Invasion and evolutionary history of a generalist fish parasite

Salma Sana

A thesis submitted in partial fulfilment of the requirements of Bournemouth University for the degree of Doctor of Philosophy

December 2016

Copyright Statement

Acknowledgements

Firstly, I would like to express my deepest gratitude to Allah who gave me the strength to complete this work.

Thank you to my supervisors: Tiantian Zhang who in the first place provided me this opportunity and her guidance and support throughout the project and my day to day supervisor Demetra Andreou who has been a constant support over the last three years. Thank you for your invaluable guidance and help throughout this project. I would also like to thank Emilie Hardouin for her help with the DNA softwares and my viva examiners Rob Britton and Trenton Garner for their valuable feedback and advice.

Dr Richard Paley who always has kindly provided RA spores and cultures whenever there has been a mishap at my end.

I am also grateful to my parents and family for their love and support at every step of my life. I can't thank you enough for all you have done for me.

Last- but in no way least- to Ata without his support and sacrifice I would have never been able to complete my PhD and Kazeem who makes me feel proud every day.

Authors declaration

I confirm that the work presented in this thesis is my own work, with the following exceptions:

Chapter 4: The work presented in this chapter was in collaboration with the Environment Agency. The fish were sampled and dissected at Environment Agency and samples (organ tissue) were provided for processing. Histopathology was also performed by the Environment Agency.

Chapter 5: Water sampling pre-*Pseudorasbora parva* eradication (2013) by Phil Davison.

Salma Sana

Invasion and evolutionary history of a generalist fish parasite

Abstract

The introduction of non-native species can lead to the introduction of non-native parasites to their introduced range which can pose significant risk to native biodiversity. The cyprinid fish species, *Pseudorasbora parva,* is a well-studied example of accidental introduction to a new range; it has been accidentally introduced from China to Europe. *Pseudorasbora parva* has been hypothesized to have also introduced the generalist fish pathogen *Sphaerothecum destruens* to Europe which has been identified as a potential threat to European fish biodiversity. Due to the management implications associated with the parasite's status (native or non-native), this work aimed at determining the *S. destruens* origin and distribution across its native and non-native *P. parva* populations, whilst also developing eDNA detection methods in order to assess the efficacy of *P. parva* eradication as a viable control measure for *S. destruens*. Due to the unique taxonomical position of *S. destruens* in tree of life, its mitochondrial DNA evolutionary history was also investigated to better decipher its phylogenetic position.

Sphaerothecum destruens presence was confirmed in 90 % of the *P. parva* sampled populations from China, with a maximum prevalence of 10 %. Furthermore, the phylogenetic and demographic analysis of both the host and the parasite support the hypothesis that *S. destruens* has been introduced to Europe through the accidental introduction of its reservoir host *P. parva*. The non-native status of *S. destruens* in Europe has important management implications for the parasite. Furthermore, *S.*

destruens was detected in 50 % of the *P. parva* samples from 7 populations in the UK and identified new potential hosts for *S. destruens* in the wild including chub *Squalius cephalus*, dace *leuciscus*, roach *Rutilus rutilus* and brown trout *Salmo trutta*. The environmental DNA method detected *S. destruens* in water samples from a *P. parva* eradicated site 2 years after its eradication which emphasizes that preventive measures against pathogen expansion should be implemented. The phylogenetic tree based on mitochondrial derived protein sequences revealed an interesting position for *S. destruens* as a sister group to Filasterea and Choanoflagellate and Metazoa group and it has the most derived mitochondrial genome among Choanozoa.

Table of Contents

List of figures

List of Tables

Chapter 1

General Introduction

In the last two decades there has been an increasing trend toward aquaculture production and fish farming which often relies on the introduction of non-native species to new environments and geographic areas (Gozlan 2008, Gozlan et al. 2010b, Peeler et al. 2011). Non-native species, defined here as species that have been translocated outside their natural range irrespective of political borders (Riley 2005), are both intentionally and unintentionally introduced for trade, food and resource production, ornamental purposes or as bio-control (Gozlan et al. 2010b). Accidental introductions, also referred to as biological pollution (Elliott 2003), account for 8 % of non-native aquatic species incidences (Gozlan et al. 2010b).

The introduction of non-native species can adversely impact native species populations, the environment and ecosystems through a number of different processes including increased predation, competitive exclusion, non-native species dominance and disease introduction (Peeler et al. 2011). Introduced species can influence the native species through increased predation, as in the case of the introduction of salmonids rainbow trout *Oncorhynchus mykiss* and brook trout *Salvelinus fontinalis* in Sierra Nevada,USA, which resulted in significant declines of the yellow legged frog *Rana mucosa* populations due to increased predation of its tadpoles (Knapp and Matthews 2000). The eradication of the two trout species led to the rapid recovery of frog populations (Vredenburg 2004). Non-native species can also compete for resources i.e. food and space with native species. For example, the red squirrel *Sciurus vulgaris* has suffered population declines in the UK, Ireland and Italy due to competition for food and space resources with the introduced North American grey squirrels *Sciurus caroliniensis.* Lower body mass and fecundity was reported in native *S. vulgaris* that cohabited with *S. caroliniensis* (Gurnell et al. 2004).

In addition to competition and predation, non-native species can change their new ecosystems, as is the case of grass carp *Ctenopharyngodon idella* that reduces natural aquatic vegetation and common carp *Cyprinus carpio* that can lead to significant increases in water turbidity due to the rooting activity during foraging (Pimentel et al. 2000). Non-native species can also have severe impacts on the ecosystem by altering the community structure, for example, the introduction of the bivalves *Potamocorbula amurensis* and *Dreissena polymorpha* in North America (Kimmerer et al. 1994, Strayer

and Smith 1996). In both cases, the invasive species became significantly dominant at the introduced sites, outnumbering the native benthic organisms which can have indirect effects on food webs and nutrient dynamics (Ruiz et al. 1997).

Non-native species can also introduce disease with their harboured parasites posing a significant risk to native biodiversity (Murray and Peeler 2005, Crowl et al. 2008, Poulin et al. 2011). These parasites can play an important role in invasiveness of the non-native species (Prenter et al. 2004) as they can facilitate the non-native species establishment and competitiveness (Mordecai 2013). For example, the crayfish plague, *Aphanomyces astaci*, introduced to the UK along with the North American signal crayfish, *Pacifastacus leniusculus,* has had detrimental effects on the native crayfish *Austropotamobius pallipes* which has facilitated a rapid expansion of the non-native *P. leniusculus* (Dunn 2009).

The introduction of new parasites can lead to disease emergence, facilitated by the increase of the parasite's geographic range and host switching (Peeler et al. 2011). Non-native parasites can have several biological features which increase their probability of successfully being translocated and becoming established into new environments. These can include direct life-cycles, generalist nature (i.e. they can infect more than one species), tolerant and long-lived environmental infectious propagules and a wide temperature tolerance (Andreou et al. 2009, Fisher et al. 2012). Their generalist nature facilitates host-switching that can lead to disease emergence in native species. Notable examples include the fungi *Batrachochytrium dendrobatidis* (chytrid fungus) which has been introduced across the world through the pet trade can infect 508 host species (Fisher et al. 2012). Another fungi *Geomyces destructans*, the causative agent for the White-Nose Syndrome (WNS), has been introduced to North America most probably through contaminated caver clothing and has been linked to population declines in multiple species of bats in North America (Blehert et al. 2009, Turner et al. 2011). *Geomyces destructans* has also been found colonizing the skin of hibernating bats in Europe without any associated deaths (Wibbelt et al. 2010) which suggests that the parasite may be native to Europe (Puechmaille et al. 2011).

Invasive parasites can also include parasites with indirect life-cycles. For example, the parasitic nematode *Anguillicola crassus* is native to the Japanese eel *Anguilla japonica* and was introduced to Europe through the aquaculture trade (Peters and Hartmann 1986). Since its introduction it has infected the European eel *Anguilla anguilla*, resulting in high mortalities (Kennedy 2007) to the extent that numerous stocks are considered vulnerable or endangered (Costa-Dias et al. 2010).

Despite their life-history traits, aquatic non-native parasites are often hard to detect due to low visibility and a high turnover of mortalities in this environment (Hudson et al. 2002, Tompkins et al. 2011, Gozlan 2012). In addition, pathogen introductions through fish movements are the driving force for the emergence of aquatic diseases worldwide and will continue to have serious consequences for wild fish populations (Perkins et al. 2008, Peeler et al. 2011). It is thus crucial to have designated monitoring and risk assessment procedures that can be used to evaluate the hazard posed by identified non-native parasites.

In England and Wales, all identified non-native fish parasites have their risk independently assessed by a panel of experts using the risk assessment developed by Williams et al. (2013). Risk is assessed using a range of criteria which combine expert opinion on the potential for spread and economic impact of the parasite as well as direct evidence of disease pathology in the fish hosts. Parasites that have been identified as high risk, as well as ones that have been identified as non-native but have not had their risk assessed due to knowledge gaps, are listed on the Environment Agency's Category 2 non-native parasite list. Once a parasite has been listed, its movement between water bodies is restricted. Monitoring for its presence is achieved through mandatory fish health checks of any legally performed fish movements between water bodies. These risk assessments determine the disease threats to native populations from the new introduced parasites and help in the formulation of rapid management decisions in terms of their control.

In order to carry out the risk assessments for a particular parasite, it is important to determine the invasive status of the parasite which can be challenging due to incomplete data on parasite ranges. This is particularly relevant to parasites that are hosted by fish of poor economic value that do not have well established parasite profiles in their native ranges (Williams et al. 2013). Such is the case of *Sphaerothecum destruens*, a generalist pathogen that has been identified as a high risk to European fish diversity (Andreou and Gozlan 2016) and the focus of this PhD thesis. *Sphaerothecum destruens* has been hypothesized to have been introduced to Europe along with its reservoir host, the freshwater fish *Pseudorasbora parva* from China (Gozlan et al. 2005). However, its status as a non-native parasite to Europe has yet to be confirmed (Gozlan et al. 2009).

Pseudorasbora parva represents a well-studied example of accidental introduction to a new range. It was introduced at River Danube as a part of trade in Chinese carp, *Ctenopharyngodon idella and Hypophthalmichthys molitrix*, for the development of aquaculture in the former Union of Soviet Socialist Republics (USSR) (Van Zon 1977). This small fresh water cyprinid fish species has spread through Europe and has reached North Africa in less than 50 years (Figure 1.1; Gozlan et al. 2010a). Gozlan et al. (2005) discovered that *P. parva* found in the UK harboured the fungal-like parasite *S. destruens* and could transmit the disease to susceptible fishes without any harm to itself - identifying it as a healthy host of *S. destruens* (Gozlan et al. 2005). Despite the parasite being identified as a potential threat, its status in Europe still remains uncertain, i.e. is it a native or a non-native parasite.

1.1 The Healthy Host, *Pseudorasbora parva*

In order to determine the potential invasion history of *S. destruens* in Europe, a good understanding of invasion history of its host (*P. parva*) is required due to the close association of the two species. *Pseudorasbora parva* commonly known as topmouth gudgeon, is a small freshwater cyprinid whose native range includes China, Japan and Korea (Pinder et al. 2005). In China, *P. parva* is present to the north and south of the River Yangtze (Figure 1.2 A) which marks the boundary of two climatic zones; to its north there is a temperate climatic zone whereas to its south there is a tropical climatic zone (Domrös and Gongbing 1988). This reflects the wide temperature tolerance of *P.*

parva (Gozlan et al. 2010a). In its native Chinese range, *P. parva* was accidentally transferred, into almost all natural lakes, reservoirs and lower Upper Mekong basin (Yunnan province); into the Upper reach of Yellow basin river (Qinghai and Gansu province); into inland waters (Inner Mongolia); almost all natural lakes, rivers and reservoirs (Xinjiang), along with the movements of Chinese carps from the east of China for aquaculture (Gozlan et al. 2010a).

Figure 1.1. Topmouth gudgeon, *Pseudorasbora parva*, is a small cyprinid species which grows to ~8 cm with a life span of approximately four years (Britton et al. 2007). It sexually matures by one year of age and can have multiple spawning events annually, between the months of April and July (Pinder and Gozlan 2003).

Pseudorasbora parva was first introduced unintentionally to mainland Europe at River Danube in 1960 and within 40 years of its introduction it rapidly inhabited mainland Europe from east to west (Pinder et al. 2005). Several known *P. parva* introductions occurred into Hungary, Lithuania, Romania and Ukraine (Gozlan et al. 2010a). Except for Lithuania, all these introductions led do the dispersal of *P. parva* into local catchments and connected reservoirs. The *P. parva* introductions around the Black sea led to a westward spread in Europe and to Turkey and Iran. *Pseudorasbora parva's* introduction into the former Czechoslovakia from Hungary led to the inter-country spread of *P. parva* in central Europe (Gozlan et al. 2010a). *Pseudorasbora parva* was introduced to Germany from Czechoslovakia and from there to Holland, Belgium and the UK (Gozlan et al. 2002). Since its first introduction, *P. parva* has invaded, on average, five countries every decade. Its rapid dispersal rate has already indicated signs of saturation in the former Czechoslovakia and the Netherlands. The main factors attributed to its primary introduction pathways are aquaculture due to *P. parva's* association with Chinese carp species and common carp *C. carpio* (65 %), recreation fishing (22 %), ornamental fish trade (9 %) and natural dispersal (1 %) that also accounts for the main secondary introduction pathway (Gozlan et al. 2010a).

The first occurrence of *P. parva* in England was observed in an ornamental pond in Chiltrens, UK (Domaniewski and Wheeler 1996). The only known introduction of this species in England has been a contaminated golden orfe *Leuciscus idus* consignment form Germany to Crampmoor fisheries, Hampshire in 1980 (Gozlan et al. 2002). Since its introduction into the UK it has been reported from at least 35 sites across England and Wales with 23 confirmed sites and 12 suspected sites Figure 1.3A (Britton et al. 2008b, GBNNSS 2015).

Pseudorasbora parva's invasion history has been extensively studied from an ecological view in Gozlan et al. (2010a) and a population genetics perspective (Simon et al. 2011, Simon et al. 2015, Hardouin et al. submitted). Hardouin et al. performed the most extensive population genetics study which involved 27 populations from the invasive range (Eurasia) and 30 populations from the species native range (China, Japan and Taiwan). Specifically, the analysis using 597 bp of the Cytochrome-b (Cyt-b) gene of *P. parva* has identified four distinct haplogroups (A, B, C and D) with hapolgroups A and B being the most prevalent (92 %). These haplogroups can be found across China with haplogroup A being most prevalent north of the River Yangtze whilst haplogroup B is most prevalent south of the River Yangtze (Figure 1.2). The genetic analysis of *P. parva's* European invasive range revealed that the two main haplogroups A and B with two identified routes of *P. parva* spread, east to south (Bulgaria, Armenia and Turkey) which had only haplogroup B and east to west (from Hungary to the UK) having *P. parva* with a mixture of A and B haplogroups (Figure 1.3 B, Hardouin et al. submitted). This result suggests that there was more than one independent accidental introduction of *P. parva* in Europe.

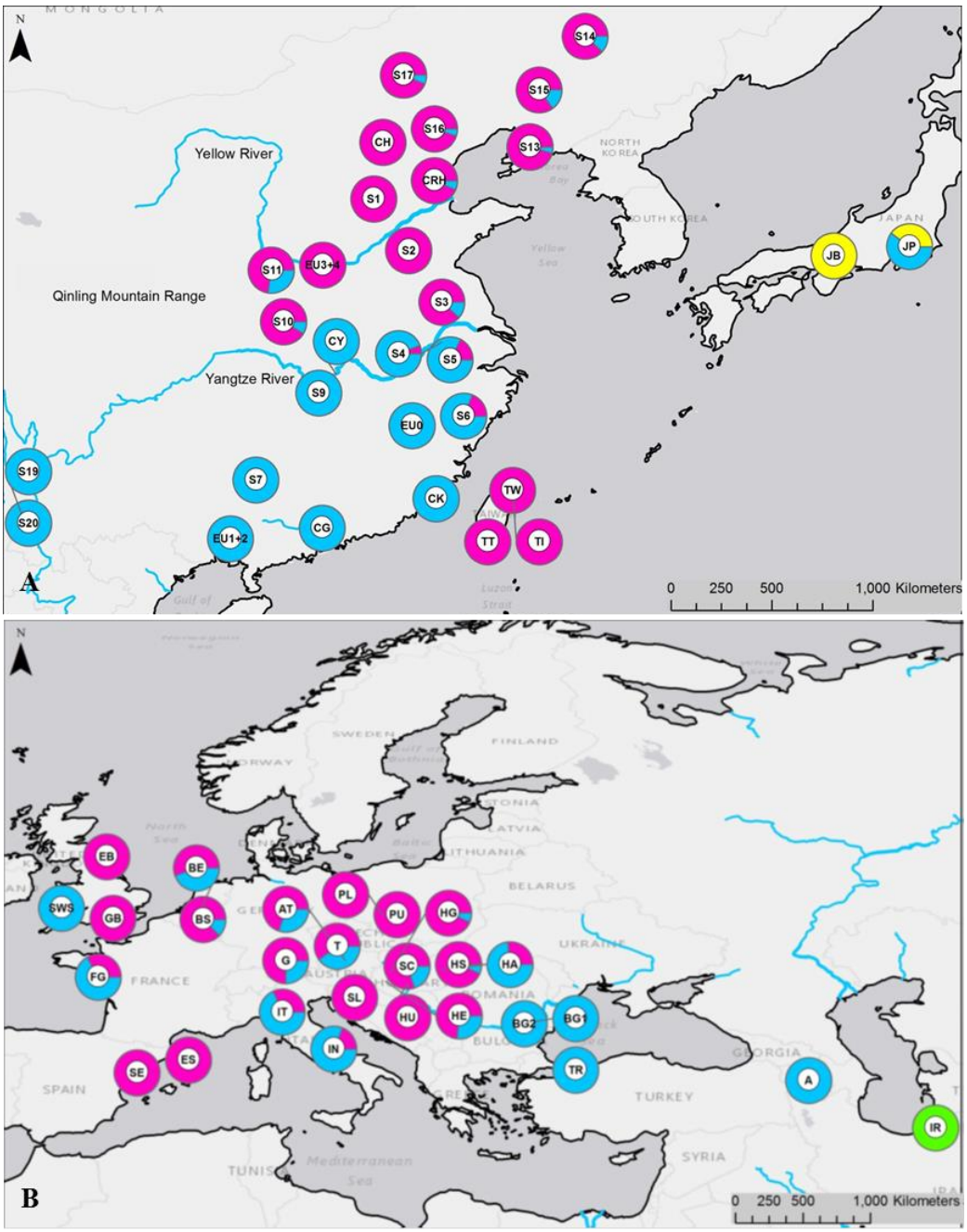


Figure 1.2. Phylogenetic analysis of cytochrome-b gene (597 bp) of *Pseudorasbora parva* from its native (A) and invasive (B) range across Eurasia. The prevalent *P. parva* haplogroups at each site are indicated by colour code. Haplogroup A (pink), Haplogroup B (blue), Haplogroup C (yellow) and Haplogroup D (green) (Hardouin et al. submitted).

The highly invasive nature of *P. parva* has been attributed to its favourable life history traits of nest guarding, batch spawning, early maturity onset and small size (Gozlan et al. 2002). It is also an opportunistic feeder and often feeds on eggs and larvae of native fish species such as zander *Sander lucioperca*, pike *Esox lucius* and perch *Perca*

fluviatilis (Pinder and Gozlan 2003). In 1992 it was recognized as an international pest species (Welcomme 1992). Its discovery as a healthy carrier of *S. destruens* raised its potential threat to native fish biodiversity (Gozlan et al. 2005).

The sites colonized by *P. parva* in the UK included lakes, enclosed still waters, aquaculture facilities and still water fisheries with a number of these having direct connections to rivers. Some of these water bodies pose the risk of rapid fluvial dispersal of *P. parva* to river catchments of high conservation value (Pinder et al. 2005). The combined ecological and disease spread risks have resulted to the Environment Agency designing an eradication plan to prevent the further spread of *P. parva* in the UK.

1.2 *Pseudorasbora parva* eradication in the UK

The significant impact on ecosystem function and biodiversity due to the introduction of invasive alien species is now well established and accepted by conservation agencies and a large number of studies are being carried out in this area (Hulme 2006). Resultantly, there is increasing pressure on government bodies, policy makers and environmental agencies to address these issues (Hulme 2006).

The eradication and removal programme of *P. parva* from the UK water bodies with imminent threat of downstream dispersal of *P. parva* into river networks, was initiated in 2005 (Britton and Brazier 2006). Three types of strategies were used in dealing with the control and removal of *P. parva* depending on the risk posed to nearby river and lakes. The first strategy of "do-nothing" was applied to sites with low risk of further dispersal of *P. parva* and low fishery and conservation value. The second approach of control and suppression involved medium-risk sites and has proven to be 99 % successful in reducing *P. parva* abundance. The third method involved the eradication of *P. parva* from the water bodies assessed as "high risk". Eradication involved the complete removal of *P. parva* population from the waterbodies by treating it with the piscicide rotenone (Britton et al. 2008a).

Six *P. parva* infested lakes were categorized as high risk in England and Wales (Britton et al. 2010). Eradication of *P. parva* included rotenone application at the sites from Cumbria, north Yorkshire, Surrey, Devon and Berkshire and drain-down and disinfection which involved the de-watering of the lake, followed by the fish removal and destroy, and drying of the lake bed prior to application of disinfectant quick lime at the West Midlands site (Britton et al. 2008a, Britton et al. 2010). The technique has proven to be successful, as no *P. parva* have been recorded after the operation (2010). As of July 2014, 15 confirmed *P. parva* sites have been successfully eradicated (Figure 1.3 B), with the aim to complete removal and eradication of *P. parva* from England by 2017 (GBNNSS 2015). Although, the presence and spread of *S. destruens* has been a concern, none of the sites which have been eradicated prior to 2012 had the *P. parva* checked for the presence of *S. destruens.*

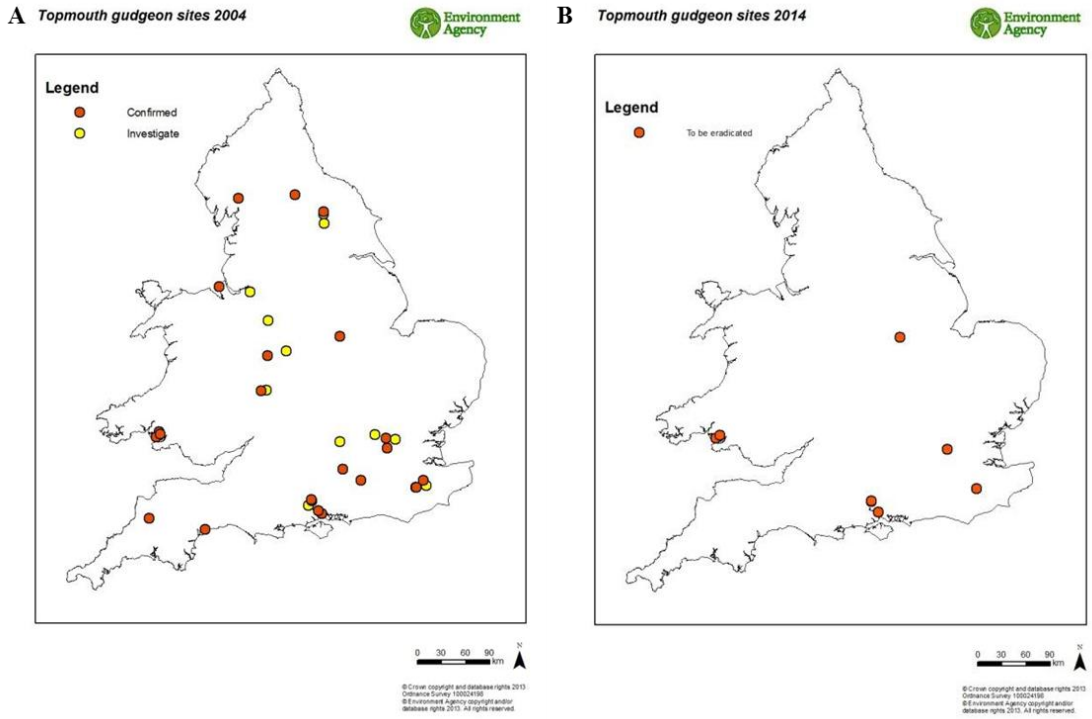


Figure 1.3. Eradication and removal program of *Pseudorasbora parva* from the UK. (A) *Pseudorasbora parva* infested water sites in the UK in 2004 and (B) Remaining *P. parva* infested sites in 2014 since the initiation of eradication program in 2005. The red circles are confirmed and yellow circles were suspected sites for *P. parva* presence (which were found to be negative) (GBNNSS 2015).

1.3 The Rosette Agent, *Sphaerothecum destruens*

Sphaerothecum destruens previously known as the rosette agent, is an animal fungal-like obligate intracellular fish parasite, with asexual reproduction (Arkush et al. 2003). *Sphaerothecum destruens* was first discovered as a cause of disease in Chinook salmon *Oncorhynchus tshawytscha* in Washington, USA, where it resulted in over 80 % mortality in 3-year-old fish (Harrell et al. 1986). *S. destruens* was later reported to cause chronic mortality in sub-adult Atlantic salmon *Salmo salar* in a Northern California farm (Hedrick et al. 1989). The third reported occurrence for the parasite in the USA was in winter-run Chinook salmon *O. tshawytscha* held at Bodega marine laboratory, where 40.1 % fish were found heavily parasitized with *S. destruens* (Arkush et al. 1998).

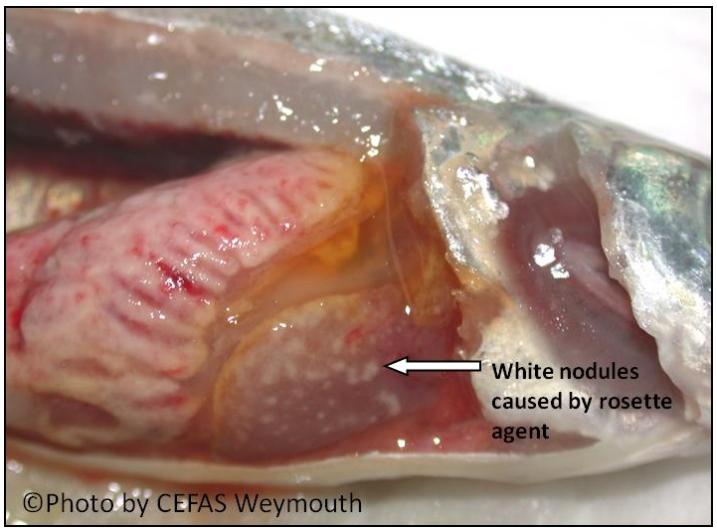


Figure 1.4. Atlantic salmon *Salmo salar* infected with *Sphaerothecum destruens* through intraperitoneal injection. The parasite results in white nodules on the surface of the liver (arrow) and haemorrhaging in the pyloric caecae (Paley et al. 2012)

Sphaerothecum destruens was first recorded in the UK in 2005 following cohabitation studies of *P. parva* and *Leucaspius delineatus*, during which the later was found to emaciate, cease reproduction and shoaling, and cause 67 % mortality (Figure 1.5; Gozlan et al. 2005). Histological examination of the emaciated *L. delineatus* revealed a parasite similar to *S. destruens* and was named the rosette-like agent. A wide range of susceptible salmonid and cyprinid fish hosts have been identified for *S. destruens* through experimental infections: chinook salmon *Oncorhynchus tshawytscha*, coho

salmon *Oncorhynchus kisutch*, rainbow trout *Oncorhynchus mykiss*, brown trout *Salmo trutta*, and brook trout *Salvelinus fontinalis* (Arkush et al. 1998), bream *Abramis brama*, carp *C. carpio* and roach *Rutilus rutilus* (Andreou et al. 2012).

