analyses were carried out on each gene region using neighbour joining with PAUP (4.0b10), Bayesian analysis using MrBayes 3.0 (Ronquist *et al.*, 2003) and the recent modification to the program Bayesian Evolutionary Analysis of Sampling Trees (BEAST, Drummond *et al.*, 2007), Species Tree Ancestral Reconstruction –BEAST (*BEAST, Heled *et al.*, 2010). Evolutionary models for each locus were determined by

using likelihood analysis, with three substitution schemes in jModelTest 0.1.1 (Posada, 2008) and MEGA5 (Tamura *et al.*, 2011). Evolutionary models were ranked based on Bayesian information criterion (BIC). Aligned sequences for the seven nuclear loci were uploaded into BEAUTi v1.6.2., which produced the input XML file required for BEAST. In BEAUTi, species tree ancestral reconstruction (*BEAST) was selected, implementing the multilocus species tree estimation function. A mapping file was imported that designated each isolate as its own hypothetical species in order to see how the individual isolates were placed in the final species tree. Evolutionary models GTR + gamma was used for Btub, G002278 and G002286. While HKY + gamma was used for G001129, G002263, ITS and OCM1. The base frequencies for each locus were set to empirical, the clock model was set to estimated random local clock, the coefficient of variance was changed to normal with a mean of 0.5 and a standard deviation of 0.2, and the weight for GTR substitution parameters was increased to 1.5. Four analyses were run concurrently, with the chain length for each analysis set for 1×10^8 generations, logging every 2 000 generations and recording branch lengths (mutations per site). The graphical and statistical software package Tracer v1.5 was used to monitor convergence (Drummond *et al.*, 2007). The tree files from the four runs were then combined with a burn-in of 12 500 generations, renumbered, with the suffix changed to .nex.t and scientific notation removed using LogCombiner 1.6.1(Drummond *et al.*, 2007).

Gene and Species Tree Visualization

Using Mr. Bayes 3.0, phylogenetic trees were generated by executing a nexus file containing only the isolate headings - which allowed the program to read the tree files generated by *BEAST. The sumt command was then used, specifying a zero burn-in

(burnin = 0), only a single run (nruns=1) and 50% majority consensus tree (contype= halfcompat). Initial nexus input and tree file names were the same, with the tree file having an additional .t at the end (i.e. filename.nex and filename.nex.t respectively).

Two additional species trees were generated in SplitsTree4 (Huson *et al.*, 2006) and BUCKy 1.4.0 (Bayesian Untangling of Concordance Knots in yeast) (Larget *et al.*, 2010). Both programs used the gene tree files with combined runs and burn-ins generated from *BEAST that were used in Mr. Bayes, but required an additional formatting step where the files were resampled every 50 generations and renumbered using LogCombiner 1.6.1. This was necessary as the tree files were otherwise too large to be opened in these programs.

Before uploading data into SplitsTree, the memory allocated to the program was increased by opening a disk operating system (DOS) window, navigating into the SplitsTree folder and typing "SplitsTree.exe –J-Xmx1000M". This had to be done each time prior to using SplitsTree. The gene tree files were uploaded by using the "Load Multiple Trees" option under Tools in the File menu. Once uploaded, Consensus Networks with thresholds of 0, 0.1, 0.25, 0.33 and 0.5 were generated and the images were exported and saved as pdf files.

BUCKy is a C++ program that runs in Linux or Mac OSX and therefore was run using an Oracle VM VirtualBox 4.1.2 with an Ubuntu 32bit operating system. A shared folder between the main computer and the virtual machine was established and the seven nuclear gene tree files were copied into it. Files were converted from BEAST file format to MrBayes file format using beast2phy linux scripts available online (http://www.abc.se/~nylander/beast2phy/index.html). File reformatting was unnecessary

for Mr. Bayes or SplitsTree as long as scientific notation was removed, but BUCKy was unable to recognize the unconverted files. The beast2phy and BUCKy executable files were directly downloaded into VirtualBox and extracted using a terminal window and installed into the bin directory. After navigating back into the shared folder the files were converted individually by typing "beast2phy.pl --format=mb --outfile=*outputfilename.t inputfilename.nex.t*" with the input file directly following the output file separated by a space.

Before BUCKy was run, the files were further formatted using the subprogram mbsum to individually convert each gene tree file into an input file with the suffix ".in". This was accomplished by typing the program name and then the file name that was to be converted, i.e. "mbsum *inputfilename.t*". After this was completed, BUCKy was run by typing "bucky *.in", which analyzed all the files in the folder with the ending ".in". The run1.concordance file was opened using Microsoft Word and the tree files were copied into a text file which was opened using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). The fourth tree, the Primary Concordance Tree with Sample Concordance Factors, was selected and branch labels were made visible (concordance factors are saved as branch lengths) before the tree was exported and saved as a pdf file. Phylogenetic trees were edited using Adobe Illustrator CS2.

RESULTS

Preliminary Screening of Isolates

The in-house BLAST of the var. *ultimum* neotype and var. *sporangiiferum* ex-type strain, and the resulting ITS1 phylogenetic tree, matched three isolates to var. *sporangiiferum* that had either been misidentified or not yet identified in our collection: BR944 (*Pythium sp.*), DAOM240290 (unknown) and CBS 122650 (*Pythium glomeratum*) (Table 1). It also matched an additional five samples to var. *sporangiiferum* that had been identified as var. *ultimum* in Barr's 1996 paper because they had failed to produce zoospores: BR776, BR781, BR783, BR841 and BR790 (Table 1). In total, twelve samples of var. *sporangiiferum*, eight of which were putatively classified by sequence comparisons, and seventeen samples of var. *ultimum* were included in this study (Table 1).

PCR and Sequencing Primer Performance

The main primer pair used for gene amplification and sequencing for the G002286 gene region was PYU1_G002286_F_245 and PYU1_G002286_R_1473_pul, while Pyt_G002286_F_207, Pyt_G002286_R_1552 and Pyt_G002286_R_1451 were used for isolates that did not initially sequence well. The later primer pair was also used for sequencing the *P. splendens* and *P. heterothallicum* outgroups.

The three G002278 primers PYU1_G002278_F_144, Oom_G002278_F_442 and Oom_G002278_R_1352 worked for amplification and sequencing of all isolates. PYU1_G002278_R_994 worked well for PCR but was unsuccessful for sequencing.

For G002263, PYU1_G002263_F_1028 and PYU1_G002263_R_2034 worked well for the seventeen isolates of var. *ultimum* but consistently generated multiple bands in var. *sporangiiferum,* with only four of the twelve isolates being successfully sequenced. The primers PYUvS_G002263_F_986 and PYU_G002263_R_2095 were designed and successfully used to amplify and sequence all twelve isolates of var. *sporangiiferum*. The *P. splendens* and *P. heterothallicum* outgroups were sequenced using PYU1_G002263_F_1057 and PYU1_G002263_R_2029.

The G001129 primers PYU1_G001129_F_391 and PYU1_G001129_R_1445 yielded multiple bands in many of the isolates. Therefore, PYU1_G001129_F_107 and PYU1_G001219_R_1386 were used for amplification and sequencing of isolates that did not work well with the initial two primers. The later primer pair produced cleaner bands, longer sequences, and was preferred. However, multiple bands occurred in the two *P. heterothallicum* isolates.

Sequencing and Sequence Editing

All genes were successfully sequenced for the 29 isolates of *P. ultimum*. The two isolates of *P. heterothallicum* (BR490 and CBS 45067) were not successfully sequenced in all genes and therefore were dropped from the final analysis. It should also be noted that even though *P. heterothallicum* was sequenced using G001129 primers, the inability to align the sequences with those from the other isolates suggested that the amplified gene region was not the target gene. The aligned sequence lengths with gaps and outgroups for each gene are as follows: ITS 911 bp, OCM1 1622 bp, Btub 1008 bp, G002286 1158 bp, G002278 1132 bp, G002263 1022 bp and G001129 1244 bp, for a total of 8097 bp.

Sequencing Ambiguities

When strong double peaks were present at the same location, in all aligned chromatograms for a particular gene and isolate, they were considered to be the result of heterozygous alleles and were left as ambiguous bases in edited sequences. Ambiguous bases were observed in OCM1, G002286, G002278, G001129 and Btub. These ambiguities were almost exclusively present in isolates of var. *ultimum,* except for one isolate of var. *sporangiiferum* in G001129.

Sequences of OCM1 had the highest number of isolates with sequencing ambiguities and the most ambiguities per isolate, with seven isolates having more than five ambiguities and four with more than thirty. Several isolates had notable numbers of ambiguities across multiple genes. These were BR825 and BR827 for three genes (OCM1, G002278 and G002286), BR583 and BR638 for four genes (OCM1, Btub, G002278 and G001129) and BR657 for five genes (OCM1, Btub, G002286, G002278, and G001129). These ambiguities were treated as unknown characters in *BEAST.

Phylogenetic Trees

Figure 1 shows seven individual gene trees plus the species tree generated using *BEAST. Each tree is a 50% majority consensus tree, with a topology representative of relationships observed in at least 50% of the trees in the posterior distribution. Branch lengths of each gene tree are mutations per site, outgroup branch lengths can be found in Table 4 and were trimmed during the formatting of the figures. The *BEAST species tree branch lengths are in demographic time (dmt), which indicate the estimated ancestral population size at the beginning of each branch.

Figure 2 shows three species trees generated in separate programs. The *BEAST species tree is the same as the one illustrated in Figure 1. The BUCKy tree is a primary concordance tree that has support values as concordance factors (CF). Concordance factors are the proportion of genes that actually support the clade and are embedded in the tree file as branch lengths. Therefore, branch lengths do not provide any additional information. The SplitsTree species tree is a Consensus Network with a 0.50 threshold and mean edge weights, these trees are used to visualize conflict between gene trees by creating a network. Support is in confidence values, which is the percentage of trees in which the node is present and branch lengths are proportional to the weight of an associated split, which are analogous to the branch lengths in a phylogenetic tree.

Figure 3 shows the *BEAST species tree in a radial layout. Although the figure is generated from the same tree file as the phylogram illustrated in Figures 1 and 2, the radial layout better visualizes the branch lengths between well-maintained clades. As described in the legend for Table 4, the branch length of the outgroup *P. splendens* in the *BEAST species tree would be four time the length of the scale bar.

The *BEAST analysis showed good convergence within 100 million generations, and a majority of parameters had an effective sample size (ESS) > 200. Based on well supported clades in the *BEAST Species 50% majority consensus tree, the 29 isolates of *P. ultimum* were separated into eight genetic groups and coded with different colours. The five genetic groups within var. *ultimum* were assigned shades of yellow and orange while blue, purple and green were given to the three genetic groups making up var. *sporangiiferum* (Table 1). These colours were used to help visualize changes and trends in phylogenetic topology across the individual gene trees (Figure 1) and between species

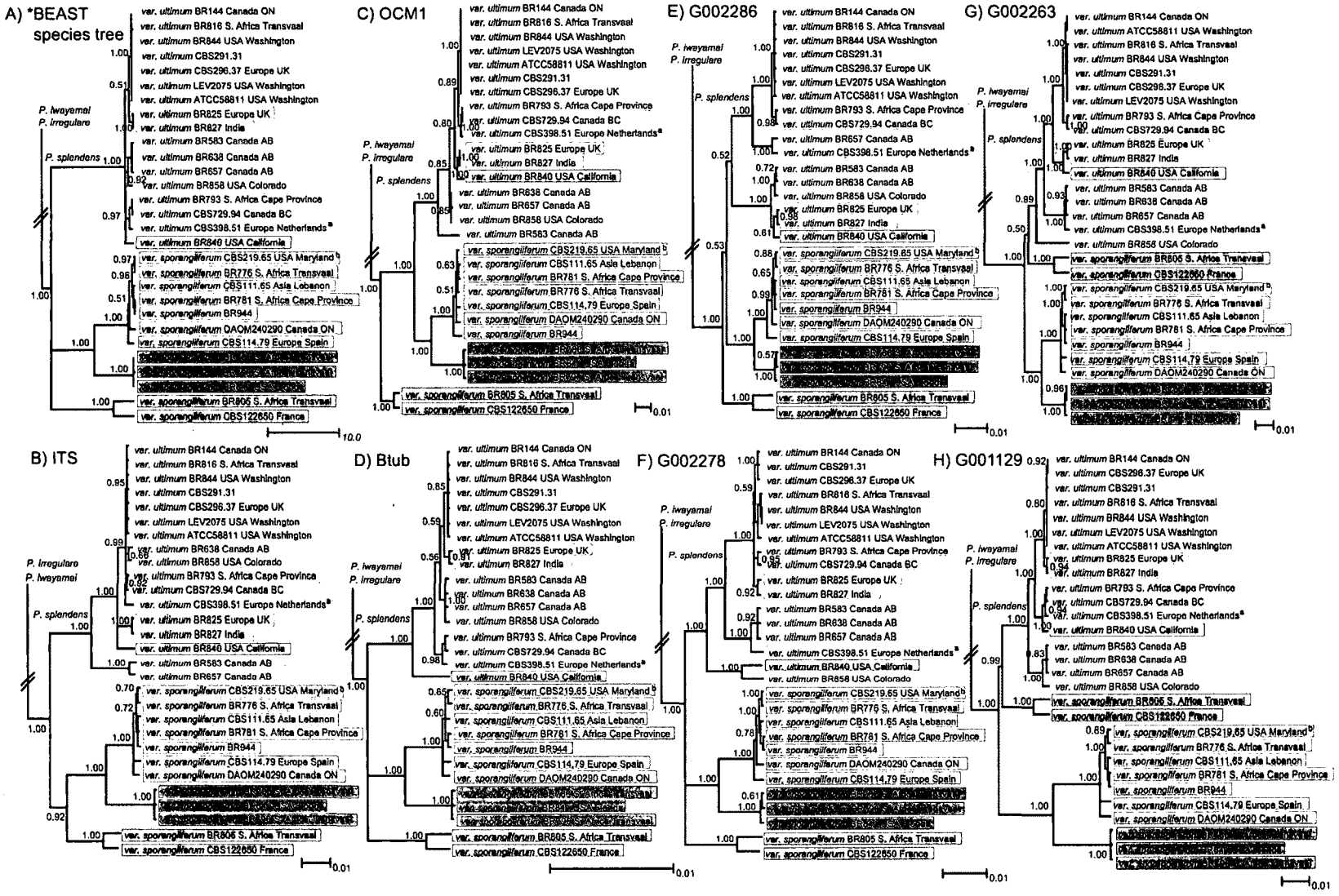

Figure 1: Bayesian 50% majority consensus trees for species **(A)** and individual nuclear gene trees **(B-H)** generated with *BEAST. Node values indicate the posterior probability, [a]neotype of *P. ultimum* var. *ultimum* (Plaats-Niterink, 1981), [b]type strain for *P. ultimum* var. *sporangiiferum* (Drechsler, 1960).

Table 4: The outgroup branch lengths for *BEAST trees. Branch lengths for the gene trees indicate mutations per site, while branch lengths for the species tree are in demographic time.

