The oomycete *Saprolegnia parasitica*: molecular tools for improved taxonomy and species identification

by

Wai Lam Leung
B.Sc., University of Victoria, 2008

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biology

Supervisory Committee

The oomycete *Saprolegnia parasitica*: molecular tools for improved
taxonomy and species identification

by

Wai Lam Leung
B.Sc., University of Victoria, 2008

Supervisory Committee

Dr. William Hintz, Department of Biology
Co-Supervisor

Dr. Paul de la Bastide, Department of Biology
Co-Supervisor

Dr. John Taylor, Department of Biology
Departmental Member

Abstract

Supervisory Committee

Dr. William Hintz, Department of Biology
Co-Supervisor

Dr. Paul de la Bastide, Department of Biology
Co-Supervisor

Dr. John Taylor, Department of Biology
Departmental Member

Saprolegnia parasitica is a pathogenic oomycete that cause saprolegniosis. Freshwater fish like salmon and trout species are particularly vulnerable to infection, which is characterized by cotton-like grayish mycelial growth on the surface of the fish. Currently, an effective treatment for this disease is not available. This pathogen has a great impact on freshwater fish species world-wide. An initial step to keep this devastating disease at bay is the ability to detect the responsible pathogen, so that appropriate actions could be taken before it becomes widespread. The development of molecular tools that will accurately and rapidly detect *S. parasitica* is the main goal of this project.

This project is divided into two main sections. The first section describes initial marker design efforts that were focused on the internal transcribed spacer (ITS) regions. Efforts were also made for the collection of field samples and the generation of our own ITS data that includes a number of *Saprolegnia* spp. Compiled sequence data allowed the identification of unknown samples and the adoption of the clade taxonomic system that other researchers had established for species designations. The accumulated sequence data helped to clarify the taxonomy within the genus *Saprolegnia* and complemented previous studies. The design of broad specificity PCR primers also allowed a quick initial detection of *Saprolegnia* spp., which could then be identified to species either by determining ITS nucleotide sequence or by a subsequent step of RFLP marker. Isolates sequence data in the compiled sample collection could be used for validation purposes in further marker development.

The second section of the project described the development of higher specificity molecular markers for the detection of *S. parasitica*. These were based on the study of three different gene loci as potential markers. These included the Pumilio RNA-binding protein (Puf), Glutathionylspermidine synthetase (Gsp) and the thiazole biosynthetic enzyme (Thi4). The nucleotide sequence of each locus was studied to develop suitable PCR primers that were then refined through testing against our isolate collection to improve their specificity for the target species. *Saprolegnia parasitica*-specific markers were developed for the Puf and Gsp loci and these were further evaluated using our field collected samples.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

%	percent
®	registered trademark
µg	microgram (s)
µL	microliter (s)
°	degree Celsius
a/c	autoclaved
Ab-GPA	glucose peptone agar added with antibiotics
bp	base pairs
dH$_2$O	distilled water
DNA	deoxyribonucleic acid
dNTP	deoxynucleotidetriphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
g	gram(s)
EtOH	ethanol
g	relative centrifugal force
GPA	glucose peptone agar
GPB	glucose peptone broth
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kb	kilobase pair (s)
LB	Luria-Bertani medium

M	Molar
min	minute (s)
mg	milligram (s)
mL	millilitre (s)
mM	millimolar
ng	nanogram (s)
PCR	polymerase chain reaction
rDNA	ribosomal deoxyribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
SDS	sodium dodecyl sulphate
sec	second (s)
TAE	tris-acetate ethylenediaminetetraacetic acid
™	trademark
Tris	tris(hydroxymethyl)aminomethane
U	unit
UV	ultraviolet
V	volts
v/v	volume to volume ratio
w/v	weight to volume ratio
X	times
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Acknowledgments

Thank you is not enough to express my gratitude for so many people that had assisted, encouraged and inspired me during this study. I am so grateful to my supervisor, “the wizard”, Dr. Hintz, for his wonderful ideas, optimism and continual support, to my co-supervisor, Paul de la Bastide, for his “endless patience” and excellent guidance, to Dr. Taylor, for his great ideas and questions. I must acknowledge my past labmates: Rebecca Jantzen, April Goebl, Sarah Truelson and many more; my current labmates: Jonathon LeBlanc, Joyce Carnerio, Cayla Naumann, Irina Kassatenko, I am greatly indebted to all of you because without your help and encouragement I cannot get to this point. So much love and laughter in the lab! I must acknowledge my Lord and Savior for HIS love and guidance, and many thanks to my wonderful husband for his support throughout this project.

Chapter 1

Introduction

Background

What is *Saprolegnia parasitica*?

Saprolegnia parasitica belongs to the oomycetes a group of heterotrophs that are saprophytes or parasites targeting a wide range of hosts (Robertson *et al.*, 2009; Phillips *et al.*, 2008). Saprolegniosis is the infection caused by *Saprolegnia* species and is characterized by external white or grey patches of cotton-like filamentous mycelial growth on the surface of a host (Phillips *et al.*, 2008; van West, 2006; Hatai and Hoshiai, 1992). Infection of fish usually starts from the head or fins, and then spreads over the whole body (Ramaiah, 2006). Infected fish usually succumb to the disease due to imbalanced osmoregulation, resulting in hemodilution (Robertson *et al*., 2009; van West, 2006). Members of the genus *Saprolegnia* are responsible for various diseases in animals (Robertson *et al*., 2009; Phillips *et al.*, 2008; van West, 2006). For example, *S. ferax* and *S. diclina* are thought to be responsible for the decline in amphibian populations in the wild (Fernández-Benéitez *et al.*, 2008; Blaustein *et al*., 1994). *Saprolegnia diclina* is believed to be a powerful pathogen of fish eggs (Thoen *et al*., 2011; Robertson *et al*., 2009; Fregeneda-Grandes *et al*., 2007). *Saprolegnia monoica* is responsible for losses in sturgeon hatcheries (Phillips *et al.*, 2008), while *S. parasitica* is believed to be the primary pathogenic fungus that causes saprolegniosis on most species of fish (Phillips *et al*., 2008; van West, 2006).

Impact of *Saprolegnia* on fish aquaculture

Saprolegnia parasitica, the focus of this study, is a devastating pathogen that infects many local freshwater fish species, as well as fish in other regions of the world. Trout and salmon species are particularly vulnerable to its attack. It is estimated that 10% of hatched salmon raised in aquaculture facilities die due to *S. parasitica* infection (Robertson *et al.*, 2009). *Saprolegnia parasitica* is also responsible for "winter kill" observed in the catfish industry (van West, 2006). In addition to fish farms, *S. parasitica* is believed to play a significant role in the decline of wild fish populations (van West, 2006).

Fish farming, also known as fin fish aquaculture, has become the fastest growing food sector as the demand for fish and shellfish has increased dramatically over the years (Robertson *et al*., 2009; van West, 2006). Fish diseases are the main cause of economic loss in aquaculture. Oomycete infections are second only to the bacterial infections, which are the most common disease-causing agents found in aquaculture industries (Almeida *et al.*, 2009; Meyer, 1991). Therefore, *Saprolegnia* infections represent a serious problem for aquaculture industries all over the world (FAO, 2010).

The life cycle of *S. parasitica* is complex (Figure 1) and consists of both sexual and asexual phases of reproduction (van West, 2006). Asexual reproduction consists of the production of short-lived primary motile zoospores which encyst shortly after release (Ramaiah, 2006; van West, 2006). These encysted zoospores then germinate to form secondary zoospores, which are more motile than the primary zoospores (Walker and van West, 2007; Ramaiah, 2006; van West, 2006). Secondary zoospores are considered to be the main dispersal phase, as well as the infective stage of *S. parasitica*, because they exhibit repeated cycles of encystment and zoospore release (referred to as polyplanetism or Repeated Zoospore Emergence - RZE) (Walker and van West, 2007; van West, 2006).

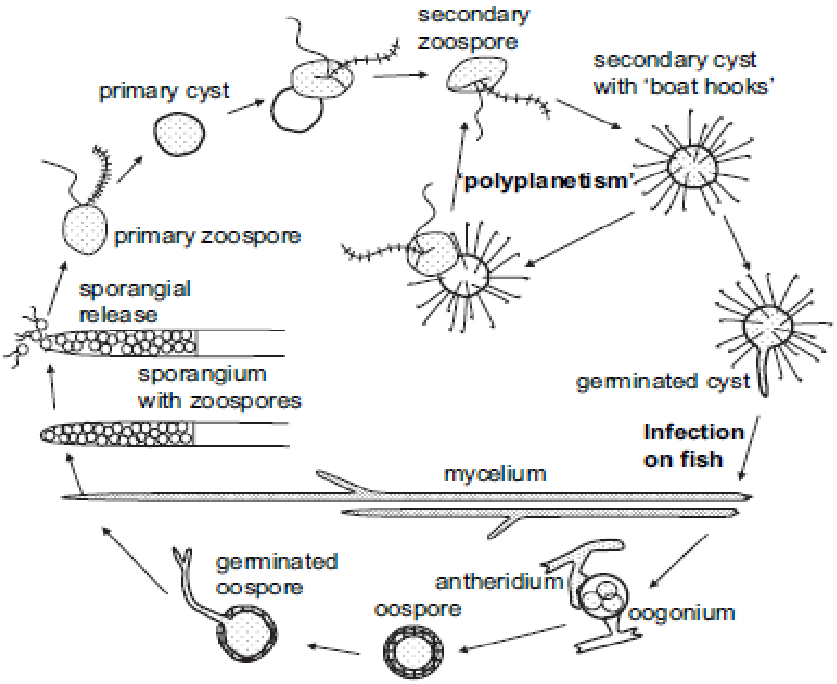


Figure 1. Schematic diagram of the life cycle of *Saprolegnia parasitica* (van West, 2006).

It has been shown that there are bundles of hooked hairs on the secondary cysts (Robertson *et al*., 2009; Willoughby, 1985) and it has been suggested that pathogenicity is related to hair number and length on these cysts (Robertson *et al*., 2009). Interestingly, Stueland's group (2005) showed both the most virulent strains of *Saprolegnia* isolates, as well as the avirulent strains tested, possessed bundles of long and hooked hairs. However, Fregeneda-Grandes's group (2000) showed that all samples derived from the lesions of fish displaying saprolegniosis produced cysts that were ornamented with fewer bundles of shorter hairs. In addition, when kept under low nutrient conditions, *S. parasitica* cysts can grow into a very distinct form that is termed indirect or retracted germination (Willoughby, 1985). All of these characteristics, namely, RZE, hair ornamentation of secondary cysts, and indirect germination are believed to increase the opportunities for *Saprolegnia* species to find a suitable host and shorten the colonization period (Robertson *et al*., 2009; Diéguez-Uribeondo *et al*., 2007).

Current treatment options for saprolegniosis

Prior to the year 2002, *S. parasitica* infections in fish hatcheries were kept under control through the use of malachite green. Since this chemical was banned due to its potential carcinogenic effects, saprolegniosis has re-emerged and once again, become a significant problem for the aquaculture industry (Robertson *et al*., 2009; Torto-Alalibo *et al.*, 2005). Chemicals such as amphotericin B (Robertson *et al*., 2009), formalin, hydrogen peroxide, copper sulfate pentahydrate, diquat bromide (Mitchell *et al.*, 2010), some derivative chitosan products (Muzzarelli *et al.*, 2001), Bronopol (Aller-Gancedo and Fregeneda-Grandes, 2007; Pottinger and Day, 1999), sodium chloride (Ali, 2005), or nikkomycin Z (Guerriero *et al.*, 2010) can all be effective at inhibiting the growth of *S. parasitica*. However, due to cost, practicality, as well as the potential toxicity of some of these chemicals, no current treatment is as effective as malachite green for the control of *S. parasitica* (Aller-Gancedo and Fregeneda-Grandes, 2007; Sudova *et al.*, 2007). Developing a conventional vaccine, or even a DNA-based vaccine, which involves injecting the gene that encodes the antigen into the fish muscle tissue, may have potential, should an appropriate antigen be discovered (Robertson *et al*., 2009).

Current research

Physiological aspects and the life cycle of *S. parasitica* had been well studied, although mechanisms underlying its pathogenicity, host specificity, and its population structure have not been described (Robertson *et al.*, 2009). Several research projects centered on *S. parasitica* have been initiated. For example, the Aberdeen Oomycete group is working on the identification and functional characterization of *S. parasitica* effector proteins, as well as creating gene expression profiles, using two kinds of interactive libraries (van West *et al*., 2010; Robertson *et al.,* 2009; University of Aberdeen, web resource). They are also currently investigating the potential usefulness of SpX protein as a vaccine against *S. parasitica*. These researchers have recently collaborated with the Broad Institute to establish a *S. parasitica* genome database portal (Broad Institute, web resource).

Identification challenge

To prevent saprolegniosis, the early detection and accurate identification of the responsible pathogen is essential. Traditional identification of the genus *Saprolegnia* is based on morphological features, for instance, documenting the method of zoospore release (Hulvey *et al.,* 2007; Leclerc *et al.,* 2000). Identification of species is even more difficult because it largely relies upon characteristics of the sexual structures, which do not usually form on fish lesions, nor do they routinely occur in fresh water (Diéguez-Uribeondo *et al.,* 2007; Fregeneda-Grandes, *et al.,* 2000). The morphological characteristics of different species within the same genus are sometimes similar (Diéguez-Uribeondo *et al.,* 2007; Fregeneda-Grandes *et al.,* 2000), and require taxonomic expertise to delineate different species. In addition, it takes time to culture samples so that these features may be observed (Ke *et al.,* 2009). Defining species strictly based on morphological characteristics is not reliable, and is at times impossible (Ke *et al.,* 2009; Diéguez-Uribeondo *et al.,* 2007).

Taxonomy confusion in the genus ***Saprolegnia***

The inability to recognize the true identity of samples can lead to misidentification. To complicate matters even further, different authors have used different names to actually refer to the same species (Diéguez-Uribeondo *et al.,* 2007; Fregeneda-Grandes *et al.,* 2007; Hulvey *et al.,* 2007). For example, *S. diclina* Type I, *S. parasitica, S. salmonis, S. diclina-S. parasitica* complex, or simply *Saprolegnia sp.* can all refer to the pathogenic isolates derived from fish lesions (Diéguez-Uribeondo *et al.,* 2007). An easier and more accurate way to identify *S. parasitica* is needed, and a molecular genetic approach seems to be the most appropriate.

Basic unit – the species

A classification system has been established so that the immense number of organisms can be studied in a manageable way. The species is the basic unit in taxonomy and this area of study is dynamic, and classification schemes are constantly changing with improved information (Spooner *et al.,* 2005). Different disciplines use different approaches to define a species. For instance, in fungal taxonomy, the focus has traditionally been on morphological features. A mycologist would classify a species based primarily on appearance, and may use sexual characteristics as a baseline for identification of species with a sexual or perfect life cycle, or conidial development and morphology for asexual or imperfect species. In contrast, a microbiologist would identify a fungal or bacterial species based on its nutritional requirements or its biochemical defence components, but may altogether ignore morphological features of the sexual stage (Guarro *et al.*, 1999).

How to define a species

There is no single universal definition of a species that every discipline can agree on, but overall there are a few species concepts used to define species (Spooner *et al.*, 2005; Guarro *et al.*, 1999). For example, the morphological species concept, for which morphological characteristics that are common or different among individuals are used to define a species; the ecological concept, which is based on adaptation to certain habitats. The biological concept, which relies on interbreeding of members within a species, is the most common definition of a species, which describes individuals/populations as the same species if they are capable of interbreeding and producing viable offspring (Spooner *et al.*, 2005; Guarro *et al.*, 1999). The polythetic concept defines species based on a combination of characteristics (Guarro *et al.*, 1999). In addition, the phylogenetic-species concept is based on the use of molecular techniques for the analysis of DNA nucleotide sequences to delineate species. The basic premise is that number of changes to particular gene sequences may be related to the time of divergence of two species from a common ancestor. This method is particularly useful for organisms that have cryptic sexual stages or lack them entirely and should be more reliable than classification based solely on morphological observation (Spooner *et al.*, 2005; Taylor *et al.*, 2000; Guarro *et al.*, 1999). In the current study, the word “species” is referring to phylogenetic species.

Genetic marker

A genetic marker is defined as “a DNA sequence with a known physical location on a chromosome. Genetic markers can help link an inherited disease with the responsible gene. DNA segments close to each other on a chromosome tend to be inherited together. Genetic markers are used to track the inheritance of a nearby gene that has not yet been identified, but whose approximate location is known. The genetic marker itself maybe a part of a gene or may have no known function.” (National Human Genome Research institute, web resource). Genetic material can be used as a tool to assess genetic relatedness among cells, individuals, members of a population or a species. There are different methods available for generating markers, all of which are based on detecting the presence of variability in the same region (Schlötterer, 2004). Each method has its strength and weakness. What kind of research questions needs to be addressed, what resolution is needed, financial constraints of the study, and what level of expertise is available are just some of the many criteria a researcher may consider when deciding which marker to use (Spooner *et al.,* 2005; Schlötterer, 2004; Sunnucks, 2000).

According to Mueller and Wolfenbarger (2000), an ideal approach to assessing genetic diversity should meet several criteria including being cost-effective and efficient, capable of generating multiple independent markers that have good resolution, and are reproducible, operates even using only a minute amount or even partially degraded DNA sample, and especially useful when there is no genome information for the organism of interest.

In general, a genetic marker can be categorized as either a multilocus or single-locus marker depending on how many loci in a genome it can recognize (Spooner *et al.,* 2005;

Mueller and Wolfenbarger, 2000; Sunnucks, 2000). It can also be classified as either a dominant or a co-dominant marker (Spooner *et al.*, 2005; Mueller and Wolfenbarger, 2000; Sunnucks, 2000). A co-dominant marker can reveal homozygote or heterozygote status of a locus. That is, it allows researchers to distinguish an AA genotype (homozygous) from Aa (heterozygous) genotype, whereas a dominant marker does not allow this kind of determination (Spooner *et al.,* 2005; Mueller and Wolfenbarger, 2000; Sunnucks, 2000). AlIozymes, restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), microsatellites, amplified fragment length polymorphisms (AFLP), just to name a few, are some popular methods researchers use to generate genetic markers (Spooner *et al.*, 2005; Schlötterer, 2004; Mueller and Wolfenbarger, 2000; Sunnucks, 2000). Allozymes are variants of enzymes coded by different alleles at the same locus (Schlötterer, 2004). Because of the amino acid charge differences, their mobility following native gel electrophoresis can be different hence the polymorphisms among alleles can be observed and used as marker for genetic differences.

When DNA is digested by a restriction enzyme (RE), RFLPs generate different lengths of specific DNA fragments, due to the presence or absence of RE cut sites in the template DNA. Such RFLPs can be generated by digesting whole genomic DNA, and the digested profile can subsequently be detected by probing with known labelled probes to the Southern blot hybridization membrane (Spooner *et al.,* 2005; Schlötterer, 2004; Mueller and Wolfenbarger, 2000). Alternatively, one can employ the Cleaved Amplified Polymorphic Sequence (CAPS) protocol, also known as PCR-RFLP, which uses a PCR method and appropriate primers to amplify the target DNA fragments, polymorphisms are detected within the target region by a restriction digest of the PCR product, and visualization by gel electrophoresis (Spooner *et al.*, 2005). CAPS or PCR-RFLP is considered superior to conventional RFLP methods as it only requires small quantities of DNA template (because it is a PCR-based method), and it can bypass the labour intensive steps of Southern blot hybridization and radioactive probe detection (Spooner *et al.*, 2005; Schlötterer, 2004).

Another approach is the use of RAPD based markers, dominant markers that are generated by using combinations of short nucleotide primers at relaxed annealing temperatures in a PCR reaction. The primers are designed as random sequence with the assumption that similar sequences will occur at sites distributed throughout the genome. Single primers will anneal to these different locations of the genome and when they occur on opposite strands of the DNA within a distance of up to approximately 7kb, the amplification products can serve to create a "quick profile" that can be used for further analysis, although the function of the amplified sequences will not likely be known. Most of the sequences remain anonymous. One of the drawbacks of this approach is its low reproducibility requiring many repetitions of RAPD amplifications (Spooner *et al.*, 2005; Mueller and Wolfenbarger, 2000; Sunnucks, 2000).

Microsatellites are simple short sequence repeats (SSR) containing a short repeat of nucleotides arranged in a tandem fashion. It is the variation in the numbers of these repeats that make it potentially useful as a marker. A higher mutation rate contributes to length polymorphisms detected by SSRs when compared to other genomic regions. Designing primers that flank the repeat regions allow their amplification by PCR and the size of the amplified product reflects the number of repeats, which may vary among isolates or species (Spooner *et al.,* 2005; Schlötterer, 2004; Mueller and Wolfenbarger, 2000; Sunnucks, 2000). Markers based on AFLPs are obtained by digesting total genomic DNA with restriction enzymes designed to generate overhanging fragments. The digested DNA fragments can then be ligated into adaptors that would match the overhanging sticky ends of the fragments. Fragments with adaptors, within a certain size range, would then be selected for amplification by PCR using primers that anneal to the adaptors. Subsequently, the separation of the amplified products will generate a profile of the target region of interest for study (Mueller and Wolfenbarger, 2000).

The Challenge

Many different organisms including microbes coexist in the water column. A number of species of the genus *Saprolegnia* can also be found in water, but not every *Saprolegnia* species is pathogenic (Fregeneda-Grandes *et al.,* 2007; Czeczuga and Muszyńska, 1997). Species composition varies depending on the conditions in the environment, and the health of the fish. Experiments have shown that certain species are more pathogenic towards fish (Stueland *et al.,* 2005; Fregeneda-Grandes *et al*., 2000; Singhal *et al.,* 1987) and these species are the focus of this study. Therefore, it is essential to develop a method that can delineate those pathogenic species from the non-pathogenic in the water column. The ability to employ a suitable tool quickly and accurately to detect *S. parasitica* will permit early intervention and minimize the impact of this disease making this approach a useful management tool. The ability to identify the causing agent will also make it easier to study the disease and identify other factors that may influence its occurrence. It is likely that presence or absence of the fungus in the water is not the only factor contributing to disease outbreaks. Good detection tools will improve management practices to reduce impact of saprolegniosis by alerting hatchery managers to the level of contamination by the pathogen.

Overall project objectives

1. Clarification of the taxonomy within the genus *Saprolegnia* by analysis of ITS sequence data from field isolates. The collection of fungi from fish lesions and water samples from fish hatcheries all over British Columbia (mainly focused on Vancouver Island) was conducted to provide a sample data reflective of the local aquaculture community. Sub-culturing of isolates, DNA isolation, PCR amplification of the ITS region, and subsequently, obtaining ITS region nucleotide sequence information from those samples were part of the process, and the information generated was used for comparative studies. Through the analysis of our ITS region sequence data, and comparison to published ITS sequence information, a manageable naming system for *Saprolegnia* species was established.
2. The primary objective is to develop *Saprolegnia*-specific molecular genetic markers for the screening of environmental samples that may contain this organism. Markers having different resolving power were developed to distinguish the various *Saprolegnia* species, with an emphasis on distinguishing *S. parasitica* from other members of the genus *Saprolegnia*.