Figure 1.5. Healthy (top) and emaciated (bottom) *Leucaspius delineates* following cohabitation with *Pseudorasbora parva*. *Sphaerothecum destruens* was detected in 67 % of the emaciated fish (Gozlan et al. 2005).

Phylogenetic studies using the 18S rRNA gene and the ribosomal internal transcribed spacer (ITS 1) gene identified the rosette like agent as *S. destruens* (Gozlan et al. 2009). Specifically, using the ITS 1 and 2, the UK and US isolates were identified as geographically isolated Figure 1.6 (Gozlan et al. 2009). Due to absence of *S. destruens* samples from *P. parva*'s native range the source of *S. destruens* in the UK has yet to be confirmed and the isolate has been designated as *S. destruens* (UK).

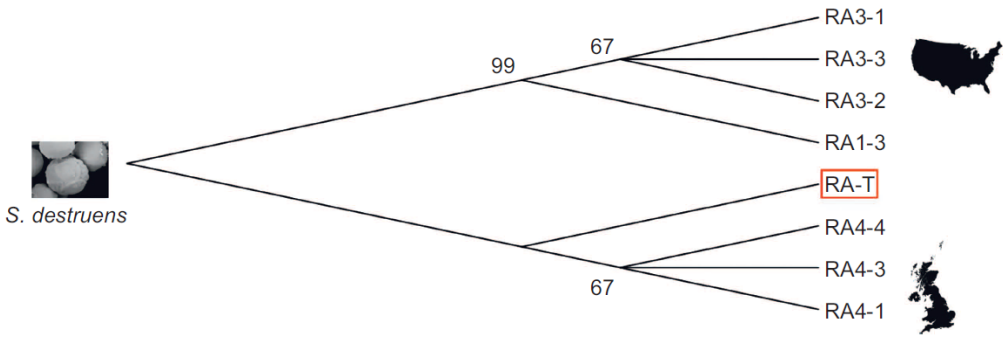


Figure 1.6. Phylogenetic tree showing geographically distinct clades of *Sphaerothecum destruens* in Europe and North America (USA) based on ribosomal ITS 1 gene (Ercan et al. 2015). RA1-3, RA3-1 – RA3-3 isolates are from US, RA4-1– RA4-3 isolates are from the UK and RA-T isolate is from Turkey.

Sphaerothecum destruens has also been reported from the Netherlands, where it was found with a prevalence of 74 % in *P. parva* populations without any clinical signs in the host (Spikmans et al. 2013), and later in Turkey and France (Charrier et al. 2016), where it has been found associated with *P. parva* and infect endemic freshwater fishes in Turkey (Figure 1.6; Ercan et al. 2015). Figure 1.7 summarises the known range of *S. destruens*. The *S. destruens* isolate from Turkey was determined to be closely related to the UK isolate but was not identical (see Figure 1.6; Ercan et al. 2015). Therefore, if *S. destruens* has been introduced to Europe with *P. parva*, more than one isolate has been introduced.

Figure 1.7. *Sphaerothecum destruens* prevalence across the Globe. *S. destruens* was detected in *Oncorhynchus tshawytscha* from Washington State (Harrell et al. 1986) and in *Salmo Salar* and *Oncorhynchus tshawytscha* from California (Hedrick et al. 1989, Arkush et al. 1998), in *Leucaspius delineatus* in the UK (Gozlan et al. 2005), in *Pseudorasbora parva* from the Netherlands and France and in Centrarchids from Turkey (Spikmans et al. 2013, Ercan et al. 2015, Charrier et al. 2016) (Abbreviations: US-United states of America, UK-United Kingdom, NL-the Netherlands, FR-France and T-Turkey).

1.4 Life history traits of *Sphaerothecum destruens*

The life cycle of *S. destruens* consists of two distinct morphological types of spherical intracytoplasmic spore stages that are 2-4 µm and 4-6 µm in diameter Figure 1.8 (Arkush et al. 2003). Spores replicate asexually through fission and can infect epithelial, mesenchymal and hematopoietic cells, eventually causing cell death. Once the spores are released they can infect further tissues or be excreted through bodily fluids e.g. bile, urine, gut epithelium, and seminal and ovarian fluids (Arkush et al. 2003). Further fish infection can occur through either ingestion or gut penetration or through skin and gills attachment (Arkush et al. 2003). Incubation in freshwater triggers the release of a minimum of 5 motile uniflagellate zoospores per *S. destruens* spore (Figure 1.8). Zoospores have an average body diameter and flagellum length of 2 µm and 10 µm respectively (Arkush et al. 2003) and have been shown to have a wide temperature tolerance 4 to 30 °C (Andreou et al. 2009). The cell wall of *S. destruens* is made up of three defined layers; an external layer of membranous structure, a central electron dense layer and an internal electrolucent layer. The cell cytoplasm consists of peripherally oriented mitochondria and both membrane-bound and non-membrane-bound vacuoles and a relatively indistinct nucleus (Figure 1.9; Harrell et al. 1986).

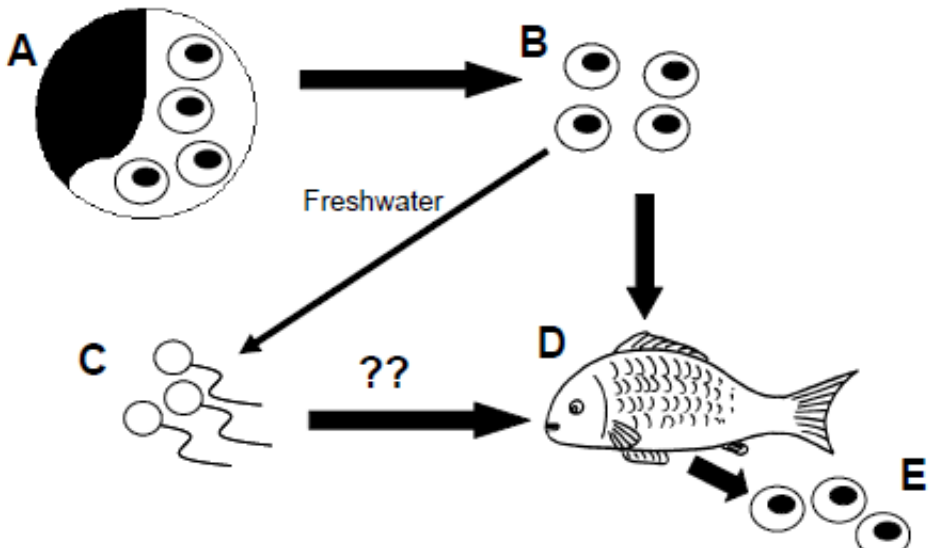


Figure 1.8. Proposed life cycle of *Sphaerothecum destruens* adapted from Arkush et al. (2003). (A) *S. destruens* spores infect host cells and divide asexually; (B) spores released through cell disruption (C) spores can propagate to zoospores following incubation in distilled water (D) released spores can infect new host fish through ingestion or gill attachment. Infection through zoospores is proposed but has not been demonstrated. (E) the infected host can release spores through bodily fluids.

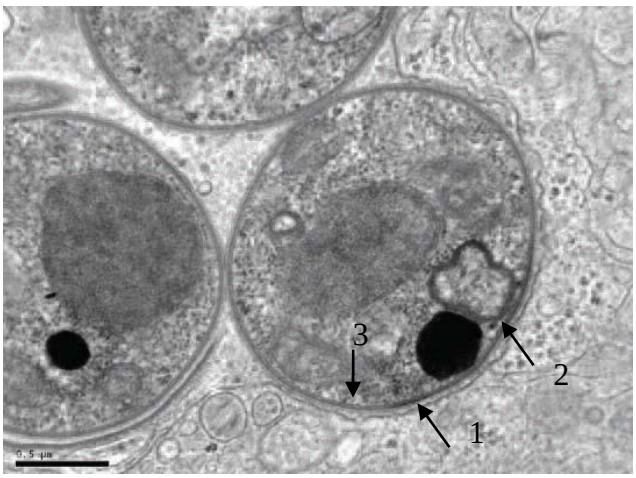


Figure 1.9. Transmission electron microscopy section of sunbleak *Leucaspius delineatus* renal tissue infected with *Sphaerothecum destruens*. The cell wall of *S. destruens* is comprised of three layers (indicated by arrows) (1) an outer layer with membranous structure; (2) a middle electron dense layer and (3) an inner electrolucent layer. Scale bar: 0.5 µm (Photo courtesy of Dr. Stephen W. Feist).

Pathological studies of *S. destruens* infection in salmonids, *O. tshawytscha* and *S. salar*, have been carried out by Arkush et al. (1998) and Hedrick et al. (1989), and in the cyprinid *L. delineatus* by Andreou et al. (2011). Two types of microscopic lesions were observed in the parasitized fish, nodular and disseminated (Arkush et al. 1998). In the nodular form of the disease the lesions exhibit a stronger host cell response evident from the granuloma formation in visceral organs such as kidney, liver and spleen. The granulomas contained numerous single parasites or multiple rosettes, replaced the normal parenchyma of the testis and the liver and were characterized by cellular debris, inflammation and numerous macrophages (Andreou et al. 2011). In the disseminated form of disease, the parasite was widely dispersed in host with occurrence in variety of cell types: hematopoietic, epithelial and mesenchymal cells and a reduced host cell response (Arkush et al. 1998). The two distinct morphotypes of *S. destruens*, 2-4 µm and 4-6 µm in diameter, were found in both types of disease (Arkush et al. 1998). Both types of lesions were reported for *L. delineatus* infection, where only the smaller spore morphotype (2 to 4 µm) was observed and *S. destruens* spores were also present within giant cells (Andreou et al. 2011). The ultrastructural characteristics of *S. destruens* in *L. delineatus* were found similar to those reported in salmonid infections (Arkush et al. 1998, Andreou et al. 2011).

The *S. destruens* spores stained with haematoxylin and eosin (H & E) appear deeply eosinophilic (Figure 1.10), where primary dye hemalum stains the chromatin material (nuclei) blue and a counter stain eosin dyes the eosinophilic bodies, including cytoplasm and extracellular proteins, in various shades of red, pink and orange. The *S. destruens* is Gram-positive and retains the primary crystal-violet stain due to the presence of a thick peptidoglycan layer in its cell wall, which appears purple coloured when observed under microscope (Figure 1.11; Arkush et al. 2003, Andreou et al. 2011, Paley et al. 2012).

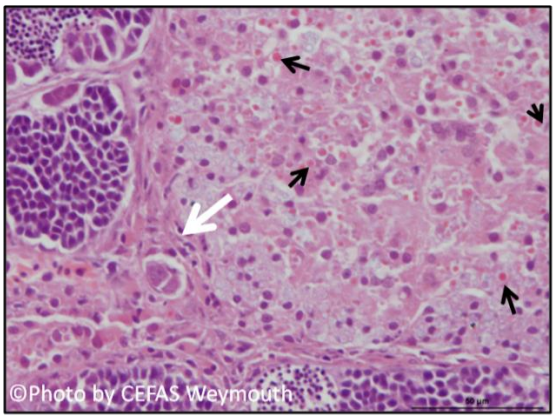


Figure 1.10. Granuloma in the testis of *Leucaspius delineatus*. The granuloma is surrounded by a thin fibroblast layer (white arrow) and numerous rosette agent spores are found within (Black arrows). H & E stain, Bar = 50 µm

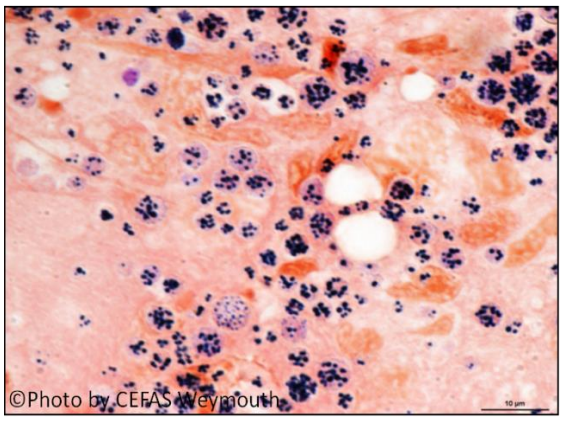


Figure 1.11. Liver lesion in *Salmo salar* showing granular Gram positive staining of cellular constituents of *Sphaerothecum destruens* cells. Gram stain, Bar = 50 µm

1.5 Classification history of *Sphaerothecum destruens*

The rosette agent was classified as *Sphaerothecum destruens* based on its two distinguishing features from its closely related genera Dermocystidium and Rhinosporidium: (i) host granulomatous response induced by parasitic spore stages; (ii) and differentiation of mature spores into multiple flagellated zoospores (Arkush et al. 2003). The genera Dermocystidium and Rhinosporidium belong to the Rhinosporideacae family within the Mesomycetozoa class which is found at the animal-fungal boundary (Mendoza et al. 2001). With the advancements in the taxonomical information based on ultrastructural and molecular phylogenetic studies, species classification in the Rhinosporideacae family has undergone many changes. Phylogenetic analyses based on small-subunit rRNA gene sequences, identified a group of eukaryotic protists that sits at the basal branch of the Metazoa, named as DRIPs clade. The clade was comprised of the *Dermocystidium spp*, the Rosette Agent, *Ichthyophonus hoferi* and *Psorospermium haeckelli* (Ragan et al. 1996). With the addition of *Rhinosporidium seeberi*, a human and animal pathogen, the acronym DRIP was replaced with the Class Mesomycetozoa to reflect their position within the Eukarya (Herr et al. 1999). The class Mesomycetozoa consisted of two orders: Dermocystida and Icthyophonida (Cavalier-Smith et al. 1998). Within the Dermocystida, *S. destruens*, *Dermocystidium* spp. and *R. seeberi* were grouped in the Rhinosporideaceae family (Mendoza et al. 2001). *Amphibiocystidium ranae*, a frog pathogen was recently added to family Rhinosporideaceae (Pereira et al. 2005).

Under the new proposed classification system for protists, published by the Society of Protozoologists, *S. destruens* stands in the super-group Opisthokonta (Mesomycetozoa: Ichthyosporea: Rhinosporideacae) (Adl et al. 2005). More recently, based on phylogenomic studies *S. destruens* was placed in a new clade termed as “teretosporea” comprised of Ichthyosporea and *Corallochytrium limacisporum* and the group was found to be the earliest Holozoan divergence followed by Filasterea and Choanoflagellatea (Figure 1.12; Torruella et al. 2015).

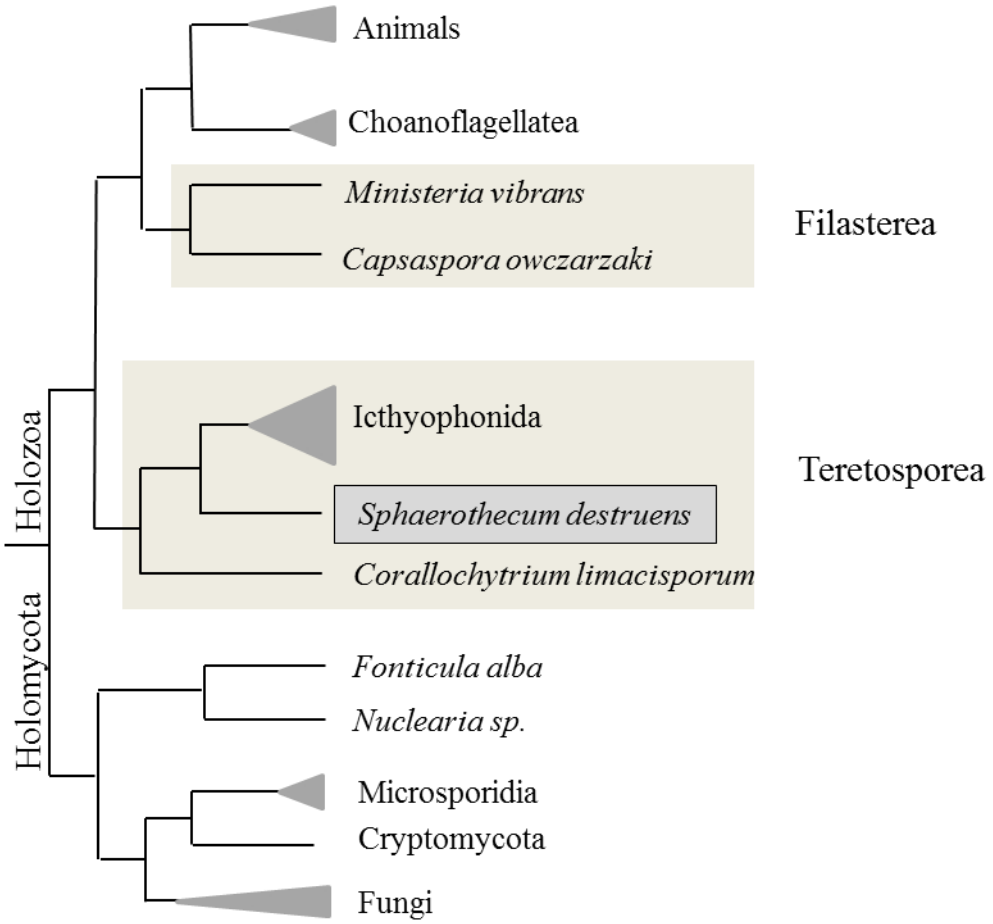


Figure 1.12. Phylogenetic relationships among Opisthokonts (Holozoa and Holomycota) based on phylogenomic data using flagellum and chitin synthases characters, adapted from Torruella et al. (2015).

1.6 Current *Sphaerothecum destruens* detection techniques

Current detection methods for the parasite *S. destruens* in its host are performed through microscopic examination of the histologic sections of visceral organs (Arkush et al. 1998) and molecular detection through the amplification of the 18S rRNA gene (Mendonca and Arkush 2004). The histological examination of sections can give insights into the host pathology and the potential impact of the parasite in the host. However, it has a low sensitivity of detection in particular for hosts where the disseminated form of the disease is most prevalent. This applied to many parasites and as a result detection of the parasite DNA through the use of the Polymerase Chain Reaction (PCR) amplification is often used (Tsui et al. 2011).

Thus, the designed PCR-based detection is a more powerful diagnostic tool for *S. destruens* detection, especially in sub clinically infected fish such as its reservoir host *P. parva* (Mendonca and Arkush 2004). The molecular detection uses a nested PCR which amplifies a 434 bp fragment of the *S. destruens* 18S rRNA gene, with a reported detection limit of 1 pg for purified *S. destruens* genomic DNA. The method was also found to be specific as *S. destruens* specific primers did not result in amplification for the related salmonid parasites *Ichthyophonus hoferi* and *Dermocystidium salmonis* of the Class Mesomycetozoa, in single-round and nested PCR assays (Mendonca and Arkush 2004). However, the PCR detection sensitivity is also affected by the DNA extraction efficiency. The *S. destruens* spores have thick cell wall which is resistant to many DNA extraction methods (Mendonca and Arkush 2004), potentially limiting the detection of the parasite in lowly infected hosts and increasing the probability of false negatives i.e. individuals are falsely identified as not carrier of the parasite. This is particularly relevant when using molecular methods for screening *P. parva* populations for the presence of *S. destruens.* Detection sensitivity in fish tissues could be improved by amplifying mitochondrial DNA regions due to the presence of multiple mitochondria in the cell compared to a single nucleus (Avise 2000). Furthermore, the rate of nucleotide substitution is higher in mitochondrial genes compared to nuclear genes, making it a potentially important phylogenetic marker (Brown et al. 1979, Avise 2000). In addition, epidemiological studies of the parasite would benefit from an environmental DNA detection method which can be used to non-invasively screen water bodies identified as 'at risk' from the parasite.

1.7 Environmental DNA as a detection tool

Environmental DNA (eDNA) is a new emerging technique that involves the study of DNA from environmental samples e.g. soil, water and sediments. The DNA source in these bodies can be through the release of faeces, saliva, urine, skin cells, and body secretions of inhabiting species (Rees et al. 2014). The eDNA technique involves the collection of environmental samples (e.g. water or soil), DNA extraction and PCR amplification of DNA of the target organism through specific primers. Application of eDNA is important in the detection of multiple or single species at a particular location. eDNA detection has gained much popularity in the fields of ecology and conservation biology (Ficetola et al. 2008), as biodiversity studies, ecological management, and

conservation programs all require an efficient detection method for rare, invasive or endangered species (Ficetola et al. 2008, Jerde et al. 2011). In addition, eDNA based detection are less time consuming and in some cases more sensitive compared to the traditional survey methods (Beja-Pereira et al. 2009). The technique has been successfully employed across a variety of habitats from terrestrial and aquatic sediments, including ice, soil, fresh water and sea water, with different approaches of sample collection methods and volumes, preservation methods and DNA extraction methods reviewed in Rees et al. (2014), Thomsen and Willerslev (2015).

The use of eDNA from water bodies for species detection was first employed to detect the American Bullfrog *Rana catesbeiana* (Ficetola et al. 2008). This is a highly invasive species which has spread around the globe (Blaustein and Kiesecker 2002, Kats and Ferrer 2003). The technique was efficient in detecting the species at selected positive sites from previous surveys (Ficetola et al. 2007a, 2007b) even at very low densities (Ficetola et al. 2008). eDNA has since been used to detect a vast range of aquatic species including amphibians (Goldberg et al. 2011, Dejean et al. 2012, Pilliod et al. 2013, 2014), fishes (Dejean et al. 2011, Jerde et al. 2011, Minamoto et al. 2012, Jerde et al. 2013, Mahon et al. 2013), arthropods (Thomsen et al. 2012) and gastropods (Goldberg et al. 2013).

eDNA can be a useful tool for detecting and quantifying parasites and their hosts in water samples. One of the emerging fungal diseases of amphibians is Chytridiomycosis. The causative agent of the disease is *Batrachochytrium dendrobatidis*, whose infectious zoospores affect the epidermis of adult amphibians and mouth parts of anuran larvae (Garner et al. 2005). Earlier surveillance strategies of amphibians involved the detection of zoosporangia in the host amphibians through microscopic or molecular techniques (Berger et al. 1998, Garner et al. 2005). Later, the pathogen was discovered outside its host and persisted in the environment for weeks to months (Johnson and Speare 2005), raises the need of a parasite detection tool in environment. eDNA has been found as an effective tool in *B. dendrobatidis* detection in environmental samples from Spain. The detection of this intracellular parasite outside its host eliminates the need of killing live hosts for pathogen detection (Walker et al. 2007). Additional parasites for which eDNA detection has been applied include the aetiological agent of crayfish plague

Aphanomyces astaci which also releases zoospores in the water (Strand et al. 2014), a number of viruses such as Cyprinid Herpesviruses and Rhabdoviruses (Minamoto et al. 2015) and a trematode *Ribeiroia ondatrae* (Huver et al. 2015).

In freshwater ecosystems, short DNA fragments can be detectable up to one month of their release in the environment (Dejean et al. 2011). However, experimental studies show that DNA degradation increases at high temperatures, high UV-B and neutral pH (Strickler et al. 2015). This indicates that eDNA persistence can vary greatly based on the environmental conditions of the aquatic habitat (Strickler et al. 2015), increasing the probability of false negative results due to low quantities of DNA (Thomsen et al. 2012). This can be overcome by increasing field samples and multiple PCRs per sample (Ficetola et al. 2008). A more reliable method employed is quantitative PCR (qPCR) which quantifies the species DNA in real time and a sample is considered positive even if one PCR replicate surpasses the fluorescence threshold (Takahara et al. 2013). Other factors that can influence the detection include the sample processing techniques and the amount of template DNA used. For example, freezing and thawing of the samples prior to filtration and larger volume of DNA template solution (5 µl) for the PCR tend to have lower detection rates compared to non-frozen samples and smaller starting DNA volumes (e.g. 2 µl) (Takahara et al. 2015). eDNA amplification can also be inhibited by humic substances co-extracted with eDNA which inhibit the functionality of Taq DNA polymerase enzyme (McKee et al. 2015). The post-treatment of the eDNA samples with 10-fold dilution or spin-column purification are effective in reducing inhibitors effects but can also affect the assay sensitivity (McKee et al. 2015). The generation of both false negatives and positives can have consequences for the monitoring and conservation programmes (Thomsen and Willerslev 2015). Despite the potential limitations associated with eDNA detection, an eDNA method for *S. destruens* detection can help in the development of epidemiological maps for the parasite and a quick screening of fish consignments before their introduction to different water bodies.

1.8 Mitochondrial DNA and its evolutionary trends

Mitochondria are double membrane-bound organelles commonly known as powerhouse of the cells whose origin within the eukaryotic cells is an important and still debated

evolutionary event. The widely accepted hypothesis explaining the origin of mitochondria in cells is the "endosymbiosis theory" which suggest that approximately 2 billion years ago mitochondria evolved only once from bacteria that were in an endosymbiotic relationship with their unicellular hosts (Gray et al. 1999, Dyall et al. 2004). Extensive biochemical and phylogenetic studies have demonstrated that mitochondria did originate from a single alpha-proteobacterial ancestor (Gray et al. 1999) from within the order Rickettsiales living inside the host cell (Wang and Wu 2015).

Mitochondrial DNA (mtDNA) in multicellular animals is usually described as a small, circular molecule with compactly arranged intron-less genes. However, studies carried out over the last decade have considerably changed this view. Variations to the standard mitochondrial genome (mt-genome) size and content were observed particularly in non-bilaterian animals. These variations ranged from remarkable diversity in mt-genome architecture (single to multiple linear and circular chromosomes), presence of extra genes (*atp9*, *polB* and *tatC*), different number of encoded tRNAs (0-25), presence/absence of introns to a large range of genome sizes (Lavrov and Pett 2016).