Outgroup	**BEAST tree*							
	species	***Btub***	***G001129***	***G002263***	***G002278***	***G002286***	***ITS***	***OCM1***
P. splendens	39.8152	0.018528	0.150119	0.141039	0.018528	0.20056	0.080509	0.284974
P. irregulare	126.5022	0.136368	0.465232	0.550084	0.130368	0.453112	0.550634	0.773012
P. iwayamai	126.6814	0.155027	0.465427	0.550368	0.155027	0.453859	0.549778	0.771694

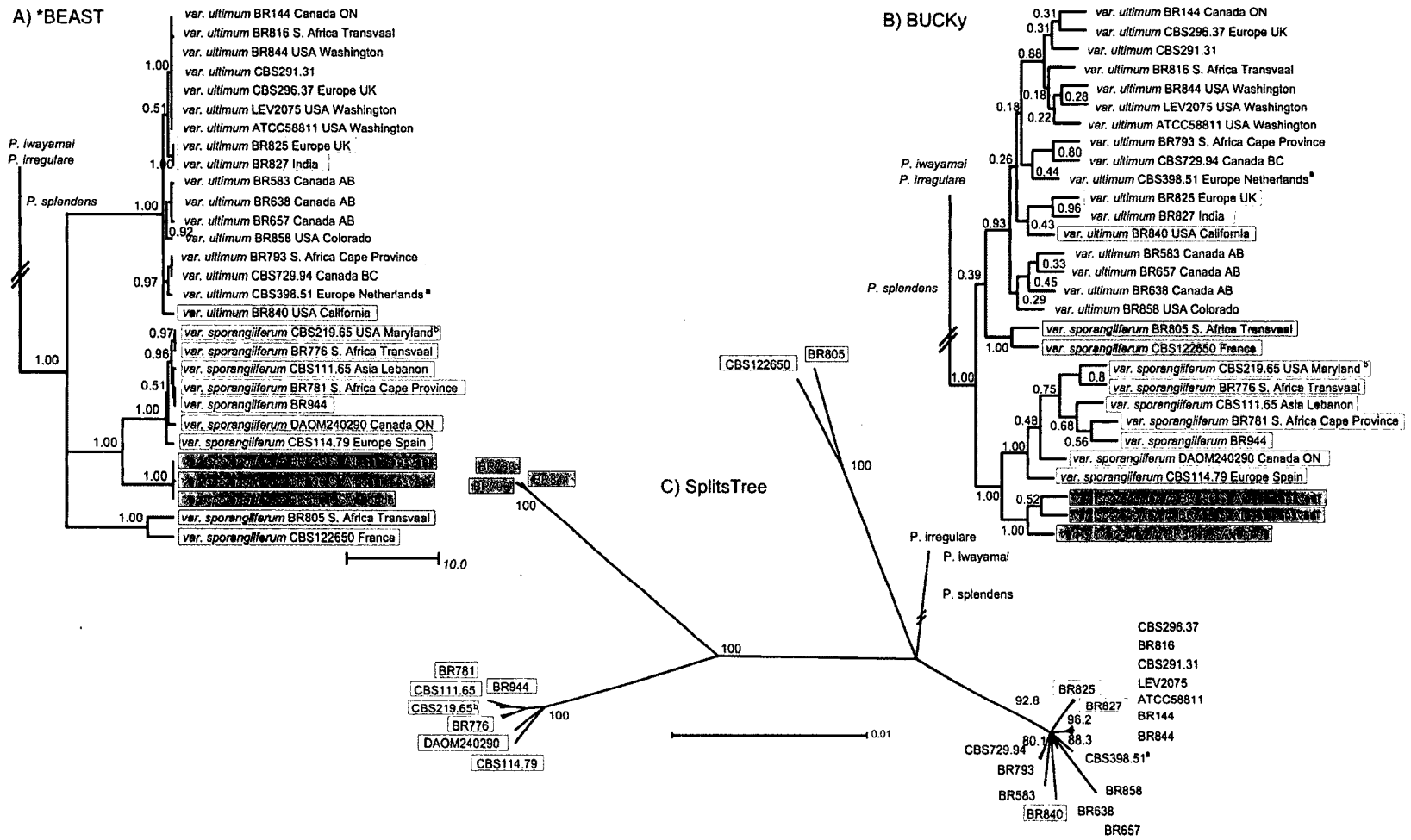

Figure 2: *Pythium ultimum* species trees generated using Bayesian tree files from seven nuclear gene regions and three different programs: **A)** 50% majority consensus tree generated in *BEAST, node values represent posterior probabilities; **B)** primary concordance tree generated in BUCKy, support values are sample wide concordance factors; **C)** consensus network generated in SplitsTree, node supports are confidence values.[a] neotype of *P. ultimum* var. *ultimum* (Plaats-Niterink, 1981),[b] type strain for *P. ultimum* var. *sporangiiferum* (Drechsler, 1960).

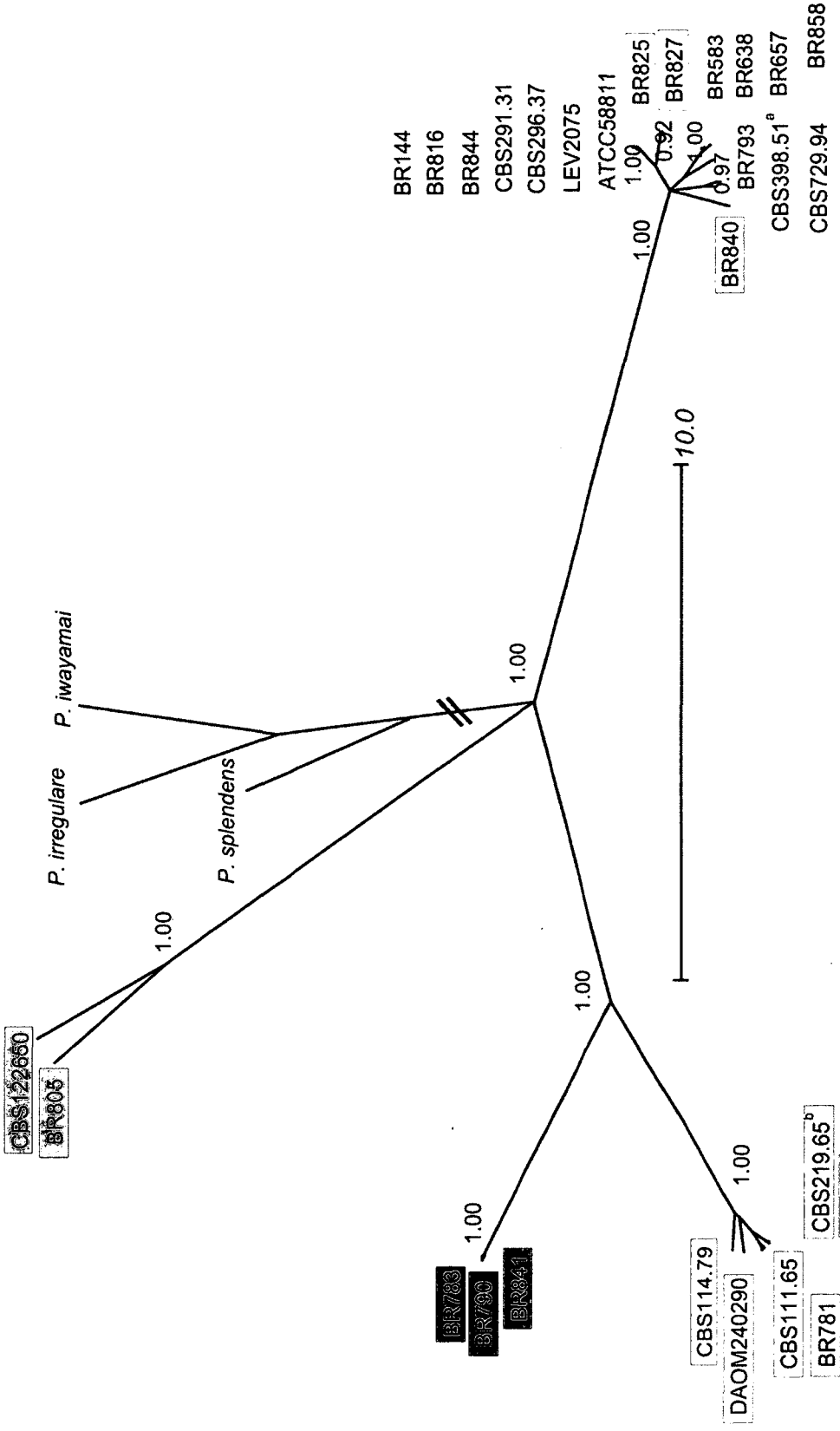

Figure 3: Unrooted radial *BEAST species tree.[a] neotype of *P. ultimum* var. *ultimum* (Plaats-Niterink, 1981),[b] type strain for *P. ultimum* var. *sporangiiferum* (Drechsler, 1960).

trees (Figure 2). Support for the eight clades in each phylogenetic tree is given in Table 5.

The five genetic groups within var. *ultimum* are maintained and distinct in the G001129 and Btub gene trees, with posterior support values of 0.80 and higher (Figure 1), and are present but not always well supported in the BUCKy species tree (Figure 2, Table 5). However, in the other five gene trees the pale orange and pale yellow groups are not maintained (Figure 1). Additionally in G002278 the external node of the bright yellow group, although present, only has a 0.59 posterior support value (Figure 1, Table 5). Only the two isolates of the medium orange group (BR825 and BR827) consistently remain together and along with the single isolate of the dark orange group (BR840) are generally well supported (Figure 1). However, since the dark orange "group" is comprised of a single isolate, it is difficult to determine how distinct it actually is, especially since it shares a common internal node with the medium orange group in three of the seven genes, the pale yellow group in G002278 and the pale orange group in G001129 (Figure 1).

In contrast, the purple, blue and green genetic groups that make up var. *sporangiiferum* are consistently distinct and have very high support values across all gene and species trees - posterior probability of 1.00 in all *BEAST trees, a CF of 1.00 in the BUCKy species tree and a confidence value of 100 in the SplitsTree species tree (Figure 1 and 2).

SplitsTree network consensus trees were examined at different thresholds: 0, 0.1, 0.25, 0.33 and 0.5, the last of which is given in Figure 2. The last two maintained networks are between the green group and the outgroups and within the yellow group. At 0.5, all networks had disappeared.

Table 5: Node support for each genetic group. *BEAST values indicate posterior probability, BUCKy values indicate sample wide concordance factors (CF) and SplitsTree produced confidence values. Blank cells indicate that the group was not maintained in the topology of the tree.

Colour Group	****BEAST***								***BUCKy***	***SplitsTree***
	species	***OCM1***	***ITS***	***Btub***	***G002286***	***G002278***	***G002263***	***G001129***		
bright yellow	1.00	1.00	0.95	0.85	1.00	0.59	1.00	0.80	0.88	88.20
pale orange	0.97	1.00		0.98				0.94	0.44	
pale yellow	0.92			1.00				1.00	0.29	
medium orange	1.00	1.00	1.00	0.91	0.98	0.92	1.00	0.94	0.96	96.20
dark orange	1.00	1.00	1.00	1.00	0.61	1.00	1.00	1.00	0.43	
blue	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
purple	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
green	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

DISCUSSION

Our analysis provides the first multi-gene based phylogeny of the *Pythium ultimum* species complex. Based on the fulfillment of reciprocal monophyly and observed genealogical concordance under the GCPSR, we propose a revision to the current classification of isolates within this species complex and describe four phylogenetically distinct species as follows. The isolates comprising the *P. ultimum* var. *ultimum* clade (yellow and orange hued groups) which also includes the neotype, become simply *Pythium ultimum*. Isolates in the blue genetic group, which contains the *P. ultimum* var. *sporangiiferum* ex-type strain be reclassified as *Pythium sporangiiferum*, and the genetically distinct purple and green groups be temporarily renamed *Pythium* sp. nov. 1 and *Pythium* sp. nov. 2, respectively.

In this study, we started by categorizing isolates by morphological variety based on the topology of ITS1 and COI phylogenies. The isolates putatively classified as var. *ultimum* were divided into five smaller groups based on well supported clades in the *BEAST species tree (Table1; Figure 3). Conflict between these groups in the individual gene trees provided evidence of gene flow within the variety (Table 5), while the larger group remained monophyletic (Figure 1). The conflict in topology among individual isolates within the variety, while maintaining monophyly as a variety, gives evidence that the putatively classified var. *ultimum* is a genetically distinct species.

Similarly, isolates of var. *sporangiiferum* were putatively identified and were divided into three smaller groups based on well supported clades in the *BEAST species tree (Table1; Figure 3). However, conflicts between these smaller groups were never observed and each group remained both distinct and monophyletic while maintaining

high support values across individual gene trees and all three species trees (Figure 1 and 2; Table 5). Therefore, all three var. *sporangiiferum* sub-groups are also proposed as genetically distinct species.

The relative genetic separation between the proposed species and other currently well accepted *Pythium* species, can be seen in the ITS based *Pythium* phylogeny from the 2004 paper by Lévesque and de Cock. This paper shows that the branch lengths (indicating total changes) between the ex-type strains of var. *ultimum* and var. *sporangiiferum* are equal to, or longer than, the branch lengths between some well recognized species pairs, for example: *P. aphanidermatum* / *P. deliense*, *P. apleroticum* / *P. aquatile*, and *P. dissimile* / *P. sulcatum* (Lévesque *et al.*, 2004).

To remove any possible geographic bias in our data, isolates of var. *ultimum* were selected to cover overlapping geographic regions with identified isolates of var. *sporangiiferum* (Table 1). Within the four proposed species, there are no obvious geographic trends and each clade contains isolates from overlapping geographic locations and various hosts. This overlap rules out geographic isolation as a cause for the distinct genetic groups we observed. The locations in which these strains were isolated also emphasises the ubiquitous and global presence of *Pythium ultimum.* Furthermore, because most (if not all) of the samples were collected from cultivated plants and soil, it is understandable that the isolates sort independently of geographical local, since cultivated plants are imported and exported around the world and *P. ultimum* is not regulated for in international trade.

The four genes identified through SNP analysis, G002286, G002278, G002263 and G001129, showed evidence of intra-varietal variation with the capacity to distinguish

between two closely related lineages (Figure 1). Primers were initially intended to be compatible with other *Pythium* species and potentially other oomycete taxa, and were, when possible, designed using alignments including sequences from other closely related species. However, due to their limited success in sister species *P. heterothallicum* and *P. splendens*, their successful application to more distantly related taxa is unlikely. Regardless, the four hyper-variable gene regions selected for use in this study have been located in other taxa within Pythiaceae. The successful use of BLAST to obtain outgroup sequences from the genomic data of additional *Pythium* and *Phytophthora* species, suggests that the gene regions, if not the primers themselves, may have potential for resolving other species complexes in this group.