Chapter 2

An examination of species boundaries within the genus *Saprolegnia* based *o*n nucleotide sequence analysis of the Internal Transcribed Spacer (ITS) of the ribosomal RNA (rDNA)

Introduction

In the past, identification of *Saprolegnia* species was mainly based on the morphology of sexual structures (Hulvey *et al*., 2007; Leclerc *et al*., 2000). Molecular genetic information concerning *Saprolegnia parasitica* was very limited. There were only thirteen nucleotides sequences and two protein sequences of *S. parasitica* available from GenBank prior to 2005 (Torto-Alalibo *et al*., 2005). As saprolegniosis drew more attention, together with advances in molecular technology, researchers started to integrate molecular approaches with the conventional morphological/physiological-approaches to study *Saprolegnia* species (Ke *et al*., 2009; Diéguez-Uribeondo *et al*., 2007; Hulvey *et al*., 2007; Bangyeekhun *et al.,* 2001; Diéguez-Uribeondo *et al*., 1996; Molina *et al.,* 1995). In addition, the work of Torto-Alalilbo *et al*., (2005) generated more than 1,500 expressed sequence tags (ESTs) from a mycelial cDNA library of *S. parasitica*, and gained more insight into genes that may play important roles in fitness and pathogenicity of *S. parasitica.*

The study of Diéguez-Uribeondo *et al*., (2007) integrated information on ITS rDNA nucleotide sequences for a collection of *Saprolegnia* spp., of mostly European origin, as well as other isolates world-wide. This extensive study and other studies that incorporated molecular approaches to address questions on phylogenetic and taxonomic aspects of the genus *Saprolegnia* paved the way for the current study. They illustrated the feasibility of using molecular techniques (ITS sequencing, restriction digest, RAPD etc.) to gain insight into *Saprolegnia* speciation and species boundaries (Ke *et al*., 2009; Diéguez-Uribeondo *et al*., 2007; Hulvey *et al*., 2007; Bangyeekhun *et al.,* 2001; Diéguez-Uribeondo *et al*., 1996; Molina *et al.,* 1995).

The current project was initiated to improve our understanding of this genus, in particular the species *Saprolegnia parasitica*, and its relative importance in freshwater aquaculture systems. This portion of the study has been focused on genetic marker development and the refinement of the taxonomy within this genus, using nucleotide sequence data. This was accomplished by designing broad specificity genetic markers to distinguish the genus *Saprolegnia*, using information available from GenBank. These markers were tested and refined, while concurrently obtaining and compiling a field sample collection. This allowed us to acquire our ITS sequence data to assist the taxonomic study and gradually build a manageable naming system for this taxonomically confusing genus. The investigation of broad specificity genetic markers can be divided into two aspects; the design of ITS region based markers and the investigation of restriction fragment length polymorphism (RFLP) patterns, for digests of the amplified ITS region.

Ribosomal RNA (rRNA) forms two subunits: the large subunit (LSU), and the small subunit (SSU). These rRNA and proteins combine to form an important structure (ribosome) where protein synthesis takes place. Biogenesis of rRNA and ribosome take place in the nucleolus. In eukaryotes the large ribosomal subunit contains the 5S, 5.8S and 28S rRNAs while the small ribosomal subunit contains only the 18S rRNA. Genes encoding rRNA occur in tandem repeats and are present as multiple copies in the genome. Each repeat (rRNA gene) consists of a promoter, external transcribed spacer

(ETS), rRNA coding sequence, internal transcribed spacer (ITS), followed by external-transcribed spacer (ETS). During the synthesis of rRNA, rDNA is transcribed into the rRNA transcript (precursor rRNA or pre-RNA) by nucleolar pol I. This pre-RNA sequence contains the 5'ETS, 18S, ITS1, 5.8S, ITS2, 28S, and followed by the 3' ETS (in general, pre-RNA and 5S RNA genes in eukaryotes are separated). Before (or during) the assembly of the small and large ribosomal subunits in the nucleolus, non-coding transcribed spacer sequences (ITSs and ETSs) are removed. Subsequently, RNA helicases and RNA chaperons assist folding and remodeling of these mature RNA molecules (18S, 5.8S, and 28S). The 5S rRNA and ribosomal proteins (synthesized separately from other rRNAs mentioned above) are then recruited to the nucleolus and combined into ribosomal subunits, which are then ready to move back to the cytoplasm for protein synthesis (Raška et al., 2004; Shaw and Jordan, 1995; Long and Dawid, 1980).

A few characteristics of rDNA (genes that encode ribosomal RNA) made it an attractive target for molecular marker design. Highly conserved regions of the 18S, 5.8S and 28S (coding) rRNAs are separated by variable, non-coding intergenic spacer (IGS) and internal transcribed spacer (ITS) regions (Richard *et al.,* 2008; Sumida *et al.,* 2004; Long and Dawid, 1980). The availability of universal primers for this region also provides a useful starting point in the study of species. Many studies, for example, of oomycetes, plants, fungi, insects, and frogs etc., have reported the use of these variable regions to study phylogeny or species identification (Schena and Cooke, 2006; Li *et al.,* 2005; Ma and Xu, 2005; Sumida *et al.,* 2004; Chen *et al*., 2001, 2000; Cooke and Duncan, 1997; Baldwin 1992). These particular sequences have worked well for identification for some species, but for other species markers having a greater resolution were needed. For instance, Chen *et al.*, (2001, 2000) reported they could identify 30 species of yeasts, comprising 98% of the clinical isolates, simply based on the combined information from the size of ITS1 and ITS2 amplicons. The high resolution they could achieve could be due to their use of automated capillary electrophoresis with fluorescently labeled PCR products under denaturing conditions (Chen *et al.*, 2001, 2000). Many studies use these regions for molecular analysis in combination with comprehensive morphological studies for the organism of interest (Ke *et al.*, 2009; Johnson *et al.*, 2008; Hulvey *et al.*, 2007; Chen *et al.*, 2001, 2000; Molina *et al.*, 1995). In studies of yeast, researchers looked at the D_2 variable region of S25, and determined that, in general, less than one percent sequence differences among tested isolates in that variable region would define those isolates as the same species (Chen *et al.*, 2001; Guarro *et al.*, 1999). This level of sequence variability could therefore be used as a guideline to evaluate whether the ITS sequence data collected in the current study supports the proposed naming system.

It is very common for researchers to use universal primer sets to first amplify regions of the ITS1, 5.8 S or ITS2 from samples of interest, determine the amplicon sequences and then, design markers or probes to detect variation in the target species (Vandersea *et al.*, 2006; Park, 2001; Gardes and Bruns, 1993). The use of 18S rDNA and ITS sequence comparisons have worked well for taxonomic and phylogenetic studies of many taxa down to the family level, especially when one also incorporates other morphological characteristics. However, at the species level, rDNA and ITS sequences are not usually variable enough to resolve species with a genetic marker based on these sequences. That is, the marker may also amplify other non-target sequences of closely-related species, which was found to be the case for *Aphanomyces astaci* (another member of the family Saprolegniaceae); there was not sufficient variation to distinguish between *A. astaci* and closely related *Aphanomyces* species (Ballesteros *et al.,* 2007; Oditmann *et al.,* 2004). Molina *et al.,* (1995) amplified 33 strains representing 18 species of *Saprolegnia* using the universal primers ITS1, ITS4, NS1, and NS8 to amplify rDNA spanning the 18S, ITS1, 5.8S and ITS2 regions. Rather than a full sequence determination, the PCR amplicons were digested with 13 different restriction enzymes that generated different RFLP patterns. They found that a *Bst*UI digest could generate the same fingerprint patterns for all *S. parasitica* tested, regardless of host and geographic origin. Their findings determined that *S. diclina* and *S. delica* were not the same species, despite a history of these two names being used interchangeably. As well, their findings confirmed that *S. diclina* and *S. parasitica* were not same species, and *S. asterophora* was obviously very different from tested *Saprolegnia* species. A similar approach in the current study may provide some useful information for the detection of *S. parasitica* and its discrimination from other species in this genus.

Establishing naming system of *Saprolegnia* spp., based on collection in this study

Through a review of the *Saprolegnia* literature it has become apparent that there is no consistent way of naming or identifying *S. parasitica* (Diéguez-Uribeondo *et al.,* 2007; Hulvey *et al.,* 2007; Leclerc *et al.,* 2000). For instance, *S. parasitica* was originally used to designate any species that failed to produce any observed sexual structures and also infected fish (Diéguez-Uribeondo *et al.,* 2007). Some researchers suggested the name of *S. parasitica* for isolates that have long hair ornamentation on the secondary cyst and exhibit indirect germination under low nutrient conditions (Willoughby 1995). Evidently there is a need for additional study of this genus to clarify species designations and provide more reliable methods of species identification for the pathogen *S. parasitica*. We initially designed genetic markers to target the ITS region of *S. parasitica*. In order to generate more information about the ITS nucleotide sequence and to validate our designed markers, we compiled our own field collection, obtained pure isolates, extracted total DNA, and determined their ITS sequences. Our results show that there is sufficient consistent variation in the ITS regions to allow us to designate phylogenetic species names for an unknown sample, in addition to designing a molecular genetic marker to detect this pathogen, I sought to establish a systematic way to delineate *Saprolegnia* species by using our collection of ITS region sequence data obtained from freshwater aquaculture facilities. This research helped to clarify the species designations and alleviate some of the taxonomy confusion that exists in this genus. Overall, this study generated much useful information in the form of an isolate collection and the nucleotide sequence data that was generated to clarify species designation within the genus and to further refine genetic marker designs. Our results will facilitate the study of this species and allow one to evaluate the health of both the natural and the artificial aquatic systems.

Materials and Methods

Field sample compilation, pure isolates collection

Environmental samples of infected fish, fish eggs, and water samples were collected from different locations in Canada, with a focus on Vancouver Island sample sites. All samples were shipped in coolers and kept at 4°C until processed.

For the culturing of isolates from fish samples, fish were first inspected for obvious lesions, usually observed as, mycelial growth on the pectoral or caudal fins. Infected areas were excised and rinsed three times with autoclaved distilled water (a/c dH_2O), and then transferred to a sterile Petri dish (100 x 15 mm) containing 15 mL of a/c dH_2O. Autoclaved hemp seeds were added as a bait substrate and were colonized by spores or mycelia within 24 to 72 hours. Single colonized hemp seeds were aseptically transferred to a glucose peptone agar (GPA, 3 g/L D-glucose, 1.25 g/L bacto peptone and 15 g/L bacto agar). To obtain pure cultures, the medium was augmented with four antibiotics (Ab-GPA) and these included Rifampicin (Calbiochem, La Jolla CA, USA) at a final concentration of 50 mg/L, Nystatin N1638 (Sigma-Aldrich, St. Louis MO, USA) at 10 mg/L, Chloramphenicol (Sigma) at 25 mg/L and Streptomycin (Calbiochem) at 10 mg/L. To obtain pure culture isolates, colonies were allowed to grow for three to five days before transfer of a mycelia plug from the edge of the colony to a fresh Ab-GPA plate; this procedure was repeated at least three times to obtain a pure single culture isolate growing on GPA.

Culturing from egg samples followed the same protocol as fish tissue samples to obtain pure cultures. To obtain pure cultures from environmental water samples, 15 to 20 mL of water was poured into a sterile petri dish; autoclaved hemp seeds were added and the same subculturing protocol as described above was followed.

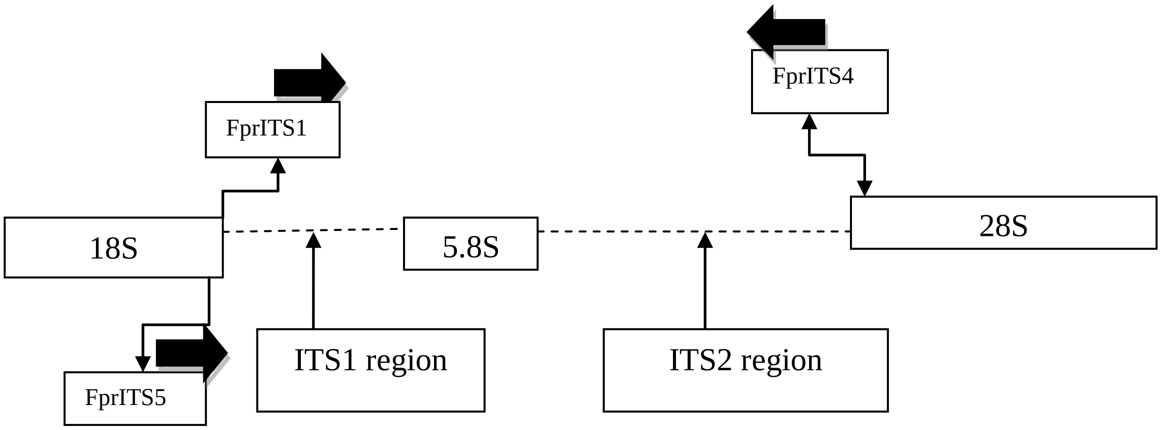
To obtain lyophilized samples of pure culture isolates, the growing edge of a mycelial colony was excised and transferred into a 125 mL Erlenmeyer flask containing 50 mL of glucose peptone broth (GPB, 3 g/L D-glucose, 1.25 g/L bacto peptone) and maintained at ambient temperature (about 25°C) until log phase was reached. Cultures were rinsed three times with a/c dH_2O, harvested by vacuum filtration, quick frozen in liquid nitrogen and immediately lyophilized for at least 48 hours.

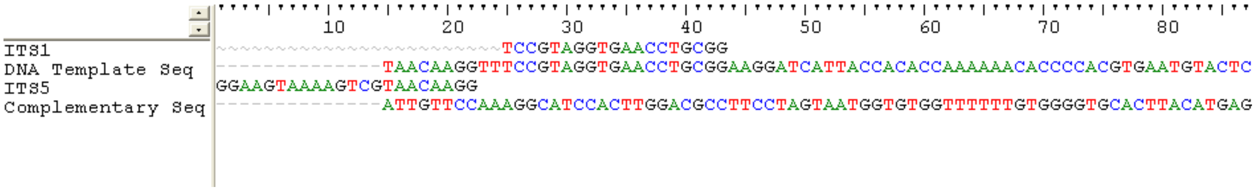
DNA isolation

For the isolation of DNA from each sample, the protocol of Möller *et al*., (1992) was followed, with some minor modifications. About 30 to 60 mg of lyophilized mycelia was ground with 100 mg of a/c zirconium/silica beads (0.5 mm diameter, Fisher Scientific, Canada) and 100 µL of TES buffer (100 mM Tris pH 8, 10 mM EDTA, 2% SDS) inside a 1.5 mL microfuge tube by use of a bead beater (MINI Beadbeater ™, Biospec Products) for 45 seconds then briefly centrifuged at 13,000 g for 30 seconds. This procedure was repeated three to five times until the mycelial sample was evenly homogenized. Total DNA extracted and re-suspended in 50 µL UltraPure™ Distilled water (GIBCO, Grand Island, New York, USA). Subsequently, samples were analyzed to determine DNA purity and concentration using the Nanodrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA), prior to preparing DNA template dilutions for use in PCR reactions.

PCR amplification of ITS region, nucleotide sequencing and isolate identification

The universal ITS region primer pair ITS5 and ITS4 (White *et al.*, 1990) were used to amplify the region between the internal transcribed spacer 1 (ITS1) and 2 (ITS2) of the rRNA cistron, including the 5.8S region. Since there are only minor differences between the forward primers ITS1 and ITS5; the ITS5 primer was primarily used in this study. The annealing sites of the two primers are close together, with the 5’end of ITS5 annealing two base pairs upstream of the 5’ end of ITS1, when using *S. parasitica* as a template (Figure 2).





```
ITS1                  ~~~~~~~~~~~~~~~~~~~~~~~~TCCGTAGGTGAACCTGCGG
DNA Template Seq      --------------TAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACCACACCAAAAAACACCCCACGTGAATGTACTC
ITS5                  GGAAGTAAAAGTCGTAACAAGG
Complementary Seq     --------------ATTGTTCCAAAGGCATCCACTTGGACGCCTTCCTAGTAATGGTGTGGTTTTTTGTGGGGTGCACTTACATGAG
```

Figure 2. Schematic showing annealing sites of universal primers ITS1, ITS5, and ITS4 to rDNA and the relative position of different ITS amplified regions. The bottom diagram shows only annealing sites for primer ITS5 and ITS1. Sequences obtained from NCBI usually started from “CACCA” which is shown at position 57 in the bottom diagram. Annealing site of ITS4 cannot be shown in the above diagram because it locates further downstream of the DNA sequence.

Each PCR reaction was performed in a 20 µL final volume using one unit of Fermentas DreamTaq DNA polymerase, a final concentration of 0.5 µM for each primer and 20.0 ng of genomic DNA per 20 µL reaction. All PCR amplification reactions using these primers were performed using the Eppendorf Mastercycler® gradient model 5331 and followed the reaction conditions described by Diéguez-Uribeondo *et al*. (2007). Following an initial denaturation (5 min at 94°), each PCR reaction was conducted by 5 cycles of denaturation (30 sec at 94°), annealing (30 sec at 55°) and extension (1 min at 72°), followed by 33 cycles of denaturation (30 sec at 94°), annealing (30 sec at 48°), and extension (1 min at 72°), a final extension (10 min at 72°) and was held at 4°until processed. A 5.0 µL volume of each PCR product was mixed with 2 µL of loading dye and loaded into each well of a 1.5% (w/v) agarose gel, separated by gel electrophoresis (1 hour and 24 min at 7 V/cm) and visualized by staining with GelRed (3X staining solution from 10,000 stock, w/v) for 30 minutes, followed by illumination under UV light. Amplified products (ITS5 and ITS4) of template DNA were initially sent (Feb to July 2010) without purification to the Macrogen direct sequencing service (Macrogen, Rockville, USA) for sequence determination. Samples subsequent to Aug 2010 were sent to Eurofins mwglOperon (Operon, Huntsville, AL, USA), for direct DNA sequencing, following purification using the QIAquick PCR purification Kit (Qiagen, Germantown, WI, USA), according to the company protocol. Sequencing results (Macrogen, Feb through July 2010; Operon, Aug 2010 to present) were visually analyzed and manipulated using the BioEdit Sequencing Alignment Editor (version 7.0.9.0) (Hall, 1999). Each sample was sequenced in two directions from opposite strands and the information was compared to ensure that the sequence was correct. Sequences were then subject to a blastn (nucleotide query/ nucleotide database search option) search, using default parameters of the National Centre for Biotechnology Information (NCBI, web resource) database, as well as the Identification Engine under the category of “Fungal identification – ITS search” of the Barcode of Life Data System v2.5 (BOLD, web resource). The most appropriate species name was provisionally assigned to each sample, based on the results obtained from both databases.

Initial design of broad range ITS primers for preliminary testing

At the beginning of the study, sequences from the rDNA region of species within the family Saprolegniaceae were downloaded from GenBank and subsequently selected, based on their sequence similarity to several putative sequences of *S. parasitica*, a subset of 18 sequences were then chosen for use in the initial ITS primer design. Regions of the ITS that were determined to have a high similarity within the genus *Saprolegnia*, but not for other genera, were used to design seven forward primers (Table 1). Since sufficient amplification selectivity should be afforded by a single well-designed primer, the seven forward primers were tested with the universal reverse primer ITS4 (White *et al.*, 1990) under the amplification profile as listed earlier.

Table 1. Primers designed for the development of the genus *Saprolegnia* markers. The first seven primers listed below were designed using the ITS nucleotide sequences from 18 submissions to the NCBI database. The last three primers shown are universal primers designed by White *et al*., (1990), which were also used in this study. Fpr and Rpr represent forward and reverse primers, respectively.

Primer name	**Primer sequence (5' to 3')**	**Expected size of amplified product when paired with ITS4 (bp)**
Saprolegnia Fpr 1A	actgatcaaaactgcagatagaaa	594
Saprolegnia Fpr 1B	actgatcaaaactgcagatagaa	594
Saprolegnia Fpr 2A	gagatgtattatttaaaggtatgcc	294
Saprolegnia Fpr 2B	atttaaaggtatgcctgcgc	283
Saprolegnia Fpr 2C	atttaaaggtatgcctgcg	283
Saprolegnia Fpr 3	caaatcgcggtagttttgc	216
Saprolegnia Fpr 4	gtatgctggtgcatttcttg	111
Universal Fpr ITS1	tccgtaggtgaacctgcgg	705
Universal Fpr ITS5	ggaagtaaaagtcgtaacaagg	707
Universal Rpr ITS4	tcctccgcttattgatatgc	----

Refinement of ITS primer specificity by targeting single nucleotide polymorphisms

One of our strategies to improve primer specificity involved designing primers that include a single nucleotide polymorphism (SNP) at the last base pair position of the 3'end. When used under a highly stringent annealing temperature during the PCR cycle, this feature should be sufficient to amplify only the target sequence, despite the existence of only one base pair difference. Representative ITS sequences from our collection of *Saprolegnia* spp., were aligned to identify potential regions that might be suitable to design forward primers based on this strategy. Candidate forward primers were designed (Table 2), and tested with the universal reverse primer ITS4 (Table 1) using the Diéguez-Uribeondo *et al.*, (2007) PCR profile. Primer sets were initially tested with a small collection of 23 isolates that included *Saprolegnia* spp., other oomycetes and fungi (Table 3), as well as a negative control that did not contain any template. Primers that appeared to preferably amplify *Saprolegnia* species were recorded. Subsequently, six of those potential forward primers (Table 2) were chosen and subject to further testing using a gradient of annealing temperatures from 50°C to 60°C (50.0, 50.2, 50.7, 51.6, 52.7, 54.0, 55.4, 56.8, 58.1, 59.2, 60.0 and 60.4°C) with four isolates designated as W90213 (ATCC reference isolate), 142, 141, and 75 which represented *S. parasitica*, *S. ferax, S. delica* I and *S. diclina* respectively. The 50° to 60° gradient PCR reaction conditions were as the following: following an initial denaturation phase (3 min at 94°), each PCR reaction received 30 cycles of denaturation (45 sec at 94°), annealing (30 sec at 50°C to 60°) and extension (1 min 30 sec at 72°), a final extension (10 min at 72°) and was held at 4°until processed.

Table 2. Forward ITS primers designed to include a 3’ end SNP.

This is likely to be *S. parasitica*-specific. All Fpr were paired with Rpr ITS4, while The two new Rpf (RprITS2_640_3(G) and RprITS2_640_3) were paired with FprITS2_510. The location of the SNP for each primer is noted in bold.

Primer name	Primer sequence (5’ to 3’)	Expected size of amplified product (bp)
FprITS1_150	gtcaatttgaatcctttttaaa**a**	600
FprITS1_187	tgatcaaaactgcagatagaaat**a**	564
FprITS2_464	gacggtacctatgcgtcct**a**	286
FprITS2_510	gcctgcgctcctttcga**a**	240
FprITS2_570	gtggcggcacacagca**c**	180
FprITS2_640	atttctgcgagtctgttgt**ca**	110
FprITS2_640_2	gatttctgcgagtctgttgt**c**	110
FprITS2_640_3	ctgcgagtctgttgtcaaag**t**	110
FprITS2_660	caaggcacgtaaggagagt**t**	90
RprITS2_640_3(G)	actttgacaacagactcgca**g**	130
RprITS2_640_3	actttgacaacagactcgc**a**	130

Table 3. Initial small sample testing of primers specific to genus *Saprolegnia*.
A total 23 templates were tested, and a water negative control.

Isolate	Isolate identity revealed from DNA sequencing and isolate determination
W (90213)	*S. parasitica* (ATCC)
18	*S. parasitica* (with extra T in ITS1)
71	*S. parasitica*
28	*S. delica* I
141	*S. delica* I
MH13	*S. delica* II
8	*S. delica* II
10	*S. delica* II
4	*S. ferax* (Canadian Collection of Fungal Cultures)
94	*S. ferax*
142	*S. ferax*
2	*S. diclina* (ATCC 56851)
73	*S. diclina*
75	*S. diclina*
1	*Leptolegnia* sp.
51	ID not clear (something in between *Saprolegnia* spp. and *Leptolegnia*)
33	*Pythium* sp.
44	*Pythium* sp.
46	*Pythium* sp.
17	*S. asterophora*
23	*S. asterophora*
52	Fungus
54	Fungus

The second phase of primer selection was a determination of whether the primer, in combination with ITS4, could amplify *S. parasitica* W90213 exclusively at higher annealing temperature. Such primer sets were subjected to another PCR reaction testing with 23 template isolates (Table 3), with an annealing temperature of 59.2°. PCR conditions were the same as gradient 50° to 60°except annealing temperature altered to 59.2°. Moreover, that same set of primer (Fpr_ITS2_640_3 with ITS4) was also tested with expanded isolates of 62 (Table 4) with an annealing temperature of 60°. The condition of PCR reactions were as the following: initial denaturing phase (3 min at 94°), each PCR reaction received 40 cycles of denaturation (45 sec at 94°), annealing phase (30 sec at 60°), and extension (40 sec at 72°), a final extension (10 min at 72°), and was held at 4° until processed.