Mt-genome diversity was observed in samples from Choanoflagellate, Ichthyosporea and Filasterea. These close unicellular relatives of Metazoa termed as Choanozoa have large mt-genomes (50 kbp-200 kbp), are spacious with long intergenic regions (greater than 100 bp) or repeat sequences, and are gene-rich. Unique mt-genome architecture was found in these organisms which ranged from single linear-chromosome to multiple linear-chromosome and single circular-chromosome (Burger et al. 2003a, Lavrov and Lang 2014). These expansionary trends in mt-genomes of protists led to the hypothesis of mitochondrial evolution from a common Holozoan (Metazoa and their unicellular relatives) ancestor along three different routes (Burger et al. 2003a). First route includes (i) the Ichthyosporea lineage (which includes *S. destruens*). The only currently available mt-genome from the Ichthyosporea belongs to *Amoebidium parasiticum* whose mt-genome underwent fragmentation and rampant expansion through the accumulation of repeat sequences (Burger et al. 2003a). The secondary route involves the expansion of mt-genome through the accumulation of long stretches of intergenic regions in Choanoflagellate lineage (*Monosiga brevicollis*). The third metazoan lineage underwent

extensive gene-loss coupled with genome contraction (Figure 1.13; Burger et al. 2003a). The compaction and gene-loss of the mt-genome was attributed to the two main events in animal evolution, the emergence of multicellularity and bilateral symmetry. However, there is no evidence if these changes in the mt-genome architecture co-occurred with the morphological transitions in animals (Figure 1.13; Lavrov 2007).

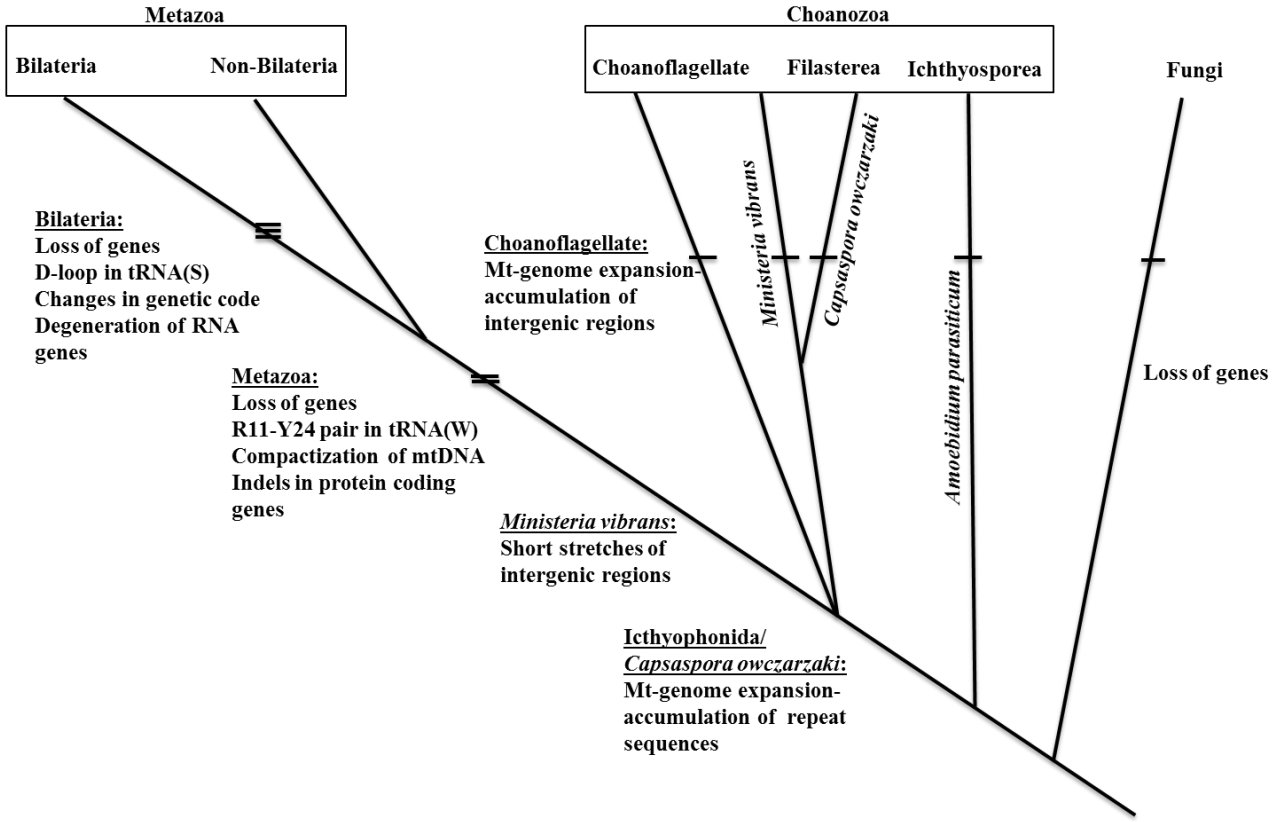


Figure 1.13. Mitochondrial evolution in the Holozoa (Metazoa + Choanozoa). The above figure illustrates the events involved in mtDNA evolution. along three main evolutionary routes (Burger et al. 2003a) with respect to the two main transitional events in animal evolution i.e. the emergence of cellularity and bilateral symmetry- adapted from Lavrov (2007). Data for Filasterea obtained from Lavrov and Lang (2014).

There had been a general assumption that animals have compact mitochondrial genomes compared to unicellular protists and fungal mtDNA. This notion was challenged with the discovery of large mitochondrial genomes ranging from 32 kbp-43 kbp in the basal animal phylum Placozoa (Signorovitch et al. 2007). Based on this discovery, it was hypothesized that the ancestral animal mitochondrion was a large, non-compacted molecule. Due to the intermediate mt-genome sizes between animals and unicellular protists, it was proposed that the mt-genome compaction occurred secondarily after the emergence of metazoans. Based on phylogenetic analysis using mitochondrial derived amino acid sequences, the phylum Placozoa was determined as the earliest offshoot from the metazoan lineage. However, the recent multigene studies have assigned Porifera as the earliest Metazoan divergence followed by Placozoa and Cnidaria (Torruella et al. 2012, Torruella et al. 2015), which challenges the above hypothesis. The phylogenetic relationships in basal animals and their unicellular relatives are subject to changes depending on the extensive taxonomic sampling and the set of genes studied for the analysis (Ruiz-Trillo et al. 2008, Shalchian-Tabrizi et al. 2008, Torruella et al. 2015).

Sphaerothecum destruens is a unicellular pathogen from Order Dermocystida (Class Mesomycetozoa) which sits at the animal-fungal boundary. To date, *S. destruens'* phylogeny has not been determined based on its mtDNA. The presence of large mt-genome (> 200 kbp) in *A. parasiticum*, which is comprised of several of hundreds of linear chromosomes (Class Ichthyosporea: Order Icthyophonida) raises the question if the same expansionary trends and peculiar mt-genome architecture is spread across the Class Ichthyosporea. Moreover, the sequenced mtDNA from Choanozoa is very limited with only four sequenced organisms to date. Ichthyosporea is considered as an earliest Holozoan divergence (Figure 1.12; Torruella et al. 2015) and mt-DNA genomic data from this group can provide valuable insights into the understanding of mitochondrial evolution.

1.9 Aims and objectives

The increasing occurrence of *S. destruens* in Europe following its initial record in 2005, in addition to its generalist nature, highlight the need to establish the origin of this

parasite to Europe (i.e. native or non-native) as this can have important implications for its monitoring and management. Thus, this PhD thesis aimed at:

1. Developing new PCR detection methods with phylogenetically informative DNA markers to investigate the phylogenetic relationships of *S. destruens* from various geographic locations (Chapter 2).
2. Investigating *S. destruens'* potential non-native status through the detection of *S. destruens* in *P. parva* across its native and non-native range (Chapter 3).
3. Creating a detailed epidemiological map for *S. destruens* presence in England and Wales (Chapter 4).
4. Developing and applying an eDNA detection technique for *S. destruens* which can be used to assess its presence in freshwater samples as well as act as a prevention tool for *S. destruens* spread between waterbodies.
5. Investigating *S. destruens* phylogeny based on its mitochondrial genome and analysing its mt-genome architecture.

1.10 Overview of the chapters

Chapter 2: Testing the specificity and sensitivity of two new DNA markers for the detection of *Sphaerothecum destruens* in fishes

In this chapter, two new DNA markers (ribosomal ITS 1 and mt Cyt-b) for improved detection of *S. destruens* in its source host *P. parva* were tested for their specificity and sensitivity. These markers were then used in Chapters 3 and 4.

Chapter 3: Global distribution of the *Sphaerothecum destruens* reveals its non-native status for Europe.

Chapter 3 investigated the global distribution of *S. destruens* by screening 21 *P. parva* populations across its native and non-native range in order to determine the parasite's origin in Europe.

Chapter 4: Epidemiology of *Sphaerothecum destruens* in Britain

In this chapter, the first epidemiological map of *S. destruens* in Britain was created and the risk of disease transfer to native fish species was assessed through histopathology and molecular detection.

Chapter 5: Development and application of environmental DNA detection assay for *Sphaerothecum destruens*

A fast and cost-effective eDNA detection method for *S. destruens* was developed and applied in this chapter and provided important insights in disease persistence in native freshwater communities.

Chapter 6: *Sphaerothecum destruens* taxonomy and mitochondrial genome organisation

The phylogenetic position of *S. destruens* was investigated using its mt-genome data with unique insights into the mtDNA genome evolution.

Chapter 7: Discussion

This chapter synthesises the main findings and makes recommendation for the management of *S. destruens* in Britain and other European countries.

Chapter 2

Testing the specificity and sensitivity of two new DNA markers for the detection of *Sphaerothecum destruens* in fishes

2.1 Introduction

The polymerase chain reaction (PCR) and other nucleic acid based assays are increasingly used for the detection of parasite infections, epidemiology and disease prevention (Weiss 1995, Mohammed et al. 2015). Their increasing use is due to the techniques being fast and more sensitive and specific compared to conventional diagnostic methods such as histology (Andree et al. 1998, Gonzalez et al. 2003). PCR can be particularly effective in detecting and classifying intracellular parasites (Mendonca and Arkush 2004). In most cases, it is difficult to physically remove intracellular parasites from host tissue, resulting in the parasite's DNA being present in a 'pool' of host DNA, further diluting its concentration. When infection is low, such as in the case of a reservoir host, the parasite DNA concentration can be below the PCR's detection limit, leading to false negatives (Taberlet et al. 1996).

Prior to using PCR as a detection method, it is important to determine the assay's detection sensitivity (or limit of detection- LoD) and specificity in order to both validate the assay and correctly interpret detection results (Burns and Valdivia 2008). Both the sensitivity and specificity of the PCR assay depends on a number of factors including target genes, primer sequences, type of PCR technique and DNA extraction procedures (Yamamoto 2002). PCR sensitivity can be determined through amplification of serial dilutions of parasite DNA in the presence and absence of host DNA, whereas PCR specificity can be determined by testing for cross amplification of closely related parasite species and host DNA.

In addition to parasite detection, the generated sequences obtained through PCR can also be used in phylogenetic studies (Lyubetsky et al. 2014, Patwardhan et al. 2014). Phylogenetic studies often use genes such as the nuclear ribosomal genes (16S rRNA, 18S rRNA and ribosomal ITS) and mitochondrial genes (COI, Cyt-b and control region) depending on the level of phylogenetic resolution required (Patwardhan et al. 2014). Mitochondrial DNA markers are often used in population level studies due to them being haploid, uniparentally inherited, can have regions that are highly variable and easily amplified (Hajibabaei et al. 2007, Dupuis et al. 2012).

The current detection methods for *Sphaerothecum destruens* in its hosts involve histology and DNA-based methods (PCR) (Mendonca and Arkush 2004). The developed 18S rRNA PCR assay can effectively detect low levels (1 pg) of *S. destruens* in its healthy host *Pseudorasbora parva* but cannot be used in population differentiation studies due to low genetic variability (Spikmans et al. 2013, Ercan et al. 2015). Phylogenetic studies have been carried out on *S. destruens* using the Internal Transcribed Spacer (ITS 1) gene (Gozlan et al. 2009) which is a highly polymorphic, non-coding region that separates the 18S and 5.8S nuclear ribosomal RNA genes and is widely used for phylogenetic and population analysis in fungi (White et al. 1990, Schoch et al. 2012). The published ITS 1 assay for *S. destruens* involved amplification of the gene from a single individual of *Leucaspius delineatus* (a highly susceptible host), and its sensitivity and specificity was not determined (Gozlan et al. 2009).

In order to address the objectives of Chapters 3 and 4 (Section 1.9), there was a need to develop detection assays for phylogenetically informative DNA markers that could detect *S. destruens* from its healthy host *P. parva*. Thus, the objectives for this chapter were to: 1) determine the sensitivity and specificity of the ITS 1 marker and redesign primers to improve these if necessary; 2) develop a detection assay for the mitochondrial DNA fragment spanning the Cytochrome b (Cyt-b) and intergenic region between Cyt-b and the cytochrome c oxidase (cox1) gene (hereafter referred to as Cyt-b) and 3) compare the LoD of the newly developed ITS 1 and Cyt-b detection assay with the gold standard for *S. destruens* DNA detection, the amplification of 18S rRNA region (Mendonca and Arkush 2004).

2.2 Materials and methods

A 2-step PCR (or nested PCR) is often used to detect low quantities of DNA. In a 2-step PCR, the amplified product is subjected to a second round of PCR with a set of nested primers located internal to the first PCR's primer set (Weiss 1995, Taberlet et al. 1996). Nested PCR can in some cases improve sensitivity by 1000 times compared to standard PCR (Yamamoto 2002). However, this can increase the risk of contamination and detection of false positives. It is thus necessary to use separate areas for PCR

preparations and amplified products, as well as adding appropriate positive and negative PCR controls (Kwok and Higuchi 1989).

In response to this, all DNA extractions from host tissues were performed in a lab dedicated to DNA extraction. PCR assays were performed in a UV-irradiated PCR hood and amplified products were handled in an area set out exclusively for gel electrophoresis. Appropriate negative controls were used throughout and included:

1) DNA extraction controls where the DNA extraction protocol was performed in the absence of host tissue (these were subjected to PCR amplification to detect any possible cross-contamination during DNA extraction);
2) PCR negative controls, where water instead of DNA was amplified in order to determine cross-contamination during the PCR and;
3) PCR positive controls where 10 ng of pure *S. destruens* DNA was amplified to ensure that the amplification was successful.

2.2.1 18S ribosomal RNA (rRNA)

The routine PCR used for *S. destruens* detection in its host is 18S rRNA (Mendonca and Arkush 2004). This already developed PCR was used as a reference PCR to compare the newly developed assays. In the 18S rRNA PCR, a 434 bp fragment is amplified using nested PCR. The sequence of the forward (Sd-1F) and reverse (Sd-1R) primers was 5'- CGA CTT TTC GGA AGG GAT GTA TT- 3' and 5'-AGT CCC AAA CTC GAC GCA CAC T-3', respectively. The first round of amplification yielded a 550 bp long amplicon. The second round assay amplified a 434 bp segment of the 18S rRNA gene; using the forward primer (Sd-2F): 5'-CCC TCG GTT TCT TGG TGA TTC ATA ATA ACT-3' and reverse primer (Sd-2R): 5'-CTC GTC GGG GCA AAC ACC TC-3'.
The reaction conditions for both PCRs were identical except for the starting template. In the first round PCR, the starting template DNA concentration was 300 ng; whereas in the second round 2 µl of the first PCR product was used. The reaction conditions were as follows; a reaction of 30 µl contained 1 X Promega Flexi buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Bioline), 0.3 µM forward and reverse primer and 0.5 U Taq polymerase (Promega). The cycling conditions included an initial denaturation cycle at 95ºC for 5

minutes followed by 35 cycles of 45 seconds at 95 ºC, 45 seconds at 60 ºC and 45 seconds at 72 ºC. A final elongation step at 72 ºC was performed for 7 minutes.

2.2.2 Internal transcribed spacer 1 (ITS 1) reaction conditions

The ITS 1 and ITS 2 markers for *S. destruens* developed by Gozlan et al. (2009) failed to amplify *S. destruens* ITS 1 and ITS 2 from *P. parva* due to strong cross-reaction with *P. parva* DNA. Thus, a new detection method was developed. The regions flanking the ribosomal internal transcribed spacer 1 (ITS 1) gene are 18S and 5.8S, which serve as good positions for primers. In order to develop primers specific to *S. destruens*, 18S and 5.8S gene sequences were obtained from GenBank for the fishes *Salmo trutta* (DQ009482), *O. mykiss* (FJ710873.1), *S. salar* (AJ427629.1) and *C. carpio* (FJ710826.1). The sequences were manually aligned with 18S and 5.8S rRNA gene sequences of deposited *S. destruens* strains (FN996945.1, AY267345.1, AY267344.1, AY388645.1 and FJ440702.1) in BioEdit, and primers for single round and nested PCR were carefully designed to avoid similar/conserved regions between fish and *S. destruens*. The designed primers were then tested against whole genomic *S. destruens* and *P. parva* DNA in PCR tests.

A list of primer combinations, PCR conditions and their output, that were tested in the development of this assay are presented in Table 2.1. The optimized primers for the 2-step PCR employed for amplification of the ribosomal ITS 1 were the forward primer Sdes2F of ss rDNA, (5'-CTT CGG ATT GGC CCT GTA C-3'), coupled with universal reverse primer NC 2, (5'-TTA GTT TCT TTT CCT CCG CT-3) in first step-PCR. The second round of PCR produced a 700 bp amplicon using the same forward primer Sdes2F and a reverse primer SD-ITS R1, 5'- CGATGCACGAGCCAAGAG-3'.

The reaction conditions were as follows: the reaction volume was 30 µl and 50 µl for first and second round of PCR respectively. The PCR reaction constituted of 1 X Promega Flexi buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.3 µM forward and reverse primer and 0.5 U Taq polymerase (Promega). The optimized cycling conditions were as follows; for the first round of PCR, an initial denaturation at 95 ºC for 3 minutes

followed by 35 cycles of 30 seconds at 95 ºC, 45 seconds at 60 ºC and 90 seconds at 62 ºC. A final extension step at 62ºC for 7 minutes. For the second step PCR, an initial denaturation at 95 ºC for 5 min followed by 35 cycles of 30 seconds at 95 ºC, 45 seconds at 59 ºC and 90 seconds at 62 ºC and a final extension step at 62 ºC for 7 minutes.

To confirm if the 700 bp amplicon which was present in *S. destruens* samples spiked with *P. parva* DNA is purely from *S. destruens* or a mixture of *S. destruens* and *P. parva* DNA; the 700 bp band was gel extracted and purified (Qiaquick Gel extraction kit, Qiagen). The purified bands from both 100 ng and 10 ng *S. destruens* DNA spiked with *P. parva* DNA were sent for sequencing (Beckman coulter genomics). The BLAST tool in the GenBank nucleotide database (NCBI) was used to confirm their identity.

2.2.3 Cytochrome b (Cyt-b)

Mitochondrial DNA is considered a suitable option when genetically exploring new species in the wild (Galtier et al. 2009). The sequence variation in *S. destruens* mitochondrial DNA is unknown, however, non-coding regions are highly likely to produce variable regions (Zuccarello et al. 1999). Accordingly, the Cyt-b-cox1 intergenic region (IGR) was targeted when designing the mitochondrial DNA assay. A cytochrome b sequence was obtained for *S. destruens* by employing the universal primers cobF424 (5'-GGWTAYGTWYTWCCWTGRGGWCARAT) and cobR876 (5'-GCRTAWGCRAAWARRAARTAYCAYTCWGG) (Burger et al. 2007). *Sphaerothecum destruens*-specific primers were then designed for single round and nested PCR. The primers were placed in positions that allowed the amplification of the Cyt-b gene fragment and Cyt-b-cox1 IGR. The primers employed for amplification of mt Cyt-b gene fragment (630 bp) plus IGR (~60 bp) were Nt-CytB-F1 (5-ATGAGTTTATGGGGAGCG) coupled with Nt-CytB-R1 (5-GCTCCAGCCAACACAGGTAAGGATAATAAC) in the first step-PCR. The second round PCR produced ~ 700 bp fragment; employing primer Nt-CytB-F2: (5-GGAGGGTTTAGTGTGGATAATGC) coupled with Nt-CytB-R1: (5-TCATCGTCAAATCCAACTCACC).

The reaction conditions included a reaction volume of 30 µl and 50 µl for first and second round of PCR respectively. The PCR reaction constituted of 1 X Promega Flexi buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.3 µM forward and reverse primer and 0.5 U Taq polymerase (Promega). The optimized cycling conditions were as follows; for first round of PCR, For first round of PCR, an initial denaturation at 95 ºC for 2 minutes, 35 x [95 ºC for 40 seconds, 56 ºC for 40 seconds, 72 ºC for 60 seconds] and a final extension step at 72 ºC for 5 min, for the second step PCR, 95 ºC for 2 minutes, 35 cycles [95 ºC for 40 seconds, 58 ºC for 40 seconds, 72 ºC for 60 seconds] and a final extension step at 72 ºC for 5 minutes.

To confirm if the amplified product was the targeted Cyt-b region, the fragments were sent for sequencing (Beckman Coulter genomics). The obtained sequences were blasted against GenBank nucleotide database to confirm their identity.

2.2.4 Determining specificity and sensitivity of the developed *Sphaerothecum destruens*-specific assays for ITS 1 and Cyt-b

In the current study, improved methods with higher sensitivity and specificity were required for improved detection of *S. destruens* in its healthy host *P. parva*. The specificity of the developed markers was tested by running PCR tests with *S. destruens* DNA and pure host fish *P. parva* DNA. The assay sensitivity was determined by amplifying serial dilutions of pure genomic *S. destruens* DNA and pure genomic *S. destruens* DNA spiked with *P. parva* DNA. Specifically, to determine and compare the detection limits of *S. destruens*-specific nested PCR, genomic *S. destruens* DNA at 50 ng/µl was serially diluted to 0.005 ng/µl for ITS 1 and 0.00005 ng/µl for Cyt-b marker in sterile UV-irradiated water. In the first PCR round, 2 µl of each dilution was used as starting template giving a range of 100 ng/µl to 0.01 ng/µl for ITS 1 and 100 ng/µl to 10 fg/µl for Cyt-b, of pure *S. destruens* genomic DNA. All amplifications were performed in duplicate. Nested PCR was performed both in absence and presence of *P. parva* genomic DNA (300 ng) to detect the specificity of PCR for *S. destruens* DNA in fish tissue.

2.2.5 Cross reactivity with other susceptible host fish species

In order to avoid false positive results (a false positive would mean that PCR yielded a positive result in the absence of parasite DNA (Whipps et al. 2006), the specificity of the Cyt-b primers was tested by carrying out the Cyt-b PCR using a number of *S. destruens* susceptible host fish species. The fish species tested for cross-amplification were carp *C. carpio*, roach *Rutilus rutilus*, minnow *Pimephales promelas*, common bream *Abramis brama*, chub *Squalius cephalus* and barbel *Barbus barbus*. PCR tests were carried out with pure fish DNA (300 ng). The cycling conditions were as described in section 2.2.3.

2.2.6 Comparison of limit of detection in light of DNA extraction efficiency

The efficiency of the DNA extraction technique can affect the sensitivity of the PCR diagnostic technique (Ghosh and Weiss 2009). The DNA extraction efficiency for *S. destruens* spores in the presence and absence of fish tissue (15 mg *C. carpio* kidney) was determined to be 500 spores and 50 spores respectively using the DNeasy Blood and Tissue kit (Qiagen) see Table 2.2 (Andreou 2010). The detection limits of *S. destruens*-specific 18S rRNA PCR were determined to be 10 pg and 1 pg in the presence and absence of *C. carpio* DNA respectively. Accordingly, in the absence of fish tissue, 50 *S. destruens* spores were considered equivalent to 1 pg of genomic DNA and likewise in the presence of fish tissue, 500 spores were considered equivalent to 10 pg of genomic DNA.

Table 2.1. List of primers, polymerase chain reaction (PCR) conditions and their output employed in the development of Internal Transcribed Spacer 1 marker

Aim	Primer combinations	PCR conditions	Results
Testing conditions published in Gozlan et al. (2009)	**Sdes2F:** (5-CTT CGG ATT GGC CCT GTA C) **NC 13R:** (5-GCT GCG TTC TTC ATC GAT)	95 ºC -3 min 35x (95 ºC for 30 s, 56 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min	No amplification. Optimized annealing temperature. At 52 ºC annealing temperature, multiple bands were obtained. The sequence reads indicated mixed amplification possibly due to cross reaction with fish DNA.
Development of nested PCR	**Nested PCR 1-2 Sdes2F /NC13R** **Nested PCR 2-2 Sdes2F/SD-ITS-R1** **SD-ITS-R1:** (5-CGA TGC ACG AGC CAA GAG-3)	95 ºC -3 min 35x (95 ºC for 30 s, 52 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min 95 ºC -3 min 35x (95 ºC for 30 s, 59 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min	Amplified *S. destruens* DNA in *Oxynoemacheilus* sp. (Ercan et al. 2015) but did not produce amplification for *S. destruens* in *P. parva*
	Nested PCR 1-2: SD-ITS-4F/SD-ITS-4R **Nested PCR 2-2: SD-ITS-5F/SD-ITS-RA**	95 ºC -5 min 35x (95 ºC for 45 s, 57 ºC for 45 s, 62 ºC for 45 s) 62 ºC for 7 min 95 ºC -3 min 35x (95 ºC for 45 s, 56 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min	Amplification produced but sequencing showed mixed read (indicating probability of fish DNA amplification).

Specificity to parasite DNA	**Nested PCR 1-2: Sdes2F/ SD-ITS-R1** **Nested PCR 2-2: SD-ITS-5F/ SD-ITS-R1** **SD-ITS-5F:** (5-AGTGAGGCTGCCGAAAAGTT)	95 ºC -5 min 35x (95 ºC for 45 s, 59 ºC for 45 s, 62 ºC for 45 s) 62 ºC for 7 min 95 ºC -3 min 35x (95 ºC for 45 s, 52 ºC, 54 ºC, 56 ºC, 58 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min	Amplification was observed for both pure *P. parva* and *S. destruens* DNA at all gradient temperatures.
	Nested PCR 1-2: Sdes2F/ SD-ITS-R1 **Nested PCR 2-2: SD-ITS-4F/ SD-ITS-R1** **SD-ITS-4F:** (5-GATTGGCCCTGTACCGCTG**)**	95 ºC -5 min 35x (95 ºC for 45 s, 59 ºC for 45 s, 62 ºC for 45 s) 62 ºC for 7 min 95 ºC -3 min 35x (95 ºC for 45 s, 52 ºC, 54 ºC, 56 ºC, 58 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min	Amplification was observed for both pure *P. parva* and *S. destruens* DNA at all gradient temperatures
	Nested PCR 1-2: Sdes2F/NC2: (5-TTA GTT TCT TTT CCT CCG CT) **Nested PCR 2-2: Sdes2F/SD-ITS-R1**	95 ºC -3 min 35x (95 ºC for 45s, 52 ºC, 54 ºC, 56 ºC, 58 ºC for 45 s, 62 ºC for 90 s) 62 ºC for 7 min 95 ºC -5 min 35x (95 ºC for 45 s, 59 ºC for 45 s, 62 ºC for 45 s) 62 ºC for 7 min	At 60 °C the amplification product at around 700bp for *S. destruens* DNA and at 500bp for *P. parva* DNA, indicates cross-amplification with *P. parva* DNA

2.3 Results

2.3.1 Development of new detection methods for *Sphaerothecum destruens*

Internal Transcribed Spacer 1

The published ITS 1 conditions (Gozlan et al. 2009) developed for the detection of *S. destruens* DNA in *L. delineatus* were found to cross-react with *P. parva* DNA. Consequently, the ITS 1 conditions were re-optimized for *S. destruens* presence in *P. parva* (Table 2.1).