Genes selected for phylogenetic analysis are assumed to have no recombination within the gene and free recombination between genes (Nichols, 2001). Genes that are situated close together are often collectively inherited, and when used in multi-gene analysis they could bias the species tree towards the evolutionary history of one section of the genome. With this in mind, after G002263, G002278 and G002286 had been sequenced, it was realized that they were located within a 54 kb region in the nuclear genome. However, the discordant topology between the individual gene trees of these loci, an indication of recombination between them, suggest that these genes are in linkage disequilibrium and therefore using them together should not bias our results. The fact that three of the longest, most variable, single copy, homologous gene regions that were annotated in the *Pythium ultimum* genome, are located so close together may indicate the presence of a hyper-variable genomic region, in which high mutation and recombination rates gives some fitness benefit. This theory is supported by the discovery of a 67kb

recombinant-rich region in *Phytophthora sojae,* which contains two avirulence genes. These genes are fundamental to plant-pathogen interactions, and allegedly have an average recombination frequency more than 11 times greater than predicted for the rest of the genome (Whisson *et al.*, 2004).

The internal transcribed spacer region (ITS1 and ITS2) of ribosomal DNA is rapidly evolving and non-coding, and has been a popular genetic marker due to its high sequence variability (Villa *et al.*, 2006). The prevalence of ITS in oomycete phylogenetics is perpetuated by the availability of universal primers, which cover a wide array of taxa, and the abundance of published data. However, it has been argued that ITS may not be an ideal phylogenetic marker, as it is present in multiple copies, is non-coding and, as observed by Robideau *et al.* (2011), it is not always congruent with formally described species. However, it was, in combination with COI, an effective way to putatively identify isolates of rare and potentially cryptic genetic groups within our database. The initial ITS1 phylogeny was able to help increase the sample size of the elusive var. *sporangiiferum,* and helped select isolates that comprehensively represented the genetic diversity of the more prolific var. *ultimum.* For *P. ultimum,* ITS was able to accurately reconstruct the phylogenetic topology of the taxa and would be a suitable marker for taxonomic identification of the four (potentially cryptic) species found in this study.

β-tubulin (Btub) was first used in phylogenetic analyses for *Phytophthora* by Kroon *et al.* (2004) and for *Pythium* and *Phytophthora* by Villa *et al.* (2006), due to its growing utility in phylogenetic studies, especially among fungi. This gene is responsible for coding for one of the two families of tubulins. Tubulins, are the base components of

microtubules, which make up the cytoskeleton, mitotic spindles and flagella of eukaryotic cells. Btub was the least variable gene used in this study, however still gave good taxonomic resolution. This indicates that even genes with relatively low variability are able to resolve the *P. ultimum* species complex. This not only gives credibility to the four proposed species, but may indicate that these four clades have been genetically distinct for some time.

OCM1 is a component protein of the characteristic tripartite tubular hairs found on the flagella of the phylogenetic group stramenopiles (which includes *P. ultimum*) and was first isolated from *Ochromonas danica* (Yamagishi *et al.*, 2007). Even though OCM1 had not yet been published as a phylogenetic marker, unpublished research by G.P. Robideau indicated that the gene contained suitable intra-specific variation for high resolution phylogenetic analysis and relevant primers were already designed. This gene was not initially identified in our study as having any intra-varietal variation since it was not fully sequenced in the genomic scaffold data from CBS 398.51. The estimated intra-varietal variation given in Table 2 was based on our sequence alignment of BR144 and CBS 398.51, and indicates that OCM1 is the second least variable gene in our analysis. However, because of the long gene fragment that was amplified, it appears to have the longest branch lengths of all the gene trees in Figure 1, and interestingly places the green genetic group so that it shares a short branch with the outgroup. This indicates that there are shared informative characters between the green genetic group and *Pythium splendens, P. iwayamai* and *P. irregulare*. Based on the results of our study, OCM1 was shown to be a good phylogenetic marker. The OCM1 primers generated a discrete band

when visualized with gel electrophoresis and amplified a relatively long gene region (1622 bp), which aligned well with the outgroups.

Our method of identifying highly variable genes was to compare the annotated genome of BR144 to the genomic scaffold data of CBS 398.51 and select genes that had high rates of SNPs/kb, an indication of intra-varietal variation. This process was based on the idea that if there was high intra-varietal variation (within one variety) there should be even higher inter-varietal variation (between varieties). However, if both isolates sampled have the same allele for a certain gene, variation would be low or non-existent. Therefore, while this method worked well for this study, it is not infallible. The least variable of our genes identified using this process- G002263 and G001129 (Table 2), were the most variable when compared to our selected outgroup species. To acquire homologous gene sequences for G002263 and G001129 from the genomes of the outgroups, a protein BLAST had to be used, while a nucleotide BLAST was sufficient for the other five genes.

This study undertakes species level analysis at such a high resolution that it borders on population genetics. Trees are presented in Figure 1 as phylograms, however radial diagrams (which are more commonly used in population genetics) were also examined. Radial diagrams provided better visualization of branch lengths making the distance between the closely related clades more noticeable. When viewed in this layout, it becomes apparent that the green group occasionally shares a common node with the outgroup (OCM1 and G002286; Figure 1). This indicates, as mentioned previously, that the green group has shared informative characteristics with the outgroup for some loci.

The two groups with the lowest support in the *BEAST species tree (pale orange and pale yellow) were also the least maintained across the individual gene trees. In the pale orange group, even though BR793 and CBS 729.94 are consistently together, CBS 398.51 appeared elsewhere in the tree topology in three of the seven genes. In the *BEAST species tree the pale orange group had a posterior probability of 0.97, while in the BUCKy species tree it only had a CF of 0.44. The four isolates of the pale yellow group are only maintained in two of the seven gene trees, with no two isolates staying together across all trees. Similarly, the *BEAST support value for the pale yellow group is much higher (posterior probability of 0.92) than the support value in BUCKy (CF of 0.29). The support values in *BEAST represent the probability of the node given the model, data and priors, while the BUCKy tree gives a support value which is a direct indication of what was represented in the provided tree files. Given our data, it would be impossible to say if the *BEAST support values are suspiciously high, especially since the process in which they are generated should be more reflective of the data and the evolutionary process, as predicted by the coalescent model. It is interesting, however, that both programs constructed the same tree topology.

The program SplitsTree analyzes and visualizes evolutionary data in order to indicate how tree-like, or bifurcating, the data set truly is. In SplitsTree, data sets that contain conflicting signal between loci (caused by hybridization events, gene transfer or introgression) generate species trees that appear as a network. When SplitsTree analyzed our data it appeared tree-like under 50% majority consensus conditions. Therefore, our data does not seem to contain conflicting evolutionary signals between loci, or evidence of a hybridization event, gene transfer or introgression in the genes analysed.

Our findings do not support either of the two current circumscriptions of the *P. ultimum* species complex as previously described by Barr *et al.* 1996 or Francis *et al.* 1994. Barr (1996) described the existence of two genetically distinct morphological varieties within the *P. ultimum* species complex, while Francis (1994) argued that the morphological varieties within *P. ultimum* are not genetically distinct and therefore represent a single species.

Eight of the ten genotypes described by Barr *et al.* (1996) were included in this study, three of which were putatively classified through the ITS1 phylogeny as var. *sporangiiferum* (U6, U8 and U9). Of these, Barr suggested that U6 alone made up var. *sporangiiferum,* as all isolates within this group had been observed to produce zoospores. Interestingly, in our study the isolates included in genotype U6 are divided into two clades. Isolate BR805 assorted in the green genetic group, while the rest of the isolates in genotype U6 - CBS 219.65, CBS 114.79 and CBS 111.65 grouped with the two samples of genotype U9 - BR776 and BR781 in the blue genetic group (Figure 3). Therefore, Barr's classification of a single var. *sporangiiferum* clade (based on zoospore production and isozyme variation) was not supported by our results. Additionally, sample BR776 (U9) was initially identified as var. *sporangiiferum* by W.J. Botha, however Barr reclassified the sample as var. *ultimum* when he was unable to induce zoospore production. If the strain was originally identified because of its ability to produce zoospores, and if the failure to produce zoospores by the two isolates of U9 in Barr's experiment was situational and not because they lacked the capacity, then the blue and green genetic groups may represent the true morphological variety that was originally described. However, the blue and purple genetic groups sorted closer together in all seven

genes and the green genetic group sorted independently to either (even sorting closer to the genetic groups of var. *ultimum* in the G002263 and G001129 phylogenies). Because of this, it would be impossible to define a new species based on the production of sporangia with two (green and blue), but not all three, of the genetic var. *sporangiiferum* groups identified in this study. Furthermore, the isolates of the purple genetic group are comprised of three isolates of the non-zoospore producing genotype U8 (BR783, BR790 and BR841) and were putatively identified as var. *sporangiiferum* based on the topology of the ITS1 and COI phylogenies (Table 1). However, this group was always close to, but never combined with, the two other genetic groups within var. *sporangiiferum,* both of which contained isolates that have been observed to produce zoospores (Table1; Figure 1). Furthermore, the inconsistent placement of the green clade in relation to the other isolates in the gene trees, also means that the three var. *sporangiiferum* groups are not monophyletic. Therefore, they cannot be lumped together as a single species under the Phylogenetic Species Concept - which defines a species as the smallest monophyletic group diagnosable by a unique combination of traits (Mishler *et al.*, 2000).

Francis *et al.* (1994) used restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNA (RAPD) and sequence characterized amplified regions (SCARs) to evaluate three morphological varieties within the *P. ultimum* species complex that differed in their sexual strategy: var. *ultimum*, var. *sporangiiferum* and group HS (hyphal swelling). They determined that these varieties were not genetically distinct and therefore should be classified as a single species. The authors incorporated only two isolates of var. *sporangiiferum* into their study, Pu33 (ATCC 58811) and the type strain Pu22 (CBS 219.65 in our study) as determined by their ability to produce

sporangia and zoospores. In their study, Pu33 was classified as var. *sporangiiferum* due its ability to produce zoospores, but genetically clustered with var. *ultimum.* This mismatch of morphological characteristics and genetic assortment was used as a basis to conclude that these two varieties were not genetically distinct. Although the authors were able to show that a third morphological variety, classified by the inability to produce oogonia (group HS), was not genetically distinct through successful outcrossing with var. *ultimum,* they did not attempt similar tests with either of the isolates classified as var. *sporangiiferum*. Furthermore, isolate Pu22 (the ex-type strain of var. *sporangiiferum*) was genetically distinct when compared to the rest of the isolates of *P. ultimum.* In our study, the genetic placement of the two isolates described was confirmed, Pu33 (ATCC 58811) assorted within var. *ultimum* and CBS 219.65 was genetically distinct and assorted within the major var. *sporangiiferum* clade. Although zoospore production was never rigorously tested, empty sporangia were observed on agar-plates of ATCC 58811. Even though our results support the findings of Francis (1994), by including more isolates into our analysis, we were able to show that the *P. ultimum* species complex is not a single species (Figure 1 and 2).

Using sporangia and zoospore production as the sole means of taxonomic classification can be problematic and there is growing support in literature that it is inappropriate for phylogenetic purposes. Lévesque and de Cock (2004) used phenetic analysis to create a molecular phylogeny of the *Pythium* genus, which incorporated 116 species and varieties of *Pythium* - 40 of which were ex-types and 20 of which were newly described species. They stated that most morphological characteristics were not correlated with evolutionary patterns and suggested that morphological classification of

species, within the *Pythium* genus, may not always be practical. The limitation of traditional classification based on morphological traits and growth characteristics is also suggested by Kroon *et al.* (2004), who used combined nuclear and mitochondrial gene phylogenies to revise the classification of *Phytophthora* species. They point out that convergent evolution often results in shared morphological traits, which do not necessarily correlate with relatedness or evolutionary history (Kroon *et al.*, 2004). Additionally, Jeger and Pautasso (2008) stated that flagellated spores have been lost, independently, multiple times in different phylogenetic lineages and had no major phylogenetic significance.

The key morphological difference between the two varieties of *Pythium ultimum* is that var. *sporangiiferum* was characterised by the fact that it "readily" produced sporangia and zoospores at room temperature. Even at the time of the varietal description, isolates of var. *ultimum* had been known, although reluctantly, to produce zoospores from over-ripened oospores under specific conditions (90-day-old maizemeal-agar plates flooded with fresh water and kept at 10°C) (Drechsler, 1952; Plaats-Niterink, 1981). Furthermore, the reliability of expressing this reproductive strategy may be dubious. A strain from a zoospore producing variety may fall back on other structures of its life cycle during times of stress (including its sexual state or growing vegetatively) and its inclination to reproduce asexually may vary over time. This is supported by Jeger and Pautasso (2008) who stated that oomycete plant pathogens show remarkable flexibility in both their life cycles and their ability to adapt to changing environmental circumstances. It has also been noted that inducing sporangia and zoospore production can be difficult when one is inexperienced in the technique, as explained by Abdelzaher *et al.* (1997; as

cited in Gherbawy *et al.*, 2005), and as experienced during the course of this research with the failure to observe and time zoospore release, even in the positive controls (data not shown). Research done by G.P. Robideau, included in Lévesque *et al.* (2010), showed that isolates within var. *ultimum* have all the known genes necessary for zoospore production - when compared to select species within the zoospore producing genus of *Phytophohora* (*Phyt. sojae, Phyt. ramorum* and *Phyt. infestans*). In addition, Robideau was able to demonstrate that most of these genes have been lost in non-zoospore producing *Hyaloperonospora* sp. (downy mildew). This indicates that, at least genetically, var. *ultimum* has the capacity to produce zoospores, as the strain ATCC 58811 (Pu33) appears to indicate, but may not do so readily under laboratory conditions. On the other hand, since the genetically similar var. *sporangiiferum* still produces zoospores, it should not be ruled out that the presence of these genes in var. *ultimum* may be evolutionary baggage and that given sufficient time, these genes may be lost.

In this study we delineate four genetically distinct species that are not well represented by the ability to produce sporangia and zoospores. We are currently unaware of any morphological, pathological, geographical or host preference that would distinguish between these four newly identified species. It has been argued that not all species are worth describing since species are functionally named to give them a handle with which they can be discussed, referred to, or for tracking and recording relevant data (Mishler *et al.*, 2000). Cryptic species especially, which are only identifiable by molecular or chemical means, are criticized for not having a justifiable reason to be described if there is no functional purpose for doing so. Even though our four phylogenetic species are well resolved under the operational methodology of the GCPSR,

we ideally would have an additional reason for splitting a species complex with no descriptive difference to distinguish them. *Pythium ultimum* is a major, global plant pathogen for which there are currently no sanctions against for international trade. Identifying and keeping track of these rarer sister species is important for safeguarding against the possibility of an unforeseen future epidemic. Even though the rarer species are currently less prolific, or at least less sampled, changes in climate (Pautasso *et al.*, 2012), dispersal to a new geographical region, and/or exposure to a new host, could all lead to unexpected outbreaks that can quickly become devastating (Gamliel, 2008). The ability to quickly and accurately identify a new threat allows for faster detection and response if it were to ever become a problem. Quick and accurate detection and identification also allows for more effective quarantine measures to be implemented, allowing more time to formulate a management strategy, while also minimizing the environmental and economic impact by slowing dispersal.