We further hypothesized that the *S. parasitica* specific primer set would have greater specificity, if both the forward and reverse primers included a SNP at 3’ end, as compared to only one SNP in the forward primer. Since Fpr_ITS2_510 and Fpr_ITS2_640_3 when tested, the former set appeared to have the ability to discriminate *S. diclina* and *S. ferax* from *S. parasitic* and *S. delica* I isolates at an annealing temperature beyond 59.2°, and when annealing temperature was higher than 58.1°, only the latter set of primer shown to be able amplifying *S. parasitica*. Therefore, two 3’ SNP reverse primers were designed (RprITS2_640_3(G) and RprITS2_640_3) and they were each tested with forward primers (FprITS2_510) (Table 2) using 62 isolates (Table 4) with PCR annealing temperature of 60° as described above.

Table 4. Advanced larger sample testing of primers.

Designed primers that were shown to be specific to *Saprolegnia* spp. were progressively refined and re-designed. A total of 62 samples were tested, and a water negative control.

Isolate	**Isolate identity revealed from DNA sequencing and isolate determination**
9	*S. parasitica*
W	*S. parasitica* (ATCC 90123)
MH1	*S. parasitica*
12	*S. parasitica*
40	*S. parasitica*
50	*S. parasitica*
62	*S. parasitica*
70	*S. parasitica*
115	*S. parasitica*
125	*S. parasitica*
154	*S. parasitica*
166	*S. parasitica*
168	*S. parasitica*
215	*S. parasitica*
226	*S. parasitica*
112	*S. parasitica*
260	*S. parasitica*
264	*S. parasitica*
271	*S. parasitica*
288	*S. parasitica*
295	*S. parasitica*

315	*S. delica* II
88	*S. parasitica*
177	*S. parasitica*
198	*S. parasitica*
319	*S. parasitica*
MH12	*S. parasitica*
SP	*S. delica* II (ATCC42062)
MH13	*S. delica* II
133	*S. delica* II
91	*S. delica* II
116	*S. delica* II
80	*S. delica* II
158	*S. delica* II
172	*S. delica* II
190	*S. delica* II
210	*S. delica* II
222	*S. delica* II
214	*S. delica* II
227	*S. delica* II
265	*S. delica* II
304	*S. delica* II
90	*S. delica* II
4	*S. ferax* (Canadian Collection of Fungal Cultures)
94	*S. ferax*
142	*S. ferax*
217	*S. ferax*
311	*S. ferax*

2	*S. diclina* (ATCC 56851)
73	*S. diclina*
35	*S. diclina*
31	*S. diclina*
17	*S. asterophora*
23	*S. asterophora*
275	*S. asterophora*
5	*S. delica* I
11	*S. delica* I
27	*S. delica* I (not 100% sure due to sequencing problem)
141	*S. delica* I
18	*S. parasitica* (with extra T in ITS1)
30	*S. parasitica* (with extra T in ITS1)
67	*S. parasitica* (with extra T in ITS1)

The generation of restriction enzyme maps for the ITS region and the evaluation of ITS RFLPs in the marker development

The use of ITS RFLPs as a genetic marker for *S. parasitica* was initially evaluated using our collection of confirmed isolates of *Saprolegnia* spp. and the restriction enzyme *Bst*UI. Template DNA was PCR amplified using the primer set Fpr1A and ITS4 (Table 1), and the Diéguez-Uribeondo *et al.*, (2007) PCR profile. Amplified PCR products were subsequently restriction enzyme (RE) digested for one hour. Digested products were visualized as described previously. Our initial analysis excluded *Bst*UI as a useful enzyme and prompted the development of a restriction enzyme map for the ITS region, using our confirmed sequences for species comparisons (Table 5). The candidate enzyme

*Bst*BI was tested using template DNA from a range of confirmed *Saprolegnia* spp., *Leptolegnia* sp. and higher fungi. The primers Fpr1A and ITS4 were used to amplify DNA by use of a revised reaction protocol that included an initial denaturation phase (3 min at 94°), followed by 30 cycles of denaturation (45 sec at 94°), annealing (30 sec at 60°) and extension (1 min at 72°), a final extension (10 min at 72°) and was held at 4° until processed. Amplified PCR products (2 µL) were subsequently digested with the enzyme *Bst*BI for two hours at 65° and digested products were visualized as described previously.

Table 5. Unique restriction site found in *S. parasitica* within ITS region. The number shown indicated the ITS nucleotide position for each restriction enzyme site. When comparing different *Saprolegnia* species, the enzyme *Bst*BI (in bold) has a unique site found only in *S. parasitica* but not in other closely related species.

Restriction enzyme tested	S. *parasitica* (W, 40, 42, 18)	*S. ferax* (4)	*S. ferax* (25)	*S. delica* II (S.P)	*S. delica* II (222)	*S. delica* I (16)	*S. delica* I (11)	*S. diclina* (2)	*S. diclina* (75)
*Acc*65I	407	405	406	405	405	405	405	404	404
*Bae*I	424	422	423	422	422	422	422	421	421
*Bae*I	391	389	390	389	389	389	389	388	388
*Ban*I	407	405	406	405	405	405	405	404	404
*Bcl*I	127	126	126	126	126	126	126	125	125
*Bgl*I	63	62	62	62	62	62	62	62	62
*Bmt*I	73	72	72	72	72	72	72	72	72
*Bsa*JI	299	298	298	298	298	298	298	297	297
*Bsa*XI	591	589	590	589	589	---	---	---	---
*Bsa*XI	621	619	620	619	619	---	---	---	---
***Bst*BI**	461	----	---	---	---	---	---	---	---
*Eco*RV	671	669	670	669	669	669	669	667	607
*Hga*I	405	403	404	---	---	403	403	402	402
*Hpy*1888III	349	---	---	---	---	---	---	---	---
*Hpy*F10VI	64	---	---	---	---	---	---	---	---
*Kpn*I	411	409	410	409	409	409	409	408	408
*Mbo*II	220	219	219	219	219	219	219	218	218
*Mfe*I	159	158	158	158	158	158	158	157	157
*Mwo*I	63	---	---	---	---	---	---	---	---
*Nhe*I	69	68	68	68	68	68	68	68	68
*Nla*IV	409	407	407	407	407	407	407	406	406
*Nsp*I	158	157	157	---	---	157	157	156	156
*Pml*I	26	26	26	26	26	26	26	26	26
PstI	140	139	139	139	139	139	140	138	138
*Sfc*I	136	135	135	135	135	135	135	134	134
*Sna*BI	231	230	230	230	230	230	230	229	229
*Sph*I	158	157	157	---	---	157	157	156	156
*Tsp*GWI	313	312	312	312	312	312	312	311	311

Development and analysis of ITS nucleotide sequence database

Following the assignment of the most appropriate species name, isolates collected in the current study were organized using the clade system of Diéguez-Uribeondo *et al.*, (2007). Nucleotide sequences from this previous study were obtained from the GenBank database for 128 isolates of *Saprolegnia* spp. and isolate assignment was verified by a current

Blastn search. Sequences from the Diéguez-Uribeonodo *et al*., (2007) study were organized using the clade system, manually inspected to determine sequence variation within each clade, and compared to the sequences from the current study by using a pairwise alignment approach. Representative samples from each clade of the published study were compared with two representative samples of each clade from the current study (a total of 35 ITS sequences) to confirm isolate identities. These new isolates were then utilized to construct a phylogenetic tree. This was completed by using the one click mode of the online phylogeny platform Phylogeny.fr, employing the default settings (Dereeper *et al.*, 2008). This phylogeny platform used the program MUSCLE for alignments, GBlock for curation, PhyML for phylogeny, and TreeDyn to render the phylogenetic tree. These serial steps allowed the assignment of a proper species name or clade to most field collected samples (Figure 3).

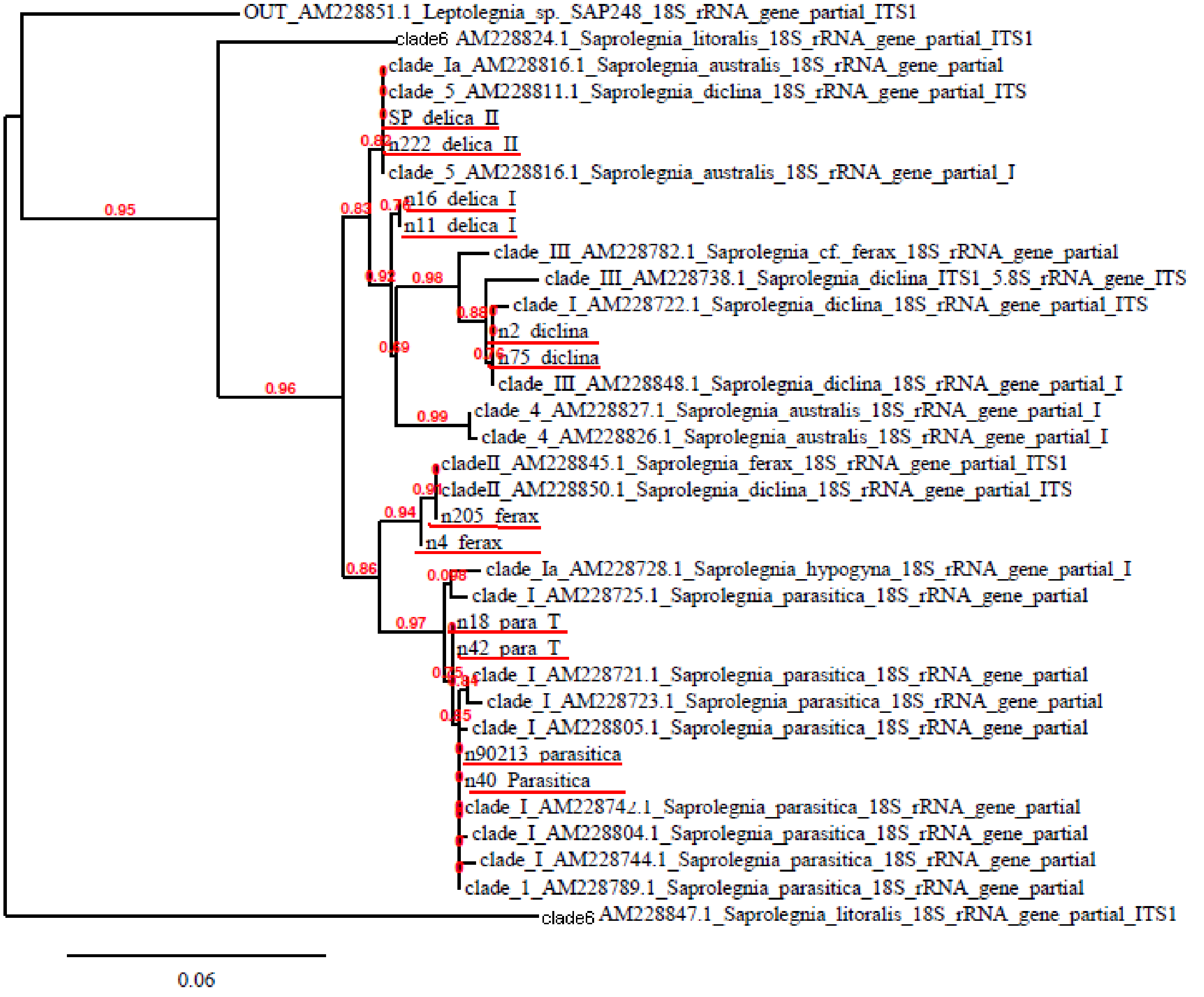


Figure 3. Phylogenetic tree constructed using Phylogeny.fr platform. **(One click mode). Tree created by choosing representative samples from each clade of Diéguez-Uribeondo *et al*., (2007) and two representative isolates of each *Saprolegnia* species in the current study. Samples from the current study are underlined in red and any samples showing NCBI accession number and clade designation are from the study of Diéguez-Uribeondo *et al*., (2007).**

Phylogenetic analysis of ITS sequence data from field collected isolates

A phylogenetic analysis of sequence data was conducted for all field isolates (457 isolates), employing different alignment tools and programs to determine if a consistent relationship could be revealed among isolates. However, due to the constraints of the analytical tools (maximum number of sequences input is 200 for nucleic acids), a subset of isolates was chosen for this detailed analysis. Most of the isolates were identified as *S. parasitica* (279) and *S. delica* II (71). Subsets of isolates were selected for comparative analysis as platforms such as Phylogeny.fr were not able to process more than 200 DNA sequence for analysis. As well, it is difficult to visualize such a large tree for interpretation. The sequence collection was screened manually to eliminate poor quality sequence data and multiple samples from the same location. A total of 122 sequences were selected and they included 32 *S. parasitica*, 17 *S. delica* II, 8 *S. delica* I, 20 *S. ferax*, 8 *S. diclina*, 4 *Aphanomycetes*, 8 *S. asterophora*, 13 *Pythium* species, 11 species of higher fungi, and 1 green algae (Figure 4).

All of the isolate information collected in the current study was compiled into a single table and associated sampling data was included for each species. This included data on the collection site, sample origin (e.g., fish lesions, eggs or water samples and species identity (Table 6).

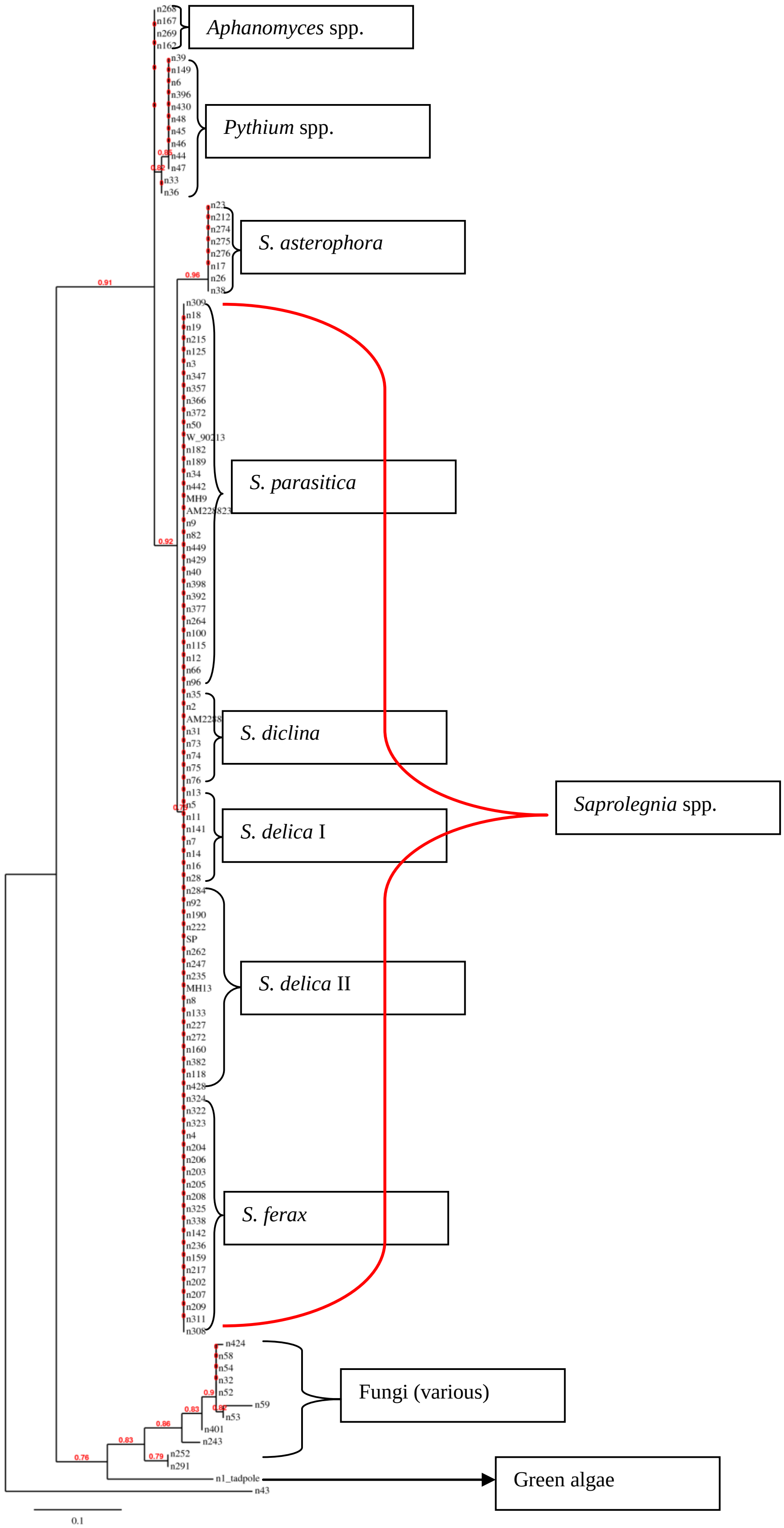

n268
n167
n269
n162
Aphanomyces spp.
n39
n149
n6
n396
n430
n48
n45
n46
0.85
n44
0.82
n47
n33
n36
Pythium spp.
n23
n212
n274
n275
n276
n17
0.96
n26
n38
S. asterophora
0.91
n309
n18
n19
n215
n125
n3
n347
n357
n366
n372
n50
W_90213
n182
n189
n34
n442
MH9
AM228823
n9
n82
0.92
n449
n429
n40
n398
n392
n377
n264
n100
n115
n12
n66
n96
S. parasitica
n35
n2
AM2288
n31
n73
n74
n75
n76
S. diclina
n13
n5
n11
n141
n7
n14
n16
n28
S. delica I
Saprolegnia spp.
n284
n92
n190
n222
SP
n262
n247
n235
MH13
n8
n133
n227
n272
n160
n382
n118
n428
S. delica II
n324
n322
n323
n4
n204
n206
n203
n205
n208
n325
n338
n142
n236
n159
n217
n202
n207
n209
n311
n308
S. ferax
n424
n58
n54
n32
0.9
n52
n59
0.82
0.83
n53
0.86
n401
n243
0.83
n252
0.76
0.79
n291
Fungi (various)
n1_tadpole
Green algae
n43
0.1

Figure 4. Phylogenetic tree created using Phylogeny.fr platform.

(one click mode) 122 sequences of various samples from the current study, as well as isolates from other genera were selected to generate this tree. Sequences selected from other genera for this tree construction included various kinds of fungi, for example, genus *Nectria*, *Penicillium*, and some unspecified soil fungi. One isolate of green algae was also included.

Table 6. Isolate sources for each identified species in the current study.

Species identities were confirmed by ITS region nucleotide sequence data. From routine sampling, the majority of isolates associated with fish, fish eggs and water were the species *S. parasitica*. Other species were also collected from some of these sources, but were not necessarily infective of pathogenic on fish or fish eggs.

Species	**Isolate origin by source (percentage of total number of isolates)**	**Number of confirmed isolates for each species (percentage of total number)**
Saprolegnia parasitica	Eggs 7 (2.3%) Fish 150 (49.8%) Water 144 (47.8%)	301 (65.6%)
Saprolegnia parasitica with T SNP	Egg 5 (55.6%) Fish 0 (0%) Water 4 (44.4%)	9 (2.0%)
Saprolegnia ferax	Eggs 0 (0%) Fish 0 (0%) Water 24 (100%)	24 (5.2%)
Saprolegnia delica I	Eggs 8 (88.9%) Fish 1 (11.1%) Water 0 (0%)	9 (2.0%)
Saprolegnia delica II	Eggs 4 (5.5%) Fish 3 (4.1%) Water 66 (90.4%)	73 (15.9%)
Saprolegnia diclina	Eggs 0 (0%) Fish 0 (0%) Water 7 (100%)	7 (1.5%)
Saprolegnia asterophora	Eggs 7 (87.5%) Fish 0 (0%) Water 1 (12.5%)	8 (1.7%)
Aphanomyces spp.	Eggs 0 (0%) Fish 0 (0%) Water 4 (100%)	4 (0.9%)
Pythium spp.	Eggs 0 (0%) Fish 0 (0%) Water 13 (100%)	13 (2.8%)
Various fungus	Eggs 1 (9.1%) Fish 5 (45.5%) Water 5 (45.5%)	11 (2.4%)
Total number of isolates =		**459**

An additional set of ITS sequences (71 sequences in total) was kindly provided by Petrisko *et al*. (2008), who had not deposited their sequence data in any public database. The naming system used in the data file provided did not agree with the published manuscript. It was therefore necessary to compile this sequence data into a single table containing all known data. The species designations were verified and updated by conducting searches of public databases (NCBI and BOLD) and subsequently compared to our own sequence database collection. A subset of representative isolates from their collection were selected and compared with some of our representative isolates in a phylogenetic analysis, again using platform Phylogeny.fr as described previously (Figure 5).

Most of the samples collected by Petrisko *et al*. (2008) were derived from the eggs of six species of amphibians; consequently, no comparison with our collection was made based on isolate origin.

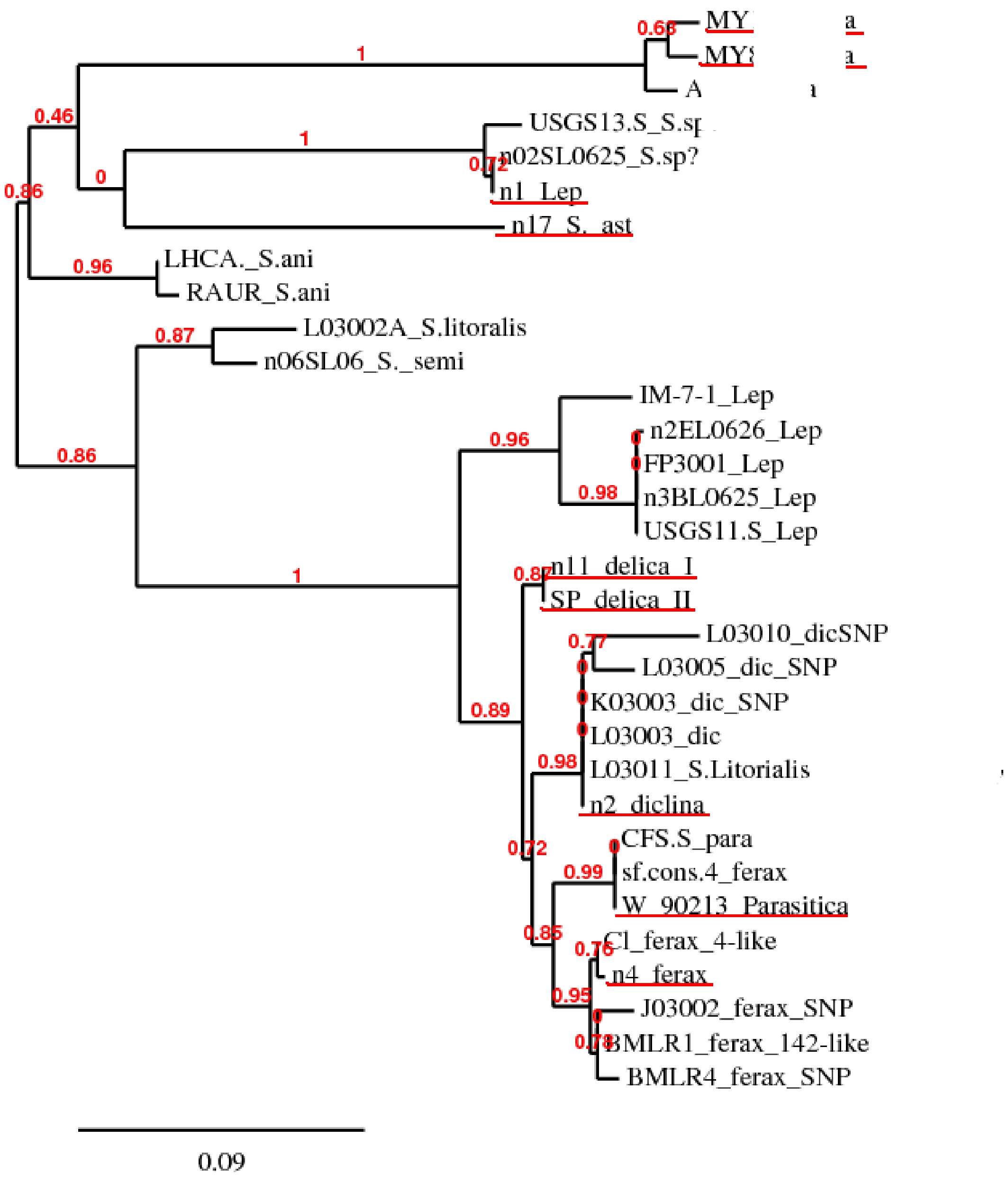


Figure 5. Phylogenetic tree constructed by using Phylogeny.fr platform.