The primer pair found successful in amplifying *S. destruens* DNA in its reservoir host was Sdes2F coupled with NC2 in the first round of PCR. The second round employed Sdes2F and SD-ITS-R1. It was impossible to amplify only *S. destruens* DNA in the presence of fish DNA due to cross amplification with fish DNA. However, using an annealing temperature of 60 °C for primer pair Sdes2F and NC2 in the first round, the size of amplified products obtained for the ITS 1 gene (in second round) were different for *S. destruens* DNA (700 bp) and *P. parva* DNA (500 bp) (Figure 2.1), allowing for the *S. destruens* specific amplicon to be gel extracted and sequenced.

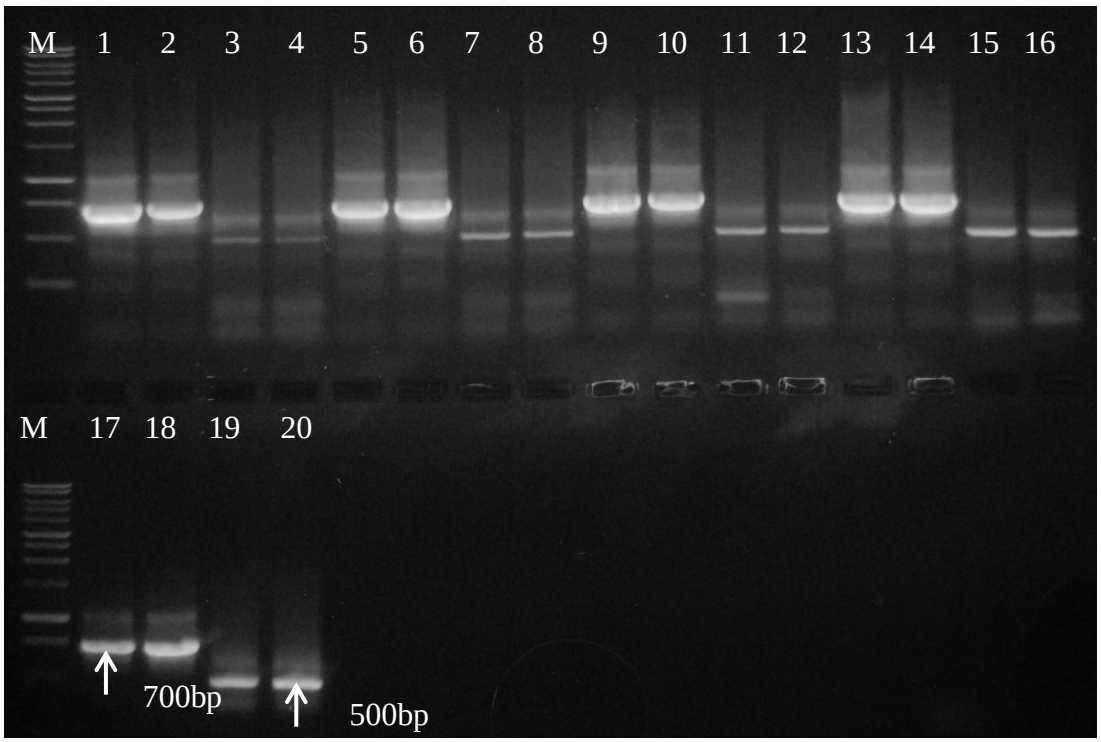


Figure 2.1. PCR amplification of ITS 1 gene by primer pair Sdes2F and NC2 in first round gradient PCR and Sdes2F and SD-ITS-R1 in second round of PCR tested with pure *Sphaerothecum destruens* and *Pseudorasbora parva* DNA. Second round PCR products were migrated on gel. Lane M: 1 Kb DNA ladder, Lane 1, 2, 5, 6, 9, 10, 13, 14, 17, 18: second round PCR products from *S. destruens* DNA at 52 ºC, 54 ºC, 56 ºC, 58 ºC and 60 ºC. Lane 3, 4, 7, 8, 11, 12, 15, 16, 19, 20: second round PCR products from *P. parva* DNA at 52 ºC, 54 ºC, 56 ºC, 58 ºC and 60 ºC. Duplicate samples were loaded sequentially.

2.3.2 Detection limits of developed markers for *S. destruens* DNA in presence and absence of fish DNA *P. parva*

2.3.2.1 Internal Transcribed Spacer 1

In order to determine and compare the efficiency of *S. destruens*-specific ITS 1 nested PCR, the reaction was performed on 10-fold serial dilution of genomic *S. destruens* DNA in presence and absence of *P. parva* genomic DNA. Using 10-fold serial dilutions of *S. destruens* genomic DNA, the amplification limit of nested PCR was 1ng of total *S. destruens* DNA. With the addition of 300 ng of total *P. parva* genomic DNA, no inhibition of *S. destruens*-specific ITS amplification (700 bp) was found. Additionally, very light bands appeared at 0.1 ng of *S. destruens* DNA spiked with *P. parva* genomic DNA (300 ng). Some non-specific amplification was found with *P. parva* DNA (~500

bp long) across several dilutions (Figure 2.2). The 700 bp bands were identified as pure *S. destruens* ITS sequence.

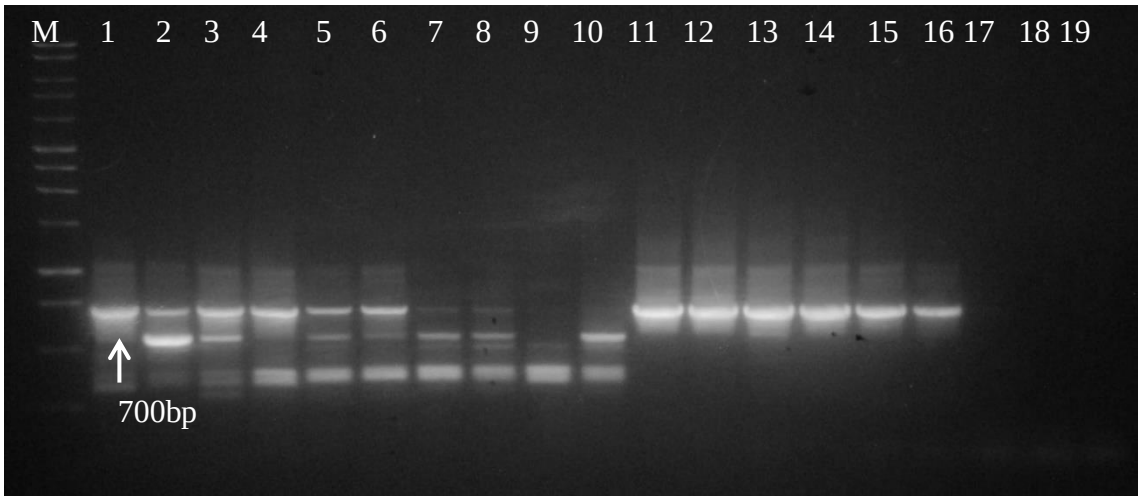


Figure 2.2. Detection limit of *Sphaerothecum destruens*-specific nested internal transcribed spacer (ITS 1) gene PCR of pure *S. destruens* genomic DNA and in presence of *Pseudorasbora parva* genomic DNA. Lane M (1 Kb DNA ladder); Lane 1-10: second round PCR products from 100 ng, 10 ng, 1 ng, 0.1 ng, 0.01 ng spiked with 300 ng *P. parva* genomic DNA. Duplicates were loaded next to one another. Lane 12-20, second round PCR products from 100 ng, 10 ng, 1 ng, 0.1 ng, 0.01 ng *S. destruens* total genomic DNA.

2.3.2.2 Cytochrome b

In order to determine the detection limit of *S. destruens*-specific Cyt-b marker, the reaction was performed in 10-fold serial dilution of *S. destruens* genomic DNA only in the absence of *P. parva* DNA, as the primers were not found to cross-react with fish DNA (Figure 2.5). The detection limit was 0.1 pg of total *S. destruens* DNA (Fig 2.3). The sequences were identified as pure *S. destruens* mitochondrial Cyt-b fragment. With the addition of 300 ng of *P. parva* genomic DNA, there was no inhibition of *S. destruens* Cyt-b amplification up to 1 pg concentration of *S. destruens* genomic DNA. However, at 0.1 pg there was positive amplification for only one replicate (Figure 2.4).

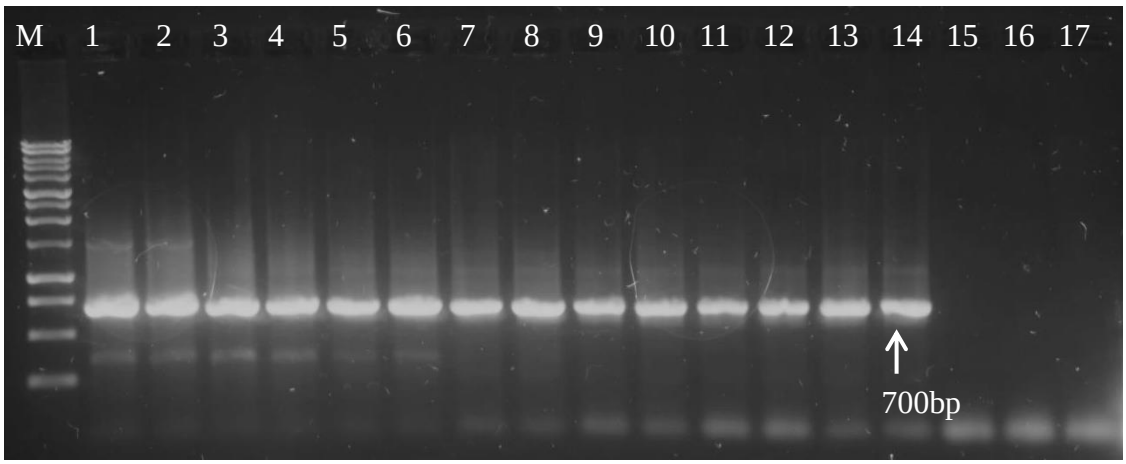


Figure 2.3. Detection limit of *Sphaerothecum destruens*-specific nested Cytochrome b (Cyt-b) gene PCR of *S. destruens* total genomic DNA. Lane M (1 Kb DNA ladder); Lane 1-16: second round PCR products from 100 ng, 10 ng, 1 ng, 0.1 ng, 10 pg, 1 pg, 0.1 pg and 10 fg of *S. destruens* total genomic DNA. Lane 17: negative PCR control. Duplicates were loaded sequentially.

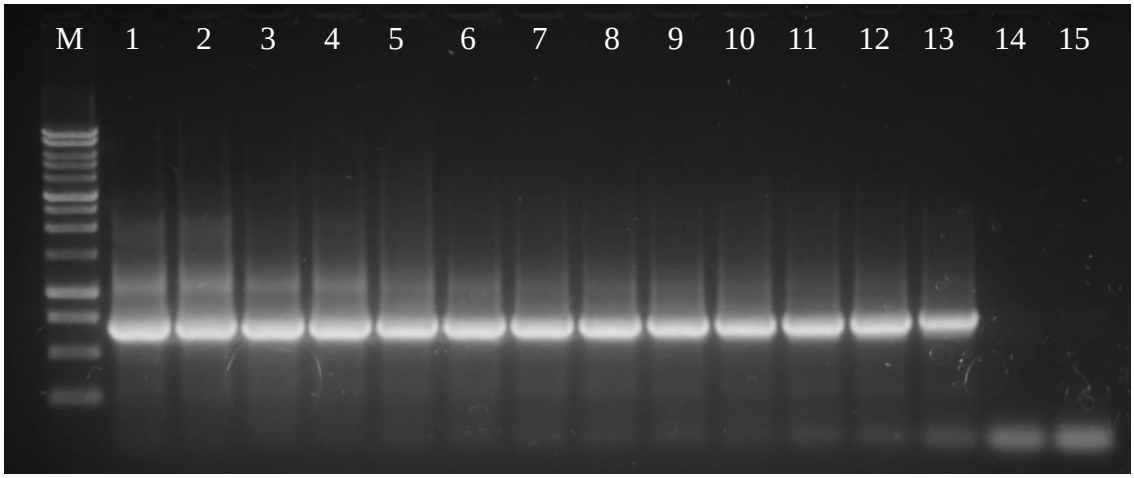


Figure 2.4. Detection limit of *Sphaerothecum destruens*-specific nested Cytochrome b gene PCR in the presence of *Pseudorasbora parva* genomic DNA. Lane M (1 Kb DNA ladder); Lane 1-13: second round PCR products from 100 ng, 10 ng, 1 ng, 0.1 ng, 10 pg, 1 pg, 0.1 pg of S. *destruens* total genomic DNA spiked with 300 ng *P. parva* DNA. Lane 15: negative PCR control. Duplicates were loaded next to one another.

2.3.3 Cross-reactivity with fish DNA

The Cyt-b primers did not cross-react with any of the tested fish species DNA (Figure 2.5). The specificity of Cyt-b assay makes it suitable to detect *S. destruens* from various host fish.

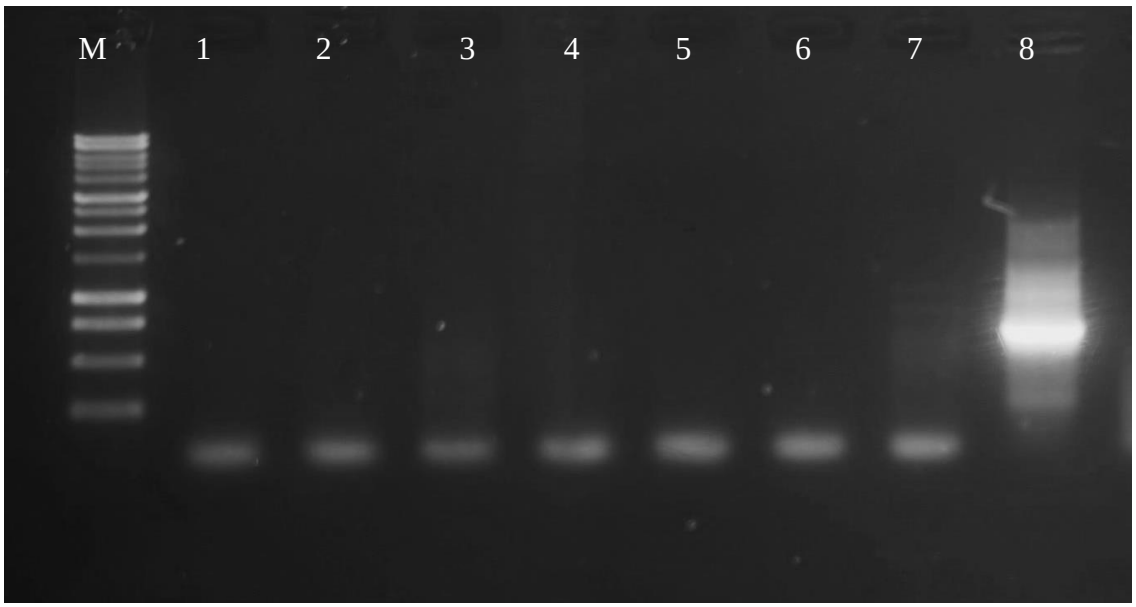


Figure 2.5. Cross-reactivity of *Sphaerothecum destruens*-specific nested Cytochrome b gene marker with fish DNA. Lane M: 1 Kb DNA ladder; Lane 1-7: second round PCR products from carp *Cyprinus carpio*, chub *Squalius cephalus*, barbel *Barbus barbus*, roach *Rutilus rutilus*, minnow *Pimephales promelas*, bream *Abramis brama* and topmouth gudgeon *Pseudorasbora parva*. Lane 8: positive PCR control.

2.3.4 Limit of detection comparison

The Cyt-b assay had the best LoD, with the assay detecting up to 0.1 pg of DNA in the presence of host DNA, followed by the 18S rRNA assay with a LoD of 1 pg and the ITS 1 with an LoD of 1 ng (Table 2.2).

Table 2.2. Comparison of the detection limit (LoD) of the developed markers ITS 1 and Cyt-b in comparison with the already developed 18S marker (Mendonca and Arkush 2004) in the absence and presence of fish DNA.

S. destruens		**Markers detection limit**					
Whole genomic DNA	**Spores equivalent**	**18S**		**ITS 1**		**Cyt-b**	
		Absence	**Presence**	**Absence**	**Presence**	**Absence**	**Presence**
1 ng	50,000	✓	✓	✓	✓	✓	✓
0.1 ng	5000	✓	✓	-	-	✓	✓
10 pg	500	✓	✓	-	-	✓	✓
1 pg	50	✓	✓	-	-	✓	✓
0.1 pg	5	-	-	-	-	✓	✓

2.4 Discussion

In this Chapter, two new assays for *S. destruens* were developed in order to be used in Chapters 3 and 4. The Cyt-b was the most sensitive and specific assay with the detection of *S. destruens* DNA at 0.1 pg compared to 1 pg for the 18S rRNA (Mendonca and Arkush 2004) and at 1ng for ITS 1 (Table 2.2). As expected the Cyt-b assay was the most sensitive and this is most probably due to the presence of multiple copies of mitochondria within an *S. destruens* cell leading to multiple copies of the targeted DNA region (Avise 2000). The assay was also highly specific in no cross-reaction with any of the fish species DNA tested. The low detection in the ITS 1 assay was most probably due to the strong cross-reaction with *P. parva* DNA. Due to the limited DNA region suitable for primer design (100 bp in 18S and 57 bp in 5.8S peripheral regions of ITS 1), a limited number of primers could be designed and the described assay was the best that could be developed with the available information. The sensitivity of the ITS 1 assay from Gozlan et al. (2009) was not determined (due to only DNA from infected host being available in that study), and could not be compared with the sensitivity reported here. Due to the low detection sensitivity of ITS 1 PCR and the inconsistency of Cyt-b PCR amplification at 0.1 pg concentration, a multiple tube approach could be applied where each positive sample (through 18S rRNA PCR) can be independently amplified

three times for the ITS 1 and Cyt-b region. This increases the probability of amplification and is often used in ancient or highly degraded DNA samples (Navidi et al. 1992, Taberlet et al. 1996).

Even though the detection of *S. destruens* DNA was improved by 10 fold using the Cyt-b assay (0.1 pg of pure genomic DNA), the detection ability of assay is still limited by the extraction efficiency of *S. destruens* spores which is limited to 500 spores in the presence of fish tissue (Andreou 2010). These extraction efficiencies are relatively comparable to 100 conidia, which are physiologically similar to *S. destruens* spores cell wall, for fungi *Geomyces destructans* per 2 mg wing tissue (Lorch et al. 2010) and inferior to 10 zoospores for *Batrachochytrium dendrobatidis* (Annis et al. 2004). Mendonca and Arkush (2004) could extract DNA from up to 10 *S. destruens* spores but this value was not consistently reproducible for their own work. The possible explanation for this could be the thick cell wall of *S. destruens* spores that is resistant to many DNA extraction methods (Mendonca and Arkush 2004). Resultantly, the number of spores that release DNA in the extraction process may account for a fraction of total number of spores initially enumerated (Mendonca and Arkush 2004).

For *S. destruens*, PCR is still a more sensitive method of detection, especially where host infection is in the disseminated form, which involves low numbers of spores with no histopathological response by the host decreasing the probability of detection through histology. However, the available PCR assays for *S. destruens* are not as sensitive as compared to the detection limits in the similar studies for *Microsporidium seriolae* (0.01 pg), *Ichthyophonus hoferi* (10^{-5} parasite spores) and *Geomyces destructans* (5 fg) (Bell et al. 1999, Whipps et al. 2006, Lorch et al. 2010) which limits the detection efficiency. Thus, when the assays are used in epidemiological work, it is highly likely that the determined prevalence will be an underestimate of the true prevalence of the parasite due to false negatives. This is particularly relevant to epidemiological studies using reservoir hosts such as *P. parva* which harbour very low infections, increasing the probability of false negatives. This will be problematic for *P. parva* populations, which test as *S. destruens* negative. In this case, it would be advisable to: a) increase the number of fish tested; b) revisit the site at different times, as *S. destruens* infections can vary seasonally with highest prevalence observed in the

spring (Ercan et al. 2015) and; c) if access to live fish is possible, cohabit the *P. parva* individuals with highly susceptible hosts (Gozlan et al. 2005).

2.5 Summary

In this study, two novel phylogenetic markers ITS 1 and Cyt-b were developed and optimized for the detection of *S. destruens* in its healthy host *P. parva*. The Cyt-b marker was the most sensitive (0.1 pg) and specific for *S. destruens* detection compared to the already developed 18S rRNA marker (1 pg) by Mendonca and Arkush (2004). The ITS 1 marker was improved for its specificity for *S. destruens* DNA in *P. parva* but its sensitivity was very low (1 ng) compared to the other two markers. The ITS 1 marker could be effectively applied for *S. destruens* detection in highly infected host fish.

Chapter 3

Global **s reveals its non-**

of the *Sphaerothe*

tive status

3.1 Introduction

The international trade of contaminated animal stocks, environmental changes and ecological disturbances have contributed to human-mediated disease outbreaks (Voyles et al. 2014). In particular, the trade of domestic, cultured and wild animals has translocated thousands of species across the globe that has resulted in a concurrent translocation of their pathogens leading to the global emergence of infectious diseases, a large proportion of these being caused by fungi and fungus-like organisms (Fisher et al. 2012). Notable examples include the chytrid fungus that is the causal agent for Chytridiomycosis that has led to the decline of over 200 amphibian species (Skerratt et al. 2007) and the white-nose syndrome in bats (Blehert et al. 2009).

Species translocations can introduce novel pathogens to new geographic areas which can lead to disease emergence in native species (Daszak et al. 2001). Fishes comprise the biggest group among the introduced aquatic animals with an estimated 624 freshwater fishes established outside their natural range at the end of 20th century (Gozlan et al. 2010b) and over one billion ornamental fish from over 100 countries are internationally traded each year (Whittington and Chong 2007). In Europe, the rise in the introduction of non-native fish species is mainly for aquaculture, recreational fishing and the aquarist trade (Peeler et al. 2011). There were at least 38 non-native freshwater fishes introduced to England and Wales by 2010 (Britton et al. 2010)

Non-native fish species introduced to Europe for aquaculture included the Japanese eel *Anguilla japonicas* which led to the introduction of the parasitic nematode *Anguillicoloides crassus* that has contributed to the decline of European eel *Anguilla anguilla* populations (Koops and Hartmann 1989, Kirk 2003). The import of fathead minnows *Pimephales promelas* to France from the USA has resulted in rainbow trout *Oncorhynchus mykiss* suffering from the enteric red mouth disease caused by bacterium *Yersina ruckeri* (Michel et al. 1986). Notably one of the world's most invasive species the topmouth gudgeon, *Pseudorasbora parva* has been accidentally introduced through the trade of Chinese carps (*Ctenopharyngodon idella, Hypophthalmichthys molitrix*) (Gozlan et al. 2002) and has potentially acted as the source of the emergent fish parasite *Sphaerothecum destruens* (Gozlan et al. 2005).

Sphaerothecum destruens, commonly known as the Rosette agent, is a novel obligate intracellular eukaryotic parasite which sits at animal-fungal boundary. *Sphaerothecum destruens* is a multi-host parasite that has resulted in steady chronic mortalities across salmonids (Harrell et al. 1986, Hedrick et al. 1989, Arkush et al. 2003) and centrarchids (Ercan et al. 2015). The first reported discovery of *S. destruens* was in the USA where it was associated with high mortalities in Chinook salmon *Oncorhynchus tshawytscha* (Harrell et al. 1986) and Atlantic salmon *Salmo salar* (Hedrick et al. 1989). The parasite was first reported in UK in 2005, associated with increased mortalities and spawning inhibition in the cyprinid sunbleak *L. delineatus* (Gozlan et al. 2005). Phylogenetic analysis of ribosomal Internal Transcribed Spacer 1 (ITS 1) gene identified the UK *S. destruens* isolate as geographically distinct from the USA isolates (Gozlan et al. 2009). Since 2005, the parasite has been reported from the Netherlands with a high prevalence (74 %) in *P. parva* (Spikmans et al. 2013) and Turkey where it was detected in *P. parva* and bass *Dicentrarchus labrax,* and has resulted in high mortalities in centrarchid species *Squalius fellowesii, Oxynoemacheilus* sp. and *Petroleuciscus smyrnaeus* (Ercan et al. 2015). The ITS 1 gene was only sequenced by Ercan et al. (2015) and has identified that the Turkish isolate is closely related to and groups with the UK isolate.

A pathogen's origin and phylogenetic history can provide valuable information in understanding the drivers of disease emergence and can be used to develop effective control strategies (Eskew and Todd 2013). Two hypotheses are often proposed for explaining a disease outbreak: (i) the novel pathogen hypothesis postulates that the disease outbreak is the result of a novel pathogen introduction to new geographic areas and (ii) the endemic pathogen hypothesis postulates that disease emergence is driven by environmental changes resulting in increased pathogenicity (Rachowicz et al. 2005). Both hypotheses could apply to the case of *S. destruens,* and in this study the novel pathogen hypothesis was tested by screening 21 *P. parva* populations across its native and non-native range for the presence of *S. destruens*. This was achieved by: (i) testing the origin of *S. destruens* using phylogenetic analysis with three markers; the 18S rRNA, ribosomal ITS 1 and mitochondrial Cyt-b gene on *S. destruens* isolates obtained across the globe; (ii) determining potential associations between specific *P. parva* haplotypes and the presence of *S. destruens;* and (iii) examining the demographic changes in *S. destruens* and its host *P. parva* populations through mismatch distribution analysis.