When *Pythium ultimum* was first described, the complete loss of sporangia and zoospore production observed by Trow was thought to be an indication that this species had completely transitioned to living in a terrestrial environment. *Pythium ultimum* was therefore considered the ultimate or final species, in a gradual progression that had been observed in species already described in the genus, to have adapted to terrestrial living by the complete loss of structures designed for an ancestral aquatic environment (Trow, 1901).

Pythium ultimum var. *sporangiiferum* was named to describe a morphological variety within the species that was readily able to produce zoospores and sporangia, a character *P. ultimum* had originally been described to lack, and later modified to say that

it occurred only rarely. Of the very few differences noticed between the two morphological varieties, besides its rarity, var. *sporangiiferum* isolates were also noted to have a slightly higher maximal cardinal temperature (37-40°C vs. 35°C in var. *ultimum*) (Drechsler, 1960; Plaats-Niterink, 1981). Although, prior to the description of the morphological variety, another sporangia producing *P. ultimum* isolate was described on orchids, but in concordance to cooler temperatures (Ark *et al.*, 1949). However, the type strain of var. *sporangiiferum* was first discovered as it was isolated at a time of relatively high temperatures, where the species as a whole is normally isolated in cooler, more temperate seasons (Drechsler, 1960). This preference for warmer temperatures, along with the inherent fondness of *Pythium* species to infect stressed plants (Lévesque *et al.*, 2010), could give isolates of var. *sporangiiferum* an advantage in coming years if climate change continues as predicted.

There are unlikely to be any significant morphological differences between the proposed species groups, although this has yet to be confirmed. However, Barr (1996) noted a few differences in isolates present in the new groups, which may be of interest for taxonomists examining the isolates of the proposed species. Isolate BR805, from the green genetic group, was observed to have the overall highest aplerotic index, or the oopsore : oogonium ratio (Shahzad *et al.*, 1992), of all samples examined (81.7%). However, Barr noted that the oospores were often irregular, which made the calculation unreliable (Barr *et al.*, 1996). Alternatively, isolate BR841 from the purple genetic group had the overall smallest aplerotic index at 56.3% and two (although it wasn't specified which two) of the isolates of U8, which subsequently makes up the purple genetic group in our study, were observed to have "vegetative bodies of exceptional shape [...]

(reniform, sometimes 40 μm long)"(Barr *et al.*, 1996). However, Barr later described the three isolates within genotype U8 as "generally fitted the range of characters for *P. ultimum*" (Barr *et al.*, 1996). These differences, although noteworthy, do not appear to have been exceptional enough to be considered outside the normal range of the existing species description. On a similarly interesting but unexceptional note, there appears to be a preference for citrus by the two novel *Pythium* sp. nov. 1 and nov. 2 (Table 1). However, there are also isolates of var. *ultimum* and var. *sporangiiferum* that were cultured from infected citrus, and with the limited number of samples from the two novel species, it is unclear if this indicates a preference for citrus hosts or simply a result of a sampling bias.

Chapter 2: Examining mitochondrial and nuclear gene tree conflict in species level analysis of *Pythium ultimum*.

INTRODUCTION

Mitochondrial genes (mtDNA) have been used extensively for high resolution species and population studies over the last thirty years. The mitochondrial genome was generally assumed to be maternally inherited, non-recombinant and selectively neutral, with a small effective population size and relatively high mutation rate (Bossu *et al.*, 2009; Galtier *et al.*, 2009). These characteristics were argued to result in more informative characters and faster lineage sorting relative to nuclear genes and therefore were considered more sensitive to evolutionary signal between closely related taxa. Undeniably, when genes were first used as a source of informative characters for the Phylogenetic Species Concept (PSC), the small genome size and multiple plasmid copies of the mitochondrial genome were better suited for the capabilities of the molecular tools of the day.

However, with increasing reports of incongruence between mitochondrial and nuclear species estimates, the use of mtDNA for phylogenetic and taxonomic studies has become a highly contentious issue - especially when used as the solitary means for species delimitation (Bossu *et al.*, 2009; Fort *et al.*, 1984; Galtier *et al.*, 2009; Rubinoff *et al.*, 2006; Rubinoff *et al.*, 2005). Part of the debate focuses on the fact that the mitochondrial genome functionally acts as a single locus, and since it is under different evolutionary pressures than the nuclear genome, it may not be representative of the organism as a whole (Bossu *et al.*, 2009). Furthermore, mtDNA is more sensitive to population bottlenecks, introgression, selective sweeps, and incomplete lineage sorting

because of its smaller population size and uniparental inheritance (Rubinoff *et al.*, 2006). These evolutionary processes can result in conflict in tree topology when compared to nuclear genes. However, horizontal gene transfer, resulting in nuclear pseudogenes of mitochondrial origin, can also result in conflicting species estimates when amplification of unintended targets produce ambiguities in sequencing results (Ballard *et al.*, 2004).

While many of the arguments against the use of mtDNA are the same as those that support it, some studies are starting to question the validity of the fundamental characteristics that these arguments are based on, such as neutrality, uniparental inheritance, recombination and homeoplasmy (Ballard *et al.*, 2004; Rubinoff *et al.*, 2006).

It is argued that the incongruence of mtDNA when compared to nuclear gene trees make it inappropriate for phylogenetic analysis (Ballard *et al.*, 2004; Rubinoff *et al.*, 2006; Shaw, 2002). It is counter-argued that although mtDNA may not reflect nuclear inheritance patterns the mitochondria is a vital part of cell survival and therefore evolutionary selection and inheritance of mitochondrial genes contain nontrivial information (Rubinoff *et al.*, 2005). Some of those involved in this debate point out that this conflict allows for greater knowledge of evolutionary history and can be used to examine some of the processes that would lead to the incongruence, such as introgression and hybridization events (Bossu *et al.*, 2009).

Papers that advocate for mtDNA never go so far as to recommend they be used alone for phylogenetic analysis and acknowledge that they are only informative if the source of incongruence is from 'true' evolutionary processes, such as hybridization and population bottlenecks (Rubinoff *et al.*, 2005). Conflicting trees resulting from

heteroplasmy (mixed plasmid genomes), nuclear pseudogenes and homoplasy (shared nuclear states due to convergent evolution) would be uninformative (Rubinoff *et al.*, 2005).

The argument against the use of mtDNA for phylogenetics has further been fueled by the proposed use of a single and relatively short (~600 bp) region of the mitochondrial barcoding gene cytochrome c oxidase subunit 1 (COI), for species level identification (Ballard *et al.*, 2004; Rubinoff *et al.*, 2006; Rubinoff *et al.*, 2005). For the purposes of this study, the use and appropriateness of COI for barcoding and barcoding itself are not addressed and are not central to the experiment. What is important is that the Barcode of Life initiative has resulted in the generation of a considerable number of COI sequences, making it a significant resource of publicly available and searchable data. Consequently, COI is routinely included in species level analysis because of primer availability and implementation in similar studies.

Whether mtDNA should be used for phylogenetic analysis is hardly clear. The more we learn about mitochondrial genomes, the more we realize that the basic assumptions most of the arguments are based on do not always hold true (Ballard *et al.*, 2004; Galtier *et al.*, 2009; Rubinoff *et al.*, 2006; Rubinoff *et al.*, 2005). It cannot be ignored that mtDNA was the platform on which current molecular analysis was built and that, in general, when combined with several nuclear genes can give high resolution estimates of species boundaries (Kouvelis *et al.*, 2008; Kroon *et al.*, 2004; Rubinoff *et al.*, 2005).

Even if the mitochondrial genome is functionally a single locus, it is still possible that sampling error could result in a gene tree that is not representative of the entire

mitochondrial genome. This possibility is supported by studies that use multiple mitochondrial loci in their analysis and note incongruence of branches between the mitochondrial genes (Kroon *et al.*, 2004).

A major criticism of COI as a barcode is that 600 bp is not long enough for phylogenetic estimates (Min *et al.*, 2007; Rubinoff *et al.*, 2006). By sequencing COI, COII (cytochrome c oxidase subunit 2) and the internal spacer, we can increase the length of the sequence and generate more informative characters for phylogenetic analysis. Min and Hickey (2007) observed that longer sequence lengths increased the statistical support of internal nodes for genus level analysis using COI, and suggested that closely related species were unlikely to be resolved with the standard barcode length of ~600 bp. Rubinoff and Holland (2005) recommended the additional sequencing of multiple mtDNA loci to confirm incongruence and reduce the chance of sampling error. While Kouvelis *et al.* (2008) advocated for the use of multiple mtDNA loci to help resolve taxonomic uncertainty in ITS for true fungi. Additionally, paternal leakage and mitochondrial recombination have been observed in hybrids in distantly related oomycetes, and these processes could result in differences between mitochondrial gene trees (Hoarau *et al.*, 2009).

This study is based on preliminary research on the *Pythium ultimum* species complex in which incongruence was observed between ITS and COI phylogenies. Combining nuclear and mitochondrial genes for phylogenetic analysis of *Phytophthora* and *Pythium* species is not uncommon (Blair *et al.*, 2012; Gómez-Alpizar *et al.*, 2007; Kroon *et al.*, 2004; Spies *et al.*, 2011; Villa *et al.*, 2006) and has been done using COI and ITS (Choi *et al.*, 2006), COI and COII (Martin *et al.*, 2003), just COII (Martin,

2000), just COI (Man In 't veld, 2007) and just ITS (Gherbawy *et al.*, 2005; Lévesque *et al.*, 2004).

Pythium ultimum has a circular mitochondrial genome that is 59 689 bp in length and is mostly comprised of a large inverted repeat (21 950 bp) that is separated by a small (2 711 bp) and large (13 078 bp) non-repeat regions. It contains 35 protein coding genes, 2 rRNA genes and 19 tRNA genes (Lévesque *et al.*, 2010). No nuclear pseudogenes of mitochondrial origin have been found in *P. ultimum* (Lévesque *et al.*, 2010).

This study acknowledges that taxonomists are using mitochondrial genes in multiple gene genealogies to resolve boundaries between closely related species and within genera. The objective of this study is to examine if changing the sampling parameters by sampling different regions of the mitochondrial genome, selecting more variable genes, or by increasing the length of the gene being sequenced, mitochondrial genes that are more capable of resolving intra-specific variation relative to each other will be found. We expect that the general incongruence between mitochondrial and nuclear gene trees will be maintained even with increased sequence length and across multiple mtDNA markers. Although we cannot know for sure what the true species tree actually is, we will assume that consistently supported branches across multiple independent nuclear genes are the most realistic estimate of species boundaries that we have available (as generated in Chapter 1).

MATERIALS AND METHODS

Materials and Methods were as described in Chapter 1, with changes as follows.

Selection of Isolates

To make comparisons between the nuclear and mitochondrial species trees, the same isolates were used for both studies (Chapter 1: Table 1).

SNP Analysis and Gene Selection

The inter-varietal variation (SNP/kb) for every gene in the *P. ultimum* mitochondrial genome was determined by doing a BLAST search using the fully annotated genome data of *P. ultimum* var. *ultimum* BR144 (Lévesque *et al.*, 2010) against the scaffold data of the genome for *P. ultimum* var. *sporangiiferum* CBS 219.65 (Hamilton, unpublished). Candidate genes were selected based on three criteria: that they had relatively high contig SNPs, were at least 1 000 bp in length and that both repeat and non repeat regions of the mitochondrial genome were sampled. The query sequence ID of gene regions with low percent identity were cross-referenced with the *P. ultimum* mitochondrial genome for BR144 (NC 014280.1) on the NCBI website (http://www.ncbi.nlm.nih.gov/). Once candidate genes were identified, corresponding FASTA files were obtained online for BR144, while analogous scaffold regions for CBS 219.65 were consolidated to create full length gene regions to be used for primer design.

Through this process, four highly variable genes over 1 000 bp in length that gave good coverage of the mitochondrial genome were identified: nad7, nad5, cob and nad11 (Table 1). The Cox1, Cox2-1 spacer and Cox2 gene region (COIICOI) were also included in the analysis.

Primer Design

Eighteen primers were designed for the genes nad7, nad5, cob, nad11 and COIICOI. Conserved regions in *P. ultimum* were identified by aligning genomic sequences from BR144, and CBS 219.65 in MegAlign (DNASTAR). Primers were designed in Oligo 7 (Molecular Biology Insights, Inc.) by using the program's primers and probes search function. Search parameters were set to find primers that encompassed a region between 900-1550 bp in length with an approximate melting temperature of 55°C, and that coincided with conserved regions in the *P. ultimum* varietal alignment. The location of the primer on the gene is indicated in the primer name.

Gene Amplification with PCR

Thermocycler profiles for the amplification of the five gene regions consisted of an initial denaturation at 95°C for 3 min, followed by 40 cycles of 95°C for 30s (1 min for COIICOI), annealing for 45s (1 min for COIICOI) at a temperature specific for each primer pair (Table 2), and extension at 72°C for 1 min 30s (2 min 10 sec for COIICOI and 1 min for cob), with a final extension at 72°C for 8 min (10 min for COIICOI).

Sequencing

Thermocycler profiles for sequencing reactions of the five gene regions had an initial denaturation at 95°C for 3 min, followed by 40 cycles at 95°C for 30s, annealing for 20s at 50°C and extension at 60°C for 2 min or 4 min for COIICOI.

Outgroups

Pythium heterothallicum (CBS 450.67 and BR490) and *P. splendens* (CBS 462.48) sequences were generated using the primers and methods described previously.

Table 1: Gene information for the six mitochondrial genes used for phylogenetic analysis of *Pythium ultimum*. Gene product names and locus tags were acquired from the NCBI website [a], while inter-varietal variation and length were obtained through an in-house BLAST search of the annotated genome of BR144 (var. *ultimum*) and the scaffold data of CBS 219.65 (type strain of var. *sporangiiferum*).

Gene	*Product name*	*Inter-varietal Variation (SNP/kb)*	*Length (bp)*	*Locus tag*
COII	cytochrome c oxidase subunit 2	23.3	774	PyulOm_p040
COI	cytochrome c oxidase subunit 1	31.8	1479	PyulOm_p042
nad5	NADH dehydrogenase subunit 5	24.1	1995	PyulOm_p064
nad7	NADH dehydrogenase subunit 7	41.6	1179	PyulOm_p036
nad11	NADH dehydrogenase subunit 11	21.0	2004	PyulOm_p002
cob	apocytochrome b	15.7	1149	PyulOm_p027

[a] http://www.ncbi.nlm.nih.gov

Sequence Editing, Alignment and Phylogenetic Analysis

Eight phylogenetic trees were generated using sequencing data from five mitochondrial genes (nad7, nad5, cob, nad11, and COIICOI) to create five gene trees and three species tree.