(One click mode) One representative isolate of each *Saprolegnia* spp. as well as two *Achlya* samples we obtained from Malaysia (designated as MY1 and MY8) were included from this current study (underlined in red), in addition to representative samples from each group of Petrisko *et al.*, (2008) were included for this analysis. Based on information from the phylogenetic tree of Petrisko *et al.*, (2008), they separated into four genera (See also Figure 8); these included *Leptolegnia*, *Saprolegnia*, *Achlya*, and *S. semihypogyna*. The *Leptolegnia* sp. isolates were under the genus *Leptolegnia; S.diclina*, *S. australis*, one *S. litoralis*, *S. ferax*, *S. parasitica*,

S. hypogyna, S. salmonis, S. oliviae, S. anomalies, S. longicaulis and S. bulbosa **were in the genus** ***Saprolegnia*****. Various kinds of** ***Achlya*** **as well as** ***S. anisospora*** **and three designated as** ***Saprolegnia*** **sp. were grouped under the genus** ***Achlya*****. Note that the abbreviations from this figure represent the following taxa: S sp. indicates** ***Saprolegnia*** **sp.; Lep indicates** ***Leptolegnia*** **sp.; S. ast indicates** ***S. asterophora*****; S. ani indicates** ***S. anisospora*****; S. semi indicates** ***S. semihypogyna*****; dic indicates** ***S. diclina*****; AA indicates** ***Achlya americana*****.**

Various ***Saprolegnia*** **spp. from the current study were grouped with the Petrisko** ***et al.*****, (2008)** ***Saprolegnia*** **group, except** ***S. asterophora*****, which was shown to be more closely related to the genus** ***Achlya*** **of Petrisko** ***et al.*****, (2008). The same was true for the putative** ***Leptolegnia*** **sp. in the current study, which was shown to be more closely related to genus** ***Achlya*** **of Petrisko** ***et al.*****, (2008). Moreover, the two** ***Achlya*** **samples collected from Malaysia more closely related to** ***Achlya americana*****.**

Evaluating the 5.8S rDNA region in marker development for *Saprolegnia* spp.

The 5.8S rDNA region was investigated using the ITS nucleotide sequence data in the current study to determine if it could be useful for distinguishing species that are not *S. parasitica*. As well, the identities of some isolates were not clear due to inconsistent sequence data, or the failure of sequencing analysis hence examination the 5.8S region was expected to clarify their phylogeny.

Using both MUSCLE (Phylogeny.fr) and the ClustalW alignment tool from MEGA 5.05 phylogenetic and molecular evolutionary analyses were conducted (Tamura *et al.*, 2011). Sequence data from 457 isolates were aligned and included the ITS1, 5.8S, and ITS2 regions. The ITS1 and ITS2 regions were trimmed out, leaving only the 5.8S region for comparison of all *Saprolegnia* spp.

Viewing the data set in the conservation plot mode of BioEdit, it was apparent that all *Saprolegnia* spp. (*S. parasitica, S. delica I, S. delica II*, *S. ferax,* and *S. diclina,* with the exception of *S. asterophora*) had identical 5.8S (absence of SNPs). In order to make the isolate collection more manageable, I selected 55 sequences that were shown to have SNPs in 5.8S region, indicating that they were not genus *Saprolegnia* (except for the possibility of being S. *asterophora*), and manually aligned, and screened them using *S. parasitica* W (90213) as a reference.

Nucleotide sequence variability and species designation for *Saprolegnia* spp.

In an effort to assess the validity of the naming system used in this study, the level of nucleotide sequence variability was evaluated between species of the genus *Saprolegnia*. The isolate W (90213) of *S. parasitica* was used as a reference and its ITS sequence was compared manually in a pairwise fashion to confirm sequences of the other reference isolates of this genus (Table 7). The percent differences in nucleotide sequence for both of the variable regions ITS1 and ITS2 were determined and used to evaluate whether the naming system was appropriate. As well, the isolates 222 and 16, designated as *S. delica* II and *S. delica* I, respectively, were also compared in this study.

Table 7. Comparison of ITS sequences among *Saprolegnia* spp.

Comparison focusing on the variable regions within ITS1 and ITS2. For each comparison, the number of SNP differences between each sequence was determined and expressed as a percentage of the total number of nucleotides in that sequence. The ITS1 (167 bp) and ITS2 (332 bp) regions were evaluated separately. The isolates included W (ATCC 90213, *S. parasitica*), 40 (*S. parasitica*), 18 (*S. parasitica* with one SNP (T) in the ITS1 region), 4 (*S. ferax*, Canadian Collection of Fungal Culture), 205 (*S. ferax*), SP (*S. parasitica* ATCC 42062, determined to be *S. delica* II in the current study), 222 (*S. delica* II), 16 (*S. delica* I), 11 (*S. delica* I), 2(*S. diclina* ATCC 56851), and 75 (*S. diclina*). This data was compiled to validate the proposed naming system in the current study. This analysis hypothesized that nucleotide differences larger than 1% qualified the compared isolates to be different species.

Pairwise comparison of isolates	**Number of SNPs found in ITS1 region**	**Percent in ITS1 region**	**Number of SNPs found in ITS2 region**	**Percent difference in ITS2 region based on total**
W vs 18	1	0.6	0	0
W vs 40	1	0.6	0	0
W vs 4	3	1.8	17	5.1
W vs 205	3	1.8	17	5.1
W vs SP	7	4.2	16	4.8
W vs 222	7	4.2	16	4.8
W vs 16	6	3.6	17	5.1
W vs 11	6	3.6	17	5.1
W vs 2	5	3.0	23	6.9
W vs 75	5	3.0	23	6.9
222 vs 16	3	1.8	3	0.9

Results

Culture collection

Over the course of this study, more than 400 pure culture isolates had been collected from environmental samples (Table 6), their genomic DNA extracted and amplified by PCR with ITS specific primers. The ITS sequence information generated in this study provided useful information about *S. parasitica* and its closely related species.

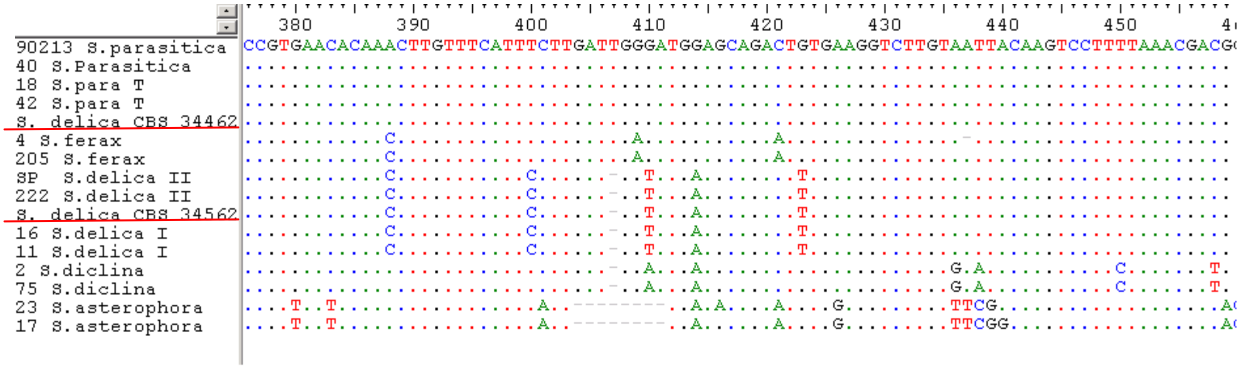
In this collection, *S. parasitica* was the most abundant species isolated from environmental samples, comprising 65.6% of all isolates, along with the *S. parasitica* containing one SNP (2.0%). The other species of interest included *S. ferax* (5.2%), *S. delica* I (2.0%), *S. delica* II (15.9%), *S. diclina* (1.5%), *S. asterophora* (1.7%), *Aphanomycetes* spp. (0.9%), *Pythium* spp. (2.8%), and higher fungi (2.4%) (Table 6). High numbers of isolates were derived from water and fish samples. Almost all isolates obtained from fish were identified as *S. parasitica* (94%). The second most abundant species recovered was *S. delica* II (15.9%), and most were derived from water samples. Almost 90% of *S. delica* I and *S. asterophora* were derived from fish eggs, although these species were only a small proportion of the isolates obtained in this study (2.0% and 1.7% respectively). The species of *S. ferax* and *S. diclina* were isolated only from water samples and did not appear to colonize on animal host.

The nucleotide sequence data compiled in this study has contributed to a sizable ITS data for the genus *Saprolegnia*. For the most part, ITS sequence was highly conserved within each of the species identified in this study, which included *S. parasitica*, *S delica* I, *S. delica* II, *S. ferax*, and *S. diclina*, and *S. asterophpra*. Based on the results of pairwise analysis of ITS1 and ITS2 regions between species (Table 7), isolates identified as *S.*

parasitica had less than 1% variation within both the ITS1 and ITS2 regions. When *S. parasitica* was compared with other *Saprolegnia* species, sequence variation ranged from 1.8% to 4.2% in the ITS1 region, and 4.8% to 6.9% in the ITS2 regions. Also, when comparing *S. delica* II (isolate 222) with *S. delica* I (isolate 16) in this study, there was a 1.8% sequence variation in the ITS1 region and 0.9% difference in the ITS2 region, which is very close to the threshold value of 1.0% for species delineations.

Based on our ITS sequence database, the 5.8S region sequence alone was sufficient to differentiate many non-*Saprolegnia* species from *Saprolegnia* species, the latter defined as *S. parasitica*, *S. delica* I , *S. delica* II, *S. ferax*, and *S. diclina* only. Non-*Saprolegnia* species, such as *Pythium* spp., *Aphanomyces* spp., and higher fungi were shown to have more than two SNPs in the 5.8 region; however, *S. asterophora* also fell into this group. Although certain SNPs patterns located in the ITS regions were not variable enough to support species-specific marker design, sequencing results revealed consistent SNPs patterns that allowed the categorization of unknown isolates into different phylogenetic species.

Searches for *Saprolegnia delica* in the NCBI database only provided five DNA sequence results; two were COI genes, one rRNA large subunit sequence, and two rRNA with ITS sequences. When the two ITS sequences were compared to our ITS database, one of the two (CBS34462) was highly similar to *S. parasitica* isolates (one SNP when compared with *S. parasitica*, and two SNPs when compared with *S. parasitica* with an extra T). The other putative *S. delica* sequence (CBS 34562) shares features with *S. delica* I in our collection (Figure 6). In addition, the species *S. delica* was not described in the study of Diéguez-Uribeondo *et al*., (2007); the designation in the current study is based upon the voucher sequences available in NCBI. (*S. delica* voucher CBS34462 – HQ643977.1 and *S. delica* voucher CBS34562 – HQ643976.1).



```
                       380       390       400       410       420       430       440       450       4
90213 S.parasitica  CCGTGAACACAAACTTGTTTCATTTCTTGATTGGGATGGAGCAGACTGTGAAGGTCTTGTAATTACAAGTCCTTTTAAACGACG
40 S.Parasitica     ....................................................................................
18 S.para T         ....................................................................................
42 S.para T         ....................................................................................
S. delica CBS 34462 ....................................................................................
4 S.ferax           ............C....................A...........A...............-......................
205 S.ferax         ............C....................A...........A......................................
SP  S.delica II     ............C...........C......-..T...A........T....................................
222 S.delica II     ............C...........C......-..T...A........T....................................
S. delica CBS 34562 ............C...........C......-..T...A........T....................................
16 S.delica I       ............C...........C......-..T...A........T....................................
11 S.delica I       ............C...........C......-..T...A........T....................................
2 S.diclina         ...............................-..A...A.....................G.A...........C.......T.
75 S.diclina        ...............................-..A...A.....................G.A...........C.......T.
23 S.asterophora    ....T..T.................A..--------..A.A....A....G.........TTCG...................A
17 S.asterophora    ....T..T.................A..--------..A......A....G.........TTCGG..................A
```

Figure 6. Alignment of representative *Saprolegnia* species in conservative plot view. **This plot view shows a partial sequence of the ITS2 region. This illustrates the high level of similarity among the DNA sequences of different *Saprolegnia* species (except *S. asterophora*, which are obviously very different), although there are unique SNP patterns that permit species delineation. Also note the two sequences of *S. delica* obtained from GenBank (underlined in red); one is highly similar to *S. parasitica* isolates, while the other highly similar to *S. delica* I in the current study.**

Primer development

Broad specificity primer sets were designed based on the nucleotide sequence information available from GenBank at the beginning of this study. In particularly, primer Fpr1A, when paired with ITS4 (Table 1), was able to resolve designation to genus level but could not resolve *Saprolegnia* species from each other. It was therefore possible to discriminate the genus *Saprolegnia* from *Pythium* or species of higher fungi, but not between *Saprolegnia* species, with the exception of *S. asterophora*.

Primers that contained a *S. parasitica*-specific SNP at the 3'end were tested for increased specificity (Table 2). It was determined that neither employing one Fpr 3' end primer with the universal Rpr ITS4, nor using both Fpr and Rpr primers both having 3' end SNPs were specific for *S. parasitica,* when tested on a collection 62 isolates (Table 4). Closely related *Saprolegnia* spp. co-amplified even at a high annealing temperature of 60^0C. Higher stringency did not improve specificity, as all *Saprolegnia* spp. showed amplified products of equal band intensity.

RFLPs

While the restriction enzyme digest of the ITS5/ITS4 PCR amplicons using *Bst*UI did generate a distinct pattern when comparing *S. parasitica* and *S. diclina*, but it was impossible to differentiate among *S. parasitica*, *S. ferax*, and *S. delica* (gel picture not shown).

Following the mapping of restriction enzymes sites in the ITS sequences of representative species (Table 5), it was determined that the enzyme *Bst*BI should be appropriate for the identification of *S. para*sitica by RFLP analysis. A restriction digest using *Bst*BI, following PCR amplification (primer Fpr1A and ITS4, annealing temperature of 60°C), cut the ITS sequence of *S. parasitica* once at nucleotide position 461 (Table 5). Restriction enzyme sites were absent from all other species tested, allowing *S. parasitica* to be distinguished from other genera, as well as from other closely related *Saprolegnia* species (Figure 7).

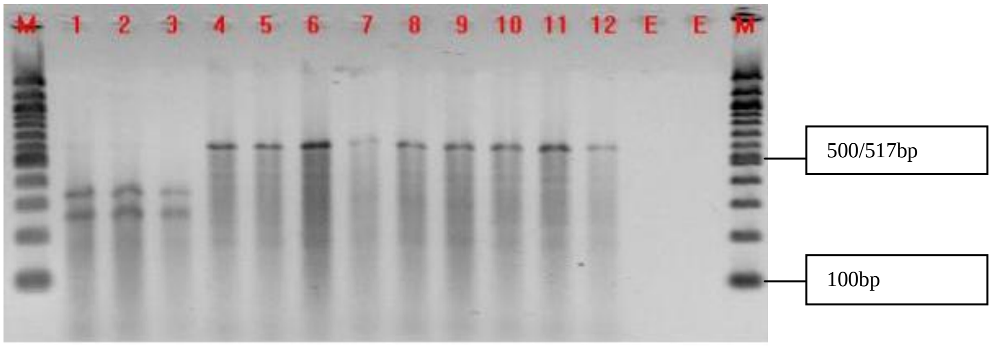


Figure 7. Restriction enzyme digest (*Bst*BI) of PCR products.

This result was generated with the primer pair Fpr 1A and ITS4, and subsequently digested PCR products with *Bst*BI. Samples include *S. parasitica* (lane 1 to 3), *S. delica* II (lane 4 to 6), *S. ferax* (lane 7 to 9), *S. diclina* (lane 10 to 12), E indicates empty lane, while M indicates 100 bp molecular marker (NEB).

When comparing RFLP analysis (*Bst*BI digest) with DNA direct sequencing protocols for the identification of pure sample isolates, both methods involve four steps, and are identical for the first two steps, which involve one round of PCR amplification, followed by electrophoresis and visualization. However, both protocols diverge from step three onwards, where the RFLP protocol involves restriction digestion with *Bst*BI and subsequent electrophoresis and visualization whereas the DNA direct sequencing protocol involves PCR purification, followed by sequence determination. The *Bst*BI digestion of the amplicons was both quicker and more cost-effective.

Phylogenetic analysis

Isolates of *Aphanomycetes* species, *Pythium* species, and various species of higher fungi were grouped into their distinct groups whereas *Saprolegnia* species, for instance, *S. parasitia*, *S. ferax*, *S. diclina*, and *S. delica* I and II were usually all fit into a bigger cluster where each species seems to equally related to each other. Moreover, *S. asterophora* species formed a distinct group which did not cluster with other *Saprolgenia* species (Figure 4).

Comparative studies

A total of 128 ITS nucleotide sequences for *Saprolegnia* spp. are available from the study of Diéguez-Uribbeonodo *et al.,* (2007), who deposited this data in the Genbank database. In their study, they categorized their sample isolates into six clades. They included 90 sequence belong to clade I; 5 sequences to clade Ia; 4 sequences to clade II; 8 sequences to clade III; 12 sequences to clade IV; 6 sequences to clade V; and 3 (including one outgroup sequences as clade 6. Clade I includes 88 *S. parasitica*, one *S. diclina*, and one *S. hypogyna*; clade Ia includes four *S. hypogyna* and one *S. australis*; clade II includes three *S. ferax* and one *S. diclina*; clade III includes six *S. diclina*, one *S. cf. ferax,* and one *S. australis*; clade IV includes 12 *S. australis*; clade V includes five *S. australis* and one *S. diclina*; clade VI includes two *S. litoralis* and one *Leptolegnia sp.* as an outgroup. Some of the designated species names had been changed since their publication; the names used in the current study were updated by a GenBank search, to verify whether there had been any subsequent changes to species designations since the publication date.

In the current study *S. parasitica*, *S. ferax, S. diclina*, *S. delica* I, and *S. delica* II were defined as various closely related *Saprolegnia* species and were equivalent to clades I, II, III and V of Diéguez-Uribeondo *et al.,* (2007) (*S. delica* I and II were in the same clade) . There were no sequences in our data set having high similarity to members of Diéguez-Uribeondo's clade IV isolates (Figure 3). However, comparing Petrisko's sequence collection to the Diéguez-Uribeondo set, one isolate, designated as sp.cons.4, was shown to have a high similarity with clade IV. This isolate was not shown on her published phylogenetic tree probably due to low bootstrap values in her study. This comparison was based only on nucleotide sequences, no physiological or morphological features were examined in the current study, or by Petrisko *et al*., (2008). Most of the samples collected by Petrisko *et al.,* (2008) were derived from the eggs of six species of amphibians; consequently, no comparison with our collection was made based on isolate origin.

Some comparisons can be made between the current study and the work of Diéguez-Uribeondo *et al.*, (2007) by extracting their data (Table 8). Their most abundant species is *S. parasitica*, which represented more than half of their total isolates (128 isolates distributed worldwide (68.0%), and more than half of those isolates were derived from fish (62.1%), which was similar to the current study conducted in Canada. Their second most abundant species was *S. australis* (18.0%). Most isolates of *S. ferax* were obtained from water samples, similar to the current study, while some *S. diclina* isolates were obtained from egg samples, unlike the current study.

Table 8. Isolate samples information extracted from Diéguez-Uribeondo *et al.*, (2007). Sample sources were categorized as fish egg, fish, water, and others [other than the previous three sources (Table 6)] and their abundance calculated in each category. Species designation was determined using the newly updated names obtained from recent NCBI blast search result.

Species	Isolate origin by source (percentage of total number of isolates)	Number of confirmed isolates for each species (percentage of total number)
Saprolegnia parasitica	Eggs 0 (0%) Fish 54 (62.1%) Water 31 (35.6%) Others 2 (2.3%)	87 (68.0%)
Saprolegnia diclina	Eggs 4 (80.0%) Fish 1 (20.0%) Water 0 (0%) Others 0 (0%)	5 (3.9%)
Saprolegnia hypogyna	Eggs 1 (20.0%) Fish 0 (0%) Water 4 (80.0%) Others 0 (0%)	5 (3.9%)
Saprolegnia australis	Eggs 7 (30.4%) Fish 13 (56.5%) Water 1 (4.3%) Others 2 (8.7%)	23 (18.0%)
Saprolegnia ferax	Eggs 0 (0%) Fish 0 (0%) Water 3 (60.0%) Others 2 (40.0%)	5 (3.9%)
Saprolegnia litoralis (outgroup)	Eggs 0 (0%) Fish 0 (0%) Water 1 (50.0%) Others 1 (50.0%)	2 (1.6%)
Leptolegnia sp. (outgroup)	Eggs 0 (87.5%) Fish 0 (0%) Water 1 (12.5%) Others 0 (0%)	1 (0.8%)
Total number of isolates =		**128**

In the study of Petrisko *et al.*, (2008) the phylogenetic analysis identified four main groups: *Leptolegnia*, *Saprolegnia*, *Achlya*, and *S. semihypogyna* (Figure 8). An updated sequence analysis of the data used in this study revealed mixed results. For *Leptolegnia*, confirmed in the NCBI database as *Leptolegnia* sp., whereas the BOLD database provided matched named *S. parasitica*. In the current study we also surveyed isolates from a fish farm in Malaysia and recovered two *Achlya* isolates (MY1 and MY8) which had 15 SNP difference between them. The *Achlya* isolate from Petrisko's group (2008), in particularly, the *Achlya americana* that Petrisko *et al*., (2008) designated as AA in their database, when compared with two *Achlya* isolates in the current study, had 21 and 19 SNPs differences respectively.

The isolates of *S. diclina*, *S. australis* (two), *S. litoralis* (one), *S. parasitica*, *S. ferax* (and a few more) constituted the group *Saprolegnia*. As well, three *Saprolegnia* sp. were identified in the *Achlya* group of Petrisko *et al*., (2008). When these *Saprolegnia* sp., sequences were compared to the NCBI database, they were simply identified as *Saprolegnia* sp., but from the BOLD database, they matched *Leptolegnia caudate*.

Isolate 1 from the current study as *Leptolegnia* sp. is actually more similar to the three *Saprolegnia* sp. that are placed in the *Achlya* group of Petrisko *et al*., (2008) and are less similar to the *Leptolegnia* group on the top of their tree (Figure 5 and Figure 8).