3.2 Material and Methods

3.2.1 Fish Dissection

Pseudorasbora parva obtained from different parts of the world were collected in July-August 2011 and were fixed in 100 % ethanol (Table 3.1). For *S. destruens* prevalence, 20 *P. parva* from each site were dissected and their kidneys were collected and frozen at -20 ºC until further testing. The fish were dissected by cutting the skin from anal fin along the belly of fish to the operculum, followed by removal of operculum and pectoral fin. Skin was cut from above the exposed gills going posteriorly along the side of the fish and then down to the anal fin. After the removal of skin, internal organs were exposed which were carefully removed and kidney were collected from the back of the fish. The dissection tools were sterilized with IMS (Industrial Methylated Spirit) after every dissection to avoid any cross-contamination.

3.2.2 DNA Extraction

DNA was extracted using the rodent tail protocol of the Qiagen DNeasy 96 Blood and tissue kit (Qiagen, Germany). All steps were performed according to manufacturer's guidelines with an overnight incubation at 55 ºC and elution volume of 200 µl. Extracted DNA from fish tissue was quantified by Nanodrop 2000 (Thermoscientific). Negative DNA extraction controls, where no tissue was added were also included.

3.2.3 *Sphaerothecum destruens* detection using PCR

Sphaerothecum destruens was detected using nested PCR which amplified a segment of the 18S rRNA gene of *S. destruens* using the oligonucleotide primers published in Mendonca and Arkush (2004) and as explained in Section 2.2.1. The PCR was performed in a PCR cabinet which was decontaminated using UV irradiation for 15 minutes prior and after each assay preparation. Negative PCR controls were inserted during PCR steps to detect cross-contamination. Due to the low detection sensitivities of the markers and low concentration of parasite DNA in the reservoir host, three independent PCRs were performed for each molecular marker.

All *S. destruens*-positive samples, through 18S rRNA PCR, were also amplified for the ITS 1 and Cyt-b regions with PCR and cycling conditions as described in Sections 2.2.3 and 2.2.4. The products of ITS 1 nested PCR was excised from the agarose gel and purified by Qiagen Gel Extraction kit. Purified fragments were cloned using TOPO TA cloning kit (Invitrogen, Life Technologies Paisley, United Kingdom) and sequenced with M13F and M13R primers (Beckman Coulter genomics). All amplified products were run on 1 % agarose gels stained with SYBR® safe DNA gel stain at 70 V for 40 min. The gel was viewed under UV transilluminator SAFE Imager (Invitrogen). In order to confirm that *S. destruens*-specific DNA was amplified all, amplified products were sent for Sanger DNA sequencing (Beckman coulter genomics).

3.2.4 Amplification of *Pseudorasbora parva* mitochondrial cytochrome b gene

PCR amplification of the cytochrome b gene was carried out on *P. parva* individuals that had tested positive for *S. destruens* using the primers L15267 (5-AAT GAC TTG AAG AAC CAC CGT-3') and H15891Ph (5-GTT TGA TCC CGT TTC GTG TA-3') developed by Briolay et al. (1998), resulted in amplified product of app. 600bp. The reaction conditions were as follows: reaction volume of 50 µl consisted of DNA template (100 ng), 1 X Promega Flexi buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.3 µM forward and reverse primer and 0.5 U Taq polymerase. The cycling conditions were as follows: an initial denaturation at 95 ºC for 15 min followed by 35 cycles of 30 seconds at 95 ºC, 90 seconds at 60 ºC and 60 seconds at 72 ºC, with a final elongation step at 72 ºC for 15 min. All amplified products were sent for Sanger DNA sequencing (Beckman coulter genomics).

3.2.5 Phylogenetic analysis

The obtained sequences for *S. destruens* were searched against GenBank nucleotide database to confirm their identity using the BLAST tool. All the sequences for the different markers (18S rRNA, ribosomal ITS 1 and Cyt-b gene) were aligned using Clustal W in BioEdit ver. 5.0.9 (Hall 1999) and examined by eye to eliminate ambiguities and to check for polymorphic sites. The haplotype diversity was calculated using DnaSP 5.10 (Librado and Rozas 2009) and was mapped by calculating network

through median joining by NETWORK (Bandelt et al. 1999); available at http://www.fluxus-engineeering.com).

A phylogenetic tree was drawn for *S. destruens* using the ITS 1 marker using Mr Bayes (Ronquist et al. 2012) using the sequences generated in the present study and those published in Gozlan et al. (2009) and Ercan et al. (2015): FJ440707.1, FJ440708.1, FJ440709.1, FJ440702.1, FJ440703.1, FJ440704.1, KT361608.1. jModeltest v 2.1.4 (Darriba et al. 2012) was used to determine the best model fitting the data.

The Cyt-b sequences for *P. parva* that had tested positive for *S. destruens* were grouped together with *P. parva* Cyt-b sequences across populations in its native range (China) and invasive range (Europe) (Simon et al. 2011). These populations included the populations tested for *S. destruens* in this study. The sequences were aligned by Clustal W in BioEdit (Hall 1999). Haplotypic diversity was calculated in DnaSP. A phylogenetic tree was constructed to identify the *P. parva* haplotypes associated with the presence of *S. destruens*. The phylogenetic analysis using Maximum likelihood was performed using Mr Bayes (Ronquist et al. 2012) and posterior probabilities were obtained after 2,500,000 generations with a burn-in of 25 %. The tree was calculated using Hasegawa-Kishino-Yano model with Gamma distribution (HKY+G) model (Hasegawa et al. 1985) determined with jModeltest v 2.1.4 (Darriba et al. 2012). The Cyt-b gene sequences from *Ictiobus bubalus* (JF799443.1), *Hypentelium nigricans* (JF799441.1), and *Danio rerio* (JN234356.1) were used as outgroups.

3.2.6 Population demographic analysis

To infer the demographic history of *S. destruens* and its host *P. parva* populations in China, the mismatch distribution analysis was carried on the 18S rRNA gene for *S. destruens* and Cyt-b gene for *P. parva*. This analysis plots the distribution of nucleotide differences between each pair of sequences and compares it to the expected values for a model of population expansion. A unimodal distribution is indicative of a population expansion in the recent past whereas a bimodal/multimodal distribution indicates a population that is at demographic equilibrium (Rogers and Harpending 1992).

Demographic changes were analysed by calculating Harpending's raggedness index (H_{ri}) which quantifies the smoothness of the observed mismatch distribution (Harpending et al. 1993) and the sum of squared deviations (SSD) between the observed and expected mismatch for the nucleotide differences (Schneider and Excoffier 1999) in Arlequin version 3.5 (Excoffier and Lischer 2010).

Three tests were used to test for population expansion of *P. parva*: Fu's *F*s test (Fu 1997), Tajima's *D* test (Tajima 1989) and Ramos-Onsins & Rozas' R_2 test (Ramos-Onsins and Rozas 2002). The latter was performed in DnaSP (Librado and Rozas 2009) and Fu's *F*s and Tajima's D test were carried out in Arlequin version 3.5 (Excoffier and Lischer 2010). Only the Tajima's D test and R_2 test were performed for *S. destruens*, as the Fu's *F*s test is not suitable for small sample sizes (Ramos-Onsins and Rozas 2002). For Tajima's *D* and *F*s, *P*-values were calculated based on a coalescent simulation algorithm and for the R_2 test the *P*-values were based on parametric bootstrapping with coalescence simulations.

3.3 Results

3.3.1 *Sphaerothecum destruens* detection and its prevalence across the sampled populations

A total of 420 *P. parva* fish representing 21 populations from across Eurasia (10 in China, 8 in Europe, and one population each from Morocco, Iran, and Japan) were screened for the presence of *S. destruens* (Figure 3.1). Out of the 10 Chinese populations, 9 were found positive for presence of *S. destruens* (Table 3.1 and Figure 3.1). The prevalence of the parasite in the Chinese populations ranged between 0 and 10 %. The overall prevalence of *S. destruens* across all sampled Chinese populations was 6 % (12/200). *S. destruens* was only found in two European populations; Spain and the United Kingdom (Figure 3.1 and Table 3.1), with a prevalence of 5 % in both populations. The overall prevalence in Europe was 1.4 % (2/140). *S. destruens* was not detected in samples from Morocco, Iran or Japan.

Table 3.1. Sampled populations of *Pseudorasbora parva* and the distribution of *Sphaerothecum destruens* in *P. parva* using molecular detection across *P. parva*'s native and non-native range. Tick (✓) indicates positive and hyphen (-) indicates no amplification with the particular marker.

Population	**Code**	**Geographical Coordinates**		**Sample size**	**Name of the Positive individuals**	**Genetic Marker amplified**		
		X	**Y**			**18S**	**Cyt-b**	**ITS 1**
Chinese site 1	S1	115.56	37.55	20	S1-15	✓	✓	-
					S1-7	✓	-	-
Chinese site 2	S2	117.12	34.81	20	S2-1	✓	-	-
Chinese site 3	S3	118.59	33.19	20	S3-16	✓	✓	✓
Chinese site 7	S7	110.32	25.27	20	0			
Chinese site 9	S9	113.11	29.15	20	S9-13	✓	-	-
Chinese site 11	S11	110.99	34.62	20	S11-20	✓	✓	-
Chinese site 12	S12	117	38.7	20	S12-10	✓	✓	-
					S12-19	✓	-	-
Chinese site 13	S13	122.52	40.10	20	S13-16	✓	✓	-
Chinese site 14	S14	124.99	45.03	20	S14-13	✓	-	-
					S14-19	✓	-	-
Chinese site 16	S16	118.27	40.90	20	S16-19	✓	-	-
Austria	A	14.72	48.19	20	0			
Bulgaria	BG	43	26	20	0			
France	F	-1.73	47.10	20	0			
Iran	IR	54.78	37.05	20	0			

Italy	IT	10	44	20	0			
Japan	JP	139.43	35.67	20	0			
Morocco	M	32.11	2.89	20	0			
Spain	SE	0.86	40.7	20	SE-7	✓	-	-
Turkey	T	30.04	40.91	20	0			
United Kingdom	UK	1	51	20	UK-1	✓	✓	-
Hungary	H	18	46	20	0			

3.3.3 Phylogenetic analysis of *Sphaerothecum destruens* using the 18S rRNA, ITS 1 and Cytochrome b sequences

A 397 bp fragment of 18S rRNA was obtained for 14 *S. destruens* individuals from 11 different populations (Table 3.1) and aligned with the published sequences from Arkush et al. (2003) and Paley et al. (2012). The overall haplotype diversity was calculated to be 0.29 and four haplotypes were identified. Only three individuals, *S. salar* (USA) and two *P. parva* (China) had different haplotypes (Figure 3.2).

Figure 3.1. *Pseudorasbora parva* sites screened in this study for *Sphaerothecum destruens* presence in the host's native range (China) and invasive range (Europe, North Africa and the Middle East). The red sites indicated sites with *S. destruens* positive *P. parva* and black sites represent sites without *S. destruens* positive individuals using the 18S rRNA detection method. Abbreviations: Europe- A: Austria; BG: Bulgaria; F: France; H: Hungary, IT: Italy; SP: Spain; T: Turkey and UK: United Kingdom. North Africa- M: Morocco. Middle east- I: Iran. Asia- China (S1, S2, S3, S7, S9, S11, S12, S13, S14, S16); JP: Japan.

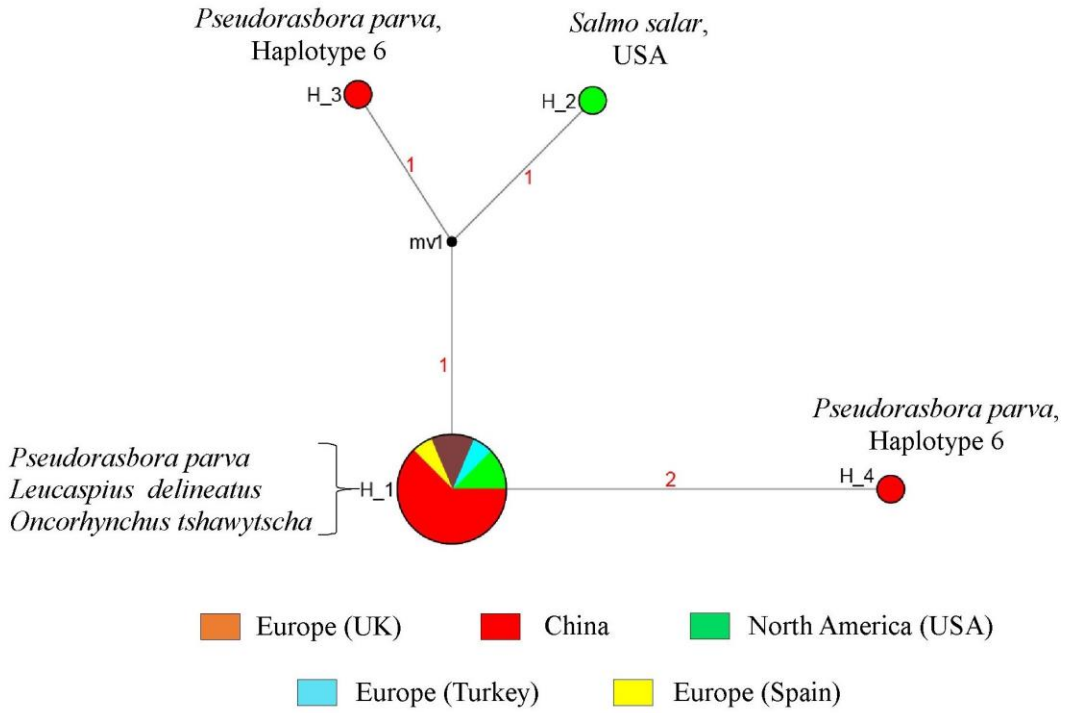


Figure 3.2. Minimum spanning network based on 18S rRNA segment (397 bp) of *Sphaerothecum destruens* strains from China and Spain together with already deposited *S. destruens* sequences. The colour code indicates *S. destruens* individuals from different locales; the size of the circle represents the number of individuals sharing a particular haplotype. Brown: UK (n=2), Green: USA (n=3), Red (n=11): China, Light blue: Turkey (n=1), Yellow: Spain (n=1). (Haplotype 1: UK-FN996945.1, USA-AY267344.1, AY267345.1; Haplotype 2: USA- AY267346.1 and Turkey)

Due to the low sensitivity of the ITS 1 assay (Section 2.3.2.1), only a single individual from the Chinese site 3 had DNA that had amplified with this marker. The 623 bp fragment of the ITS 1 region obtained from China was compared with the published sequences from Europe (UK and Turkey) and North America (USA) (Figure 3.3; (Gozlan et al. 2009, Ercan et al. 2015). Two main clades were present, with all the samples from the USA grouping together and samples from Europe grouping with the sample from China. Individuals originating from the UK and China clustered together, and the Turkish samples were more closely related to the UK and China samples compared to the USA samples (Figure 3.3).

The *S. destruens* Cyt-b was amplified from six populations (5 from China and 1 from UK). Interestingly, no variation was found between the Chinese and the UK *S. destruens* samples (Appendix 1).

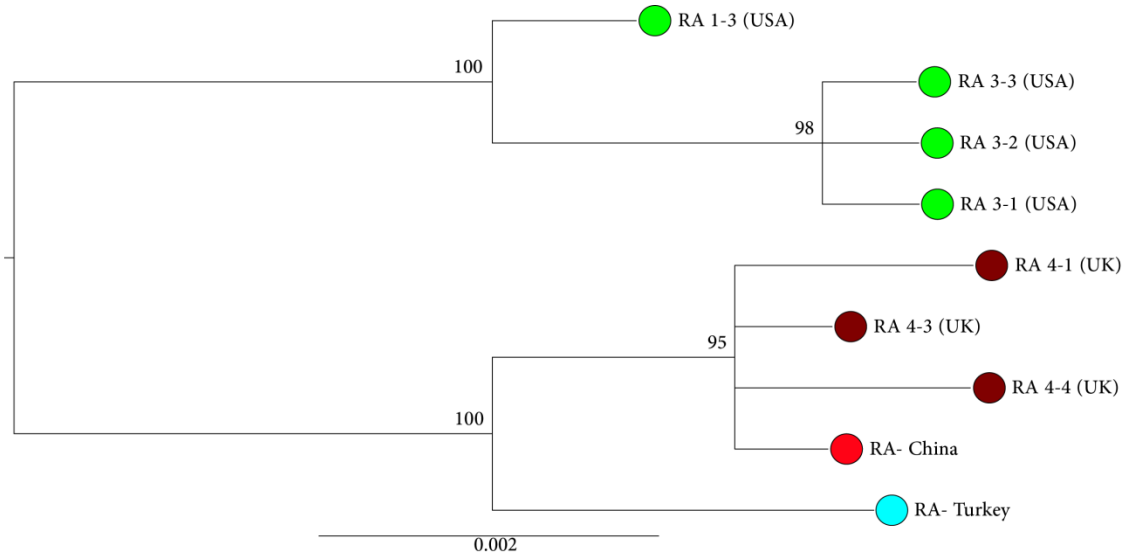


Figure 3.3. Phylogenetic tree resulting from maximum likelihood method based on Hasegawa-Kishino-Yano model (Hasegawa et al. 1985) analysis of the ribosomal ITS 1 sequences of *Sphaerothecum destruens*. Isolate origin and GenBank accession numbers are: RA 1-3 (FJ44070.1), RA 3-1 (FJ440708.1), RA 3-2 (FJ440709.1), RA 3-3 (FJ440710.1), RA 4-1 (FJ440702.1), RA 4-3 (FJ440703.1), RA 4-4 (FJ440704.1), RA-Turkey (KT361608.1) and RA-China (to be deposited).

3.3.4 Amplification of *Pseudorasbora parva* mitochondrial cytochrome b gene

A total of 91 haplotypes from 949 *P. parva* individuals were identified in the dataset of Cyt-b sequences of *P. parva* populations (62) across the world. The most abundant haplotypes across the world were Hap_1, Hap_2, Hap_3, Hap_4, Hap_5, Hap_6, Hap_12 and Hap_17 with occurrence in more than one country (Figure 3.4). Five Cyt-b haplotype in *P. parva* were found associated with the presence of *S. destruens* (Hap_1, Hap_ 4, Hap_6, Hap_7, Hap_12 and Hap_55; Figure 3.4). The highest number of *P. parva* individuals (n=7) found positive for *S. destruens* had Cyt-b haplotype Hap_6. The remaining haplotypes had one *P. parva* individual positive for *S. destruens*.

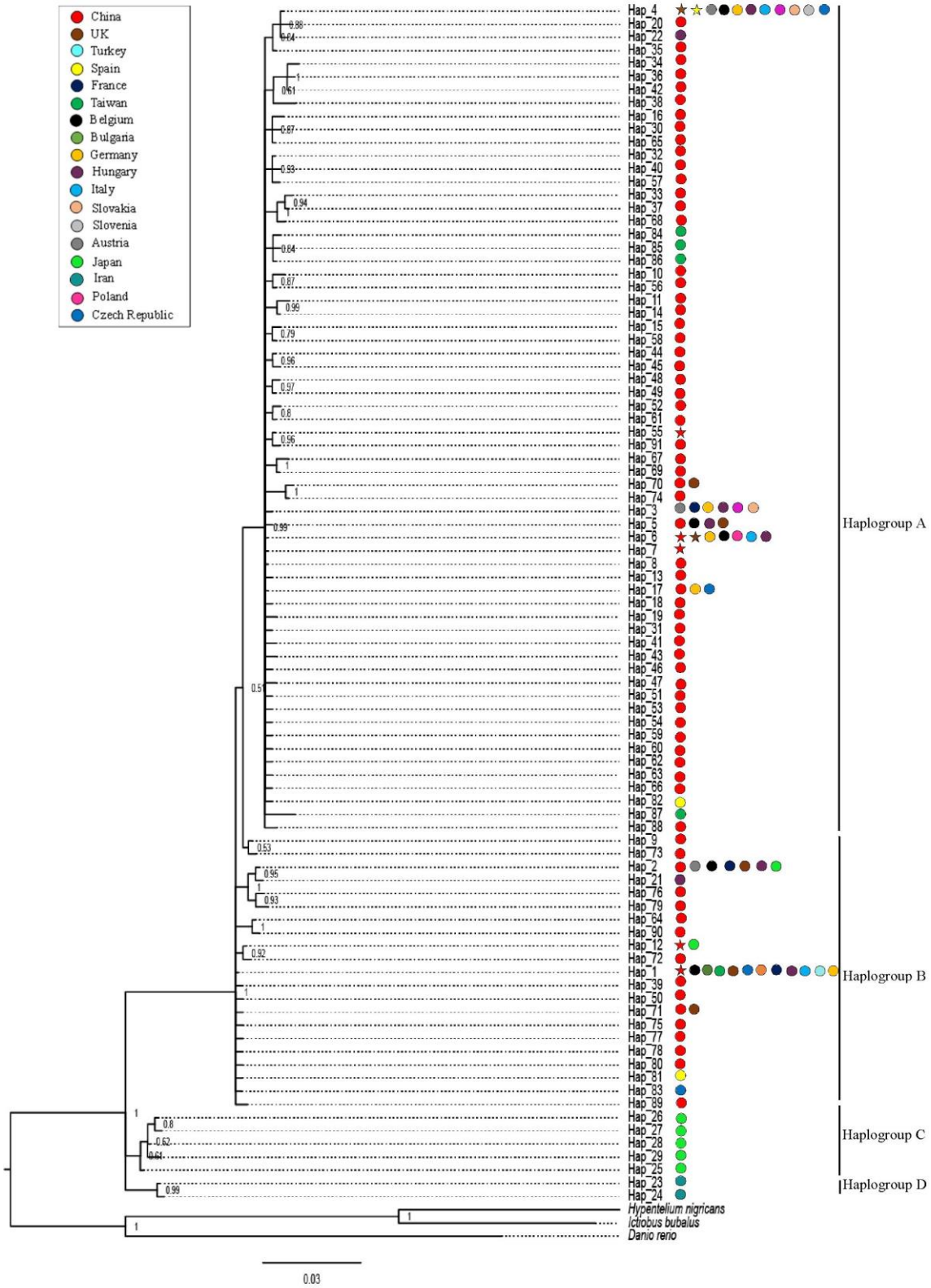


Figure 3.4. Molecular phylogenetic analysis of cytochrome b haplotypes of *Pseudorasbora parva* populations across the globe. The tree was inferred from maximum likelihood method based on the Hasegawa-Kishino-Yano model (HKY) (Hasegawa et al. 1985) with gamma distribution in Mr Bayes (Ronquist et al. 2012). The coloured circles indicate the countries that each haplotype has been found in and the coloured stars indicate *Sphaerothecum destruens* positive *P. parva* haplotypes in that country.

3.3.5 Demographic analysis of *Pseudorasbora parva* and *Sphaerothecum destruens*

The demographic analysis of the host, *P. parva* and parasite *S. destruens*, supported a potential recent population expansion. The sum of squared differences (SSD) and Harpending's raggedness index (H_{ri}) were not significant for both species, indicating the data is relatively good fit for population expansion (Figure 3.5). Both species also had significant negative values for the Tajima's *D*, test further supporting population expansion. *Pseudorasbora parva's* population expansion was further supported by the R_2 test. The R_2 test however was not significant for the *S. destruens* population which was in contrast with the Tajima's *D* negative value and the unimodal mismatch distribution.

The observed mismatch distribution for *P. parva* when all the populations are considered (Figure 3.5 B) is bimodal. The bimodality of the mismatch distribution for *P. parva* could be due to the presence of different haplogroups (Maltagliati et al. 2010). In China there are two established haplogroups A and B (Figure 3.4) despite the statistical tests indicating that the population as a whole has undergone a recent population expansion. When the data was split by haplogroups, the mismatch distribution was unimodal (Figure 3.5 C, D), further supporting that both haplogroups have undergone recent population expansion.

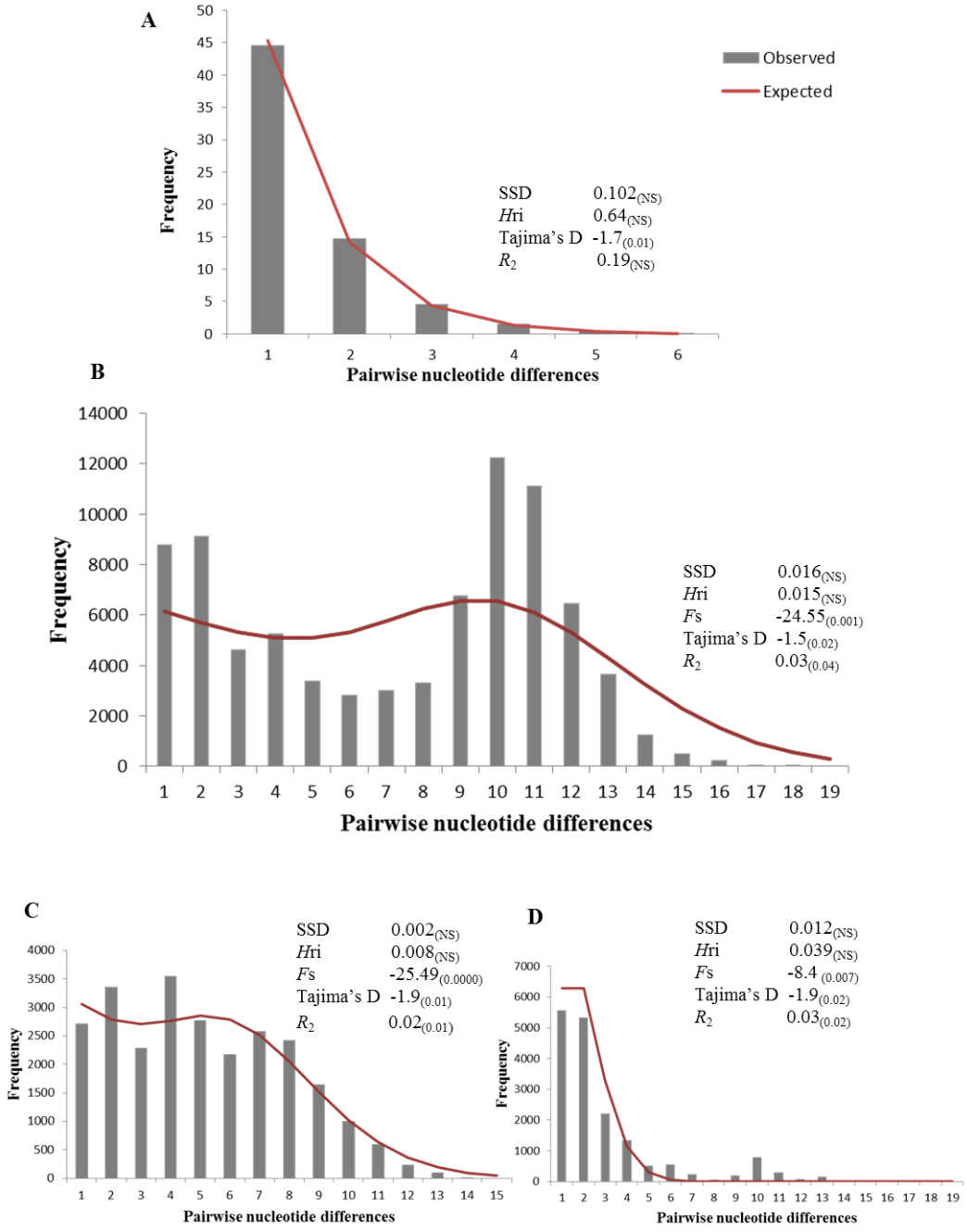


Figure 3.5. Frequency distribution of number of pairwise mismatch nucleotide differences between: (A) 18S rRNA sequences of *Sphaerothecum destruens* populations in China; (B) cytochrome b sequences of *Pseudorasbora parva* populations in China; (C) *P. parva* haplogroup A, and (D) *P. parva* haplogroup B. The solid line is the theoretical distribution under the hypothesis of population expansion. Sum of squared differences (SSD), Harpending's raggedness index (H*ri*), Fu's *F*s, Tajima's *D* and Ramos-Onsins & Rozas' R_2 statistics are listed next to each dataset. The *P*-values for each statistical test can be found in parenthesis. Significance was set at a *P*-value of 0.05 except for *F*s which was set at 0.01.