Sequencing results were edited using Lasergene SeqMan software (DNAStar) and aligned using the online multiple alignment program MAFFT v. 6 (http://mafft.cbrc.jp/alignment/server/) (Katoh *et al.*, 2008). Aligned sequences were then visually checked in BioEdit Sequence Alignment Editor Version 7.0.9, where the sequences were sorted by title with outgroups moved to the bottom. Alignments were converted from FASTA to nexus files with SeaView version 4.3.2 (Gouy *et al.*, 2010). Phylogenetic analyses were carried out on each gene region using the program *BEAST (Heled *et al.*, 2010). Evolutionary models for each locus were determined using the three substitution schemes in jModelTest 0.1.1 (Posada, 2008) and MEGA5 (Tamura *et al.*, 2011). Aligned sequences for the five mitochondrial loci were uploaded into BEAUTi v1.6.2, where species tree ancestral reconstruction (*BEAST) was selected. A mapping file was imported that specified each isolate as its own species in order to see how the individual isolates were placed in the final species tree. The evolutionary model GTR + invariant sites were used for nad7, nad5 and cob, while HKY + gamma was used for nad11, and HKY + invariant sites + gamma was used for COIICOI. The base frequencies for each locus were set to empirical, the clock model was set to estimated random local clock and ploidy type was set to mitochondrial. Five analyses were run concurrently, the chain length for each analysis was set for 150 000 000 generations, logging every 10 000 generations and recording branch lengths. The graphical and statistical software package

Tracer v1.5 was used to monitor convergence (Drummond *et al.*, 2007). The tree files from the five runs were then combined with a burn-in of 3 750 generations, renumbered, with the suffix changed to .nex.t and scientific notation removed, using LogCombiner 1.6.1. Using Mr. Bayes 3.0, phylogenetic trees were then generated by executing a nexus file containing only the isolate headings. The sumt command was used, specifying no burn-in (burnin = 0), only a single run (nruns=1) and 50% majority consensus tree (contype= halfcompat). SplitsTree and BUCKy species trees were generated as previously described (Chapter 1). Phylogenetic trees were edited using Adobe Illustrator CS2.

Pairwise Distances

Outgroup sequences were removed and distance matrices were generated from each alignment using the uncorrected pairwise ("raw") in R (R Development Core Team, 2011). The bottom half of the matrix was extracted and the distribution was plotted using the package "ggplot2" (R Development Core Team, 2011).

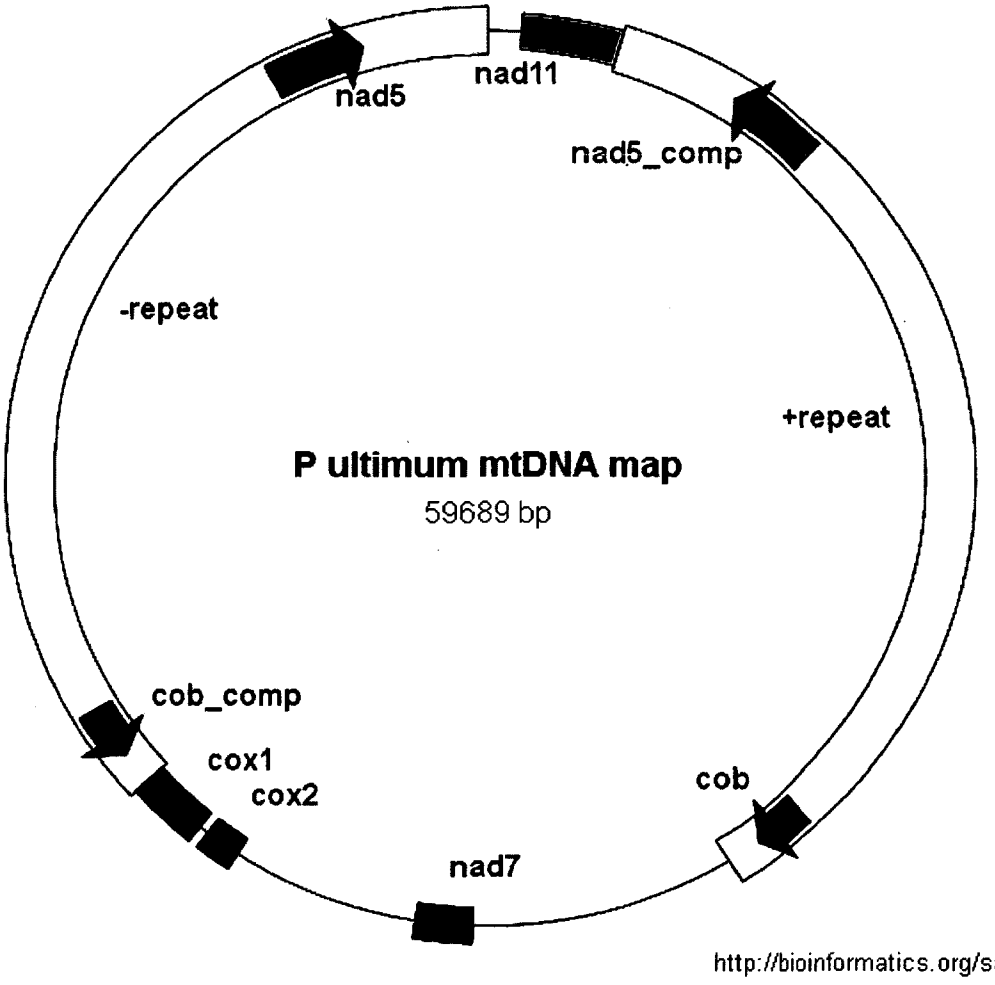

Figure 1: Diagram of the circular mitochondrial genome of *Pythium ultimum*, showing placement and length of the six mitochondrial gene regions used for phylogenetic analysis. Genes were selected to encompass good coverage of the genome and represent both repeat and non-repeat regions. Arrowed boxes in the repeat region indicate the direction that the gene was annotated in the published *P. ultimum* genome. The image was created using the online tool Scalable Vector Graphics and Plasmid Map version 0.1 from Savvy at http://bioinformatics.org/savvy/.

RESULTS

Gene Selection

Four highly variable genes over 1 000 bp in length that gave good coverage of the mitochondrial genome were identified: nad7, nad5, cob and nad11 (Figure 1; Table 1). These were used with the Cox1, Cox2-1 spacer and Cox2 gene region (COIICOI), which was included because of its prevalence in literature and the prolific amount of data available. In the mitochondrial genome, COIICOI and nad7 are located in the large non-repeat region. Nad11 comprises most of the small repeat region, and nad5 and cob are located on opposite flanking ends of the repeat region (Figure 1).

Primer Design

The nad11 gene fragment in var. *sporangiiferum* found during a BLAST search was 584 bp shorter than the annotated full-length gene in the var. *ultimum* genome. However, even though the reverse primer Pul_vu_nad11_R1584 was designed using the none-overlapping region in var. *ultimum,* the primer also worked well for all isolates of var. *sporangiiferum.* This suggests that the 3' end of the scaffold for nad11 in CBS 219.65 was probably missing from the data set or was too fragmented to be located during the BLAST search, but is present *in vivo.*

PCR and Sequencing Primer Performance

The following primer sets were used for all isolates for both amplification and sequencing: Pul_nad11_F192, Pul_vu_nad11_R1584 and Pul_nad11_F292; Pul_nad5_F299 and Pul_nad5_R1830; Pul_nad7_F38 and Pul_nad7_R1072, and lastly Pul_cob_F161 and Pul_cob_R1008. The primers Pul_nad11_R1128, Pul_nad5_F467 and

Pul_nad5_R1372 were only used for isolates that either failed to produce amplicons or produced short or low quality sequencing reads with the initial primers. Four primers were designed but never tested: Pul_nad7_F58, Pul_nad7_R989, Pul_cob_R1008 and Pul_vs_nad11_R1222, these primers were designed as backups in case the preferred primers, which gave longer sequences, failed. All primers are listed in Table 2.

The amplified COIICOI gene region was 2 036 bp in length and required ten primers to sequence the entire region. Four of these primers were used for both PCR amplification and sequencing, while the remaining six were only used for sequencing (Table 2).

Sequencing and Sequence Editing

All genes were successfully sequenced for all 29 isolates of *P. ultimum* as well as the *P. splendens* and *P. heterothallicum* ex-type strains (CBS 462.48 and CBS 45067 respectively) and an additional isolate of *P. heterothallicum* (BR490). The aligned sequence lengths with gaps and outgroups for each gene are as follows: COIICOI 2098 bp, nad7 938 bp, nad5 1448 bp, cob 790 bp, and nad11 1289 bp, with a total length of 6563 bp.

For nad5, the *P. splendens* ex-type strain did not work well with Pul_nad5_F_299. When used in combination with the other reverse primers, the isolate either produced three bands (Pul_nad5_R1830) or did not amplify at all (Pul_nad5_R1372). This resulted in a final sequence length significantly shorter (875 bp long) than the rest of the alignment.

Outgroups

The mitochondrial genomic scaffolds for the other *Pythium* genomes were unavailable when this chapter was being completed. Therefore, only *P. splendens* and *P. heterothallicum* were used as outgroups.

Phylogenetic Trees

The colour groups used in Chapter 1 were transposed on to the individual gene and species trees for the mitochondrial data. This was done in order to visualize changes and consistencies between the phylogenetic topology of the two genomes.

The most striking difference was with the var. *sporangiiferum* genetic groups, with the separation of two isolates from the blue group (CBS 114.79 and CBS 111.65) and subsequent placement with the two isolates from the green group (BR805 and CBS 122650), this change was maintained and well supported across all trees (Figure 2 and Figure 3). The reduced blue group, purple group and now blue/green group were also distinct and well supported across all trees.

A few relatively subtle changes were also observed within the var. *ultimum* samples. The single isolate from the dark orange group, BR840, was embedded within the bright yellow group, while the pale yellow and pale orange groups were split up. Only the medium orange group (BR825 and BR827) remained distinct and well supported across all trees.

It should be noted, that between trees there was very little change in phylogenetic topology. The only major differences were the clade arrangement concerning BR657 in nad5 and the inclusion of the green/blue group with the rest of var. *sporangiiferum* in nad7 (Figure 2).

Table 2: Primers used for PCR and sequencing reactions of mitochondrial genes for isolates of *Pythium ultimum.* Annealing temperatures (Ta) are only provided for PCR primers. Dashes indicate that the primer was designed but never tested for PCR or sequencing.

Gene Region	*Primer name*	*Primer sequence*	*Ta*	*Reference*	*Fragment Length*
COIICOI	PYU_COII_F_36	TTTAGATATGCCTGAACCTTG	52		2153
	PYU1_COI_R_2113	ATACATAATGAAAATGTGCTAC	52		
	FM35	CAGAACCTTGGCAATTAGG	56	Martin, 2000	
	FM55	GGCATACCAGCTAAACCTAA	56	Martin, 2000	
	FM78_Pyt [a]	ACAAATTTCACTACATTGACC		modified from Martin, 2003	
	FM80 [a]	AATATCTTTATGATTTGTTGAAA		Martin *et al.*, 2003	
	FM50_Pyt [a]	GTTTACTGTAGGTCTAGATG		modified from Martin, 2003	
	FM85-mod [a]	RRHWACKTGACTDATRATACCAAA		Robideau *et al.*, 2011	
	Oom-CoxI-lev-lo [a]	CYTCHGGRTGWCCRAAAAACCAAA		Robideau *et al.*, 2011	
	FM52 [a]	GTTGTGCTAATTCCATTCTAA		Martin, 2000	
nad7	Pul_nad7_F38	TTGGTCCACAACATCCAG	55		969
	Pul_nad7_R1072	CTAAAGCTTGTAAATGTGCAAA	55		
	Pul_nad7_F58	GCTCATGGTGTTTTAAGACT	-		
	Pul_nad7_R989	CATATTCACCTTTAGGRGCTT	-		
nad5	Pul_nad5_F299	ATTCTTCTCAATATATGGCACR	55		1460
	Pul_nad5_R1830	TAAAACMAYTTGTCGGCTT	55		
	Pul_nad5_F467	TTACACGTTTACAAGCMAA	55		
	Pul_nad5_R1372	TCTAAAGGAACATCATGTGC	55		
cob	Pul_cob_F161	CACCACATATTGATTTAGCTT	50		830
	Pul_cob_R1034	ACACCTATTTCTGTATAAGGAT	50		

	Pul_cob_R1008	CAGGCATTTGACCTACCCAA	-	
nad11	Pul_nad11_F192	GCCTGTTATGCCWAATATGAAA	55	1326
	Pul_vu_nad11_R1584	AGCATCTTCAGTAAAATGATGT	55	
	Pul_nad11_R1128	ACCTTCYTGACGAATATCACTA	55	
	Pul_nad11_F292	TGTCCTATATGTGATCAAGC	55	
	Pul_vs_nad11_R1222	CAACAGGATATGTAAAATCAAT	-	

[a] primers used only for sequencing

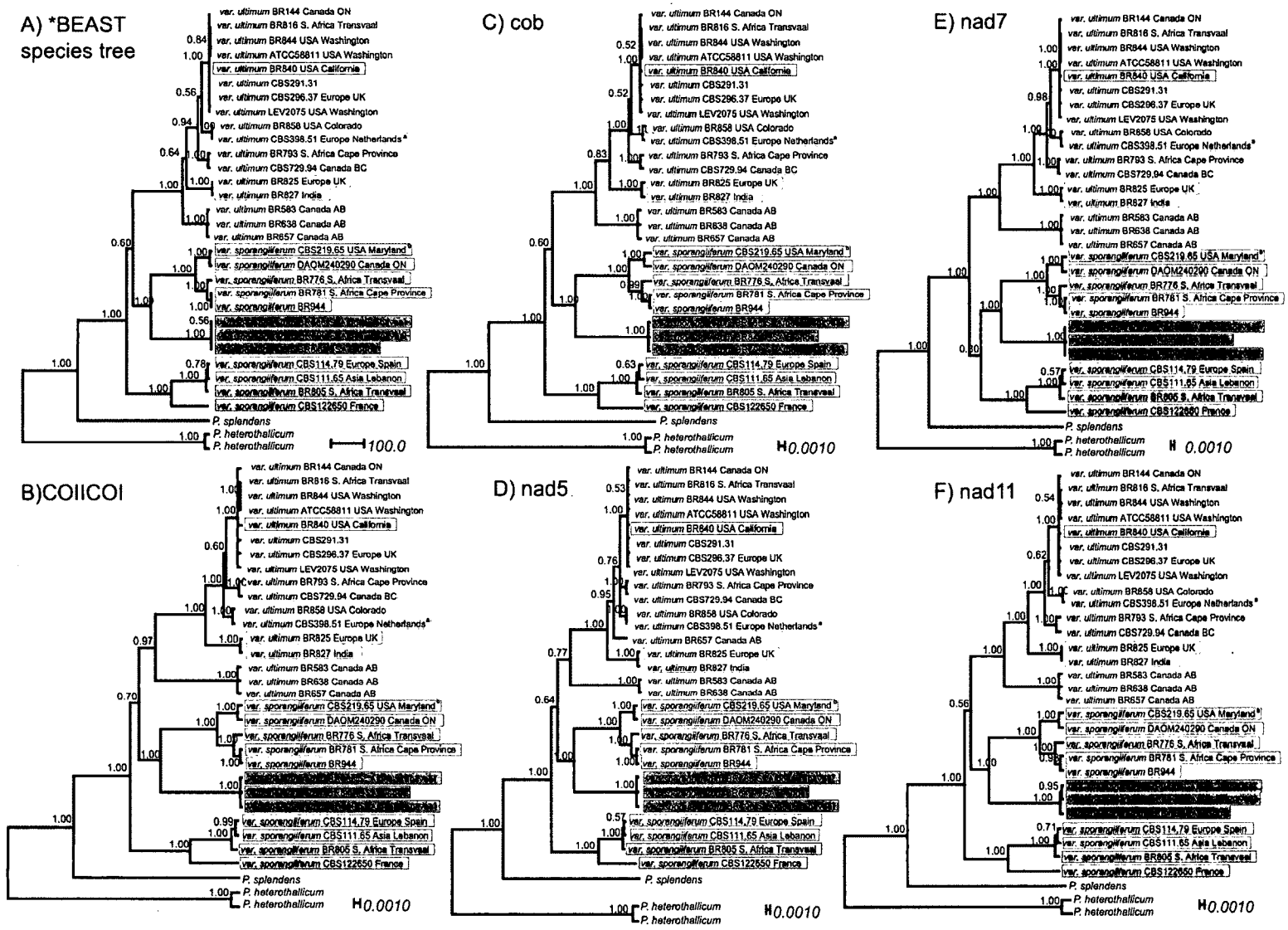

Figure 2: Bayesian 50% majority consensus trees for species **(A)** and individual mitochondrial gene trees **(B-F)** generated with *BEAST. Node values indicate the posterior probability, [a] neotype of *P. ultimum* var. *ultimum* (Plaats-Niterink, 1981), [b] type strain for *P. ultimum* var. *sporangiiferum* (Drechsler, 1960).