The study of Petrisko *et al*., (2008) also collected three samples from fish/ fish eggs, including IS-11-1 (CFS.S scrape from Chinook Hatchery), IS-12-1 (m.cons.4 from Rainbow trout eggs), and IS-13-1 (11.S from dead Rainbow trout from the wild). Based on ITS sequences comparison, these samples were equivalent to isolates of *S. parasitica*, *S. ferax*, (with 2 SNPs), and *S. delica* II (with 1 SNP), respectively, in the current study. These samples are underlined in red in Figure 8.

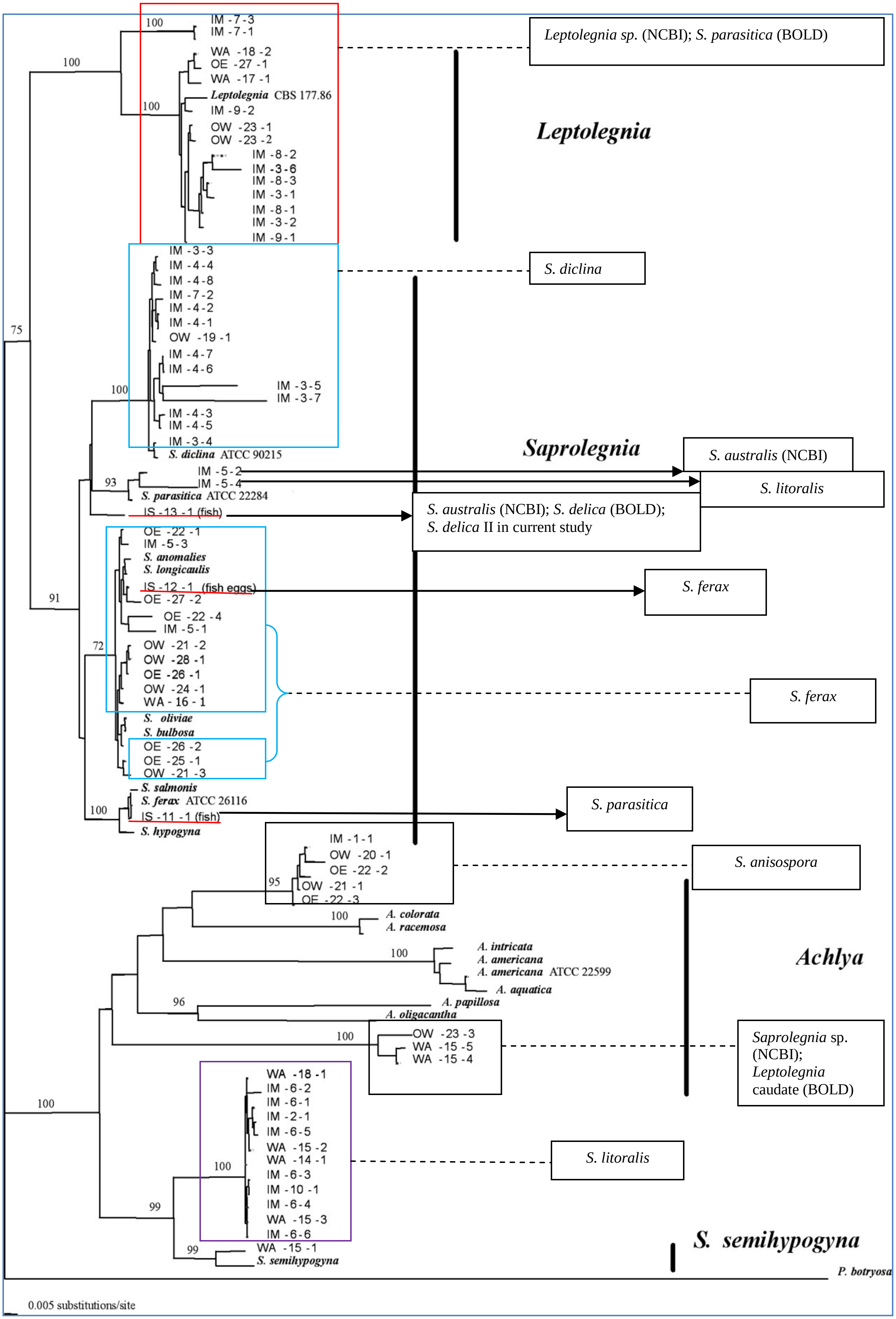

100
IM -7-3
IM -7-1
100
WA -18-2
OE -27-1
WA -17-1
Leptolegnia CBS 177.86
100
IM -9-2
OW -23-1
OW -23-2
IM -8-2
IM -3-6
IM -8-3
IM -3-1
IM -8-1
IM -3-2
IM -9-1
Leptolegnia sp. (NCBI); S. parasitica (BOLD)
Leptolegnia
75
IM -3-3
IM -4-4
IM -4-8
IM -7-2
IM -4-2
IM -4-1
OW -19-1
IM -4-7
IM -4-6
100
IM -3-5
IM -3-7
IM -4-3
IM -4-5
IM -3-4
S. diclina ATCC 90215
S. diclina
Saprolegnia
S. australis (NCBI)
93
IM -5-2
IM -5-4
S. parasitica ATCC 22284
S. litoralis
IS -13-1 (fish)
S. australis (NCBI); S. delica (BOLD);
S. delica II in current study
OE -22-1
IM -5-3
S. anomalies
S. longicaulis
91
IS -12-1 (fish eggs)
OE -27-2
S. ferax
OE -22-4
IM -5-1
72
OW -21-2
OW -28-1
OE -26-1
OW -24-1
WA -16-1
S. ferax
S. oliviae
S. bulbosa
OE -26-2
OE -25-1
OW -21-3
S. salmonis
S. ferax ATCC 26116
100
IS -11-1 (fish)
S. hypogyna
S. parasitica
IM -1-1
OW -20-1
OE -22-2
95
OW -21-1
OE -22-3
S. anisospora
100
A. colorata
A. racemosa
100
A. intricata
A. americana
A. americana ATCC 22599
A. aquatica
Achlya
96
A. papillosa
A. oligacantha
100
OW -23-3
WA -15-5
WA -15-4
Saprolegnia sp. (NCBI); Leptolegnia caudate (BOLD)
100
WA -18-1
IM -6-2
IM -6-1
IM -2-1
IM -6-5
WA -15-2
WA -14-1
100
IM -6-3
IM -10-1
IM -6-4
WA -15-3
IM -6-6
S. litoralis
99
99
WA -15-1
S. semihypogyna
S. semihypogyna
P. botryosa
0.005 substitutions/site

Figure 8. Jukes-Cantor Neighbor joining distance tree for Saprolegniaceae isolates. Isolates derived from amphibian eggs (n=65), fish scrapings or eggs (n=3) and reference (n=20). Figure from Petrisko *et al*., (2008) has been modified to include blastn (NCBI) search and BOLD search results for their sequence data (labelled boxes). Isolates underlined in red were the fish scrapings or fish egg samples that Petrisko *et al.*, (2008) collected during their study. Boxes labelled with solid arrows indicate species names designated to individual isolates, and boxes labelled with dotted lines indicate a group of isolates that were designated as the same species name.

Discussion

To resolve questions surrounding the naming and phylogenies of *Saprolegnia* species, we focused on the design of genetic markers utilizing ITS sequence, the compilation of field collected isolates and the analysis of ITS nucleotide sequence data from the genus *Saprolegnia*. Much useful information was generated from this data, which provided both an isolate collection and nucleotide sequence information for the ITS region. Isolate identity was related to isolate origin information and compared to other studies (Petrisko *et al*., 2008, Diéguez-Uribeondo *et al.,* 2007) to obtain in a broader picture of the overall occurrence of *S. parasitica* and related species from different sources and locations. As well, the ITS database revealed consistent SNPs in the ITS1 and ITS2 regions, allowing us to designate species names to unknown isolates and compare them to information gathered from searches of the NCBI and BOLD databases.

The study of Petrisko *et al*., (2008) examined egg samples from six different species of amphibians in Pacific Northwest, USA (Idaho-Montana, Idaho-South, Oregon-East, Oregon-West, and Washington). It was encouraging that none of her samples derived from amphibian eggs had ITS sequence that matched *S. parasitica* in the current study.

However, when comparing three of their samples that were derived from Chinook salmon hatchery, a rainbow trout egg, and a rainbow trout from the wild (underlined in red in Figure 8), the ITS sequences were highly similar to *S. parasitica*, *S. ferax*, and *S. delica* II, respectively as designated in the current study. Among 33 sequences they identified as belonging to the genus *Saprolegnia*, only one sequence, derived from a fish isolate matched *S. parasitica*, and another derived from wild rainbow trout was highly similar to *S. delica* II in the current study. Species abundance data from our entire collection (Table 6) showed that most isolates derived from fish lesions were *S. parasitica* (65.6%), and the second most abundant species was *S. delica* II. The high abundance of *S. parasitica* indicates that this species plays an important role as an infective agent of fish, particularly in a fish farm setting.

Saprolegnia diclina was believed to be highly pathogenic towards fish eggs (Robertson *et al.*, 2009). More than 50% (out of a total of 74 isolates) of the colonized egg cultures were identified as *S. diclina* by Fregeneda-Grandes *et al.*, (2007). In contrast, the current study (Table 6) found that isolates determined to be *S. delica* I and *S. asterohpora* were almost exclusively derived from fish eggs, whereas all *S. diclina* isolates were obtained from water samples. This discrepancy may arise simply due to former misidentification which was based solely on morphology in previous studies. However, it should be noted that the sample sizes of *S. diclina* and *S. asterophora* were very small in the current study. It is difficult to determine whether these *Saprolegnia* species are just saprophytes that happened to be found on dead fish egg samples or are in fact, pathogenic towards fish eggs.

The current study supports Molina *et al.,* (1985) who proposed that *S. delica* and *S. diclina* are quite different. *Saprolegnia dicina* and *S. parasitica* are, undoubtedly, not the same species. Moreover, *S. asterophora* is very different from all the samples that were designated as *Saprolegnia* species.

Diéguez-Uribeondo *et al.,* (2007) investigated phylogenetic and taxonomic relationships of *Saprolegnia diclina-Saprolegnia parasitica,* and suggested that all isolates of *S. parasitica* should be assigned to clade I. The creation of this clade for *S. parasitica* isolates is in agreement with the results of the current study (Figure 3). One of the common characteristics of samples in this Clade is that all *Saprolegnia* isolates obtained from lesions on salmonid fish fell into clade I (Diéguez-Uribeondo *et al.,* 2007). Similarly, the investigation of Fregeneda-Grandes *et al.,* (2007) also identified *S. parasitica* from all of their broodfish samples. A disproportionally large number of *S. parasitica* being isolated from fish samples in the current study indicates a very close association of this species with fish farms. Although no test had been done in this study to investigate whether *S. parasitica* is the primary pathogen toward fish, it is undeniable that *S. parasitica* is ubiquitous in this system. The study of Stueland *et al.,* (2005) supports this statement; they examined seven *Saprolegnia* spp. for their pathogenicity towards Atlantic salmon, and found two strains were significantly pathogenic towards challenged fish. The current and previous studies provide evidence that the presence of *S. parasitica* is likely to be tightly associated with freshwater fish aquaculture. This is also reflected by the identity of isolates collected by Petrisko *et al.,* (2008), who opportunistically collected samples from a fish hatchery.

Based on the ITS sequence comparison studies on Petrisko *et al*., (2008) and Diéguez-Uribeondo *et al*., (2007) there is still some taxonomic confusion of species in the genus *Saprolegnia*. A simple database search to establish the identity of an unknown ITS sequence suspected to represent a member of the genus *Saprolegnia* often results in conflicting results depending on the database queried (i.e BOLD Vs NCBI). Two database searches yield two very different species names for the same query sequence (Figure 8).Sometimes, one database can yield two (or more) different species names with the same degree of identity. This is not unexpected as it has been suggested that up to 20% of sequences in public databases like GenBank are misidentified (Bridge *et al*., 2003). Complication of species identification in *Saprolegnia* could also be attributed to the initial identification of *Saprolegnia* species based on morphology which were not revised following molecular analysis. This confusion was apparent when I attempted to search NCBI by species name, for instance "*S. delica*", in attempt to validate the adoption of this name for my *S. delica* (I and II) isolates. Combining the identification by sequence similarity with the clade designation generated by Diéguez-Uribeondo *et al.,* (2007) seemed to be quite reliable for discerning member of *S. parasitica*. For instance, if an unknown isolate had high sequence similarity to *S. parasitica* ITS sequences and group to clade I and was furthermore isolated from an infected fish, I recommend assigning that isolates as *S. parasitica*. This way, one can avoid misidentification and misunderstanding as sequence information is easier to communicate.

Based on the composition of our culture collection *Saprolegnia* species, other than *S. parasitica* and *S. delica* II, seem unlikely to be infective on fish, as they were usually found in low abundance and were not associate with fish lesions. If we could gather additional information about the conditions at the time of sampling for confirmed isolates, for example, the water temperature or any potential stress factors like vaccination of the fish taken place during the day of sampling etc., we might obtain a clearer picture of the correlation between the presence of *S. parasitica* and its activity as a primary infection agent against fish.

The development of genetic markers based on the ITS regions were found not to be sufficiently specific to resolve individuals at the species level, due to the limited sequence variation among different *Saprolegnia* species. Reliable discriminatory SNPs were located but their positions were not close enough to design a sensitive primer set that would amplify only the target species. In a similar vein, Oidtmann *et al.*, (2004) designed primers based on the ITS sequences for detecting *A. astaci*, but it was later pointed out by Ballestero *et al*., (2007) that this primer set was not specific enough and co-amplified other closely related species. A similar problem was also seen when distinguishing between closely related taxa of *Phytophthora* species (Schena and Cooke, 2006) when the ITS region was used to differentiate among closely related species. The primer sets designed based on ITS region in the current study are still be useful as they can serve as broad specificity markers to resolve samples down to the genus level. When combining these primers (Fpr 1A and ITS4), and a restriction digest (*Bst*BI) of the PCR products, this two-step diagnostic method can delineate *S. parasitica* from other closely related species. In addition, this RFLP is significantly more time and cost effective, when compared with ITS direct sequencing. To confirm the identity of an unknown sample as *S. parasitica*, this protocol may be completed within the same day, much faster than direct sequencing.

Nucleotide sequence information from the 5.8S region investigation showed that all closely related *Saprolegnia* species have identical sequences (except *S. asterophora*), whereas any other species have at least two or more SNPs within this region. This pattern was also seen in all the phylogenetic trees that were created using 5.8S sequence data and different methods of analysis (data not shown). Closely related species of the genus *Saprolegnia* (except *S. asterophora*) grouped together, while isolates from other genera fell into separate and groups. Sequence information from the 5.8S region also confirmed that *S. asterophora* was very distinct from a number of *Saprolegnia* species.

The large ITS sequences collection generated in this study will be added to the existing sequence databases (NCBI), and can contribute to DNA barcoding projects should researchers agree on using the ITS region as one of the genes for DNA barcoding. The use of DNA barcoding is an attempt to standardize the way we identify species (Stoeckle, 2003). The approach is to find a region (gene) that can be easily amplified by employing broad range primers for all organisms (Robideau *et al.*, 2011; Kress and Erickson, 2008; Stoeckle, 2003). Subsequently, sequences of that amplified region can be compared to easily accessible global reference library. Such library would be freely available. As the accessible library expands in size, the easier and the more accurate the identification of the unknown species would become; it could also be used to analyze the evolutionary relationship within and among groups (Robideau *et al.*, 2011; Kress and Erickson, 2008; Stoeckle, 2003). As well, as more people use it, the cost of operation decreases and the quality of data increases. A problem faced by many researchers is the choice of marker. Researchers agreed on the use of cytochrome c oxidase subunit 1 of the mitochondria region (COI) as a barcode for animal species (Robideau *et al.*, 2011; Hebert *et al*., 2004a;

2004b). In contrast, there is no consensus on what gene to use for plants. Hollingsworth *et al.*, (2009) suggested using a combination of Ribulose-bisphosphate carboxylase [rbcL] and maturase K gene [matK] as barcode for plants. As for organisms like *S. parasitia*, there is no consensus on which gene is preferable to use. Based on investigations by Robideau *et al.*, (2011), they recommended using both ITS and COI for the identification of oomycetes as the two datasets can complement each other.

The choice of ITS region for the barcoding of fungal species appears to be an appropriate starting point because there are already exists an extensive ITS sequence database (both in BOLD and NCBI), for many fungi and oomycete species. In addition, the results obtained in this project reveal that there is sufficient variation in ITS region to group these oomycete isolates into species. This indicates that DNA barcoding for identifying an unknown species is feasible once a comprehensive library be established. The public availability of more ITS sequences would facilitate identification of *S. parasitica* and other species of this genus.

The ITS region, particularly the ITS2 regions, seems to contain additional information for researchers to mine for information regarding species identification. Recently, researchers found that this nucleotide variable region also has a highly conserved two-dimensional structure that is conserved throughout eukaryotes. Combining the information of fast evolving nucleotide sequences in addition to more constrained and slowly evolving secondary structures, we can understand more about species boundaries, as well as interrelationship at higher levels (Koetschan *et al.,* 2010; Schultz *et al.,* 2006). An ITS2 database has recently been established and is easily accessed by researchers worldwide. This database has also implemented various tools for one to explore information obtained by secondary structures. For example CBCanalyzer, (CBC stands for compensatory base change), which can be used to estimate species boundaries (Koetschan *et al.,* 2010; Wolf *et al.*, 2005). Koetschan *et al.*, (2010) stated that using secondary structures for constructing phylogenies could potentially improve the stability of the constructed trees, as well as making the phylogenetic estimation more accurate. It maybe worthwhile to explore this idea for the expanded ITS2 data compiled in the current study.

In conclusion, we successfully cultured more than 400 field isolates comprising a diversity of oomycete and fungal species. Direct sequencing of those samples targeting ITS regions provided us with information that was used to clarify the taxonomy of *S. parasitica* and resolve individual species. The classification of *Saprolegnia* species based on ITS sequence results provides highly reliable and non-subjective approach to species identification which no longer relied on, sometimes transient, morphological features. In addition, through analysis of the *Saprolegnia* ITS sequence database, we devised a two-step protocol for quick identification of *S. parasitica*. This method allowed us to determine whether an unknown pure culture isolate was *S. parasitica* at low cost and within a single working day.

Chapter 3

Development and optimization of species-specific genetic markers for the monitoring of *Saprolegnia parasitica*, a pathogen of fresh-water fish

Introduction

In order to respond to an infection quickly, an accurate and rapid method of identification for the responsible pathogen is essential (Boyle *et al.,* 2004; Chen *et al.,* 2001; Chen *et al.,* 2000). Species identification based on morphology (Hulvey *et al.,* 2007; Leclerc *et al.,* 2000; Taylor *et al.,* 2000), mating analysis (Taylor and Fisher, 2003), or based on biochemical characteristics (Yuasa and Hatai, 1996) may be accurate, but these methods are time-consuming, require specific expertise, and may not always provide satisfactory results (Cordier *et al.,* 2007; Ke, *et al*., 2009; Oditmann *et al.,* 2004; Chen *et al.,* 2000; Molina, 1995). For example, immunoassays are highly specific, but are difficult and expensive to develop (Symondson, 2002).

The main objective of this project was to design *Saprolegnia*-specific markers that permitted the detection and differentiation of pathogenic *S. parasitica* from other closely related species. This relied on the development of appropriate markers to detect and identify pathogenic *S. parasitica* so that effective management protocols can be implemented. Such a detection method would also help researchers to understand the dynamics of this disease and monitor its occurrence under different physical conditions encountered in fish farms and contained aquaculture systems. In addition to being useful for species identification, DNA barcoding has the potential for deciphering taxonomic confusion for this genus (Kress and Erickson, 2008; Hebert *et al.,* 2004 a, b). This relies on a comprehensive library for specific gene loci to be established. An ITS2 ribosomal RNA database has been used to reveal both sequence and structural information

(Koetschan *et al.,* 2010; Schultz *et al.,* 2006). However, all of these methods require DNA sequencing of an unknown sample. It would be highly advantageous to be able to detect and positively identify *S. parasitica* without the need of performing DNA sequencing.

Our investigation of the *Saprolegnia* ITS regions revealed that there was sufficient nucleotide sequence variability within ITS regions to clarify some taxonomic confusion in the genus *Saprolegnia.* As well, it permitted the identification of unknown field collected isolates to species level, based on consistent single nucleotide polymorphisms (SNPs) within each species. However, as determined by the findings of Oidtmann *et al.,* (2006), the co-amplification of closely related species by PCR may be inevitable, especially when DNA template are high in concentration and there are only a few nucleotide differences between template DNA and designed-primers. We determined that, in order to unequivocally differentiate *S. parasitica* from other closely related species, a two-step diagnostic protocol, including amplification and sequence analysis was necessary. To improve upon this two-step method, it was necessary to assess other regions of the *S. parasitica* genome for variability to design better markers for the detection of *S. parasitica*, and subsequently test these markers by using our compiled isolate collection to validate their reliability.

One of many reasons the rDNA region was an attractive target for species-specific marker design is that it contained both coding and non-coding regions (Sumida *et al.,* 2004; Gardes and Bruns, 1993; Baldwin, 1992).The coding regions, being under greater selective pressure, are highly conserved and it is therefore helpful to use these regions for registering the alignment when comparing sequences (Sumida *et al.,* 2004). In contrast, sequences within the non-coding regions (ITS) are under more relaxed evolutionary pressure, and as a result, have developed more nucleotide variability (Sumida *et al.*, 2004; Gardes and Bruns, 1993; Baldwin, 1992). Information nested within this kind of highly variable region in the genome could be useful to distinguish *S. parasitica* from other *Saprolegnia* species.

Schena and Cooke (2006) evaluated the Intergenic Spacer Region (IGS) of the mitochondria rDNA (mt-IGS), the IGS region of the nuclear rDNA (rDNA-IGS), and ras-related protein (*Ypt*1) as loci for the development of diagnostic molecular markers for pathogenic species of *Phytophthora spp.* All of these loci contain both coding and non-coding regions and are therefore suitable for marker development. They found that the mt-IGS was good for diagnostic and phylogenetic studies. The rDNA-IGS was, comparatively, the least variable locus, while the intronic region of the*Ypt*1 gene was the most polymorphic and therefore useful for marker development. As well, it contained several conserved exon regions separated by variable introns that exhibited sufficient variation to allow the development of different molecular markers for discrimination at both genus and species level. The *Ypt*1 gene was found to only be suitable for *Phytophthora spp.*, and not suitable for the genus *Pythium* (Moorman *et al.,* 2002), for which the ITS region was more useful for identification. With this background as a guideline, three independent gene loci were selected from the genome of *S. parasitica* to determine which would be suitable for detecting *S. parasitica* in a simple, one step procedure that would not require DNA sequencing for species identification. The gene loci chosen for the development of the *S. parasitica* specific markers included the

Pumilio-family RNA binding repeat (Puf), Glutathionylspermidine synthetase (Gsp), and Thiazole biosynthetic enzyme (Thi4).

Pumilio-family RNA binding repeat (Puf locus)

The Puf locus includes a family of RNA-binding proteins that repress translation by binding to the 3' untranslated region (3' UTR) of a target mRNA (Spassov and Jurecic, 2003). This family of proteins can be found in a wide range of organisms that include yeast, fruit fly, zebra fish and humans (Spassov and Jurecic, 2003). Oomycete species like *Phytophthora infestans* also possess this gene (Cvitanich and Judelson, 2003). The Puf protein produced in *S. parasitica* (Puf1) has been shown to regulate the expression of phases between germination and secondary cyst formation (Andersson and Cerenius, 2002). The arrangement of highly conserved and highly variable regions of the Puf1 locus makes it an attractive marker because the conserved region can be used for identifying the target gene (Puf) whereas the variable region can be used for designing *S. parasitica*-specific maker. Puf has a highly conserved C-terminal RNA-binding domain containing eight Pumilio tandem repeats that are flanked by two sequences, Csp1 and Csp2. In addition, it has a highly variable N-terminal region that could potentially be useful for designing *Saprolegnia*-specific markers.