3.4 Discussion

In this study, the novel pathogen hypothesis was investigated for the emergent *S. destruens* parasite that has been closely associated with the invasive non-native *P. parva*. In the host's native range, China, *S. destruens* was detected in 9 out of the 10 tested *P. parva* populations with a maximum prevalence of 10 %. In *P. parva'*s non-native range of Eurasia, *S. destruens* was detected in 2 populations out of 11 populations (in the UK and Spain with a maximum prevalence of 5 %). Further to the presence of *S. destruens* being confirmed in China for the first time, the phylogenetic analysis and demographic analysis of both the host and the parasite support the hypothesis that *S. destruens* has been introduced to Europe through the accidental introduction of its reservoir host *P. parva*.

The phylogenetic analysis using the ITS 1 marker was performed on the only positive sample (due to the techniques low detection sensitivity). The *P. parva* population in which the parasite was found (site 3) is located north to the River Yangtze in China. Phylogenetic analysis has indicated that in the UK the most prevalent *P. parva* haplotypes are the same as those found north of the River Yangtze which can explain the similarity between the Chinese and UK *S. destruens* isolates. The Turkish *S. destruens* isolate, although more closely related to other European isolates, was still different and grouped on its own within the overall European and China clade (Figure 3.3). This could be explained by the phylogenetic history of the host species, as *P. parva* in Turkey was introduced from Bulgaria and has haplotypes which are often found in the south of the River Yangtze (Simon et al. 2011). In Europe, *P. parva* was intially introduced in the River Danube, Romania, and successively to Germany and in southern parts of UK, where *S. destruens* was orginally discovered in the UK (Simon et al. 2011, Gozlan et al. 2005).

The wide distribution of the *P. parva* populations infected by *S. destruens* across China suggests that the parasite could be coexisting with its *P. parva* reservoir host. The distance between the most southerly and northerly *S. destruens* positive *P. parva* populations is ~ 18, 000 kilometres which is 7 times greater than the distance between its European populations (Turkey to UK;~ 2,500 kilometres). The emergence and

association of fungal parasites with reservoir hosts over a long evolutionary history has been recently demonstrated for the chytrid fungus *Batrachochytrium salamandrivorans* (Martel et al. 2014). This pathogen has likely originated and coexisted with its reservoir hosts for millions of years in Asia before being introduced across the world with the trait of its reservoir hosts. Similarly, to *B. salamandrivorans*, the data presented here suggest that *S. destruens* has been introduced to Europe via the accidental introduction of its reservoir host.

The demographic analyses of *P. parva* and *S. destruens* partially suggest that both species have undergone a recent population expansion. A recent population expansion for *P. parva* is supported with all tests. However, for *S. destruens*, this is supported using SSD, H*ri* values and Tajima's *D* test, but not by the R_2 test. The partial congruence between the demographic history of the two species is suprising especially in light of the true generalist nature of *S. detruens* and its ability to use a number of different hosts. Recent work, however, has indicated that following the establishment of a generalist parasite in a community, its population dynamics are driven via intra-host transmission rather than inter-host transmission (Fenton et al. 2015). This could explain the observed similarity in the population demographic history of the two species.

The mitochondrial studies on *P. parva* have identified that Haplogroup A is prevalent north of the River Yangtze, an area with a temperate climate. Haplogroup B is mostly found to the south with tropical climatic conditions. Most *S. detruens* positive individuals had haplotypes belonging to haplogroup A (Figure 3.5). Studies on *S. destruens* have identified that low temperatures of 4 ºC and 15 ºC are correlated with higher spore and zoospore survival compared to higher temperatures of 25 ºC and 30 ºC (Andreou et al. 2009), suggesting that the parasite could be more adapted to temperate climates.

Currently, there are 14 known susceptible species to *S. destruens*, including valuable aquaculture species (salmon, carp and bass; (Andreou et al. 2012, Paley et al. 2012, Ercan et al. 2015) and endemic fishes to Europe that are of high conservation value (Gozlan et al. 2005, Ercan et al. 2015). The close association of the reservoir host, *P.*

parva with aquaculture facilities (due its accidental introduction along with carp from China; (Gozlan et al. 2010a) and the predicted ability of the parasite to establish in online adjacent freshwater communities within a year of its introduction in an aquaculture facility (Al-Shorbaji et al. 2016) increases the risk of disease to native fishes.

This study represents a first screening of native and invasive *P. parva* populations for the presence of *S. destruens*. It is important to note that the prevalence values recorded in positive populations are very likely to be underestimates of the true prevalence of this parasite, as only the kidney was sampled. *S. destruens* infects multiple organs and it does so unequally. This lack of infection localisation can lead to an underestimate of its prevalence. Thus, populations detected as negative for *S. destruens* in this study must be treated with caution as it cannot be excluded that the parasite might be present in these and other populations in the country. For example, the Turkish population sampled in this study was found to be negative for *S. destruens* whereas the parasite has been detected in another *P. parva* population in Turkey (Ercan et al. 2015). Thus, in countries where *P. parva* is present, an extensive sampling of its established populations would be necessary to determine whether *S. destruens* has been co-introduced.

In the last 20 years, aquaculture production has increased exponentially to support economic growth with its expansion being highly reliant on non-native species (Peeler et al. 2011). The introduction of non-native species can be detrimental both to ecosystem services and native communities through direct competition and disease introduction (Pelicice et al. 2014). Here, we have documented the introduction of an emergent generalist pathogen through the accidental introduction of its highly invasive host. The potential threats associated with aquaculture production and the resultant fish movements highlight the importance of risk assessments to identify emergent parasites. Horizon scanning for potential emergent diseases will be critical in informing strict biosecurity controls in order to prevent disease introduction.

3.5 Summary

The present study provided the first evidence of the possible introduction of *S. destruens* to Europe through accidental introduction of its host *P. parva* from Asia. The hypothesis was supported from the two nuclear (18S rRNA and ribosomal ITS 1) and one mitochondrial (Cyt-b) marker that the two isolates i.e. from UK and China are not geographically distinct. It also confirmed the presence of *S. destruens* in China and expanded the confirmed range of *S. destruens* to more locations in Europe (Spain). The demographic analysis showed that *S. destruens* and its host *P. parva* populations have potentially undergone similar demographic expansion in its native range (China) providing support for the hypothesis that these two species have been historically closely associated. The evidence of non-native status of *S. destruens* in Europe will have management implications for *S. destruens* in the UK.

Chapter 4

Epidemiological studies of *Sphaerothecum destruens* in the UK

4.1 Introduction

The UK is the first European country where *S. destruens* has been identified in *P. parva* populations (Gozlan et al. 2005). *Pseudorasbora parva* was first recorded in the UK in an aquaculture facility in southern England in 1996 (Domaniewski and Wheeler 1996) and has rapidly spread and colonized up to 23 UK water bodies (Britton et al. 2008a). Most *P. parva* populations in the UK have been associated with aquaculture or recreational fisheries, with no recorded established populations in wild habitats such as streams, rivers or lakes. In response to the potential threats posed by *P. parva* (Britton et al. 2007) a national programme of eradication for *P. parva* has been designed and administered (Britton et al. 2010). The programme aimed at complete eradication of *P. parva* from high risk sites (with high risk sites identified based on the conservation and fishery value of the adjacent water body) or containment in case of medium risk sites (Britton et al. 2008a). By 2014, 15 out of 23 confirmed *P. parva* sites had been eradicated, with a further 6 sites to be eradicated in England by 2017 (Britton et al. 2010, GBNNSS 2015).

Despite *P. parva* having no wild populations in the UK, a number of sites invaded by *P. parva* have water effluents which flow into wild freshwater habitats. This can have important implications for the parasite's transmission, as epidemiological modelling has indicated that *S. destruens* can spread and establish to connected downstream communities through environmental transmission of their infective spores and zoospores within one year post introduction of infected *P. parva* (Al-Shorbaji et al. 2016). Furthermore, the same work has indicated that *S. destruens* can establish in new hosts and maintain its transmission in the absence of the initial reservoir host - in this case *P. parva* (Al-Shorbaji et al. 2016). As *S. destruens* is a true generalist, it is thus highly probable that adjacent downstream communities to established *P. parva* populations positive for *S. destruens* have established infections (Andreou and Gozlan 2016).

In Chapter 3, China was identified as the origin of *S. destruens* in Europe and the UK. The work indicated that due to the low detection in reservoir hosts because of low *S.*

destruens infection, the prevalence of *S. destruens* in its European range was most probably underestimated. In addition, the lack of histopathological evidence of infection, due to the samples being preserved in ethanol, has prevented an evaluation of the potential impact of the parasite for the fish host.

Thus, the objectives of this chapter were to: (i) Determine the distribution and presence *S. destruens* pathology in the UK *P. parva* populations across a wide geographical range. *Sphaerothecum destruens* epidemiology has never been determined on this scale before. The *P. parva* populations were screened for *S. destruens* through PCR methods (Chapter 2) followed by histopathology; (ii) Assess the risk of disease transfer to adjacent native fish communities by sampling fish species that have been in close proximity to reservoir host *P. parva* populations; and (iii) Infer the demographic history of *S. destruens* and its host *P. parva* through the mismatch distribution analysis and check for congruence in order to test the hypothesis that *S. destruens* has spread through the movement of *P. parva* in the UK.

4.2 Material and Methods

4.2.1 Populations screened for *Sphaerothecum destruens*

Seven extant *P. parva* populations were sampled from England and Wales prior to their eradication in 2013-2015 by the Environment Agency (Figure 4.1) and were euthanized following the Home Office (HO) guidelines. The fish were dissected and liver and kidney were sampled, with half the organ fixed in 100 % ethanol for molecular detection and the remaining half in 10 % neutral buffered formalin for histopathology. The sampled water bodies included six enclosed still water fisheries and two fisheries with outlets to streams (Figure 4.1; Table 4.1). Samples from freshwater fishes (*Salmo trutta*, *Rutilus rutilus*, *Squalius cephalus* and *Leuciscus leuciscus*) from the Tadburn Lake stream adjacent to Site 1a were also collected by electrofishing and euthanized following HO guidelines (Table 4.1).

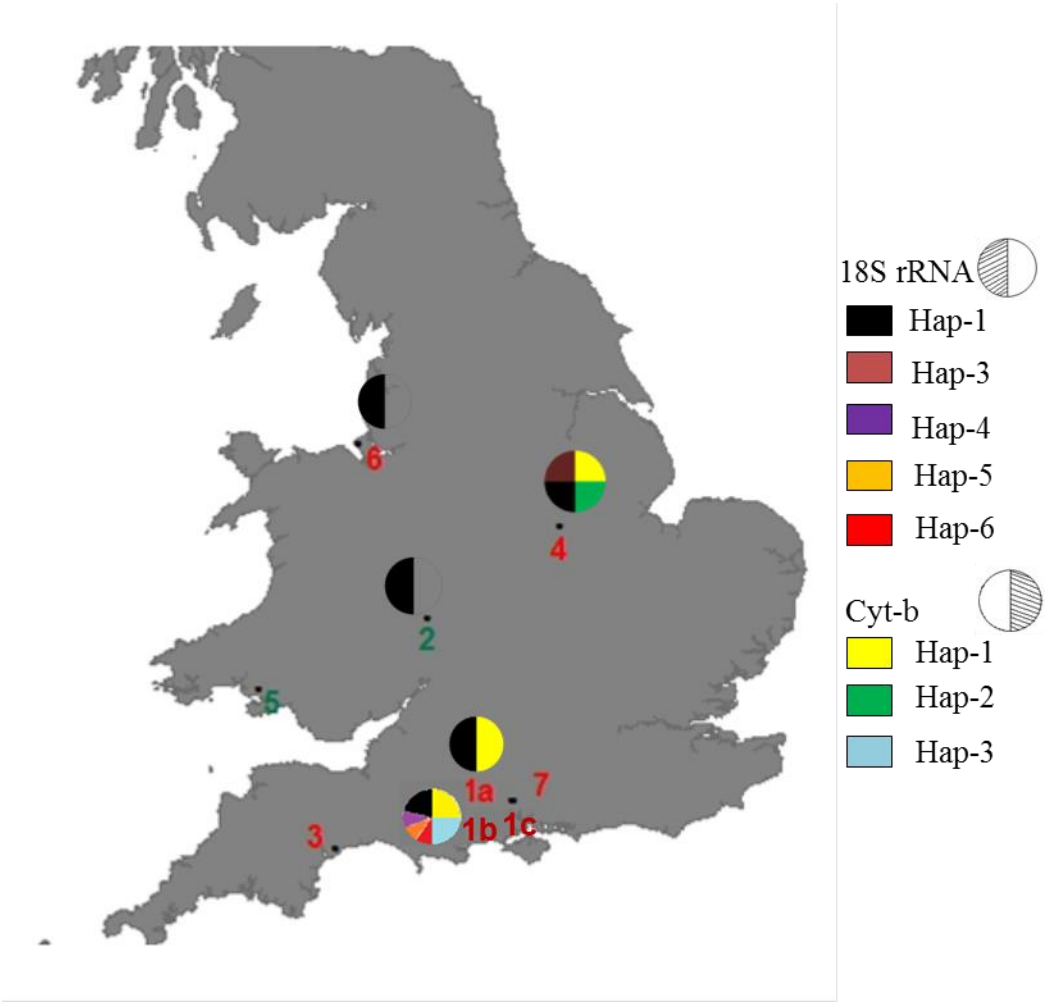


Figure 4.1. Distribution of sampled *Pseudorasbora parva* populations across the UK. Population 1a is the hypothesised first *P. parva* population in UK to have been introduced in mid-80s (Domaniewski and Wheeler 1996). Details of each sampled population can be found in Table 4.1. The red and green numbering for each population represents the two genetically different metapopulations of the host *P. parva* in the UK (Blake et al. unpublished). The 18S rRNA haplotypes for *S. destruens* are represented on the left of the circle, the Cyt-b haplotypes are on the right of the circle. The different colours represent different haplotypes.

Table 4.1. Geographical location and *Sphaerothecum destruens* prevalence for all fish populations screened using two DNA markers – 18S rRNA and Cyt-b.

Population	Water type	Sampling year	Geographical Coordinates	Sampled Fish species	Sample size	Positive Fish species for *S. destruens*	Prevalence of *S. destruens* (%)	Genetic marker of *S. destruens*	
								18S rRNA	Cyt-b
1a	Disused aquaculture facility, online to river	2013	NGR: SU3822 S. England	*P. parva*	30	*P. parva*	3.33 (1/30)	✓ (Hap_1)	✓ (Hap_1)
1b	River adjacent to site 1-part B (slower flowing section of river adjacent to site 1a)			*S. trutta* *S. cephalus* *R. rutilus* *L. leuciscus*	3 4 2 5	*S. trutta*	33.3 (1/3)	✓ (Hap_1)	-
						S. cephalus	50 (2/4)	✓ (Hap_4)	✓ (Hap_1)
						S. cephalus		✓ (Hap_1)	✓ (Hap_1)
						R. rutilus	100 (2/2)	✓ (Hap_1)	✓ (Hap_1)
						R. rutilus		✓ (Hap_1)	✓ (Hap_3)
						L. leuciscus	60 (3/5)	✓ (Hap_1)	-
						L. leuciscus		✓ (Hap_6)	-
						L. leuciscus		✓ (Hap_5)	-

1c	River adjacent to site 1-part A (high flowing stream directly linked to site 1a			*S. trutta*	30	-	0	
2	Enclosed still water fishery	2013	NGR: SO7657 Midlands	*P. parva*	30	-	0	
				R. rutilus	5	*R. rutilus*	20 (1/5)	✓ (Hap_1)
3	Ornamental pond with outlet to stream	2013	NGR: SY0786 SE England	*R. rutilus*	10	-	0	
4	Enclosed still water fishery	2014	NGR: SK7425 Midlands	*P. parva*	30	*P. parva*	6.66 (2/30)	✓ (Hap_3) ✓ (Hap_1)
								✓ (Hap_1) ✓ (Hap_2)
5	Reservoir	2014	NGR: SN5104 S. Wales	*P. parva*	30	-	0	
6	Enclosed still water fishery-1	2014	NGR: SJ2487	*P. parva*	30	*P. parva*	6.66 (2/30)	✓ (Hap_1) ✓ (Hap_1)
								- -
	Enclosed still water fishery-2	2015		*P. parva*	30	-	0	
7	Enclosed still water fishery	2014	NGR:SU3922	*P. parva*	30	-	0	

4.2.2 *Sphaerothecum destruens* detection through PCR and histopathology

Kidney and liver DNA was extracted from freshwater fishes to investigate the presence of *S. destruens* in those tissues, as the parasite is usually in high prevalence in these organs (Andreou et al. 2011). The liver and kidney samples of each fish were pooled together prior to DNA extraction using the rodent tail protocol of the Qiagen DNeasy Blood and tissue kit (Qiagen, Germany). All steps were performed according to manufacturer's guidelines. Extracted DNA from fish tissue was quantified by Nanodrop 2000 (Thermoscientific) and 300 ng DNA were used in subsequent PCR analysis. The extracted DNA was stored at -20 °C until further testing.

The *S. destruens*-specific nested PCR was performed for the amplification of a segment of 18S rRNA using specific primers published in Mendonca and Arkush (2004), as described in Section 2.2.1. In order to carry out phylogenetic analysis, all the positive samples for 18S rRNA, were also amplified for the ITS 1 and Cyt-b regions with PCR and cycling conditions as described in Sections 2.2.2 and 2.2.3. There was a lag time of 6 months between the application of 18S rRNA and Cyt-b marker on the DNA samples.

All *S. destruens* positive fish samples were investigated using histopathological analysis. Tissue (kidney and liver) were fixed in 10 % neutral buffered formalin (NBF) for 24 hours before transferring to 70 % industrial methylated spirit (IMS). Samples were infiltrated with paraffin under vacuum using standard protocols (Bancroft et al. 1996). Sections were cut to a thickness of 3-5 µm, mounted onto glass slides, and stained with haematoxylin and eosin (H&E) or Gram's stain. Five transverse stained sections were analysed by light microscopy (Nikon Eclipse E800); digital images and measurements were obtained using the Lucia™ Screen Measurement System (Nikon, UK).

4.2.3 Amplification of *Pseudorasbora parva* mitochondrial cytochrome b gene

PCR amplification of the Cyt-b gene of *P. parva* individuals tested positive for *S. destruens* was carried out as explained in Section 3.2.6.

4.2.4 Phylogenetic analysis

The sequences generated for 18S rRNA and Cyt-b gene for *S. destruens* were aligned using Clustal W in BioEdit ver. 5.0.9 (Hall 1999) and visually checked to eliminate ambiguities and identify polymorphic sites. A phylogenetic network was generated using DnaSP version 5.10 (Librado and Rozas 2009) and Network Publisher (Bandelt et al. 1999); available at http://www.fluxus-engineeering.com). The sequences used are the ones generated in the present study and from Chapter 3, as well as all the published sequences available for the 18S rRNA marker (AY267344.1, AY267345.1, AY267346.1, FN996945.1 (Arkush et al. 2003, Paley et al. 2012).

The Cyt-b sequences obtained for *P. parva* were aligned with the sequences obtained for *P. parva* populations in the UK (Blake et al. unpublished), using Clustal W in BioEdit ver. 5.0.9 (Hall 1999). A phylogenetic tree was constructed to identify the haplotypes of *P. parva* individuals positive for *S. destruens*. The phylogenetic analysis using maximum likelihood was performed in MEGA 7 (Kumar et al. 2016). The tree was calculated using Hasegawa-Kishino-Yano model (Hasegawa et al. 1985) with 10,000 bootstraps. The Cyt-b gene sequences from *Ictiobus bubalus* (JF799443.1), *Hypentelium nigricans* (JF799441.1) and *Danio rerio* (JN234356.1) were used as outgroups.

To analyse the demographic history of *S. destruens* and *P. parva* populations in the UK, the mismatch distribution analysis was carried out for the 18S rRNA and Cyt-b genes as described in Section 3.2.6. The expected values for a model of population expansion were calculated and plotted against the observed values. Demographic changes were analysed by calculating Harpending's raggedness index ($H_{ri)}$ (Harpending et al. 1993) and the sum of squared deviations (SSD) (Schneider and Excoffier 1999), carried out in Arlequin version 3.5 (Excoffier and Lischer 2010). To test for the hypothesis of population expansion in *S. destruens* and its host *P. parva*, only Ramos-Onsins & Rozas' R_2 test (Ramos-Onsins and Rozas 2002), which is a powerful test for small sample sizes, was performed for *S. destruens*. For *P. parva*, three tests: Fu's (Fu 1997) *F*s, and Tajima's *D* (Tajima 1989) test by Arlequin and Ramos-Onsins & Rozas' R_2 test (Ramos-Onsins and Rozas 2002) by DnaSP (Librado and Rozas 2009) were conducted.

For the Tajima's *D* and *F*s tests, *P*-values were calculated based on the coalescent simulation algorithm. The R_2 test *P*-values were based on parametric bootstrapping (10, 000) with coalescence simulations.

4.3 Results

4.3.1 *Sphaerothecum destruens* prevalence in UK water bodies

Sphaerothecum destruens was detected using molecular tools in 14 individuals out of the 269 sampled across 7 locations (Table 4.1, Fig 4.1). Fifty percent of the *P. parva* sites tested positive for *S. destruens* with the prevalence ranging from 3 to 6 %. Two of the 3 populations were from enclosed still water fisheries (populations 4 and 6, with 6.7 % prevalence (Figure 4.1). The third *S. destruens* positive *P. parva* population was 1a which represented the first accidental *P. parva* introduction in 1996 and tested positive for *S. destruens* with a prevalence of 3 % (Figure 4.1). This population was from a disused aquaculture facility whose effluents are discharged in the adjacent Tadburn lake stream (populations 1b and c) which connects with the River Test in Hampshire. A number of native species were sampled from this stream (Table 4.1) and tested positive for *S. destruens*. These included chub *Squalius cephalus*, dace *Leuciscus leuciscus*, brown trout *Salmo trutta* and roach *Rutilus rutilus* (population 1b) with an overall prevalence across all species of 57 %. In population 2, both *P. parva* and *R. rutilus* were sampled with *S. destruens* being present in *R. rutilus* with a prevalence of 20 %.

The histopathology of the fish samples tested positive for *S. destruens* through PCR did not show any signs of disease.

4.3.2 Phylogenetic analysis of *Sphaerothecum destruens* using the 18S rRNA and cytochrome b sequences

Five *S. destruens* haplotypes were detected for the 18S rRNA gene in 14 individuals from 5 UK water bodies. Nine variable sites were recorded on the 18S rRNA Haplotype 1 (Figure 4.2 A) that was also the most abundant with 10 individuals from this study and 3 individuals from previous studies having this haplotype (Arkush et al. 2003, Paley et al. 2012). All of the isolates obtained from the UK water bodies grouped with the

already deposited 18S rRNA sequences except for four isolates. Haplotype 3 was found in *P. parva* from site 4, Haplotype 4 was found in *S. cephalus* and Haplotype 5, 6 were found in *L. leuciscus* all from site 1b (Figure 4.1).

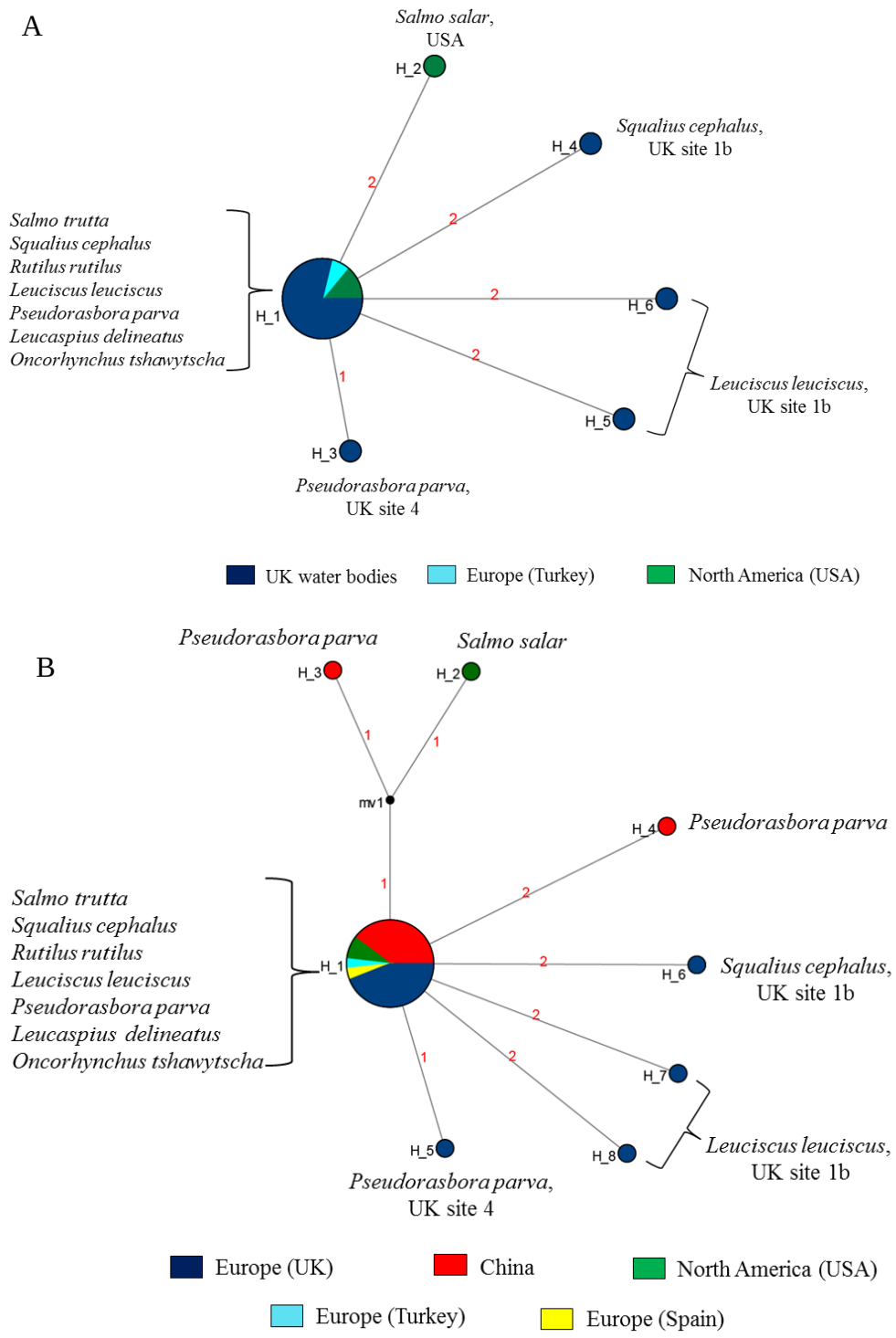


Figure 4.2. Minimum spanning network based on 18S rRNA sequences (397 bp) of *Sphaerothecum destruens* isolated from (A) UK water bodies (n=14); (B) across world; n= 14 (Chapter 3) and UK water bodies n= 14 (this Chapter), together with *S. destruens* sequences (Haplotype 1: UK-FN996945.1, USA-AY267344.1, AY267345.1; Haplotype 2: USA- AY267346.1 and Turkey).