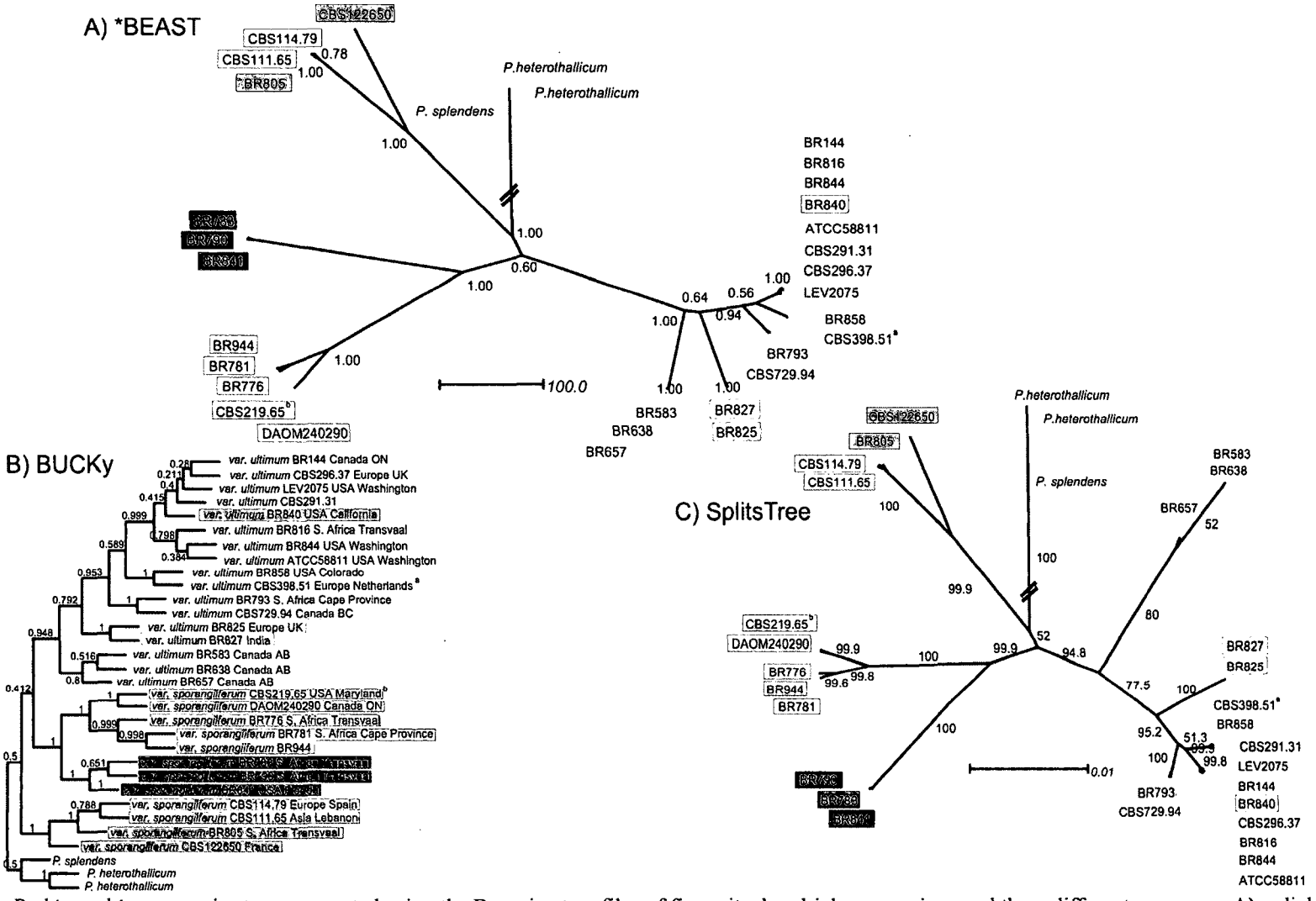

Figure 3: *Pythium ultimum* species trees generated using the Bayesian tree files of five mitochondrial gene regions and three different programs: **A)** radial 50% majority consensus tree generated in *BEAST, node values represent posterior probabilities; **B)** primary concordance tree generated in BUCKy, support values are sample wide concordance factors; **C)** unrooted radial consensus network generated in SplitsTree, node supports are confidence values. [a] neotype of *P. ultimum* var. *ultimum* (Plaats-Niterink, 1981), [b] type strain for *P. ultimum* var. *sporangiiferum* (Drechsler, 1960).

SplitsTree network consensus trees were examined at different thresholds: 0, 0.1, 0.25, 0.33 and 0.5, the last of which is given in Figure 3. The last two maintained networks are between the green group and the outgroups and within the yellow group. At the 0.33 threshold all networks had disappeared.

The phylogenetic gene trees from the mitochondrial genes generated a different topology from those produced from nuclear data. These differences were consistent across all genes sampled and the length and the variability of the genes used did not have an impact on the topology presented.

DISCUSSION

The five mitochondrial genes, regardless of length or variability, generated almost identical phylogenetic trees, all of which conflicted with those obtained using nuclear data (Chapter 1). Since the genes were selected to sample the entire mitochondrial genome and each gene presented an almost identical topology, the assumption that the mitochondrial genome within the *P. ultimum* species complex is non-recombinant, single copy and functionally acts as a single locus is supported. Evidence of heteroplasmy in the mitochondrial genome, which would show up as SNPs in the sequencing files, was not detected. However, it is possible that low level heteroplasmy could go undetected due to the very fact that there are multiple copies of the mitochondrial genome in each cell. The random nature of primer attachment would statistically favour a more abundant genotype, which in turn would be amplified exponentially in the PCR reaction and would drown out the signal of a sparse secondary genotype. These findings are consistent with the paper by Martin (1989) who found maternal inheritance and non-recombinant mtDNA in the progeny of sexual crosses of *Pythium sylvaticum*. Martin used Restriction Fragment

Length Polymorphisms (RFLP) with Southern blots hybridized with ^{32}P-labelled probes in his analysis and estimated that this technique would detect heteroplasmy, if present, in greater than 8% of the mtDNA copies .

All mitochondrial genes tested showed conflict with the nuclear phylogenetic topology, with the inclusion of CBS 114.79 and CBS 111.65 (monophyletic within the blue genetic group for all nuclear genes) in the green genetic group (Figure 2). Maternal inheritance and the absence of recombination can result in mitochondrial genes being inherited differently than nuclear genes, while also being constrained by different evolutionary forces, these processes can result in conflict in estimated phylogenetic topology between mtDNA and nuclear gene trees. Avise (1987) noted that "it is biologically quite plausible that some individuals may truly be more closely related to members of another species than they are to conspecifics, owing solely to particular patterns of maternal lineage survival and extinction accompanying the speciation process". Due to the relatively short branch length separating CBS 114.79 and CBS 111.65 from BR805, it seems unlikely that the presence of similar mitochondrial genomes is a result of maintained ancestral polymorphism. It is more likely indicative that CBS 114.79 and CBS 111.65 are the result of a hybridization event that occurred significantly more recently than the divergence from a shared common ancestor of the four proposed species in Chapter 1, but far enough back that all nuclear signal has been lost. It could also be possible that nuclear genes are present that would be consistent with the mitochondrial gene trees, but they were simply not sampled in our study. However, the maintained introgression of a mitochondrial genome of another species, following a hybridization events late in the speciation process, accompanied by the subsequent loss of

introduced nuclear alleles, has been cited as a potential cause for discordance between mtDNA and nuclear gene trees (Ballard *et al.*, 2004; Bossu *et al.*, 2009; Rubinoff *et al.*, 2006; Shaw, 2002).

Sometimes hybridization between species or subspecies can result in mtDNA leakage in offspring resulting in heteroplasmy, which can lead to recombination in the mitochondrial genome of hybrids (Wilson *et al.*, 2012). Paternal leakage and mitochondrial recombination have been observed in hybrids of distantly related oomycetes (Hoarau *et al.*, 2009). There is no evidence of recombination in our study, which is supported by the absence of networking in the mitochondrial SplitsTree species tree at an edge weight threshold of 0.33, as well as the consistent topology across all gene and species trees (Figure 2 and 3).

In contrast to Chapter 1, the novel primers designed for COIICOI, nad7, nad5, nad11 and cob worked well for both *P. heterothallicum* and *P. splendens* isolates and may work in other closely related species. However, since there are few ambiguities incorporated into the primer sequences and mitochondrial genes tend to be variable at the third codon position, it is unlikely that they could be used for even moderately distant *Pythium* species.

All three species trees in Figure 3 indicate the presence of a short shared branch between the combined green/blue group and the selected outgroup. This was also observed between the green group and outgroup in two of the nuclear genes (OCM1 and G002286) and indicates shared derived characteristics with the outgroup. This could be the result of several things, firstly the green genetic group could have maintained ancestral character states, it may indicate that this genetic group is closer to a common

ancestor, or alternatively, that within the green clade there have been multiple instances of convergent evolution for nucleotide state – although the last one seems unlikely. It is also possible that this node, barely above the 50% threshold, is an artifact of outgroup selection (Figure 3). The more distantly related species included in the outgroup for the nuclear analysis, were not included in this chapter because the genomic mitochondrial data was unavailable. Therefore, only *P. heterothallicum* and *P. splendens* were used, both of which are sister species to *P. ultimum*.

Three strains of var. *ultimum* were selected to be included in this study based on their assortment in the initial COI phylogeny, which was used in Chapter 1 for isolate selection (BR583, BR638 and BR657). These isolates from Alberta, along with an isolate from Colorado (BR858), made up the pale yellow nuclear genetic group. Two of these (BR583 and BR657) are identified as being from *P. ultimum* "group G" (no sexual structure), while BR638 has been observed to produce oogonia (Huang *et al.*, 1992). Previous studies observed that BR583 had highly unusual mitochondrial genes (Huang *et al.*, 1992; Martin *et al.*, 1990 as cited by Barr, 1996 #12). Unsurprisingly, since all mitochondrial gene topologies were consistent, these three isolates clustered together and closer to the central internal node than the rest of the isolates of var. *ultimum* in the mitochondrial gene and species trees (Figure 3). Although they appear distinct, they are still more closely related to var. *ultimum* when compared to any of the other proposed species in Chapter 1, and gene flow observed in the nuclear gene trees indicate that they are not a genetically distinct species.

The reported rapid evolution of mitochondrial gene sequences in animals, which were estimated to be up to10 times faster than single-copy nuclear genes, was one of the

proposed advantages for using mtDNA in systematics and population genetics (Avise *et al.*, 1987). Interestingly, the highly variable mitochondrial genes selected for this study were less variable than the identified nuclear genes, even though they were selected using a similar methodology and criteria. Superficially, variation rates for the mitochondrial genes based on the comparison of two genomes appear higher when comparing Chapter 1: Table 2 and Chapter 2: Table 1. However, the variation rates for nuclear genes were determined by comparing two isolates of var. *ultimum,* while for mitochondrial genes variation rates were estimated based on one isolate of var. *ultimum* and the ex-type strain of var. *sporangiiferum* (this was done as the mitochondrial scaffold data for CBS 398.51 was not available). Using the aligned gene sequences of BR144 and CBS 398.51 as a proxy for the variation of each gene region, the mitochondrial gene in our study with the greatest intra-varietal variation was nad11 with 7.6 SNPs/kb, which was lower than most of our nuclear genes. Also, the variation in selected nuclear genes, as shown by pairwise distances in Figure 4, was generally higher than selected mitochondrial genes. Although unexpected, there are several possible explanations for why variation was lower in mitochondrial genes when compared to nuclear genes. Firstly, the basis for assumed rapid mitochondrial evolution is largely based on animal studies and may not be applicable to the distantly related oomycetes, which are found in a completely different kingdom. This is supported by a study by Vawter (1986) who found an approximately equal rate of evolution between the mitochondrial and nuclear DNA in sea urchins. This finding was in contrast to the accepted relative rates of evolution for vertebrate animals, where mtDNA was reported to diverge 5-10x faster than single-copy nuclear DNA. This difference was