Glutathionylspermidine biosynthetic pathway genes

Glutathionylspermidine synthetase (Gsp) is a bifunctional (both synthetase and amidase) enzyme (Bollinger *et al.*, 1995). It is responsible for catalyzing the synthesis of glutathionlyspermidine by conjugating one molecule of spermidine and one molecule of glutathione as well as the reverse reaction, hydrolysis of glutathionylspermidine into glutathione and spermidine (Figure 9). To date, Gsp activity has been reported only in Trypanosomatidae and *Escherichia coli* (*E. coli*). In pathogenic trypanosomatids, this pathway (Figure 9) also has a unique subsequent reaction that catalyzes the biosynthesis of trypanothione, which involves conjugating one unit of glutathionylspermidine and glutathione (Fyfe *et al.,* 2008; Oza *et al.,* 2002; Smith *et al.,* 1995).

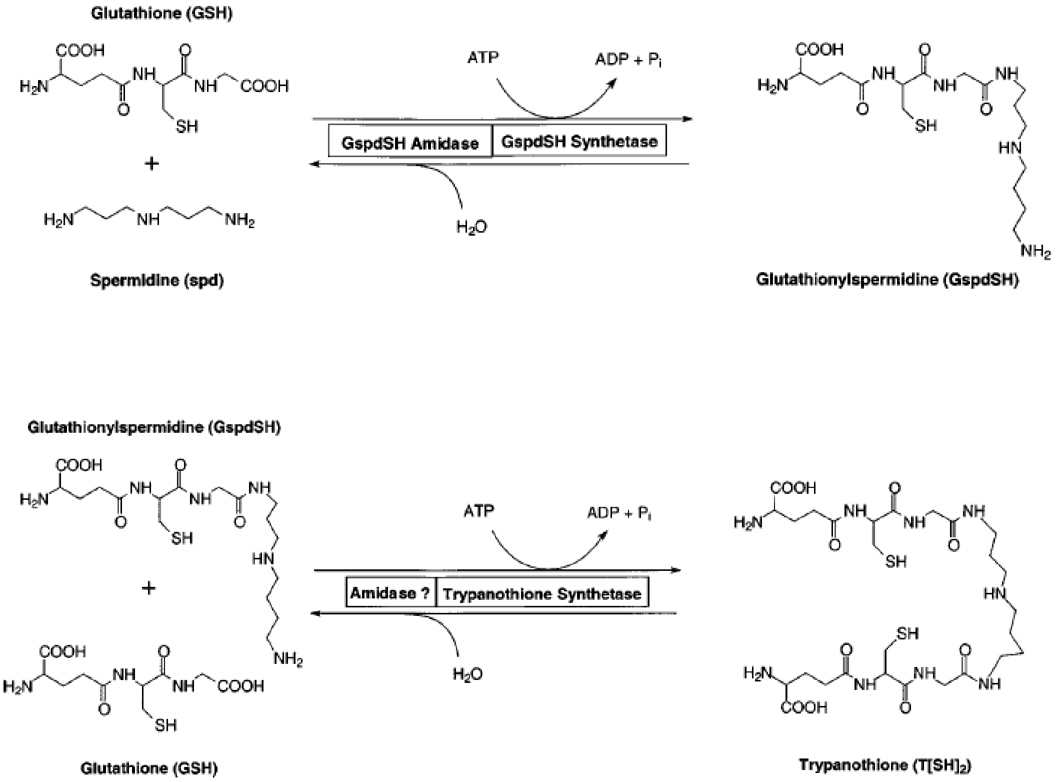


Figure 9. Glutathionylspermidine and trypanothione biosynthesis pathway.

This figure shows the glutathionylspermidine biosynthesis pathway (top), which can be found in both Trypanosomatidae and prokaryotes (e.g., *Escherichia coli*). The trypanothione biosynthesis pathway is shown on the bottom. This subsequent step is only found in Trypanosomatidae, and has been the focus of investigations that are attempting to use this step as a target for developing suitable drugs against the group of pathogens (Oza *et al*., 2002).

Some studies reported that trypanothione synthetase in *Trypanosoma cruzi* and *Crithidia fasciculate* can catalyze both the synthesis of glutathionylspermidine as well as trypanothione (Comini *et al.,* 2005; Oza *et al.,* 2002).These compounds are important for maintaining a redox-balanced intracellular environment, minimizing oxidative stress, and regulating the levels of low molecular weight compounds (Bollinger *et al.,* 1995; Fairlamb and Cerami, 1992).

The rationale for studying Gsp was based upon the previous detection of this gene in the species *Aphanomyces euteiches* (another member of the family Saprolegniaceae), but not in any *Phytophthora* species (family Pythiaceae), another major group of pathogenic oomycetes (Gaulin *et al.,* 2008).This gene was putatively involved in the biosynthetic pathway of trypanothione (Gaulin *et al.,* 2008). As the product of this unusual biosynthetic pathway has been shown to be crucial for survival and virulence of *Trypanosoma brucei brucei* and *Leishmania donovani* (Comini *et al.,* 2005; Oza *et al.,* 2002: Bollinger *et al.,* 1995), this gene, if present, could play an important role in *S. parasitica*. The more variable intronic regions of this gene may therefore be useful for diagnostic in *S. parasitica*.

Thiamine biosynthetic pathway gene

Sequence information for the thiazole biosynthetic enzyme gene locus could also provide the basis for designing a *S. parasitica* specific marker. Unlike species of the genus *Phytophthora*, which are reported to be thiamine auxotrophs, putative expressed sequence tags (ESTs) encoding thiamine biosynthetic enzymes for both *S. parasitica* and *A. euteiches* have been reported (Gaulin *et al.,* 2008; Torto-Alalibo *et al*., 2005). The Thi4 gene encodes the first enzyme responsible for thiazole formation in the thiamine biosynthesis pathway (Croft *et al.,* 2007), and encodes a short thiamine pyrophosphate element (TPP) that contains a very short conserved region and a short stretch of variable sequence, located 46 base pairs upstream of the putative *S. parasitica* thiazole biosynthetic gene (designated as SPRG_04526.1). This region also has a short intergenic region that potentially could be useful for designing markers, as there tends to be more sequence variation in the intronic and intergenic regions of genes than in coding regions (Schena and Cooke, 2006).

It was deemed that all of the above gene loci had potential nucleotide regions that could be *S. parasitica* specific, and therefore could be suitable for designing DNA-based markers.

Assessment of environmental samples for the presence of *S. parasitica*

From acquiring the initial sample to successfully obtaining a pure isolate takes both a lot of time and resources. It is preferable to devise a strategy for more rapid monitoring of this pathogenic species so that adequate management procedures could be taken in a timely fashion. With the identification of an appropriate *S. parasitica*-specific marker, our goal was to forgo all the culturing steps required for isolate identification and to amplify directly from environmental samples. It is quite common for researchers to utilize commercial kits for DNA extraction, whether it is from soil, fecal or water samples (Jones *et al.*, 2011; Kirshtein *et al.*, 2007; Ibekwe *et al.*, 2002). The filtration of environmental water samples, followed by total DNA extraction techniques was assessed in conjunction with our specific primers for the detection of *S. parasitica* to determine the feasibility of this procedure.

One of our main objectives was the development of S*aprolegnia*-specific molecular genetic markers for screening of environmental samples that may contain this organism. Markers having high resolution were designed based on Puf, Gsp, and Thi4 gene loci. It was anticipated that these markers would be able to distinguish *S. parasitica* from other close related *Saprolegnia* spp. and permit the development of a quick screening method for sampling for *S. parasitica* in water samples. This would enable close to real-time monitoring of aquaculture facilities for this pathogen.

Materials and Methods

Higher resolution markers development - Gsp, Puf and Thi4 gene loci.

Development of Gsp markers

Expressed sequence tags (ESTs) of *Aphanomyces eutechies* putatively identified as glutathionylspermidine synthetase were obtained from the databank for *Aphanomyces* ESTs (AphanoDB) (Madoui *et al.*, 2007). These EST sequences were subject to further examination using the NCBI and Universal Protein Resource (Uniprot, web resource) database to search for similar sequences among bacteria and some trypanosomatids, and for comparison with the *S. parasitica* CBS 223.65 genome portal (Broad Institute, web resource) to detect similar hypothetical proteins. ClustalW2 alignment tools (ClustalW2, web resource) were used to align the chosen representative protein sequences. The protein ClustalW output of the three alignments was then used as input for the Blocks multiple alignment processor that is built in the Consensus-Degenerate Hybrid Oligonucleotide Primers program (CODEHOP, web resource) to create conserved blocks that are necessary for the subsequent use by CODEHOP to generate degenerate primers from protein multiple sequence alignments. At the same time, the aligned multiple protein sequences were manually inspected to search for conserved regions among the different organisms, as potential primer sites. Designed primers were initially inspected for the compatibility of melting point and annealing temperatures by use of the Oligo Analysis tool and Oligo Analyzer 3.1 from Eurofins mwg|Operon (Operon, web resource) and Integrated DNA technologies (IDT, web resource), respectively. Selected primers were tested with representative isolate of *S. parasitica* (isolate MH5) and a closely related non-*S. parasitica* isolate (SP, later determined to be *S. delica* II) using a gradient PCR profile of 55°C to 65°C. A certain set of Gsp degenerate primer was shown to preferentially amplify *S. parasitica* (isolate MH5) at all tested annealing temperatures but not isolate SP, using a gradient of annealing temperatures from 55° to 65°(55.0, 55.2, 55.7, 56.6, 57.8, 59.1, 60.5, 61.8, 63.1, 64.2, 65.0, and 65.5°C). Following an initial denaturation phase (3 min at 94°), each PCR received 30 cycles of denaturation (45 sec at 94°), annealing (30 sec at 55° to 65°) and extension (1 min at 72°), a final extension (10 min at 72°), and was held at 4° until processed.

One set of the Gsp degenerate primers was shown to preferentially amplify only *S. parasitica* (MH5) at all tested annealing temperatures. In order to collect valid Gsp sequence information from *S. parasitica* isolates for the further refinement of primers, 22 isolates (Table 3) were tested with the same degenerate primers (Table 9) using the following PCR conditions. Following an initial denaturation phase (5 min at 94°), each PCR reaction received 5 cycles of denaturation (30 sec at 94°), annealing (30 sec at 55°), and extension (1 min at 72°), followed by 33 cycles of denaturation (30 sec at 94°), annealing (30 sec at 48°) and extension (1 min at 72°), a final extension (10 min at 72°), and was held at 4° until processed. Successful PCR amplicons were either sent to EuroFins mwg|Operon (Operon) (Huntsville, AL) for DNA direct sequencing when a clean, single PCR product was obtained, or cloned into the pGEM®-T vector system (Promega, Madison, WI, USA) prior to DNA sequencing.

Table 9. Initial primers designed targeting Gsp regions.

Degenerate primers designed and tested for the amplification of Gsp gene sequences in *S. parasitica*. Primers were initially designed based on putative Gsp gene sequence from an EST library of *A .euteiches*, and two hypothetical Gsp proteins of *S. parasitica*. When primer set were paired as 1/C, 1/1, 1/3, 2/C, 2/1, 2/3, 3/C, 3/1, 3/3, 4/C, 4/1, 4/3, 5/1, 5/3, and 5/C in a PCR reaction with template DNA of *S. parasitica*, their expected band sizes are approximately 900 bp, 950 bp, 1,200 bp, 530 bp, 540 bp, 820 bp, 520 bp, 510 bp, 790 bp, 360 bp, 350 bp, 630 bp, 300 bp, 600 bp, and 330 bp, respectively.

Name of Primer	**Sequence 5' to 3'**	**Symbol representation**
Fpr_Gsp_syn_1	GAYATTGTSGAGCAGAACGTSGAC	1
Fpr_Gsp_syn_2	GCGCTCGAGTAYGCGACVGAS	2
Fpr_Gsp_syn_3	ATGTTYCTCGACGCGACCGACT	3
Fpr_Gsp_syn_4	GGYCTCAAGGTGTACGAGTACAA	4
Fpr_Gsp_syn_5	MCCACGTCTTCCAGACRTT	5
Rpr_Gsp_syn_1	GCCGACTCGGCGTCGTG	1
Rpr_Gsp_syn_3	GGYGTGAGCKCAAACGACGA	3
Rpr_Gsp_syn_CodeHop_1	CGCCAGGCCCAGGTYTTCCANAC	C

Successfully amplified PCR products were separated by electrophoresis. Agarose gels containing expected PCR products were then excised and purified by using the QIAquick Gel Extraction Kit (Qiagen, Germantown, MD,USA), according to manufacturer's instructions. Subsequently, gel extracted sample DNA concentrations were determined using the Nanodrop®ND-100 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). For each sample, approximately 90 ng of DNA was added to each ligation reaction (5µL 2X buffer, 1µL T4 ligase, 1µL pGEM®-T vector) and incubated overnight at 4°C.

Electrocompetent *E. coli* 10G Elite cells (Lucigen, Middleton, WI, USA) were transformed by combining 3µL of the ligation mixture with 25 µL of electrocompetent cells, followed by incubation on ice for 5 minutes, and electroporation (1.3 V) in a chilled 2 mm cuvette (Fisher Biotech, Waltham, MA, USA). A 975 µL volume of recovery medium (Lucigen) was added immediately after electroporation, followed by incubation for 1 hour (37°C) with agitation (250 rpm). Undiluted cell culture (20 µL), as well as 20 µL to 150 µL of 1:10 diluted cell culture were spread on LB agar plates augmented with Ampicillin (0.05mg/µL), 100 µL ChromoMax^{TM} IPTG/X-gal Solution (FisherBioreagents, Waltham, MA, USA) plates, and incubated overnight at 37°C. Successfully transformed colonies of each sample were picked and incubated in LB broth with Ampicillin (Amp/LB 2µL/mL) at 37°C overnight with agitation (250 rpm). Subsequently, plasmid DNA extraction was done by using the QIAprep® Spin Miniprep kit, according to the manufacturer's instructions. Finally, to confirm plasmids contained an insert of the expected size, 1 µL of each purified plasmid DNA was digested with *Sac*I and *Sac*II (1 µL each), Buffer 4 (1.5 µL), 10X BSA (1.5 µL), a/c ddH$_2$O (up to a final volume of 15 µL) and incubated at 37°C for one and a half hours. Digested products were visualized by gel electrophoresis. Three to six clones containing inserts were sent to EuroFins mwg|Operon (Operon) for sequencing, using vector specific primers T7 and SP6.

Information obtained from sequencing results and the *S. parasitica* CSB223.65 genome portal were used for further analysis, with the aid of Primer 3 Input (V.0.4.0) (Primer 3, web resource), to refine degenerate primers and improve their specificity for *S. parasitica* (Table 10). The same isolate collection (Table 3) was used for testing, as well as the Diéguez-Uribeondo *et al*., (2007) PCR profile. Primer sets were tested and refined until they shown to preferably amplifying *S. parasitica*. These refined primer sets were then tested with an expanded isolate collection, including 62 DNA templates plus a negative control (Table 4), under a progressively more stringent annealing temperature.

Table 10. Re-designed prospective Gsp primer sets.

Prospective Gsp primers were re-designed based on information from initial degenerate Gsp primer tests (Table 9). With the assistance of Primer 3, primer sets were designed manually by considering their similarity to available Gsp sequences from GenBank (prokaryotes and Trypanosome sequences). Selected new primer sets are located in sequence regions that are not similar to the sequences available from GenBank. Subsequently, primer characteristics, (melting temperature, GC content, and the likelihood of primer dimer formation) were estimated using Oligo tools from Operon and IDT. The pairing of forward and reverse primer sets to test was based on the compatibility of their melting temperature. The table below only list Gsp primers that were shown to have promising results. Primers amplifying other *Saprolegnia* species equally well were not shown.

Name of Primer	**Sequence 5' to 3'**
Fpr_Gsp_839	TCTCGACGCGACCGACTA
Fpr_Gsp_1613	GTCGTTTGAGCTCACGCC
Rpr_Gsp_2012	CCGATACCCGTAGAATCCAA

Development of Puf markers

Sequences of the *S. parasitica* Puf gene were obtained from different databases (Uniprot, NCBI, and *S. parasitica* CSB223.65 portal) and examined manually to identify candidate regions for primer development. Primers were designed based on the hypothetical protein designated as SPRG_15362.1 from the *S. parasitica* genome portal. It had various intronic-like regions and had 162 nucleotide intergenic-like sequence upstream of the putative coding region, which was lacking in the other Puf sequences that were examined. This putative *S. parasitica* Puf sequence was subjected to a Blastn search (NCBI database) to identify low similarity regions, which were detected in the first 750 bp of SPRG_15362.1. By focusing on the low similarity region and the putatively *S.*

parasitica-specific region, Puf primers were designed with the assistance of Primer 3 (V.0.4.0). The new primers were analyzed by the use of Oligo analyzing tools and Oligo Analyzer 3.1 and were refined using a strategy similar to that described for the Gsp degenerate primers (Table 10).

Table 11. Refinement of Puf primers.

Puf primer were designed based on initial investigation from Uniprot, NCBI, and *S. parasitica* CBS 223.65 genome portal, with the assistance of Primer 3, Oligo analysis tools from IDT and Operon, as described for Table 10. After multiple tests with a small number of isolate templates (Table 3), a larger number of isolate templates (Table 4), a progressive higher stringency of annealing temperature and additional information revealed from PCR product sequencing, Puf primers were progressively refined and designed. Table below includes Puf primers that were shown to be reliable as *S. parasitica*-specific, other Puf primers determined not to be *S. parasitica*-specific are not included in this table.

Name of Primer	**Sequence 5’ to 3’**
Fpr_puf_130	ATACTATAAGCCGCGCCAAC
Fpr_puf_108	ACTTGTTGTACCATAGCCACTG
Fpr_puf_112	GTTGTACCATAGCCACTGTATC
Rpr_puf_310	CATCTCGATGCTGTTCTTGC
Rpr_puf_329	ACGGACTCTGCGAGAGAAAC

Development of Thi4 markers

Only a single putative Thi4 (SPRG_04526.1) could be found in the *S. parasitica* genome portal. This sequence was manually inspected, as well as its upstream sequences. About 46bp upstream of this hypothetical protein, there was a thiamine pyrophosphate riboswitch-like element (TPP). The entire putative Thi4 gene, as well as a 200bp upstream sequence was subject to a Blastn search (NCBI) and it was determined that the initial 140 nucleotides were unique to *S. parasitica*. Forward and reverse Thi4 primers were subsequently designed to amplify the initial 500 bp of this sequence, with the forward primer annealing site located in the first 140 bp of the sequence. Primers sets were initially tested with a small collection of 23 isolates (Table 3); the PCR products were purified using QIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions, and the purified DNA products were subsequently sent to Operon for DNA direct sequencing.

Total DNA extraction from environment water samples and initial primer tests

Environmental water samples (500 mL to 1 L volume) were collected from aquaculture facilities and filtered through a 0.45 µm or 0.22 µm filter column, according to manufacturer's instructions (MoBio, Carlsbad, CA, USA). Total DNA of each sample was extracted from filters containing the water filtrate by the use of the Powerwater® DNA isolation kit, according to the manufacturer's instructions (MoBio). Subsequently, the DNA concentrations of extracted samples were determined by use of the Nanodrop® spectrophotometer.

Template DNA samples, as well as positive and negative controls, obtained by total DNA extraction methods, were tested with primer sets for the Gsp (Gsp1613/2012) and the Puf (Puf 130/329) loci. Each PCR reaction included 10.0 ng DNA template, 0.2 mM dNTPs, 10% (v/v) 10X DreamTaq^{TM} Green buffer (Fermentas), 0.5µM of each primer, 0.5 U of DreamTaq^{TM} DNA polymerase (Fermentas), and ddH$_2$O to final volume of 10 µL. The PCR reaction profile used for the Gsp 1613/2012 primer set followed an initial denaturation phase (3 min at 94°), each PCR reaction received 30 cycles of denaturation (45 sec at 94°), annealing (30 sec at 61°), extension (40 sec at 72°), a final extension (10 min at 72°), and was held at 4° until processed. The Puf 130/329 primer set used a similar profile except the annealing temperature was different (30 sec at 60°, designated as WEBBY60) and with PCR reaction profile designated as WEB5860, which following an initial denaturation phase (5 min at 94°), each PCR reaction initially received 5 cycles of denaturation (30 sec 94°), annealing (30 sec at 58°), extension (1 min at 72°), followed by 33 cycles of denaturation (45 sec at 94°), annealing (30 sec at 60°) and extension (1 min at 72°), a final extension (10 min at 72°), and was held at 4°until processed.

Duplex PCR that employed a combination of both Gsp and Puf primer sets (Gsp 839/2012) and (Puf 130/329) had also been tested. Each PCR reaction included 10.0 ng DNA template, 0.2 mM dNTPs, 10% (v/v) 10X DreamTaq^{TM} Green buffer (Fermentas), 0.25 µM of each primer, 0.5 U of DreamTaq^{TM} DNA polymerase (Fermentas), and ddH$_2$O to final volume of 10 µL. The PCR reaction profile used for these duplex primer sets followed an initial denaturation phase (3 min at 94°), each PCR reaction received 30 cycles of denaturation (45 sec at 94°), annealing (30 sec at 60°), extension (40 sec at 72°), a final extension (10 min at 72°), and was held at 4° until processed.

Validation of marker design and comparison with ITS nucleotide sequence data

Following the optimization of Puf primer design for *S. parasitica*-specific amplification, selected pair, Puf 112/310 was evaluated with a group of confirmed pure culture isolates from the field collection. In addition, five of the test isolates were randomly picked to perform a PCR reaction for amplification of the ITS region, as described in the Materials and Methods section of Chapter 2, using the Diéguez-Uribeondo *et al*., (2007) PCR profile. The ITS amplicons were purified using the QIAquick PCR purification kit (Qiagen), according to manufacturer's instructions, followed by DNA directed sequencing by Operon. The results of the two methods were then compared for these isolates.

Results

Optimization of the single-locus primers

The Thi4 primers amplified *S. parasitica*, other *Saprolegnia* spp., as well as other more distantly related isolates, such as, *Leptolegnia* sp., and species of true fungi equally well. Partial direct sequencing results of PCR products amplified using Thi4 primers revealed that the Thi4 locus was highly conserved among all of the sequenced samples. As was discovered with the ITS regions, there are certain SNP patterns for distinguishing *S. parasitica* from non *S. parasitica* isolates. However, the SNPs found in the Thi4 locus were sparse and not located in close enough proximity to permit good primer design. In light of these findings, the design of primers for the Thi4 locus of *S. parasitica* was discontinued.

The initial tests with a small collection of isolates (Table 3) demonstrated that the Gsp and Puf loci primers were superior to the Thi4 for their ability to preferably amplify isolates of *Saprolgenia* spp. rather than more distantly related isolates. Primer sets for these loci were therefore progressively redesigned and refined, with the assistance of DNA sequencing data and alignment results. The refined primers were subsequently tested with a larger set of confirmed isolates (Table 4), as well as across a range of annealing temperatures.

Initially, results for the Gsp 1613/2012 primer set were found to be promising; when tested with 62 isolate templates (Table 4) and annealing temperature of 60°C, only *S. parasitica* isolates and some *S. delica* II were amplified by this primer set. The *S. parasitica* isolates produced prominent bands, whereas the *S. ferax* isolates produce fainter bands from the expected size. No products were amplified for *S. diclina*, *S. asterophora*, and *S. delica* I. Subsequently, the Gsp1613/2012 primer set was then tested using annealing temperature of 61°C, which further eliminated amplification of non *S. parasitica* isolates, although there was still the amplification of some larger PCR products for *S. parasitica* (Figure 10). The Puf 130/329 primer set, when testing using 5 cycles of annealing temperature at 58°C and subsequently 30 cycles of annealing temperature of 60°C, was also found to have potential, as it was shown to amplify most *S. parasitica* with a distinct band of about 400 bp; it also produced a series of consistent bands when amplifying with isolates of *S. delica* II (gel picture not shown). However, these primer sets were not perfect in the sense of being *S. parasitica*-specific.