In the sequence alignment of *S. destruens* isolates from this study together with the sequences obtained across *P. parva's* native and invasive range (Chapter 3), 12 variable sites were recorded with one parsimony site. Apart from Haplotype 1, none of the 18S rRNA haplotypes in the samples from UK were shared with samples from China (Figure 4.2B).

None of the 18S rRNA *S. destruens* positive individuals amplified with *S. destruens*-specific ITS 1 PCR. Seven out of 14 18S rRNA *S. destruens* positive individuals were amplified for the Cyt-b. The Cyt-b PCR was performed 6 months after the 18S rRNA PCR with the extracted DNA having been stored at -20 °C. This could explain the discrepancy in amplification success using the Cyt-b gene despite both the 18S rRNA and Cyt-b markers having similar detection efficiencies (Chapter 2). Three haplotypes were detected for Cyt-b (700 bp) in six individuals with a total of 22 variable sites (Figure 4.3). The *S. destruens* Cyt-b sequences obtained from Chinese *P. parva* populations showed no DNA polymorphism. These sequences were identical to the ones found in UK water bodies (site 1a, 1c and site 4) except for two isolates from *P. parva* at site 4 and from *R. rutilus* site 1b.

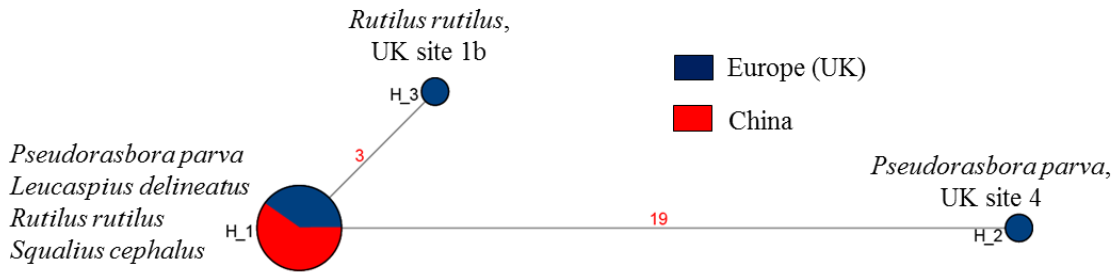


Figure 4.3. Minimum spanning network based on mitochondrial cytochrome b sequences of *Sphaerothecum destruens* obtained from *Pseudorasbora parva* (China; n=5 and UK water bodies; n=8).

4.3.3 Amplification of *Pseudorasbora parva* mitochondrial cytochrome b gene

The genotyping of *P. parva* host populations in the UK identified a total of 4 haplotypes (Haplotype 1, 3, 4 and 6; Figure 4.5). The *P. parva* populations from the UK grouped with populations across its native and invasive range (Chapter 3; Hardouin et al. submitted) and clustered with *P. parva* Haplotypes 1, 3, 4 and 6 (Figure 4.4), with

Haplotype 4 and 6 being positive for *S. destruens* presence. The highest number of *P. parva* individuals positive for *S. destruens* had *P. parva* Cyt-b Haplotype 4. Interestingly, this haplotype is present across samples from Europe and Japan but absent in samples from China. Moreover, this haplotype is positive for *S. destruens* in samples across two European countries UK and Spain. This indicates that the *P. parva* sampling in China failed to capture *P. parva* population(s) with Cyt-b Haplotype 4 which is the most common in the species invasive range. Another interesting observation is the unique Cyt-b Haplotype 2 for *S. destruens* from the UK site 4 (Figure 4.3) which was also obtained from a *P. parva* population with Cyt-b Haplotype 4.

4.3.4 Demographic analysis of *Pseudorasbora parva* and *Sphaerothecum destruens*

The mismatch distributions for the *S. destruens* 18S rRNA and Cyt-b haplotypes were multimodal (Figure 4.6), which can indicate that the population is either at demographic equilibrium or experiencing decreasing population sizes (Rogers and Harpending 1992). The observed distribution deviated significantly from the expected distribution curve for both 18S rRNA and Cyt-b gene. For the Cyt-b sequences two distinct peaks were observed (Figure 4.4). The Tajima's *D* non-significant positive value for 18S rRNA and Ramos-Onsins & Rozas' R_2 non-significant value for Cyt-b indicate that there was no demographic expansion in the UK *S. destruens* population.

The mismatch distribution plot for the *P. parva* UK populations also had a multimodal distribution (Figure 4.4 C) which can be interpreted as a signature of populations at demographic equilibrium or decline. The SSD and raggedness index values were statistically significant, indicating the data does not fit the model of population expansion. The Tajima's *D* positive value suggests that either balancing selection or decrease in population size could have occurred (Pichler 2002) but was statistically not-significant. The *F*s and R_2 values were also statistically not-significant rejecting the demographic expansion hypothesis and supporting that the population is at demographic equilibrium.

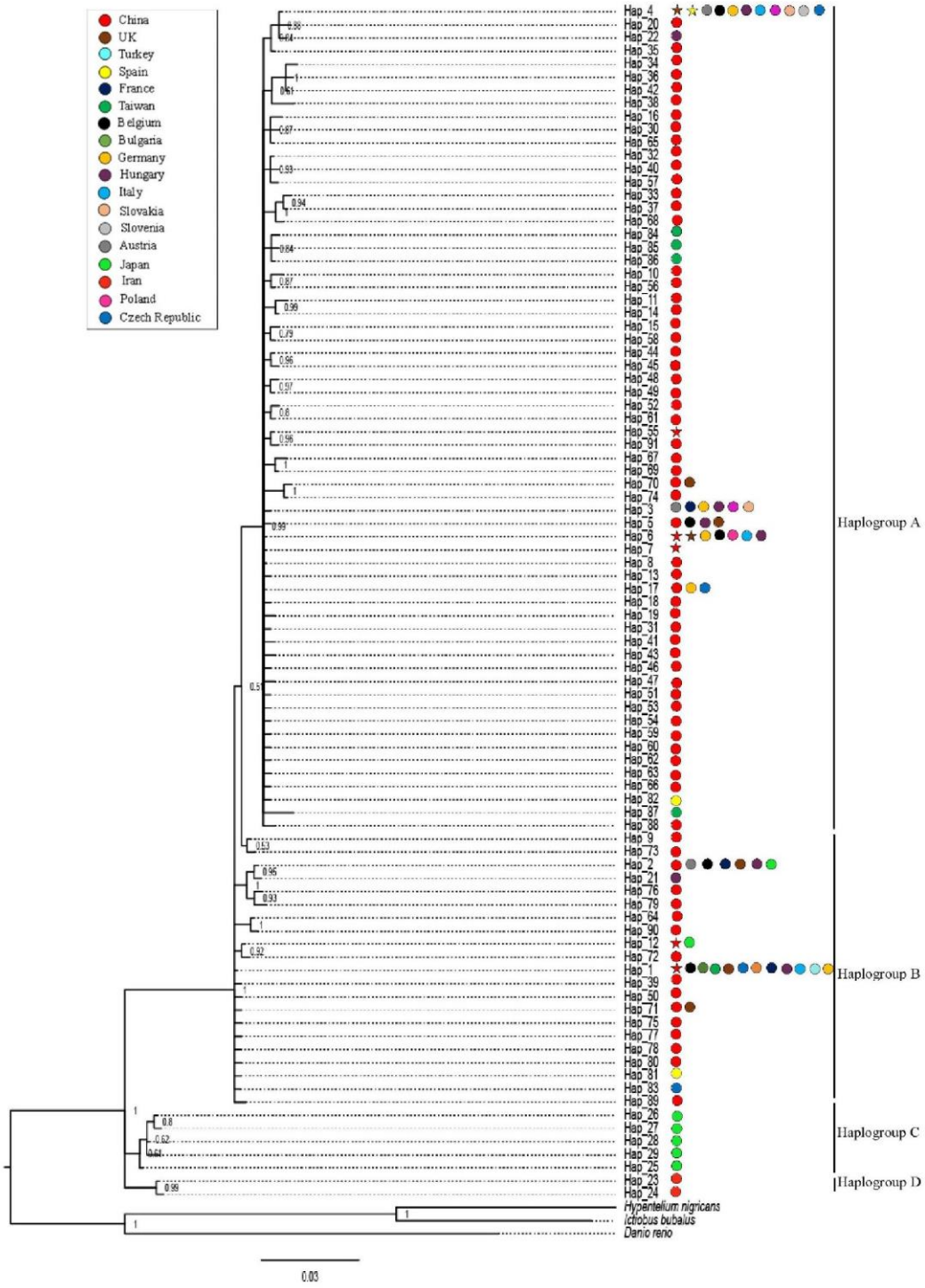


Figure 4.4. Molecular phylogenetic analysis Cyt-b haplotypes of *Pseudorasbora parva* populations across the species' native and non-native range. The tree was build using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model (Hasegawa et al. 1985) with Gamma distribution in Mr Bayes (Ronquist et al. 2012). The coloured circles indicate the countries that each haplotype has been found in and the coloured stars indicate *Sphaerothecum destruens* positive haplotypes in that country.

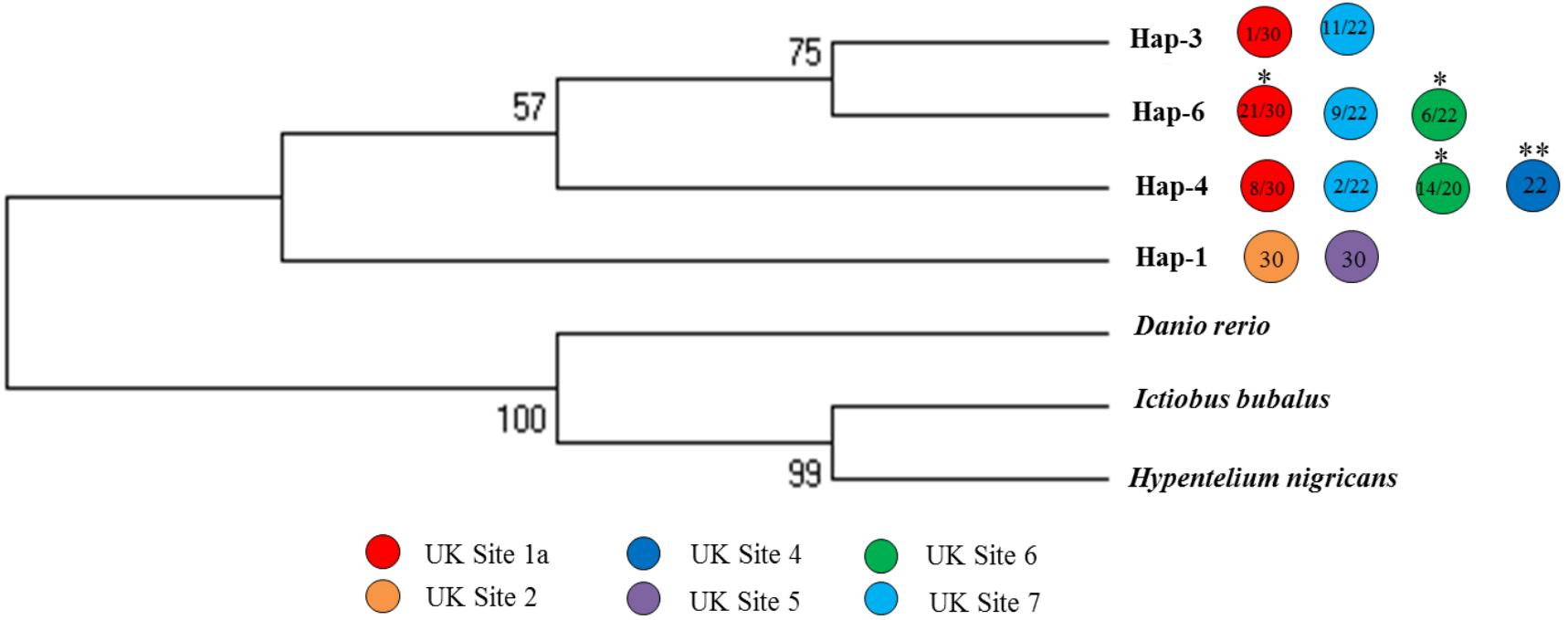


Figure 4.5. Cyt-b Haplotypes of *Pseudorasbora parva* across sampled UK sites (Blake et al. unpublished). The tree was built using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model (Hasegawa et al. 1985) with Gamma distribution analysis of mitochondria mitochondrial cytochrome b gene of *Pseudorasbora parva* in MEGA 7 (Kumar et al. 2016). *Danio rerio* (JN234356.1), *Ictiobus bubalus* (JF799443.1) and *Hypentelium nigricans* (JF799441.1) were used as outgroups. Within the circles, the proportion of each haplotype is indicated for each population. The number of asterisks on each site specifies the number of *P. parva* individuals positive for *Sphaerothecum destruens*. All *S. destruens* positive *P. parva* had Cyt-b haplotypes 1 and 2.

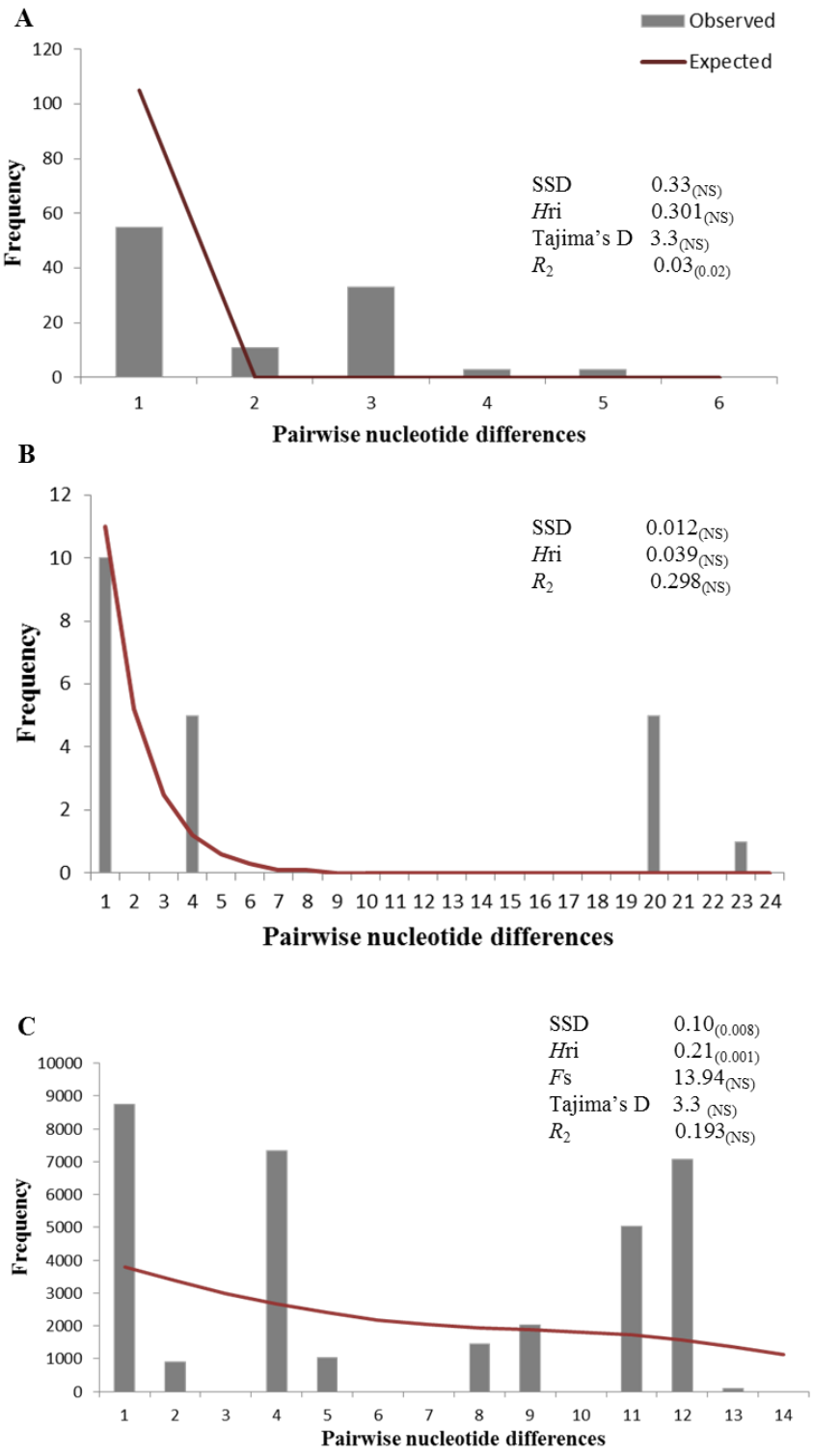


Figure 4.6. Frequency distributions of the number of pairwise nucleotide differences (mismatch) between (A) 18S rRNA and (B) Cyt-b sequences of *Sphaerothecum destruens* in the UK, and (C) Cyt-b sequences of *Pseudorasbora parva* populations in the UK. The solid line is the theoretical distribution under the hypothesis of population expansion. The values for Sum of squared differences (SSD), Harpending's raggedness index (*H*ri), Fu's *F*s, Tajima's *D* and Ramos-Onsins & Rozas' R_2 statistics are listed next to each dataset. The *P*-values for each statistical test are presented in the parenthesis and significance was set at $P < 0.05$ except for *F*s $P < 0.01$.

4.4 Discussion

In this study, the first epidemiological survey for *S. destruens* in the UK was carried out through the screening of its reservoir host populations of *P. parva*. *Sphaerothecum destruens* was detected in 50 % of the 7 *P. parva* populations sampled, with a wide geographical distribution with similar prevalence per population as in China (5-10 %). This finding provides additional support to the hypothesis that *P. parva* can act as a reservoir host of *S. destruens*. This is the first confirmed record of *S. destruens* in *P. parva* from the UK. Concomitant work suggests that there were at least two independent *P. parva* introductions to the UK (Blake et al. unpublished), with the tested populations in this study spanning populations for both fish sources (Group 1: sites 1, 3, 4, 6 and 7; Group 2: sites 2, 5). Populations from both introduction groups have been found positive for *S. destruens*.

In addition to determining the epidemiology of *S. destruens*, this study also aimed at determining whether disease transfer has occurred to native fishes in adjacent communities. This was achieved by monitoring the adjacent river communities to the first recorded introduction of *P. parva* to the UK (Site 1 Figure 4.1). The work has indicated that *S. destruens* was present in a number of native fish species adjacent to site 1 and these included salmonids (*S. trutta*) and cyprinids (*R. rutilus*, *S. cephalus* and *L. leuciscus*). This finding supports the concerns raised by Al-Shorbaji et al. (2016) and Andreou and Gozlan (2016) regarding the potential of parasite transfer to adjacent communities. The 18S rRNA and Cyt-b analysis of the *S. destruens* positive fishes at site 1 suggest the potential of a radiation of the parasite to new hosts. *S. cephalus* and *L. leuciscus* have 3 new haplotypes for the 18S rRNA gene (Haplotype 6, 7 and 8) and new Cyt-b haplotype (Haplotype 3) for *R. rutilus*. The Fenton et al. (2015) study on generalist parasites has indicated that once transmission to a new host occurs, parasite dynamics would be driven through intra-species transmission rather than inter-species transmission. This was independently shown for *S. destruens* through disease modelling which has indicated that *S. destruens* dynamics were driven through intra-species transmission following an initial inter-species transmission (Al-Shorbaji et al. 2016). Thus, the new *S. destruens* haplotypes associated with these new host species could be the result of co-evolution with new hosts. However, equally they could represent the original *S. destruens* diversity in the *P. parva* population in site 1. *P. parva* was

eradicated from Site 1 in July 2015 (personal communication Dr Rob Britton), and future work should include revisiting these native populations and testing for the presence of *S. destruens* in order to confirm that they can indeed sustain *S. destruens* in the absence of *P. parva*.

The phylogenetic analysis between *S. destruens* obtained across its native and invasive range and UK water bodies identified 8 unique haplotypes of 18S rRNA gene in 27 individuals. Haplotype 1 was the most abundant with 11 individual from China, 1 individual from Spain and Turkey, 10 individuals from UK waterbodies (this study) and 3 individuals from previous studies (Arkush et al. 2003, Paley et al. 2012).

This work was also able to test the mt Cyt-b marker developed for *S. destruens* (Chapter 2) for its phylogenetic resolution. The comparison of the Cyt-b sequences of the two unique haplotypes of *S. destruens* isolates (Haplotype 1 and 2) showed twenty-two nucleotide differences in 689 bp region. This suggests that the marker can be phylogenetic informative at the population level. However, this can only be confirmed by testing this marker with *S. destruens* strains from North America and Turkey.

The 18S rRNA and Cyt-b markers were more diverse in the UK compared to China (Chapter 3). This is despite the geographical distance between the most northerly and southerly positive *P. parva* populations in China being approximately 18,000 km compared to the UK sites 1b and 4 (~257 km apart) which showed the highest variation in 18S rRNA and Cyt-b sequences. Most of the variation could be driven by different hosts of *S. destruens* at UK site 3 and 8 but could also be the result of a high mixing of *P. parva* that has occurred in its continental European range prior to its introduction to the UK from Germany (Gozlan et al. 2010a). In addition, since the *P. parva* sampling was not exhaustive, the UK *S. destruens* diversity would suggest that not all the potential sources of *P. parva* to Europe were captured (as also supported by the P. parva Cyt-b haplotype data). However, the lack of amplification of the ITS 1 gene of *S. destruens* isolates obtained in the present study limits these conclusions as this is the

only gene that has been used previously to indicate geographical isolation of the *S. destruens* isolates (Gozlan et al. 2009).

In the UK, no population expansion was observed for *S. destruens* and its host *P. parva* whereas in its native range, China, the *P. parva* showed evidence of population expansion (Chapter 3). Mismatch analysis showed a multimodal distribution for both 18S rRNA and Cyt-b genes for *S. destruens* and for its healthy host *P. parva* populations in the UK. Multimodal/bimodal distributions could be due to: a) recent population declines, which suggest that *S. destruens* initially had a bigger population that suddenly contracted or b) balancing selection equally acting on two distinct *S. destruens* isolates resulting in bimodal distribution (Figure 4.4 B). The introduced *P. parva* populations in the UK had a multimodal distribution suggesting the UK *P. parva* population is either at long term stability or has gone through population bottleneck (Rogers and Harpending 1992). A positive Tajima's *D* value indicates a decrease in population size (Pichler 2002), but for the UK *P. parva* populations it is statistically non-significant, suggesting that the population is at equilibrium. When populations are at equilibrium, the theoretical curves are free of waves as in Figure 4.4 (Rogers and Harpending 1992). The *F*s and R_2 values were also statistically non-significant which is consistent with the suggestion that population is at demographic equilibrium.

S. destruens associated histopathology was not detected in any of the *S. destruens* positive fishes (through molecular detection). Thus, there was no evidence that the presence of the parasite has any detrimental effects for the infected fishes. However, it is important to note that histopathology is a far less sensitive detection method as it only looks at a small proportion of the whole organ. For parasites such as *S. destruens* that can infect multiple organs this can lead to limited detection. As the study by Ercan et al. (2015) has indicated, the best method by which a population impact can be detected is through close monitoring of population site coupled with disease testing in affected populations. This is due to the chronic pattern of mortality associated with *S. destruens* which is difficult to detect with single point sampling (Andreou et al. 2011)

The increasing number of hosts identified in the wild confirms the ability of *S. destruens* to use a broader phylogenetic range of hosts and resultantly maximise its survival and range expansion even in the absence of its reservoir host (Krasnov et al. 2008, Andreou and Gozlan 2016). Fish populations with low prevalence of *S. destruens* can become reservoirs for infection themselves (Peeler et al. 2011) and can result in the re-emergence of the disease if multiple reservoir host populations cross the epidemic threshold – i.e. the level above which *S. destruens* can spread significantly and cause an epidemic. The results from the present study provide valuable insights into the distribution of *S. destruens* in UK waters that can contribute in formulating management options for the parasite. It is evident that multiple *P. parva* populations must be sampled to determine the presence of *S. destruens* and that native fish communities associated with any *P. parva* populations must be monitored for the presence of *S. destruens*. This is particularly relevant to multiple European countries where *P. parva* has spread to river catchments such as Romania, Hungary, Ukraine, Slovakia, Germany, Austria and France (Gozlan et al. 2010a). Monitoring needs to include host population estimates coupled with *S. destruens* prevalence on a yearly basis as in Ercan et al. (2015).

4.5 Summary

The present study has provided valuable insights into the distribution of *S. destruens* in the UK with important management implications. The detection of *S. destruens* in *P. parva* populations emphasizes the monitoring of the adjacent native fish communities. Furthermore, the work has identified new potential fish hosts for *S. destruens* in the wild. The transfer of the parasite to new hosts stresses the implementation of effective measures to control *S. destruens'* further spread in the UK. The Cyt-b marker analysis identified two distinct haplotypes for *S. destruens*, with one haplotype similar to *S. destruens* from China and the other one unique to the UK. The demographic analysis showed that both *S. destruens* and its host *P. parva* populations are at demographic equilibrium in UK. This congruence between *S. destruens* and its host further supports the hypothesis that *P. parva* act as *S. destruens reservoir* host in the UK.