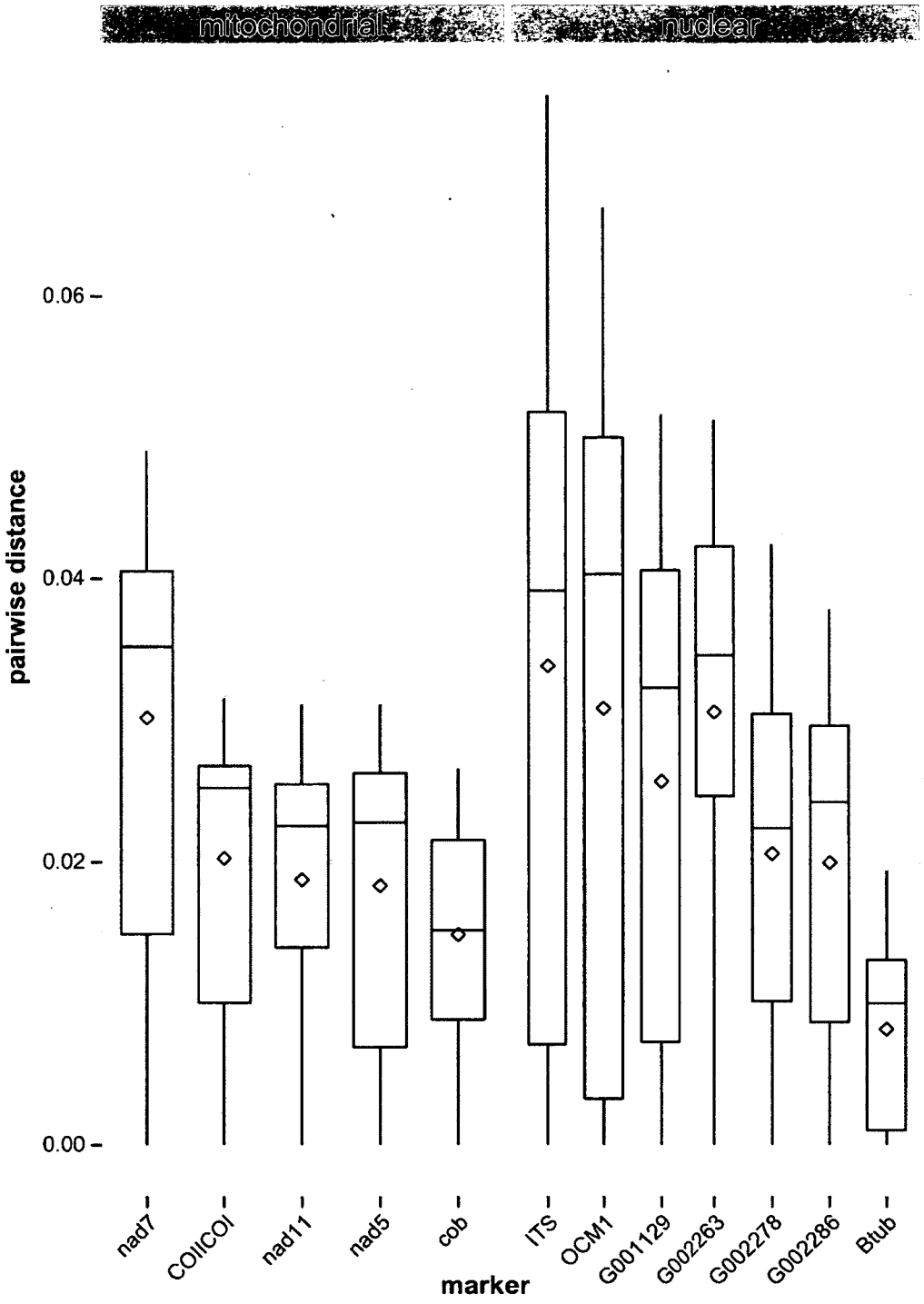

Figure 4: Variability of both nuclear and mitochondrial gene regions as quantified by overall pairwise distance. The diamond symbol indicates the average variability of each gene, the horizontal line within each box is the median, the bottom and top of each box are the 25th and 75th percentiles and the whiskers are 1.5 times the interquartile range above and below the box limits. There were no outliers beyond ± 2.7 standard deviations. Analysis was done in R. No outgroups were included in the analysis.

attributed to an increased rate of evolution in nuclear DNA and demonstrates that even within the animal kingdom higher mitochondrial evolution rates are not always maintained, as may be the case for *P. ultimum*. Secondly, access to genomic data allowed for the selection of the most variable genes within the genome that met our criteria. It may be that overall the mitochondrial sequences are evolving faster, but the longer mitochondrial genes are more conserved than equivalently long single-copy nuclear genes in our species. Thirdly, the selected genes from the mitochondria genome code for key proteins in the electron transport chain and deleterious mutations would have a significant impact on organismal fitness. On the other hand, the selected nuclear genes may be under less selection pressure, or they may be more tolerant of non-silent mutations. Regardless, the selected mitochondrial genes still have enough variability to distinguish between isolates and while having (relatively) lower gene variability by itself would not make mtDNA inappropriate for inclusion in phylogenetic studies, these finding contradict a widely cited reason for including mtDNA in analysis due to its greater gene variability (Avise *et al.*, 1987; Fisher-Reid *et al.*, 2011; Villa *et al.*, 2006).

Even though mitochondrial genes were relatively less variable when compared to the selected hyper-variable nuclear genes, the branches between genetic groups (especially those within var. *ultimum*) had significantly longer internal branchs when compared to the nuclear radial *BEAST and SplitsTree species trees (Chapter 1: Figure 2 and 3; Chapter 2: Figure 3). This is most likely a function of the non-recombinant and uniparental inheritance of mitochondrial genomes, which is supported by Ballard and Rand (2005). They noted that mtDNA became monophyletic faster when compared to a

single nuclear gene and did so far more rapidly when compared to several nuclear genes, which was supported by our data.

The maintained topology between mitochondrial gene trees make it impossible to apply the GCPSR, as only one isolate (BR657 in the nad5 gene phylogeny) was observed to show conflict with the rest of the individual gene trees (Figure 2). If one cannot observe conflict among the different trees, then it is impossible to determine the point of concordance. This renders multiple mitochondrial genes unsuitable for phylogenetic analysis. This is at odds with a paper by Kroon et al. (2004), who analysed the phylogenetic topology of 48 species within the *Phytophthora* genus and after observing discordance between the mitochondrial gene trees, advocated that multiple mitochondrial genes be incorporated with nuclear genes for phylogenetic analysis. In their paper they attribute this to potential incidents of paternal leakage and recombination in species hybrids (Kroon *et al.*, 2004). Their findings emphasise that the results of this species complex study should not be used to generally extrapolate on mitochondrial inheritance and mitochondria-based phylogenies in other closely related genera (such as *Phytophthora*), or potentially even in other *Pythium* species.

Fisher-Reid and Wiens (2011) examined the consequences of combining nuclear and mitochondrial data for phylogenetic analysis in salamanders and 13 other published vertebrate clades. They found widespread discordance between mitochondrial and nuclear gene trees and although mitochondrial genes were more variable, their tree topologies were generally not as well supported. Furthermore, nuclear genes often shared more nodes with the combined-data analysis. However, for the *Plethodon* salamanders (for which they found well supported discordance in the three mitochondrial gene trees)

the mitochondrial topology prevailed in the combined-data analysis, even though it was analysed with nine nuclear genes. This emphasises that including multiple mitochondrial genes (aside from the obvious draw backs of over sampling a single functional locus), if well supported, could significantly bias a species tree, even when combined with three times as many nuclear genes.

Cytochrome c oxidase I and II (COICOII) code for the enzyme that catalyzes the terminal step in the electron transport chain, which transfers an electron from cytochrome c to oxygen (Villa *et al.*, 2006). NADH dehydrogenase 5, 7 and 11 (nad5, nad7 and nad11) code for the enzyme complex which catalyzes the first step in the mitochondrial electron transport chain, which transfers an electron from NADH to coenzyme Q. These five genes along with apocytochrome b (cob) are all protein coding genes involved in the electron transport chain (Griffiths, 1996). Due to the importance of these genes for proper cell functioning, it would be difficult to call these genes neutral, as non-silent mutations could negatively affect energy conversion and organismal fitness (Avise *et al.*, 1987). Also, having all of the sampled genes coding for proteins involved in the electron transport chain could result in some bias. It is interesting that these genes were the longest and most variable within the genome.

An early review article promoting the use of mtDNA for population genetics and systematic studies for higher animals, quoted Wainscoat (1987) as saying "We inherit our mitochondrial DNA from just one of our sixteen great-great grandparents, yet this maternal ancestor has only contributed one-sixteenth of our nuclear DNA.". Therefore, the marker would need to be used in composite with other "molecular genealogies, including those from nuclear genes, each of which in any generation could have been

transmitted through male or female parents" (Avise *et al.*, 1987). They indicate that some of the significant limitations of using mtDNA in phylogenetics were appreciated early and even then they realized the need to include multiple markers of nuclear origin in order to reconstruct an accurate organismal phylogeny (vs. a molecular genealogy).

Even though mtDNA does not reflect nuclear inheritance patterns the mitochondria is a vital part of cell survival and therefore selection and inheritance of mitochondrial genes contain nontrivial information (Rubinoff *et al.*, 2005). Some argue that the conflict between mtDNA and nuclear genes allows for greater knowledge of the evolutionary history of the organism and can be used to examine some of the processes that would lead to the incongruence such as introgression and hybridization events (Bossu *et al.*, 2009). Papers that advocate for mtDNA never go so far as to recommend they be used alone for phylogenetic analysis and acknowledge that they are only informative if the source of incongruence is from a 'true' evolutionary processes - such as hybridization and population bottlenecks, while conflicting trees resulting from heteroplasmy (mixed plasmid genomes), nuclear pseudogenes (presence of mitochondrial genes in the nuclear genome) and homoplasy (shared nucleotide states due to convergent evolution) would be uninformative (Rubinoff *et al.*, 2005).

Clearly there are some disadvantages to using multiple mtDNA markers for phylogenetic analysis. Based on the findings of this study, using multiple mitochondrial markers for multi-gene analysis to determining species boundaries is not recommended, and is advised against. We demonstrate that a single mitochondrial marker with moderate variability can reconstruct the phylogenetic topology as predicted by the entire organelle. If one mitochondrial gene was to be included COI is a good candidate due to the large

amount of publicly available data, its use in species identification (Robideau *et al.*, 2011), and because it was recently nominated to be the standardized tool for molecular taxonomic identification (Galtier *et al.*, 2009). However, a major criticism of COI is that the barcode length of 600 bp is not long enough for phylogenetic estimates (Min *et al.*, 2007; Rubinoff *et al.*, 2006). To address this issue we instead propose the inclusion of the Cox1, Cox2-1 spacer and Cox2 gene region (COIICOI) - as suggested by Min and Hickey (2007), since it has a significantly longer sequence length (2 153 bp) and therefore has more informative characters for phylogenetic analysis. This also reduces the chance of observing homoplasy or converged nucleotide states in the third codon position, which can result in low bootstrap values when doing genus level analysis (Schroeder *et al.*, 2012).

THESIS CONCLUSIONS

Our analysis provides the first multi-gene based phylogeny of the *Pythium ultimum* species complex. Based on the fulfillment of reciprocal monophyly and observed genealogical concordance under the GCPSR, we proposed four phylogenetically distinct species: *Pythium ultimum*, *Pythium sporangiiferum*, *Pythium* sp. nov. 1 and *Pythium* sp. nov. 2.

This study also included the first mitochondrial multi-gene phylogeny for the *P. ultimum* species complex. Based on the observed lack of conflict between the tree topology of the five sampled mitochondrial gene regions and the absence of sequencing ambiguities, the mitochondrial genome of the *P. ultimum* species complex is believed to be single copy, non-recombinant and uniparentally inherited. Based on our findings, we also concluded that using multiple mitochondrial genes for phylogenetic analysis is inappropriate as the mitochondrial genome functionally acts as a single locus, which could severely bias species estimates.

APPENDIX A

Table1: Primers used for each isolate. Grey cells indicate isolates that were not used in final data sets.

	G002286					*G002278*				*G002263*						*G001129*				*Btub*					*OCM1*			
Sample ID #	*PYU1_G002286_F_245*	*PYU1_G002286_R_1473_pul*	*PYU1_G002286_F_207*	*Pyt_G002286_R_1451*	*Pyt_G002286_R_1552*	*Oom_G002278_F_442*	*Oom_G002278_R_1352*	*PYU1_G002278_F_144*	*PYU1_G002278_R_994*[P]	*PYU1_G002263_F_1007*[S]	*PYU1_G002263_R_2029*[S]	*PYUvS_G002263_F_986*	*PYU_G002263_R_2095*	*PYU1_G002263_F_1028*	*PYU1_G002263_R_2034*	*PYU1_G001129_F_107*	*PYU1_G001129_R_1386*	*PYU1_G001129_F_391*	*PYU1_G001129_R_1445*	*Oom-Btub-Up415*	*Oom-Btub-Lo1401*[P]	*Oom-Btub-lo1402*[S]	*OomPy-BTub-Up901*[S]	*OomPy-BTub-Lo954*[S]	*OCM1_up130*	*OCM1_lo1842*	*OCM1_up631*[S]	*OCM1_lo643*[S]
Pythium ultimum																												
BR144	x	x	x	x	x	x	x	x	~	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
BR816	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR844	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
CBS291.31	x	x				x	x	x	~					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
CBS296.37	x	x				x	x	x	~					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
LEV2075	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
ATCC58811	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
CBS398.51	x	x	x	x		x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
CBS729.94	x	x						x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR793	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR638	x	x				x	x	x	~					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
BR858	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR583	x	x				x	x	x	~					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
BR657	o	o	x	x	x	x	x	x	~					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
BR825	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR827	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR840	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
Pythium sporangiiferum																												
CBS219.65	x	x				x	x	x	~			x	x	o	o	x	x	x	x	x	x	x	x	x	x	x	x	x
CBS114.79	x	x				x	x	x	~			x	x	o	o	x	x	x	x	x	x	x	x	x	x	x	x	x
CBS111.65	x	x				x	x	x	~			x	x	o	o	x	x	x	x	x	x	x	x	x	x	x	x	x
BR776	x	x				x	x	x	~			x	x	o	o	x	x	x	x	x	x	x	x	x	x	x	x	x
BR781	x	x				x	x	x	~			x	x	o	o	x	x	x	x	x	x	x	x	x	x	x	x	x
BR944	x	x				x	x	x	~			x	x	o	o			x	x	x	x	x	x	x	~	~	x	x
DAOM240290	x	x	x	x	x	x	x	x	~			x	x	o	o			x	x	x	x	x	x	x	x	x	x	x
***Pythium* sp. nov. 1**																												
BR783	x	x				x	x	x	~			x	x	o	o			x	x	x	x	x	x	x	x	x	x	x
BR841	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
BR790	x	x				x	x	x	~			x	x	o	o			x	x	x	x	x	x	x	x	x	x	x
***Pythium* sp. nov. 2**																												
BR805	x	x				x	x	x	~					x	x			x	x	x	x	x	x	x	x	x	x	x
CBS122650	x	x				x	x	x	~			x	x	x	x			x	x	x	x	x	x	x	x	x	x	x
Outgroups																												
CBS462.48			x	x	x	x	x			x			x			x	x	x	x	x	x	x	x	x	x	x	x	x
BR490			x	o	x	x	x											^	^	x	x	x	x	x	x	x	x	x
CBS45067			x		x	o	o			x	x	x	x					^	^	x	x	x	x	x	x	x	x	x

x successful amplification and sequencing

~ successful amplification but failed to sequence

o failed to amplify gene region

^ successful amplification and sequencing, does not appear to be the correct gene

[S] primer used only for sequencing

[P] pcr primer used only for PCR

REFERENCES

Abdelzaher, HMA, Elnaghy, MA, Fadl-Allah, EM, and Zohri, SS. 1997. Some physical and chemical factors affecting asexual reproduction of three *Pythium* spp. Cryptogamie: Mycol. **18**:267-277.