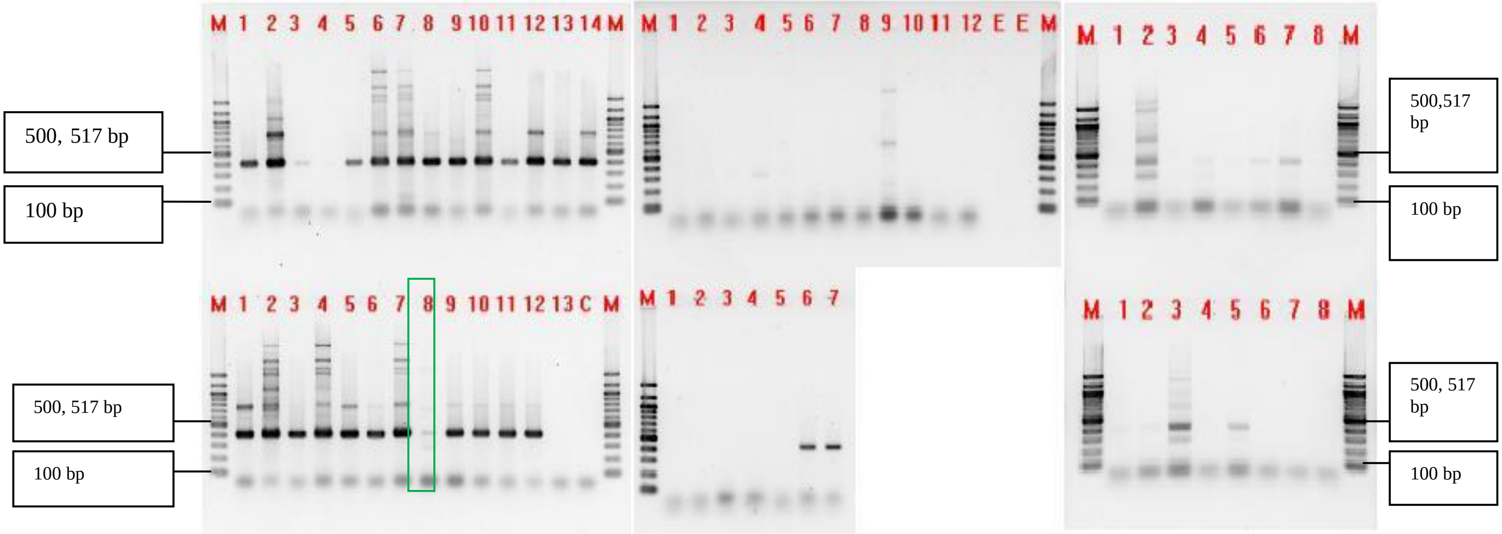


Figure 10. Amplification using primer set Gsp 1613/2012.

This PCR used annealing temperature of 61° C (Profile designated as WEBBY61); M indicates 100 bp NEB marker; C indicates water negative control; E indicates as empty lane. Gel picture in the far left are all *S. parasitica* samples, except lane 8 on the bottom of the gel, which is *S. delica* II (circled in green). Gel picture in the middle contains (top) *S. ferax* (lane 1 to 5); *S. diclina* (lane 6 to 9); *S. asterophora* (lane 10 to 12); (bottom) *S. delica* I (lane 1 to 4); *S. parasitica* SNP (T) (lane 5 to 7). Gel picture on the far right contains only *S. delica* II samples.

When the Gsp 839/2012 primers were used in duplex PCR with Puf 130/329 primers and the annealing temperature of 60°C, distinct band patterns could be found in most *S. parasitica*. An approximately 1,200 bp band, as well as 450 bp double band could be seen in most *S. parasitica* isolates; while most of the *S. delica* II isolates produced distinct 150 bp bands (Figure 11).

The primer sets of Puf 108/310 and Puf 112/310 were found to provide the optimal results when tested with the PCR reaction profile of annealing temperature at 60°and 62 isolate templates (Table 4). In particular, the Puf 112/310 primer pair amplified only the confirmed isolates of *S. parasitica* in our collection and did not appear to amplify any of the other test isolates of the genus *Saprolegnia* (Figures 12 and 13).

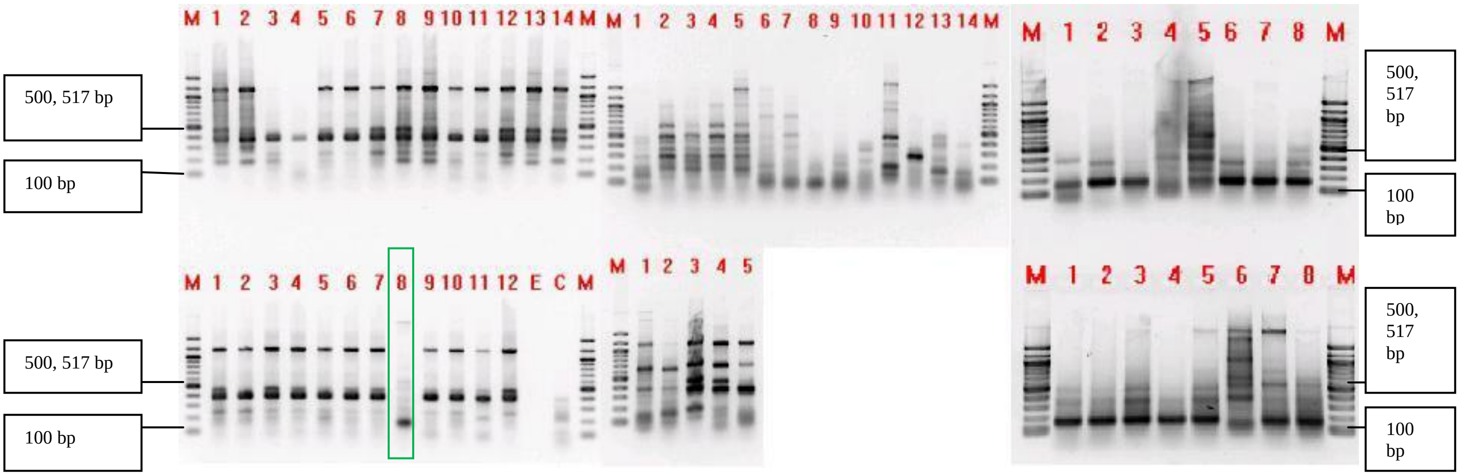


Figure 11. Amplification of duplex Gsp 839/2012 and puf 130/329 primer sets. Annealing temperature of this PCR was set at 60°C (PCR profile designated as WEBBY60); M indicates 100 bp NEB marker; C indicates water negative control; E indicates empty lanes. Gel picture on the far left are all *S. parasitica* samples, except lane 8 on the bottom of the gel, which is *S. delica* II (circled in green). Gel picture in the middle contains (top) *S. ferax* (lane 1 to 5); *S. diclina* (lane 6 to 9); *S. asterophora* (lane 10 to 12); *S. delica* I (lane 13to 14); (bottom) *S. delica* I (lane 1 to 2); *S. parasitica* with SNP (T) (lane 3 to 5). Gel picture on the far right contains only *S. delica* II samples.

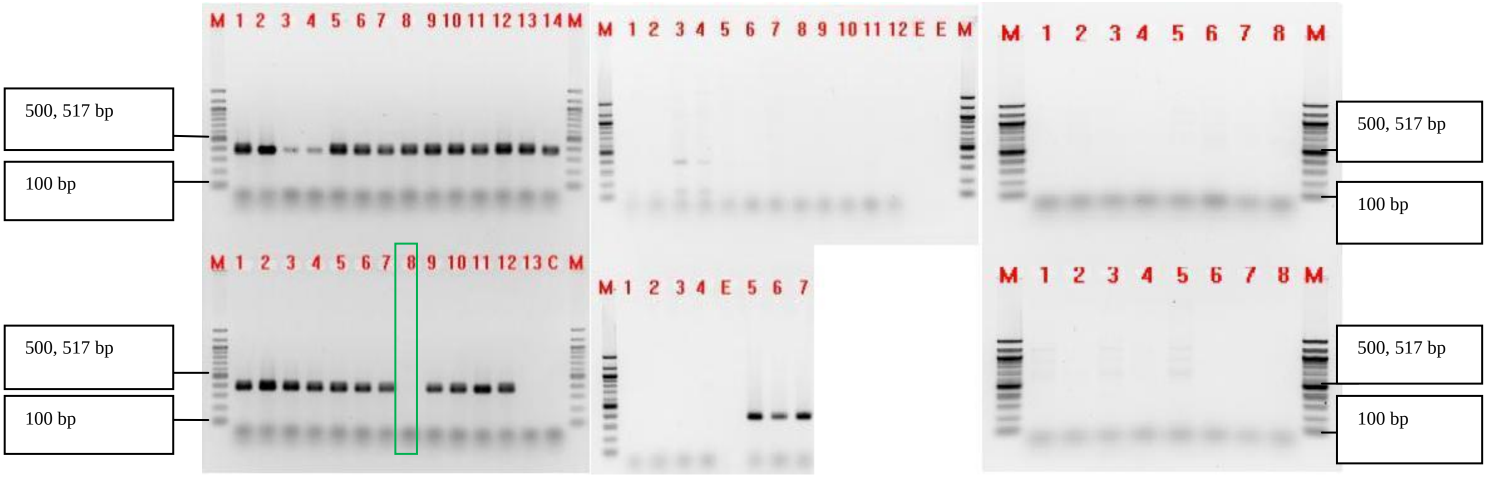


Figure 12. Amplification of primer set puf 108/310.

Annealing temperature of this PCR was set at 60°C (PCR profile designated as WEBBY60); M indicates 100 bp NEB marker; C indicates water negative control; E indicates empty lanes. Gel picture on the far left are all *S. parasitica* samples, except lane 8 on the bottom of the gel, which is *S. delica* II (circled in green). Gel picture in the middle contains (top) *S. ferax* (lane 1 to 5); *S. diclina* (lane 6 to 9); *S. asterophora* (lane 10 to 12); (bottom) *S. delica* I (lane 1 to 4); *S. parasitica* with SNP (T) (lane 5 to 7). Gel picture on the far right contains only *S. delica* II samples.

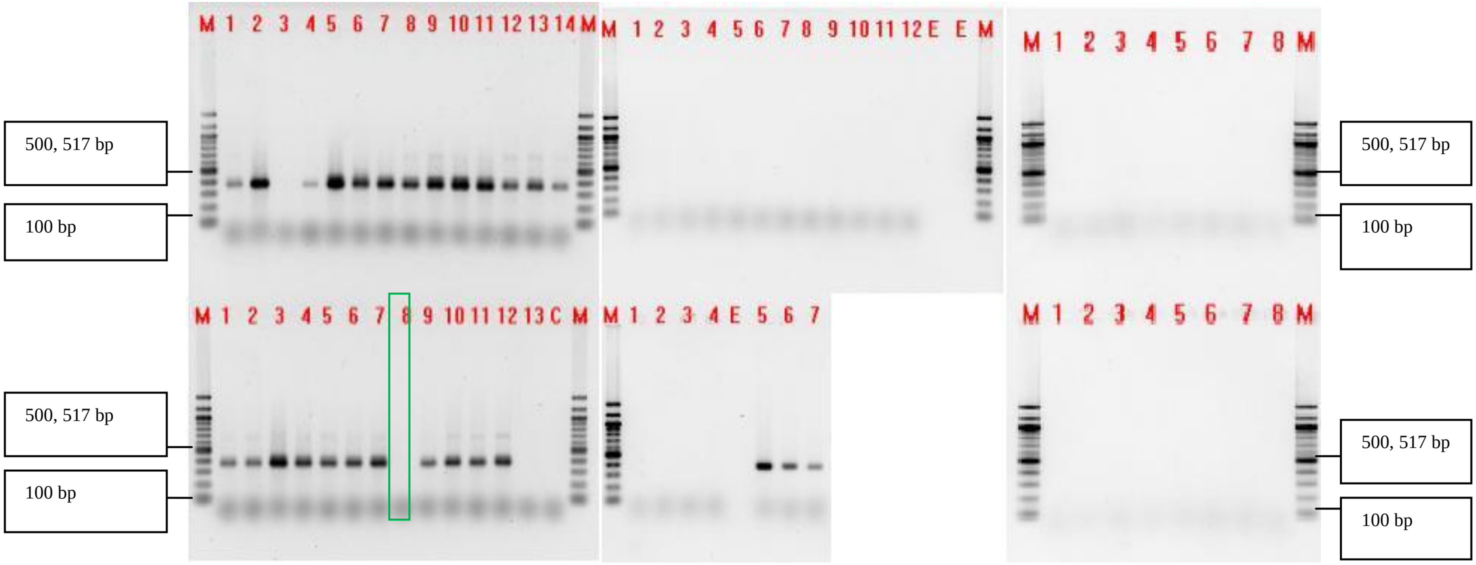


Figure 13. Amplification of primer set puf 112/310.

Annealing temperature of this PCR was set at 60°C (PCR profile designated as WEBBY60); M indicates 100 bp NEB marker; C indicates water negative control; E indicates empty lanes. Gel picture on the far left are all *S. parasitica* samples, except lane 8 on the bottom of the gel, which is *S. delica* II (circled in green). Gel picture in the middle contains (top) *S. ferax* (lane 1 to 5); *S. diclina* (lane 6 to 9); *S. asterophora* (lane 10 to 12); (bottom) *S. delica* I (lane 1 to 4); *S. parasitica* with SNP (T) (lane 5 to 7). Gel picture on the far right contains only *S. delica* II samples.

Water sampling protocol development and testing environmental samples for the presence of *S. parasitica*

All of the water samples collected from fish farms were filtered and total DNA was extracted. The DNA concentration of all samples was subsequently measured and determined to range from approximately 2 ng/µL to 120 ng/µL. All of the extracted samples had been eluted with 100 µL of PW6, as provided by the manufacturer. All steps of this procedure, including water filtration, total DNA extraction, and DNA concentration measurement, could be completed within one working day. Occasionally, some water samples required use of more than one filter due to clogging, although there did not appear to be suspended debris in the water samples, as visible debris would have been filtered prior to this procedure.

For the initial screening of water samples, the Gsp 1613/2012 and the Puf 130/329 primer sets had been used to perform PCR testing of the total DNA extraction samples (gel picture not shown), but no conclusive identification of unknown samples could be obtained using these primer sets.

Once primer design and PCR conditions had been refined, the set Puf 112/310 was found to provide the most accurate identification by amplifying only *S. parasitica* isolates. A small number of field collected pure culture samples (19 of them together with a positive and negative water control) were selected for PCR amplification using Puf 112/310 primers, and among them, five were randomly selected for PCR amplification of the ITS regions, followed by DNA direct sequencing. All of the five samples were shown to have the expected size (365 bp) of PCR amplification product for Puf when amplified with the Puf 112/310 primer pair, confirming them to be *S. parasitica*. In addition, these identifications were consistent with the results obtained from DNA direct sequencing, which confirmed the nucleotide sequences of all samples to be *S. parasitica*.

Discussion

Reliable markers developed based on Gsp and Puf loci

Reliable markers based on the Gsp and Puf loci, were developed for the identification of *S. parasitica*. Among all of the primers tested, the Puf 112/310 primer pair was shown to have the highest specificity for *S. parasitica*. When tested against a larger number of isolates (Table 4), these primers selectively amplified *S. parasitica* DNA, and not that of other closely related species (Figure 13). The results for analysis of culture isolates from randomly picked samples were also encouraging, as the Puf 112/310 primers provided the same answer as determining nucleotide sequence for these unidentified samples.

Use of commercial kits for water filtration and total DNA extraction for the purpose of testing species-specific markers

The procedures of filtering environmental water samples, followed by total DNA extraction using commercial available kits (MoBio) were found to be both time-saving and straight forward, when compared with the laborious isolation and subculturing of procedures required for pure culture isolates. Unfortunately, I did not have a chance to test any MoBio extracted samples with Puf 112/310 primer pair to assess its specificity on total DNA extractions of environmental samples. The inconclusive results obtained from the previously tested primer sets of Gsp 1613/2012 and Puf 130/329 for environmental samples were likely due to the lower specificity of these two primer sets, when compared to Puf 112/310. This lower specificity would be complicated by the possible presence of DNA template from more than one *Saprolegnia* spp. in the same total DNA sample. As well, the total DNA concentration of samples was found to vary greatly, which would likely affect the detection limit of the markers, if there was competing template DNA from other micro-organisms.

However, certain sets of the higher resolution *Saprolegnia* markers (for example, Gsp 1613/2012, Gsp 839/2012, and Puf 130/329) are still very useful for confirming the identity of an unknown pure culture, without the need of performing DNA sequencing. When a broader range of isolates were evaluated, different *Saprolegnia* spp. tested produced distinct band patterns, depending on which set of primers were used. For instance, if an unknown pure isolate is tested with Puf 112/310 and shows no amplification, and is subsequently tested with Gsp 839/2012 and shows distinct double bands from 400 to 500 bp, it would be most likely that the unknown sample is *S. delica* II.

Further validation of Puf 112/310 and the evaluation of inoculum potential

An evaluation of the inoculum potential required to elicit a saprolegniosis outbreak would be useful for us to further understand and manage this disease. It is established that certain *Saprolegnia* species can be pathogenic (Fernández-Benéitez *et al*., 2008; Ke *et al*., 2009; Stueland *et al*., 2005; Diéguez-Uribenodo *et al*., 1994; Signhal *et al*., 1987). This pathogenicity can however be variable, since infected fish do not necessary die from the disease. Hussein and Hatai (2002) challenged fishes with two strains of *S. parasitica* and determined that the fish challenged with low doses, (2 x 10^2 spores/L) did not die following exposure. Stueland *et al*., (2005), who tested seven strains of *Saprolegnia* species for their pathogenicity towards Atlantic salmon; found that strains E and K, caused 89% and 31%, respectively, cumulative mortality in challenged salmonids. They used isolate ATCC 90213 as a positive control, which also caused 18% mortality. This isolate was also used in the current study as a reference isolate. Although isolate ATCC 90213 was found not to be as virulent as strain E and K, I believe that, if the concentration of ATCC 90213 is high enough, it can still be disease causing. Also note from the publication of that study did not provide any ITS sequence information for strains E and K, hence I cannot confirm their identity, or determine if they are different species.

To characterize the exact strain and measure the amount of inoculum of *S. parasitica* required for the outbreak of saprolegniosis, we can follow the strategy of Fernández-Benéitez's group (2008) to adopt Koch's postulate, combined with quantitative real time PCR (qPCR) methods.

Koch's postulates can be used to determine whether our target, *S. parasitica*, is responsible for saprolegniosis and mortality of fish in fish farms. Koch's postulate, according to Fernández-Benéitez (2008), has the following requirements: (1) the target (*S. parasitica*) is constantly associated with the disease; (2) it was isolated from infected individuals (fish) and could be grown in pure culture; (3) the original disease and symptoms (saprolegniosis) occur in healthy individuals when they were inoculated with the pathogen from a pure culture; and (4) the same strain of pathogen could then be re-isolated from challenged individuals (fish) under experimental conditions. It should be noted in the current study that the *S. parasitica* isolates we obtained from our field collection fulfill the first two conditions of these postulates. If we can proceed to perform challenge tests on fish to test the next two conditions, it can be demonstrated clearly that

S. parasitica is responsible for saprolegniosis at our study sites. In addition, when combined with a quantification method, we can measure the amount of *S. parasitica* inoculum required for the infection to take place and possibly determine the threshold level above which symptoms are observed. In order to assess the sensitivity of the primer set Puf 112/310, it should be tested with a range of known concentrations of DNA templates. For instance, known concentrations of *S. parasitica* zoospores could be prepared by serial dilutions and, subsequently used to test the primer set Puf 112/310 to determine its detection limit. Alternatively, a known concentration range of DNA template samples could be prepared by serial dilution and used to determine primer sensitivity. With the knowledge of the DNA content of a single zoospore, it should be possible to estimate the minimum number of zoospores that can be detected by PCR in a known volume of water.

qPCR and future directions

Quantitative PCR (qPCR) is considered to be superior to conventional PCR, in that it can detect and quantify target DNA templates simultaneously. In addition, qPCR does not require post PCR manipulation, thus saving time and avoiding sample contamination (Hochwimmer *et al.,* 2009; Heid *et al.,* 1996). The use of qPCR should provide higher specificity because it should be revealed from the melting curve of the amplified product if there is any mismatch between the primer and the target sequences (Ballesteros *et al.,* 2007). The down side of this method is that the probes are comparatively expensive, and it does require more sophisticated instrumentation, since a qPCR thermocycler is required. With the development of a *Saprolegnia*-specific marker, the routine use of a total DNA extraction protocol for water filter samples and qPCR would be very useful for disease monitoring. Detection and quantification of *S. parasitica* in water samples from fish farms would allow appropriate disease management protocols to be in place. Once the specificity of Puf 112/310 primer set has been further validated, these primers may be used to design a qPCR probe for the quantification of inoculum levels.

General discussion, conclusions and future directions

The ultimate objectives of this project were the development of genetic markers that can identify *Saprolegnia parasitica* and distinguish this species from other closely related species in the same genus, as well as the clarification of the taxonomy of the genus *Saprolegnia*. To fulfill these objectives, field samples were collected from salmon hatcheries and aquaculture facilities on Vancouver Island and an extensive ITS sequence compiled. This data provided information which permitted the designation of species name to unknown *Saprolgenia* samples, based on the occurrence of specific SNP patterns in the ITS regions. The high abundance of *S. parasitica* found in our study sites confirmed our hypothesis of the close association of *S. parasitica* and the infection of fish and fish eggs in aquaculture facilities. We also found that *S. asterophora* was not closely related to any our studied *Saprolegnia* species. In addition, a relatively quick and cost effective RFLP protocol was established that could identify *S. parasitica* from other closely related *Saprolegnia* species. To further refine our *S. parasitica* detection method, three gene loci other than the ITS regions were also examined. A highly *S. parasitica*-specific marker, namely, Puf 112/310 was developed, and a quick water screening method that will facilitate the *S. parasitica* detection procedure had also been investigated and was deemed to be practical.

The ITS database established in this study will be added to the public accessible databases (NCBI), and is anticipated to contribute to projects such as DNA barcoding and ITS2 database. This work also allows us to distinguish *S. parasitica* from other closely related species, thus avoiding the problems associated with subjective morphological assessments, since identification solely based on morphology is not reliable and prone to misidentification of species.

Reliable identification is the first step towards understanding pathogenicity of the devastating fish pathogen *S. parasitica*. With the development of Puf 112/310, a quantitative method, such as qPCR probe, can be designed and used to evaluate the inoculum threshold required for the development of saprolegniosis, as well as being used for monitoring pathogen conditions in the water column. In aquaculture, many different factors may contribute to the susceptibility of fish to disease, including stress to the fish immune system by temperature change, maturity of the fish, diets, overcrowding and vaccinations. We can now isolate each of these components and screen them to identify threshold levels of *Saprolegnia* spp. to test for interactions and develop a better management protocol.