Abdelzaher, HMA, Morikawa, T, Ichitani, T, and Elnaghy, MA. 1995. Classification of *Pythium* 'group F' based on mycelial protein and isozyme patterns. Mycoscience **36**:45-49.

Altschul, SF, Gish, W, Miller, W, Myers, EW, and Lipman, DJ. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403-410.

Archibald, JK, Mort, ME, and Crawford, DJ. 2003. Bayesian inference of phylogeny: A non-technical primer. Taxon **52**:187-191.

Ark, PA and Middleton, JT. 1949. Pythium Black Rot of *Cattleya*. Phytopathology **39**:1060-1064.

Avise, JC, Arnold, J, Ball, RM, Bermingham, E, Lamb, T, Neigel, JE, Reeb, CA, and Saunders, NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. **18**:489-522.

Bakkeren, G, Kronstad, JW, and Lévesque, CA. 2000. Comparison of AFLP fingerprints and ITS sequences as phylogenetic markers in Ustilaginomycetes. Mycologia **92**:510-521.

Ballard, JWO and Rand, DM. 2005. The population biology of mitochondrial DNA and its phylogenetic implications. Annu. Rev. Ecol., Evol. Syst. **36**:621-642.

Ballard, JWO and Whitlock, MC. 2004. The incomplete natural history of mitochondria. Mol. Ecol. **13**:729-744.

Barr, DJS, Warwick, SI, and Désaulniers, NL. 1996. Isozyme variation, morphology, and growth response to temperature in *Pythium ultimum*. Can. J. Bot. **74**:753-761.

Bilodeau, GJ, Lévesque, CA, De Cock, AWAM, Brière, SC, and Hamelin, RC. 2007. Differentiation of European and North American genotypes of *Phytophthora ramorum* by real-time polymerase chain reaction primer extension. Can. J. Plant Pathol. **29**:408-420.

Blair, JE, Coffey, MD, and Martin, FN. 2012. Species Tree Estimation for the Late Blight Pathogen, *Phytophthora infestans*, and Close Relatives. PloS one **7**:e37003.

Bossu, CM and Near, TJ. 2009. Gene trees reveal repeated instances of mitochondrial DNA Introgression in orangethroat darters (percidae: etheostoma). Syst. Biol. **58**:114-129.

Choi, Y-J, Hong, S-B, and Shin, H-D. 2006. Genetic diversity within the *Albugo candida* complex (Peronosporales, Oomycota) inferred from phylogenetic analysis of ITS rDNA and COX2 mtDNA sequences. Mol. Phylogenet. Evol. **40**:400-409.

Drechsler, C. 1952. Production of zoospores from germinating oospores of *Pythium ultimum* and *Pythium debaryanum*. B. Torrey Bot. Club **79**:431-450.

Drechsler, C. 1960. Two root-rot fungi closely related to *Pythium ultimum*. Sydowia **14**:106-114.

Drummond, AJ and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. **7**:art. no. 214.

Edgar, RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792-1797.

Fisher-Reid, MC and Wiens, JJ. 2011. What are the consequences of combining nuclear and mitochondrial data for phylogenetic analysis? Lessons from Plethodon salamanders and 13 other vertebrate clades. BMC Evol. Biol. **11**:art. no. 300.

Fort, P, Bonhomme, F, Darlu, P, Piechaczyk, M, Jeanteur, P, and Thaler, L. 1984. Clonal divergence of mitochondrial DNA versus populational evolution of nuclear genome. Evol. Theory **7**:81-90.

Francis, DM, Gehlen, MF, and St Clair, DA. 1994. Genetic variation in homothallic and hyphal swelling isolates of *Pythium ultimum* var. *ultimum* and *P. ultimum* var. *sporangiferum*. Mol. Plant-Microbe Interact. **7**:766-775.

Galtier, N, Nabholz, B, GlÉmin, S, and Hurst, GDD. 2009. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. Mol. Ecol. **18**:4541-4550.

Gamliel, A. 2008. High Consequence Plant Pathogens. Pages 25-36 *in* M. L. Gullino, J. Fletcher, A. Gamliel, and J. P. Stack, editors. Crop Biosecurity. Springer Netherlands.

Gherbawy, YAMH, Abdelzaher, HMA, Meens, J, and El-Hariry, H. 2005. Morphological and molecular identification of some closely related *Pythium* species in Egypt. Arch. Phytopathol. Plant Protect. **38**:193-208.

Gómez-Alpizar, L, Carbone, I, and Ristaino, JB. 2007. An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. Proc. Natl. Acad. Sci. U. S. A. **104**:3306-3311.

Gouy, M, Guindon, S, and Gascuel, O. 2010. Sea view version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. **27**:221-224.

Griffiths, AJF. 1996. Mitochondrial inheritance in filamentous fungi. J. Genet. **75**:403-414.

Heled, J and Drummond, AJ. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. **27**:570-580.

Henk, DA, Eagle, CE, Brown, K, Van Den Berg, MA, Dyer, PS, Peterson, SW, and Fisher, MC. 2011. Speciation despite globally overlapping distributions in *Penicillium chrysogenum*: The population genetics of Alexander Fleming's lucky fungus. Mol. Ecol. **20**:4288-4301.

Hoarau, G, Coyer, JA, and Olsen, JL. 2009. Paternal leakage of mitochondrial DNA in a *Fucus (phaeophyceae)* hybrid zone. J. Phycol. **45**:621-624.

Huang, HC, Morrison, RJ, Muendel, H-H, Barr, DJS, Klassen, GR, and Bochko, J. 1992. *Pythium* sp. "group G", a form of *Pythium ultimum* causing damping-off of safflower. Can. J. Plant Pathol. **14**:229-232.

Huson, DH and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23**:254-267.

Jeger, MJ and Pautasso, M. 2008. Comparative epidemiology of zoosporic plant pathogens. Eur. J. Plant Pathol. **122**:111-126.

Katoh, K and Toh, H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. **9**:286-298.

Kouvelis, VN, Sialakouma, A, and Typas, MA. 2008. Mitochondrial gene sequences alone or combined with ITS region sequences provide firm molecular criteria for the classification of *Lecanicillium* species. Mycol. Res. **112**:829-844.

Kroon, LPNM, Bakker, FT, van den Bosch, GBM, Bonants, PJM, and Flier, WG. 2004. Phylogenetic analysis of *Phytophthora* species based on mitochondrial and nuclear DNA sequences. Fungal Genet. Biol. **41**:766-782.

Larget, BR, Kotha, SK, Dewey, CN, and Ané, C. 2010. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. Bioinformatics **26**:2910-2911.

Leavitt, SD, Johnson, LA, Goward, T, and St. Clair, LL. 2011. Species delimitation in taxonomically difficult lichen-forming fungi: An example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. Mol. Phylogenet. Evol. **60**:317-332.

Lemey, P, Salemi, M, and Vandamme, AM, editors. 2009. The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. 2 edition. Cambridge University Press, Cambridge, UK.

Lévesque, CA, Brouwer, H, Cano, L, Hamilton, JP, Holt, C, Huitema, E, Raffaele, S, Robideau, GP, Thines, M, Win, J, Zerillo, MM, Beakes, GW, Boore, JL, Busam, D, Dumas, B, Ferriera, S, Fuerstenberg, SI, Gachon, CMM, Gaulin, E, Govers, F, Grenville-Briggs, L, Horner, N, Hostetler, J, Jiang, RHY, Johnson, J, Krajaejun, T, Lin, H, Meijer, HJG, Moore, B, Morris, P, Phuntmart, V, Puiu, D, Shetty, J, Stajich, JE, Tripathy, S, Wawra, S, West, Pv, Whitty, BR, Coutinho, PM, Henrissat, B, Martin, F, Thomas, PD, Tyler, BM, De Vries, RP, Kamoun, S, Yandell, M, Tisserat, N, and Buell, CR. 2010. Genome sequence of the necrotrophic plant pathogen, *Pythium ultimum*, reveals original pathogenicity mechanisms and effector repertoire. Genome Biol. **11**:art. no. R73.

Lévesque, CA and de Cock, AWAM. 2004. Molecular phylogeny and taxonomy of the genus *Pythium*. Mycol. Res. **108**:1363-1383.

Man In 't veld, WA. 2007. Gene flow analysis demonstrates that *Phytophthora fragariae* var. *rubi* constitutes a distinct species, *Phytophthora rubi* comb. nov. Mycologia **99**:222-226.

Martin, FN. 1989. Maternal inheritance of mitochondrial DNA in sexual crosses of *Pythium sylvaticum*. Curr. Genet. **16**:373-374.

Martin, FN. 1995. Electrophoretic karyotype polymorphisms in the genus *Pythium*. Mycologia **87**:333-353.

Martin, FN. 2000. Phylogenetic relationships among some *Pythium* species inferred from sequence analysis of the mitochondrially encoded cytochrome oxidase II gene. Mycologia **92**:711-727.

Martin, FN and Kistler, HC. 1990. Species specific banding patterns of restriction endonuclease digested mitochondrial DNA in the genus *Pythium*. Exp. Mycol. **14**:32-46.

Martin, FN and Tooley, PW. 2003. Phylogenetic relationships among *Phytophthora* species inferred from sequence analysis of mitochondrially encoded cytochrome oxidase I and II genes. Mycologia **95**:269-284.

Min, XJ and Hickey, DA. 2007. Assessing the effect of varying sequence length on DNA barcoding of fungi: Barcoding. Mol. Ecol. Notes **7**:365-373.

Mishler, BD and Theriot, EC. 2000. The Phylogenetic Species Concept (sensu Mishler and Theriot): Monophyly, Apomorphy, and Phylogenetic Species Concept. Pages 44-54 *in* Q. D. Wheeler and R. Meier, editors. Species Concepts and Phylogenetic Theory: A Debate. Columbia University Press, New York.

Moller, EM, Bahnweg, R, Sandermann, H, and 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 Res. **20**:6115-6116.

Moralejo, E, Clemente, A, Descals, E, Belbahri, L, Calmin, G, Lefort, F, Spies, CFJ, and McLeod, A. 2008. *Pythium recalcitrans* sp. nov. revealed by multigene phylogenetic analysis. Mycologia **100**:310-319.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. **16**:358-364.

Pautasso, M, Döring, TF, Garbelotto, M, Pellis, L, and Jeger, MJ. 2012. Impacts of climate change on plant diseases-opinions and trends. Eur. J. Plant Pathol. **133**:295-313.

Plaats-Niterink, AJ. 1981. Monograph of the genus *Pythium*. Stud. Mycol. **21**:1-242.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. **25**:1253-1256.

R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rintoul, TL, Eggertson, QA, and Lévesque, CA. 2012. Multigene Phylogenetic Analyses to Delimit New Species in Fungal Plant Pathogens. Pages 549-569 *in* M. D. Bolton and B. P. Thomma, editors. Plant Fungal Pathogens: Methods and Protocols. Humana Press.

Robideau, GP, De Cock, AWAM, Coffey, MD, Voglmayr, H, Brouwer, H, Bala, K, Chitty, DW, Désaulniers, N, Eggertson, QA, Gachon, CMM, Hu, CH, Küpper, FC, Rintoul, TL, Sarhan, E, Verstappen, ECP, Zhang, Y, Bonants, PJM, Ristaino, JB, and André Lévesque, C. 2011. DNA barcoding of oomycetes with cytochrome c oxidase subunit I and internal transcribed spacer. Mol. Ecol. Resour. **11**:1002-1011.

Ronquist, F and Huelsenbeck, JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572-1574.

Rubinoff, D, Cameron, S, and Will, K. 2006. A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. J. Hered. **97**:581-594.

Rubinoff, D and Holland, BS. 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. **54**:952-961.

Sarver, BAJ, Ward, TJ, Gale, LR, Broz, K, Corby Kistler, H, Aoki, T, Nicholson, P, Carter, J, and O'Donnell, K. 2011. Novel *Fusarium* head blight pathogens from Nepal and Louisiana revealed by multilocus genealogical concordance. Fungal Genet. Biol. **48**:1096-1107.

Schroeder, KL, Martin, FN, de Cock, AWAM, Lévesque, CA, Spies, CFJ, Okubara, PA, and Paulitz, TC. 2012. Molecular detection and quantification of *Pythium* species - evolving taxonomy, new tools and challenges. Plant Dis.:(in press).

Shahzad, S, Coe, R, and Dick, MW. 1992. Biometry of oospores and oogonia of *Pythium* (Oomycetes): the independent taxonomic value of calculated ratios. Bot. J. Linn. Soc. **108**:143-165.

Shaw, KL. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. U. S. A. **99**:16122-16127.

Spies, CFJ, Mazzola, M, Botha, WJ, Langenhoven, SD, Mostert, L, and McLeod, A. 2011. Molecular analyses of *Pythium irregulare* isolates from grapevines in South Africa suggest a single variable species. Fungal Biol. **115**:1210-1224.

Tamura, K, Peterson, D, Peterson, N, Stecher, G, Nei, M, and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731-2739.

Taylor, JW, Jacobson, DJ, Kroken, S, Kasuga, T, Geiser, DM, Hibbett, DS, and Fisher, MC. 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genet. Biol. **31**:21-32.

Townsend, TM, Alegre, RE, Kelley, ST, Wiens, JJ, and Reeder, TW. 2008. Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: An example from squamate reptiles. Mol. Phylogenet. Evol. **47**:129-142.

Trow, AH. 1901. Observations on the Biology and Cytology of *Pythium ultimum*, n. sp. Ann. Bot. **15**:269-312.

Vawter, L and Brown, WM. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science **234**:194-196.

Villa, NO, Kageyama, K, Asano, T, and Suga, H. 2006. Phylogenetic relationships of *Pythium* and *Phytophthora* species based on ITS rDNA, cytochrome oxidase II and β-tubulin gene sequences. Mycologia **98**:410-422.

Wainscoat, J. 1987. Out of the garden of Eden. Nature **325**:13.

Whisson, SC, Basnayake, S, Maclean, DJ, Irwin, JAG, and Drenth, A. 2004. *Phytophthora sojae* avirulence genes *Avr4* and *Avr6* are located in a 24 kb, recombination-rich region of genomic DNA. Fungal Genet. Biol. **41**:62-74.

Wiens, JJ, Kuczynski, CA, and Stephens, PR. 2010. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: Implications for speciation and conservation. Biol. J. Linn. Soc. **99**:445-461.

Wilson, AJ and Xu, J. 2012. Mitochondrial inheritance: Diverse patterns and mechanisms with an emphasis on fungi. Mycology **3**:158-166.

Yamagishi, T, Motomura, T, Nagasato, C, Kato, A, and Kawai, H. 2007. A tubular mastigoneme-related protein, Ocm1, isolated from the flagellum of a chromophyte alga, *Ochromonas danica*. J. Phycol. **43**:519-527.