Moreover, we can continue the third and fourth postulates of Koch's to determine whether the isolate that we identified as *S. parasitica* is responsible for mortality of fish in a fish farm. Information from other gene loci in this genus may expand our understanding of the relationship between various highly infectious and relatively benign *Saprolegnia* species, or even their relationship with other oomycetes. Lastly, I agree with Diéguez-Uribeondo *et al*., (2007) of designating the name *S. parasitica* to all isolates of clade I. I recommend reserving the name *S. parasitica* exclusively to isolates that are demonstrated to be pathogenic towards fish (derived from fish lesions or from the water column of a fish farm) and for isolates that have the same (or high similarity) ITS nucleotide profile as the isolates we investigated. It is hoped that information gained in this study not only helps to keep this pathogen at bay, but also contributes to a better understanding of *Saprolegnia* species and the roles they play in the ecosystem.

Bibliography

Ali, E.H. (2005). Morphological and biochemical alternations of oomycete fish pathogen *Saprolegnia parasitica* as affected by salinity, ascorbic acid and their synergistic action. Mycopathologia. 159, 231-243.

Aller-Gancedo, J.M., and Fregeneda-Grandes, J.M. (2007). Comparative efficacy of Pyceze ® (bronopol) in controlling mortality of brown trout *Salmo trutta* eggs. Aquaculture Research. 38, 618-624.

Almeida, A., Cunha, A., Gomes, N.C.M., Alves, E., Costa, L., and Faustino, M. (2009). Phage therapy and photodynamic therapy: low environmental impact approaches to inactivate microorganisms in fish farming plants. Marine drugs. 7, 268-313.

Andersson, M.G., and Cerenius, L. (2002). Pumilio Homologue from *Saprolegnia parasitica* specifically expressed in undifferentiated spore cysts. Eukaryotic Cell. 1, 105-111.

Baldwin B.G. (1992). Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Molecular Phylogenetics and Evolution. 1, 3-16.

Bangyeekhun, E., Quiniou, S.M.A., Bly, J.E., and Cerenis, L. (2001). Characterization of *Saprolegnia* sp. isolates from channel fish. Disease of Aquatic Organisms. 45, 53-59.

Ballesteros, I., Martín, M.P., Cerenius, L., Söderhäll, K., Tellería, M.T., and Diéguez-Uribeondo, J. (2007). Lack of specificity of the molecular diagnostic method for identification of *Aphanomyces Astaci.* Bulletin Français de la Pêche et de la Pisciculture. 385, 17-24.

Blaustein, A.R., Hokit, D.G., O'Hara, R.K., and Holt, R.A. (1994). Pathogenic fungus contributes to amphibian losses in the Pacific Northwest. Biological Conservation. 67, 251-254.

Bollinger, A.R., Kwon, D.S., Huisman, G.W., Kolter, R., and Walsh, C.T. (1995). Glutathionylspermidine metabolism in *Escherichia coli.* Journal of Biological Chemistry. 270, 14031-14041.

BOLD. The Barcode of Life data system. Available at http://www.boldsystems.org

Boyle, D.G., Boyle, D.B., Olsen, V., Morgan, J.A.T., and Hyatt, A.D. (2004). Rapid quantitative detection of chytridiomycosis (*Batrachochytrium dendrobatidis*) in amphibian samples using real-time Taqman PCR assay. Diseases of Aquatic organisms. 60, 141-148.

Bridge, P.D., Roberts, P.J., Spooner, B.M., and Panchal, G. (2003). On the unreliability of published DNA sequences. New Phytologist. 160, 43-48.

Broad Institute. (2010). *Saprolegnia parasitica* database. Available at http://www.broadinstitue.org/annotation/genome/Saprolegnia_parasitica/Info.html

Chen, Y.C., Eisner, J.D., Kattar, M.M., Rassoulian-barrett, S.L., Lafe, K., Bui, U., Limaye, A.P., and Cookson, B.T. (2001). Polymorphic Internal Transcribed Spacer Region 1 DNA Sequences Identify Medically Important Yeasts. Journal of Clinical Microbiology. 39, 4042-4051.

Chen, Y.C., Eisner, J.D., Kattar, M.M., Rassoulian-barrett, K.L., LaFe, K., Yarfitz, S.L., Limaye, A.P., and Cookson, B.T. (2000). Identification of medically important yeasts using PCR-based detection of DNA sequence polymorphism in the internal transcribed spacer 2 region of the rRNA genes. Journal of Clinical Microbiology. 38, 2302-2310.

ClustalW2. A general purpose multiple sequence alignment program for DNA or proteins. Available at: http://www.ebi.ac.uk/Tools/msa/clustalw2

CODEHOP. Consensus Degenerate Hybrid Oligonucleotide Primers. Available at: http://www.bioinformatics.weizmann.ac.il/blocks/codehop.html

Comini, M., Menge, U., Wissing, J., and Flohé, L. (2005). Trypanothione synthesis in crithidia revisited. The Journal of Biological Chemistry. 280, 6850-60.

Cooke, D.E.L., Duncan, J.M. (1997). Phylogenetic analysis of *Phytophthora* species based on ITS1 and ITS2 sequences of the ribosomal RNA gene repeat. Mycological Research. 6, 667-677.

Cordier, C., Edel-Hermann,V., Martin-Laurent, F., Blal, B., Steinberg, C., and Alabouvette, C. (2007). SCAR-based real time PCR to identify a biocontrol strain (T1) of Trichoderma atroviride and study its populations dynamics in soils. Journal of Microbiological Methods. 68, 60-68.

Croft, M.T., Moulin, M., Webb, M.E., and Smith, A.G. (2007). Thiamine biosynthesis in algae is regulated by riboswitches. Proceedings of the National Academy of Sciences of the United States of America 104. 20770-20775.

Cvitanich, C., and Judelson, H.S. (2003). A gene expressed during sexual and asexual sporulation in *Phytophthora infestans* is a member of the Puf family of translational regulators. Eukaryotic Cell. 2, 465-473.

Czeczuga, B., and Muszyńska, E. (1997). Aquatic fungi growing on the eggs of some anadromous fish species of the family Clupediae. Acta Ichthyologica et Piscatoria. 1, 83-93.

Dereeper, A., Guignon, Y., Blanc, G., Audic, S., Buffet, S., Chevenel, F., Dufayard, J.F., Guindon, S., Lefort, Y., Lescot, M., Claverie, J.M., and Gascuel, J.M. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Research. 36, W465-469.

Diéguez-Uribeondo, J., Fregeneda-Grandes, J. M., Cerenius, L., Pérez-Iniesta, E., Aller-Gancedo, José Miguel, Tellería, M.T., Söderhäll, K., and Martín, M.P. (2007). Re-evaluation of the enigmatic species complex *Saprolegnia diclina-Saprolegnia parasitica* based on morphological, physiological and molecular data. Fungal Genetics and Biology. doi:10.1016/j.fgb.2007.02.010

Diéguez-Uribeondo, J., Cerenius, L., and Söderhäll, K. (1996). Physiological characterization of *Saprolegnia* isolates from brown trout. 140, 247-257.

Diéguez-Uribeondo, J., Cerenius, L., and Söderhäll, K. (1994). *Saprolegnia parasitica* and its virulence on three different species of freshwater crayfish. Aquaculture. 120, 219-228.

Fairlamb, A. H., and Cerami, A. (1992). Metabolism and functions of trypanotione in the Kinetoplastida. Annual Review of Microbiology. 26, 695-729.

FAO. (2010). Food and Agriculture Organization of the United Nations. Fisheries and Aquaculture Department. Available at : http://www.fao.org/fishery/en

Fernández-Benéitez, M.J., Ortiz-Santaliestra, M.E., Lizana, M., and Diéguez-Uribeobdo, J. (2008). *Saprolgenia diclina:* another species responsible for the emergent disease '*Saprolegnia* infections' in amphibians. FFMS Microbiology Letters. 279, 23-29.

Fregeneda-Grandes, J.M., Rodríguez-Cadenas, F., and Aller-Gancedo, J.M. (2007). Fungi isolated from cultured eggs, alevins and broodfish of brown trout in a hatchery affected by saprolegniosis. Journal of Fish Biology. 71, 510-518.

Fregeneda-Grandes, J.M., Fernández Díez, M., and Aller Gancedo, J.M. (2000). Ultrastructural analysis of *Saprolegnia* secondary zoospore cyst ornamentation from infected wild brown trout, *Salmo trutta* L., and river water indicates two distinct morphotypes amongst long-spined isolates. Journal of Fish Diseases. 23, 147-160.

Fyfe, P.K., Oza, S.L., fairlamb, A.H., and Hunter, W.N. (2008). *Leishmania* trypanothione synthetase-amidase structure reveals a basis for regulation of conflicting synthetic and hydrolytic activities. Journal of Biological Chemistry. 283, 17627-17680.

Gardes, M., and Bruns, T.D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. Molecular Ecology. 2, 113-118.

Gaulin, E., Madoui, M.A., Bottin, A., Jacquet, C., Mathé, C., Couloux, A., Wincker, P., Dumas, B. (2008). Transcriptome of *Aphanomyces euteiches*: New Oomycete putative pathogenicity factors and metabolic pathways. PLoS ONE. 3, e1723. doi:10.1371/journal.pone.0001723

Guarro, J., Gené, J., and Stchigel, A.M. (1999). Developments in fungal taxonomy. Clinical Micribiology reviews. 12,454-500.

Guerriero, G., Avino, M., Zhou, Q., Fugelstad, J., Clergrot, P.H., Bulone, V. (2010). Chitin Synthases from Saprolgenia are involved in tip growth and represent a potential target for anti-oomycete drugs. PLoS Pathogens. 6(8): e1001070.doi:10.1371/journal.ppat.1001070

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 41, 95-98.

Hatai, K., and Hoshiai, G. (1992). Mass mortality in cultured Coho salmon (*Oncorhynchus kisutch*) due to *Saprolegnia. parasitica* Coker. Journal of Wildlife diseases. 28, 532-535.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., and Hallwachs,W. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America. 101,14812-14827.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004b). Identification of Birds through DNA Barcodes. PLoS Biology. 2, 1657-1663.

Heid, C. A., Stevens, J., Livak, K.J., and Williams, P.M. (1996). Real time quantitative PCR. Genome Research. 6, 986-994.

Hochwimmer, G., Tober, R., Bibars-Reiter, R., Licek, E., and Steinborn, R. (2009). Identification of two GH18 chitinase family genes and their use as targets for detection of the crayfish-plague oomycete *Aphanomyces astaci.* BMC Microbiology. 9, 184. doi:10.1186/1471-2180-9-184

Hollingsworth *et al*., (2009). A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the United states of America. 106, 12794-12797.

Hulvey, J.P., Padgett, D.E., and Bailey, J.C. (2007). Species boundaries within *Saprolegnia* (*Saprolegniales*, Oomycota) based on morphological and DNA sequence data. Mycologia. 99, 421-9.

Hussein, M.M.A, and Hatai, K. (2002). Pathogenicity of *Saprolegnia* species associated with outbreaks of salmonid saprolegniosis in Japan. Fisheries Science. 68,1067-1072.

Ibekwe, A.M., Watt, P.M., Grieve, C.M., Sharma, V.K., and Lyons, S.R. (2002). Multiplex Fluorogenic Real-Time PCR for Detection and Quantification of *Escherichia coli* O157: H7 in Dairy Wastewater Wetlands. Applied and Environmental Microbiology. 68, 4853-4862.

IDT. Integrated DNA technologies. Oligo Analyzer 3.1. Available at: http://www.idtdna.com/analyzer/Applications/OligoAnalyzer

Johnson, J.E., Belmont, S. F., and Wagner, R.T. (2008). DNA barcoding as a means to identify organisms association with amphibian eggs. Herpetological Conservation and Biology. 3, 116-127.

Jones, M.D.M., Forn, I., Gadelha, C., Egna, M.J., Bass, D., Massana, R., and Richards, T.A. (2011). Discovery of novel intermediate forms redefines the fungal tree of life. Nature. 474, 200-203. doi:10.1038/nature09984

Ke, X.L., Wang, J.G., Gu, Z.M., Li,M., and Gong, X.N. (2009). Morphological and molecular phylogenetic analysis of two *Saprolegnia* sp. (Oomycetes) isolated from silver crucian carp and zebra fish. Mycological research. 113, 637-44.

Kirshtein, J., Anderson, C., Wood, J., Longcore, J., and Voytek, M. (2007). Quantitative PCR detection of *Batrachochytrium dendrobatidis* DNA from sediments and water. Diseases of Aquatic Organisms. 77, 11-15.

Koetschan, C., *et al*., (2010). The ITS2 database III - sequences and structure for phylogeny. Nucleic Acids Research. 38, D275-279.

Kress, W.J., and Erickson, D.L. (2008). DNA barcodes: genes, genomics, and bioinformatics. Proceedings of the National Academy of Sciences of the United States of America. 105, 2761-2762.

Leclerc, M.C., Guillot, J., and Deville, M. (2000). Taxonomic and phylogenetic analysis of Saprolegniaceae (Oomycetes) inferred from LSU rDNA and ITS sequence comparisons. Antonie van Leeuwenhoek 77, 369-77.

Li, C., Lee, J.S., Groebner, J.L., Ki., H.C., Klein, T.A., O'Guinn, M.L., and Chen, C. (2005). A newly recognized species in the *Anopheles* Hyrcanus group and molecular identification of related species from the republic of South Korea (Diptera:Culicidae). Zootaxa. 939, 1-8.

Long, E., and Dawid, I.B. (1980). Repeated genes in eukaryotes. Annual Review of Biochemistry. 49, 727-74.

Ma, Y., and Xu, J. (2005). The Hyrcanus group of Anopheles (Anopheles) in China (Diptera: Culicidae): species discrimination and phylogenetic relationships inferred by ribosomal DNA internal transcribed spacer 2 sequences. Journal of Medical Entomology. 42, 610-619.

Madoui, M.A., Gaulin, E., Mathé, C., San Clemente, H., Couloux, A., Wincker, P., and Dumas, B. (2007). AphanoDB: a genomic resource for *Aphanomyces* pathogens. BMC Genomics. 8, 471. doi:10.1186/1471-2164-8-471

Meyer, F.P. (1991). Aquaculture disease and health management. Journal of Animal Science. 69, 4201-4208.

Mitchell, A.J., Straus, D.L., Farmer, B., and Carter, R. (2010). Comparison of percent hatch and fungal infestation in channel Catfish eggs after copper sulfate, diquat bromide, formalin, and hydrogen peroxide treatment. North American Journal of Aquaculture. 72, 201-206.

Molina, F.I., Jong, S.C., and Ma,G. (1995). Molecular characterization and identification of Saprolegnia by restriction analysis of genes coding for ribosomal RNA. Antonie van Leeuwenhoek 68, 65-74.

Möller, E.M., Bahnweg, G., Sanderman, H., and Geiger, H.H. (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Research. 20, 6115-6116.

Moorman, G.W., Kang, S., Geiser, D.M., Kim, S.H. (2002). Identification and characterization of Pythium species associated with greenhouse floral crops in Pennsylvania. Plant Disease. 86, 1227-1231.

Mueller, U.G., and Wolfenbarger, L.L. (2000). AFLP genotyping and fingerprinting. Trends in Ecology and Evolution. 14, 389-394.

Muzzarelli, R., Muzzarelli, C., Tarsi, R., Miliani, M., Gabbanelli, F., and Cartolari, M. (2001). Fungistatic activity of modified chitosans against *Saprolegnia parasitica*. Biomacromolecules. 2, 165-169.

National Human Genome Research Institute. Talking Glossary of genetics terms. Available at: http://www.genmoe.gov/glossary/?id=75

NCBI. National Centre for Biotechnology Information. Available at http://www.ncbi.nlm.nih.gov

Oidtmann, B., Geiger, S., Steinbauer, P., Culas, A., Hoffmann, R.W. (2006). Detection of *Aphanomyces astaci* in North America crayfish by polymerase chain reaction. Diseases of Aquatic Organisms. 72, 53-64.

Oidtmann, B., Schaefers, N., Cerenius, L., Söderhäll, K., and Hoffmann, R.W. (2004). Detection of genomic DNA of the crayfish plague fungus *Aphanomyces astaci* (Oomycete) in clinical samples by PCR. Veterinary Microbiology. 100, 264-282.

Operon. eurofins mwg|operon. Oligo analysis tool. Available at:http://www.operon.com/tools/oligo-analysis-tool.aspx

Oza, S.L., Tetaud, E., Ariyanayagam, M.R., Warnon, S.S., and Fairlamb, A.H. (2002). A single enzyme catalyses formation of Trypanothione from glutathione and spermidine in *Trypanosoma cruzi*. The Journal of biological Chemistry. 277, 35853-35861.

Park, C.S., Kakinuma, M., and Amano, H. (2001). Detection and quantitative analysis of zoospores of *Pythium porphyrae*, causative organism of red rot disease in *Porphyra*, by competitive PCR. Journal of Applied Phycology.13, 433-441.

Petrisko, J.E., Pearl, C.A., Pilliod, D.S., Sheridan, P.P., Williams, C.F., Peterson, C.R., and Bury, R.B. (2008). Mycologia, 2, 171-180.

Phillips, A.J., Anderson, V.L., Robertson, E.J., Secombes, C.J., and West, P. van (2008). New insights into animal pathogenic oomycetes. Trends in microbiology. 16, 13-19.

Pottinger, T.G., and Day, J.G. (1999). A *Saprolegnia parasitica challenge* system for rainbow trout: assessment of Pyceze as an anti-fungal agent for both fish and ova. Diseases of Aquatic Organisms. 36,129-141.

Primer 3 (v. 0.4.0) Pick primers from a DNA sequence. Primer 3. Available at: attp://www.frodo.wi.edu/primer3/

Ramaiah, N. (2006). A review on fungal diseases of algae, marine fishes, shrimps and corals. Indian Journal of Marine Sciences. 35, 380-387.

Raška, I., Koberna, K., Malínský, J., Fidlerová, H., and Mašata, M. (2004). The nucleolus and transcription of ribosomal genes. Biology of the Cell. 96, 579-594.

Richard, G.F, Kerrest, A., and Dujon, B. (2008). Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. Microbiology and Molecular Biology Reviews. 72, 686-727.

Robertson, E.J., Anderson, V.L., Phillips, A.J., Secombes, C.J., Diéguez-Uribeondo, J., van West, P. (2009) *Saprolegnia* - fish interactions. *In* Oomycete Genetics and Genomics: Diversity, Interactions, and research tools. *Edited by* Kurt Lamour and Sophien Kamoun. John and Wiley & Sons, Inc., Hoboken, New Jersey, 407-424.

Robideau, G.P. *et al*. (2011). DNA barcoding of oomycetes with cytochrome c oxidase subunit I and internal transcribed spacer. Molecular Ecology Resources. 11, 1002-1011.

Schena, L., and Cooke, D.E.L. (2006). Assessing the potential of regions of the nuclear and mitochondrial genome to develop a “molecular tool box” for the detection and characterization of *Phytophthora* species. Journal of Microbiological Methods. 67, 70-85.

Schlötterer, C. (2004). The evolution of molecular markers - just a matter of fashion? Nature Reviews Genetics. 5, 63-69.

Schultz, J., Müller, T., Achtziger, M., Seibel, P. N., Dandekar, T., and Wolf, M. (2006). The internal transcribed spacer 2 database--a web server for (not only) low level phylogenetic analyses. Nucleic acids Research. 34, W704-W707.

Shaw, P.J., and Jordan, E.G. (1995). The nucleolus. Annual reviews of Cell and developmental Biology. 11, 93-121.

Singhal, R. (1987). Experimental transmission of *Saprolegnia* and *Achlya* to fish. Aquaculture. 64, 1-7.

Smith, K., Borges, A., Ariyanayagam, M.R., and Fairlamb, A.H. (1995). Glutathionylspermidine metabolism in *Escherichia coli.* The Biochemical Journal. 312, 465-469.

Spassov, D.S., and Jurecic, R. (2003) The Puf family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function? International Union of Biochemistry and Molecular Biology. 55, 359-366.

Spooner, D., Treuren, R.V., and Vicente, M.C.D. (2005). Molecular markers to genebank management. IPGRI Technical Bulletin No. 10 International Plant Genetic resources Instiute, Rome, Italy.

Stoeckle, M. (2003). Taxonomy, DNA, and the Bar Code of Life. BioScience. 53, 2-3.

Stueland, S., Hatai, K., and Skaar, I. (2005). Morphological and physiological characteristics of *Saprolegnia* spp . strains pathogenic to Atlantic salmon , *Salmo salar* L . Journal of Fish Diseases. 28,445-453.

Sudova, E., Machova, J., Svonodova, Z., and Vesely, T. (2007). Negative effects of malachite green and possibilities of its replacement in the treatment of fish eggs and fish: review. Veterinarni Medinina. 12, 527-539.

Sumida, M., Kato, Y., and Kurabayashi, A. (2004). Sequencing and analysis of the internal transcribed spacers (ITSs) and coding regions in the *Eco*R I fragment of the ribosomal DNA of the Japanese pond frog *Rana nigromaculat*a. Genes and Genetics Systems. 79,105-118.

Sunnucks, P. (2000). Efficient genetic markers for population biology. Trends in Ecology and Evolution. 15, 199-203.

Symondson, W.O.C.(2002). Molecular identification of prey in predator diets. Molecular Ecology. 11, 627-641.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA 5: Molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary distance, and Maximum Parsimony methods. Molecular Biology and Evolution. 28, 2731-2739. doi: 10.1093/molbev/msr121

Taylor, J.W., and Fisher, M.C. (2003). Fungal multilocus sequence typing — it's not just for bacteria. Current Opinion in Microbiology. 6, 351-356.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S., and Fisher, M. C. (2000). Phylogenetic species recognition and species concepts in fungi. Fungal genetics and biology. 31, 21-32.

Thoen, E., Evensen, Ø., and Skaar, I. (2011). Pathogenicity of *Saprolegnia* spp. to Atlantic salmon, *Salmo salar* L., eggs. Journal of fish diseases. 34, 601-608.

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

Uniprot. Universal Protein Resource. (2012). Available at: http://www.uniprot.org

University of Aberdeen. Aberdeen Oomycete Laboratory. Available at:http://www.abdn.ac.uk/oomycetes/Home.html

Vandersea, M. W., Litaker, R.W., Yonnish, B., Sosa, E., Landsberg, J.H., Pullinger, C., Moon-Butzin, P., Green, J., Morris, J.A., Kator, H., Noga, E. J., and Tester, P.A. (2006). Molecular Assays for Detecting *Aphanomyces invadans* in Ulcerative Mycotic Fish Lesions. Applied and Environmental Microbiology. 72, 1551-1557.

van West, P., de Brujin, I., Minor, K.L,, Phillips, A.J., Robertson, E.J., Wawra, S., Bain, J., Anderson, V.L., and Secombes, C.J. (2010). The putative RxLR effector protein SpHtp1 from the fish pathogenic oomycete Saprolegnia parasitica is translocated into fish cells. FEMS Microbiology letters. 310, 127-37.

van West. (2006). *Saprolegnia parasitica* , an oomycete pathogen with a fishy appetite: new challenges for an old problem. Mycologist. 20, 99-104.

Walker, C.A., and van West, P. (2007). Zoospore development in the oomycetes. Fungal Biology Reviews. 21, 10-18.

White, T.J., Bruns, T., Lee, S., and Taylor, J.W. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* PCR protocols: A guide to methods and applications. *Edited* by Innis, M.A., D.H. Gelfand, J.J. Sninsky and T.J. White. Academic Press. Inc., NY, 315-322.

Willoughby, L.G. (1995). Rapid preliminary screening of *Saprolegnia* on fish. Journal of Fish Diseases. 8, 473-476.

Wolf, M., Friedrich, J., Dandekar, T., and Müller, T. (2005). CBCAnalyzer: inferring phylogenies based on compensatory base changes in RNA secondary structures. In Silico Biology. 5, 291-294.

Yuasa, K., and Hatai, K. (1996). Some biochemical characteristics of the genera *Saprolegnia*, *Achlya* and *Aphanomyces* isolated from fishes with fungal infection. Mycoscience. 37, 477-479.