Chapter 5

Environmental DNA detection of *Sphaerothecum destruens* using real-time PCR

5.1 Introduction

Parasite detection using environmental DNA (eDNA) is a powerful tool in disease ecology and epidemiology as it allows fast detection (Walker et al. 2007, Huver et al. 2015). It is particularly powerful when monitoring intracellular parasites, where traditional detection techniques involve the sacrifice of the host. *Sphaerothecum destruens* is an intracellular parasite that has been identified as a potential threat to freshwater fish biodiversity (Gozlan et al. 2005), with the recommendation that its prevalence should be closely monitored (Andreou and Gozlan 2016). Furthermore, the parasite has been identified as a non-native parasite to Europe (Chapter 3), having been introduced with the highly invasive fish *Pseudorasbora parva*, potentially increasing its impact to naïve fish communities (Chapter 4). An extensive eradication program for *P. parva* has been designed and executed since 2005 in the UK (Britton and Brazier 2006).

Theoretical work has indicated that eradication of the host, in this case *P. parva*, does not prevent the establishment of *S. destruens* in adjacent fish communities (Al-Shorbaji et al. 2016) due to the environmental transmission of the parasite. *S. destruens'* spores infect host cells in which they multiply asexually, eventually causing cell death in the host. Following cell death, the spores can infect new cells and/or be released in the environment through bodily fluids such as urine, bile or reproductive fluids (Arkush et al. 2003). In the environment, *S. destruens* spores divide, releasing up to 5 flagellated zoospores per spore (Arkush et al. 2003). The ability of this parasite to persist in the environment (Andreou et al. 2009) and its indirect transmission through contact with spores or zoospores present in its surroundings, increases the probability of transmission to new geographic areas (as indicated in Chapter 4). In order to reduce the risk of disease spread, it is thus important to create epidemiological maps of *S. destruens* (see Chapter 4; Figure 4.1 for the epidemiological map for the UK) and fish movements should be screened for the presence of *S. destruens*.

eDNA detection offers a versatile detection tool that can be used to construct epidemiological maps for the parasite, as well as to establish presence/absence in fish consignments. For example, all water bodies where *S. destruens* is detected through

eDNA will have to be confirmed as positive using the traditional method of detection which involves DNA-based detection and microscopic examination of host tissue (Andreou et al. 2011). Here, an eDNA tool was developed and validated in order to monitor *S. destruens'* presence in both the wild and in fish consignments. The specific objectives were to: (i) develop and validate an eDNA detection assay for *S. destruens*, using a controlled experimental set-up and environmental samples; and (ii) use the eDNA detection assay to monitor the effectiveness of the *P. parva* eradication programme in also eradicating *S. destruens*. This was achieved using the original site of introduction of *P. parva* and *S. destruens* in the UK.

5.2 Material and methods

5.2.1 Spore collection

The *S. destruens* isolate used in this study was originally isolated from sunbleak *L. delineatus* (Paley et al. 2012). *Sphaerothecum destruens* was cultured in *Epithelioma papulosum cyprini* (EPC) cells incubated at 15 °C in minimal essential medium eagle (MEM) with sodium bicarbonate supplemented with 10 % foetal bovine serum (FBS), penicillin 100 IU/ml, streptomycin 100 µg/ml, gentamycin 50 µg/ml and 2 mM L-glutamine (Paley et al. 2012).

Twenty days following the last passage cell associated spores were collected from the infected cell monolayer. Prior to collection, the cell monolayer was washed with 5 ml MEM medium to remove any cell free spores. The infected cell layer was scraped with a cell scraper and suspended in 10 ml of MEM and transferred to 50 ml sterile tubes. The cells were centrifuged at 1,200 x g for 10 minutes and the supernatant was decanted. The cells were re-suspended in 10ml autoclaved distilled water and vortexed vigorously to release spores from host cells. The released spores were washed twice with sterile water and centrifuged at 1,200 x g for 5 minutes between each wash step. The supernatant was discarded between steps and the spore pellet was suspended in sterile water. The spores were enumerated using haemocytometer (Sigma). Cell associated spores were prepared in sterile distilled water at concentrations of 9.28×10^5 spores/ml and 9.28×10^4 spores/ml and were used for subsequent spiking experiments.

5.2.1.1 Experimental Design

To determine the detection limit for *S. destruens*, three spore concentrations representing High (1,500 spores/ml), Medium (150 spores/ml) and Low (50 spores/ml) were spiked in two water conditions - distilled water and turbid water. The turbid water treatment was created to represent natural conditions and included 10 g of un-autoclaved soil, 200 ml of aquarium water in 1,800 ml of distilled water. The experiment was set in 3 L plastic tubs which were covered with cling film with holes for air circulation. Each treatment was triplicated and both water treatments were maintained at 18 °C for 20 days. Untreated controls consisted of distilled water and were include for each sampling point.

5.2.1.2 Water collection and filtration in the experimental setup

Water was sampled from both water treatments at three time points: 6, 13 and 20 days post incubation. At each sampling point, 100 ml of water was collected (Figure 5.1) and was filtered through a 0.45 µm cellulose nitrate filter membrane (Whatman^{TM}). Prior to the sample collection, water was disturbed (to disperse the spores) with a glass pipette moved five times lengthways and sideways with water samples being collected from the centre of the tub. The centre of each tub was marked on the outside of each tub.

5.2.1.3 Disinfection procedure for filter housing

Sodium hypochlorite (NaOCl) is an effective disinfectant and can denature nucleic acids. After every filtration, the filtration setup was immersed in 0.5 % sodium hypochlorite solution for 5 minutes. All components were then washed and flushed with tap water followed by two washes with distilled water. Controls which included distilled water were used between dilutions and were run through the filtration system to detect cross-contamination. The filter papers were then removed using sterile forceps and were subjected to DNA extraction.

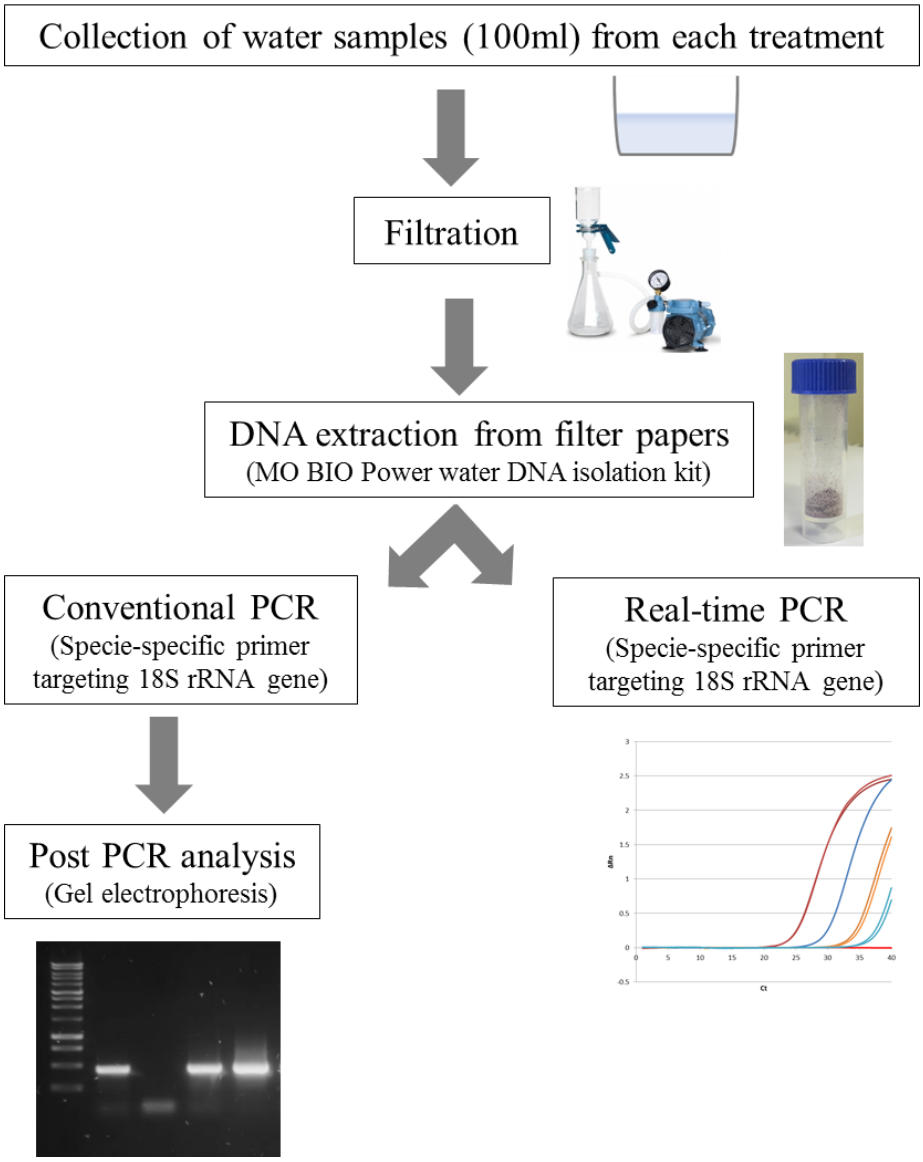


Figure 5.1. An overview of environmental DNA (eDNA) workflow for *Sphaerothecum destruens* detection in the laboratory setting.

5.2.1.4 DNA extraction

The DNA was extracted from filter papers using Power Water DNA Isolation Kit (MO BIO, Inc). Following filtration, each filter membrane was placed into a power bead tube (Figure 5.1). All the steps were performed according to manufacturer's guidelines and DNA was eluted in 100 µl elution buffer. DNA was stored at -20 °C until further use.

The extracted DNA was screened for *S. destruens'* DNA presence using the real-time PCR developed in this study. As a control, all experimental samples were also amplified

using the standard nested PCR for the detection of *S. destruens* (Section 2.2.1; Mendonca and Arkush 2004) as the amplified product from this PCR can be sequenced to confirm the parasite identity.

5.2.2 Real-time PCR

5.2.2.1 Design of real-time PCR primers and probe

Sequences from 18S rRNA gene of *S. destruens* (AY267344.1, AY267345.1, AY267346.1, and FN996945.1) and of fish species *Salmo trutta* (DQ009482.1), *R. rutilus* (AY770580.1), *Oncorhynchus mykiss* (FJ710874.1) and *C. carpio* (FJ710827.1) were retrieved from the GenBank sequence database and were aligned with Clustal W in BioEdit (Hall 1999). The primers and probe specific to *S. destruens* 18S rRNA gene segment were designed with the Primer Express 2.0 software (Applied Biosystems). The Taqman MGB probe was labelled with the fluorescent reporter dye FAM at the 5′-end and a non-fluorescent quencher MGBNFQ at the 3′-end. The unlabelled PCR primers and Taqman probe were purchased from Applied Biosystems.

Table 5.1. Real time PCR primers and probe.

Primer /probe	**Sequence (5′→3′)**	**Melting temperature (°C)**
Forward primer	ACTTTGCGAATCGTATGACATTTTGTC	62.11
Reverse primer	CCACTACCTTACCATCGAAAGTTGA	61.68
Probe	ACGATGATTCATTCAAATTTC	72.31

5.2.2.2 Real-time PCR reaction conditions

The TaqMan® Gene Expression Master Mix-UDG was used for this assay (Invitrogen). The reaction conditions consisted of 20 µl reaction volumes containing 10 µl TaqMan® Gene Expression Master Mix-UDG, 1 µl assay mix (primers and probe) and 2 µl of DNA template (undiluted). All reactions were performed in the StepOne real time PCR machine (Applied Biosystems) and analysed by StepOne software v 2.0. In all the analysis, the software defined baseline was automatically set to 0.02, and ROX was

selected as a passive reference. Cycling conditions consisted of a holding stage at 50 °C for 2 min to allow UDG enzymatic activity and initial denaturation at 95 °C for 10 min followed by 40 cycles of denaturation at 95 °C for 15 seconds and annealing at 60 °C for 1 min. The extraction blank controls from the DNA extraction were included along with the no DNA template PCR controls.

5.2.2.3 Limits of detection

In order to quantify the genomic *S. destruens* DNA in samples and to determine the sensitivity of the assay, a calibration curve was generated using genomic DNA extracted from *S. destruens* spores. The curve was obtained by plotting DNA concentration (ng/µl) against cycle threshold (C_t) values. A ten-fold serial dilution of *S. destruens* genomic DNA was prepared in UV-irradiated sterile water to give a template concentration of 10 ng/µl to 1 fg/µl. The standards were run in triplicate in order to test the repeatability of the quantification using the real-time PCR assay. Negative controls used in this assay consisted of sterile water. The detection limit was defined as the lowest genomic *S. destruens* DNA concentration detected at least 95 % of the times by the RT-PCR assay.

5.2.2.4 Real-time PCR assay specificity

The *S. destruens*-specific 18S rRNA assay was tested for cross-reactivity with pure fish DNA. The fish species tested for cross-reactivity were carp *Cyprinus carpio*, roach *Rutilus rutilus*, minnow *Pimephales promelas*, common bream *Abramis brama*, chub *Squalius cephalus*, barbel *Barbus barbus* and topmouth gudgeon *Pseudorasbora parva*. The real-time PCR primers specificity was also tested using a PCR amplified *Dermocystidium salmonis* 18S rDNA section inserted pGEM® –T (Promega) provided by Dr Richard Paley (CEFAS laboratory, Weymouth, UK). A serial dilution of all the tested fishes DNA and *D. salmonis* DNA was created in sterile UV irradiated water ranging from 5 ng/ µl to 0.0005 ng/µl. Two microliter of each dilution was used as a starting material for real time PCR giving a total genomic DNA range of 10 ng to 0.01 ng.

5.2.3 Assessment of the eDNA detection technique in the field

For the detection of *S. destruens* in environmental samples, three independent water bodies in the UK were targeted (Table 5.2): (a) a decommissioned ornamental fish farm (referred as section 1a, SU3862) where *P. parva* has been present since mid-1980s, and *S. destruens* has been detected in sampled fish (Chapter 4); (b) the stream above the decommissioned fish farm (referred as 'section 1d') as well as the stream below the decommissioned fish farm (referred to as 'section 1b' and 'Section 1c" and where *S. destruens* positive fish have been detected; Figure 5.2) (c) the Bourne stream (SZ0679) and (d) the River Teme (SO8335, SO7237).

(a, b) Sampling the decommissioned ornamental fish farm and the adjacent stream

P. parva was first recorded in the UK in the outflow stream of this ornamental fish farm in 1996 (Domaniewski and Wheeler 1996). The site was considered as a high-risk site based on the conservation and fishery value of the adjacent waters and complete eradication of *P. parva* was initiated in summer 2014. Following rotenone application, the eradication programme was completed in March 2015 (personal communication Dr Rob Britton). The ornamental fish farm consists of numerous artificial ponds that were used to breed golden orfe *Leuciscus idus*. The fishery has an outflow into the Tadburn Lake stream which flows into the River Test approximately 6 km downstream of the fishery. Water sampling at this location occurred prior and post *P. parva*'s eradication. The *P. parva* sampled from this fish farm (section 1a) and native fish species (Chub *Squalius cephalus*, Dace *leuciscus leuciscus*, brown trout *Salmo trutta* and Roach *Rutilus rutilus*) from the adjacent streams (section 1b and 1c; Figure 5.2) in 2013 confirmed *S. destruens* presence (Chapter 4, Table 4.1).

(i) Sampling of the decommissioned ornamental fish farm pre-*P. parva* eradication

The site was first sampled in 2013, whilst *P. parva* was present in the fishery. Within the ornamental fish farm section 1a, two ponds Pond 12 (P 12) and Pond 14 (P 14) and the overflow pond (OF) were sampled (Figure 5.2). All ponds were populated with *P.*

parva at varying densities Table 5.2. Six 1 L samples were collected from each pond (P12, P14, OF), from sampling points spread equidistantly around the pond edge (approximately 15 m apart in P12 and P14, and 20 m apart in OF).

(ii) Sampling of the decommissioned ornamental fish farm post *P. parva* eradication

In order to monitor the effectiveness of the eradication programme in controlling *S. destruens*, water samples were collected in 2016 from multiple locations associated with the decommissioned ornamental fish farm. These included the stream feeding the ornamental fish farm (section 1d, Figure 5.2); the fishery (Pond 1, section 1a, Figure 5.2) and the Tadburn Lake stream (sections 1b and 1c, Figure 5.2). Details on the water sampled are in Table 5.2. One litre water samples were collected from section 1c and section 1d at three sampling points and from section 1b at six sampling points (see Table 5.2). The water samples were immediately stored on ice and were filtered within 24 hours.

(c, d) Sampling the Bourne stream and River Teme

In order to test the specificity of the technique for *S. destruens* detection in natural conditions, water samples were collected from Bourne stream and River Teme (Table 5.2). These two water bodies have no known introduction or sighting of *S. destruens*' healthy host *P. parva*. One litre water samples were collected from the Bourne stream at three points approximately 356 m apart and were stored on ice and filtered within 24 hours. One litre water samples were also collected from two locations at Knightwick and below Powick weir in River Teme (Table 5.2) and were stored on ice and filtered within 24 hours.

(iii) Fish sampling from downstream of decommissioned ornamental fish farm.

The native fish species sampled from these downstream sections in 2013 pre-*P. parva* eradication were found positive for *S. destruens* through PCR (Chapter 4). Electric

fishing, using a back-mounted Smith-Root LR-24 Backpack, was performed post *P. parva* eradication in 2016 to sample fish from downstream of the fish farm referred as "section 1b" and "section 1c" respectively. The fish sampled were stone loach *Noemacheilus barbatulus*, bullhead *Cottus gobio* and sticklebacks *Gasterosteus aculeatus* (Table 5.2). All the brown trout were returned to the water without processing due to permission restriction by the Environment Agency. The sampled fish were euthanized through anaesthetic overdose on site following Home Office guidelines and were transferred to the laboratory on ice.

In the laboratory, all the fish were preserved in 100 % ethanol until further processing. The fish were dissected and kidneys were collected as explained in Section 3.2.1. The DNA extraction was performed using Qiagen DNeasy Blood and tissue kit (Qiagen) and screened for *S. destruens* presence through nested Cyt-b PCR as presented in Section 2.2.3.

5.2.3.1 Collection and filtering of environmental samples

The water samples were collected in 1 L sterile plastic bottles attached to rods and the bottles were submerged in a way that a vertical column of water was collected. The sampling equipment was changed between each sampling point and disposable gloves were used for every site. Two negative controls were inserted during field sampling (one at the start and one at the end). Field negative samples consisted of 1 L sterile plastic bottles filled with sterile water which were treated in the field in the exact manner as sample collection bottles. In the lab, the water was pre-filtered using a 200 µm filter to remove coarse material. An 80-500 ml subsample was further filtered using a 0.45 µm cellulose nitrate filter membrane (Whatman^{TM}) which was then subjected to DNA extraction (Section 5.2.2.3) and RT-PCR (Section 5.2.3.2).

Table 5.2. Field sampling for the validation of the environmental DNA (eDNA) technique for *Sphaerothecum destruens*.

Site	Ponds /stream	Sampling points	Geographical coordinates	Fish composition	Volume of water filtered (ml)	Fish sampled (number)
Decommissioned ornamental fish farm						
(i) Pre-eradication of *P. parva* (2013)	Section 1a Pond 12 (52 m x 7 m)	Six 1 L samples (12-1–12-6) around the pond edge app. 15 m apart		Low density of *P. parva*		
	Section 1a Pond 14 (52 m x 7 m)	Six 1 L samples (14-7 –14-12) around the pond edge app. 15 m apart	SU3862	High density of *P. parva*, Stickleback *Gasterosteus aculeatus* and Signal Crayfish *Pacifastacus leniusculus*.	80	-
	Section 1a Over Flow Pond (pond running east-west, to south of fishery pond row). (65 m x 15 m)	Five 1 L samples (OF-13-OF17) around the pond edge app. 20 m apart		Intermediate density of *P. parva*, *C. cyprio*, *L. idus*, *T. tinca*, *G. aculeatus* and *P. leniusculus*.		-

(ii) Water sampling post-eradication of *P. parva* (2016)	Section 1d (60 m stretch to inlet to fishery)	Three 1 L samples app. 20 m apart	SU3932			-
	Section 1a Pond 1 (52 m x 7 m)	Two 1 L samples (1a-1b) from two extremes of the pond.	SU3862	Carp (*Cyprinus carpio*)		-
	Section 1b- (122 m downstream of fishery)	Three 1 L samples (1b1- 1b3 along the stream stretch app. 40 m apart	SU3862	-	500	
	Section 1b (360 m downstream of fishery)	Three 1 L samples (1b4-1b6) along the stretch app. 50 m apart	SU3848	Stone loach *Noemacheilus barbatulus*, bullhead *Cottus gobio*, stickleback *G. aculeatus*, Signal crayfish, sea lamprey *Petromyzon marinus*		Stone loach *Noemacheilus barbatulus* (3), bullhead *Cottus gobio* (5), stickleback *G. aculeatus* (9)
	Section 1c (500 m downstream of fishery)	Three 1 L samples (1c1-1c3)	SU3842			Bullhead *C. gobio* (2)
Bourne stream	Site 1 (stagnant water)					-
	Site 2 (Fast flowing water)	1 L water samples at each site	SZ0689	Rudd *Scardinius erythrophthalmus*, *C. carpio*, Minnow	1000	-
	Site 3 (this site					

	was further downstream to site 1 & site 2)		SZ0679	*Phoxinus phoxinus*, *G. aculeatus* and *S. cephalus*		-
River Teme	Powick below	1 L water samples	SO8335	*S. cephalus*, barbel	1000	-
	Knightwick		SO7237	*Barbus barbus*, shad *Alosa fallax, P. marinus*		-

Figure 5.2. An overview of the decommissioned ornamental fish farm ponds (section 1a) and Tadburn Lake stream (blue). Water samples were collected from four ponds from Section 1a; picture insert- Pond 1 (P1), Pond 12 (P12), Pond 14 (P14) and Overflow pond (OF). After the eradication of *P. parva* from this fish farm, water samples were collected from the Tadburn Lake stream section feeding into the farm referred as "Section 1d", and the Tadburn Lake stream section receiving the facility's outflow termed as "Section 1b" and 500 m after the facility towards the Tadburn Lake stream's end "Section 1c". Blue arrows (→) indicate flow direction. The red lines represent connection between the Tadburn Lake stream and the decommissioned ornamental fish farm with continuous lines (▬) representing above ground connections and discontinuous lines (---) representing below ground connections.

5.3 Results

5.3.1 Validation of the eDNA detection method of *Sphaerothecum destruens* using real-time PCR.

Using ten-fold serial dilution of *S. destruens* genomic DNA, the limit of detection of Taqman assay was 1 pg/µl (Table 5.3). The C_t-values with standard genomic DNA dilutions in the late cycle (> 36) which corresponded to 0.1 pg/µl were unreliable as the probability of detection was < 95 % (Burns and Valdivia 2008). Therefore, the C_t-values >36.55 were scored as negative or below the detection limit, in line with other studies in development of eDNA method for parasite detection (Kirshtein et al. 2007, Huver et al. 2015). In the assay, PCR negatives had no C_t readings. The RT-PCR was also highly specific to *S. destruens* with all tested fishes and *D. salmonis* yielding no C_t values following amplification with the RT-PCR primers.

In the experimental validation set-up, *S. destruens*-specific DNA was detected by real-time PCR in both natural and turbid water conditions until 6, 13 and 20 days across all spore concentrations (Table 5.4). The real-time assay was further checked by amplifying all the samples in the experimental validation experiment with the *S. destruens* nested PCR (Mendonca and Arkush 2004). The real-time assay detected *S. destruens*-specific DNA at day 20 in the low spore concentrations whereas the nested PCR only detected *S. destruens* DNA in the lowest concentration treatment at day 6 (Table 5.5; and Appendix 2).

Table 5.3. C_t values and percentage (%) detection for the ten-fold serial dilution of *Sphaerothecum destruens* DNA using the real-time PCR assay

Standards /dilutions [a]	**Mean C_t-values (n=3)** [b]	**% detection** [c]	**DNA (ng/µl)** [d]	**ng DNA (2 µl) in RT-PCR** [e]
1/10	20.67	100	5	10
$2/10^{-1}$	23.7	100	$5x\ 10^{-1}$	1
$3/10^{-2}$	27.30	100	$5x\ 10^{-2}$	$1x\ 10^{-1}$
$4/10^{-3}$	31.53	100	$5x\ 10^{-3}$	$1x\ 10^{-2}$
$5/10^{-4}$	36.55	100	$5x\ 10^{-4}$	$1x\ 10^{-3}$
$6/10^{-5}$	38.34	66.6	$5x\ 10^{-5}$	$1x\ 10^{-4}$
$7/10^{-6}$	38.43	33.3	$5x\ 10^{-6}$	$1x\ 10^{-5}$
$8/10^{-7}$	Undetected	0	$5x10^{-7}$	$1x10^{-6}$

[a] A total of 7 diluted standards were made from a ten-fold dilution series of DNA stock with a measured concentration of 5 ng genomic DNA/µl.
[b] Mean C_t-values are based on the RT-PCR replicates of each standard.
[c] The percentage of RT-PCR replicates yielding positive replicates (detection) for each standard.
[d] Theoretical content of DNA in ng/µl for each standard calculated from concentration assigned to the DNA stock (5 ng/µl).
[e] Quantity of template DNA in each RT-PCR replicate.

Table 5.4. Detection limit of *Sphaerothecum destruens* DNA in distilled and turbid water over 20 days using real-time PCR. Results are displayed as number of positive replicate/total number of replicates for each treatment.

Spore count	**1,500,000/L (High)**		**150,000/L (Medium)**		**50,000/L (Low)**	
Incubation point	**Sterile water**	**Turbidity**	**Sterile water**	**Turbidity**	**Sterile water**	**Turbidity**
Day 6	3/3	3/3	3/3	3/3	2/3	3/3
Day 13	2/3	1/3	2/3	1/3	3/3	1/3
Day 20	3/3	1/3	3/3	1/3	3/3	2/3

Table 5.5. Comparison of detection limit of *Sphaerothecum destruens* DNA in distilled and turbid water over the course of 20 days by conventional and real-time PCR. Black and red arrows indicate conventional and real-time PCR respectively.

Spore count	**1,500,000/L (High)**				**150,000/L (Medium)**				**50,000/L (Low)**			
Incubation point	**Sterile water**		**Turbidity**		**Sterile water**		**Turbidity**		**Sterile water**		**Turbidity**	
Day 6	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Day 13	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	-	✓
Day 20	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓

5.3.2 Assessment of the eDNA technique using environmental samples

(i) Decommissioned ornamental fish farm

A) Sampling of the decommissioned ornamental fish farm pre-*P. parva* eradication

Three ponds with varying densities of *P. parva* were investigated for *S. destruens* presence. *Sphaerothecum destruens* was not detected in extractions of water samples from Pond 12. However, *S. destruens* was amplified from water samples collected in Pond 14 and the overflow pond. Pond 14 had low C_t-values with comparatively higher C_t-values for the overflow pond which corresponds to high and low levels of *S. destruens* presence respectively (Table 5.6; Figure 5.3).

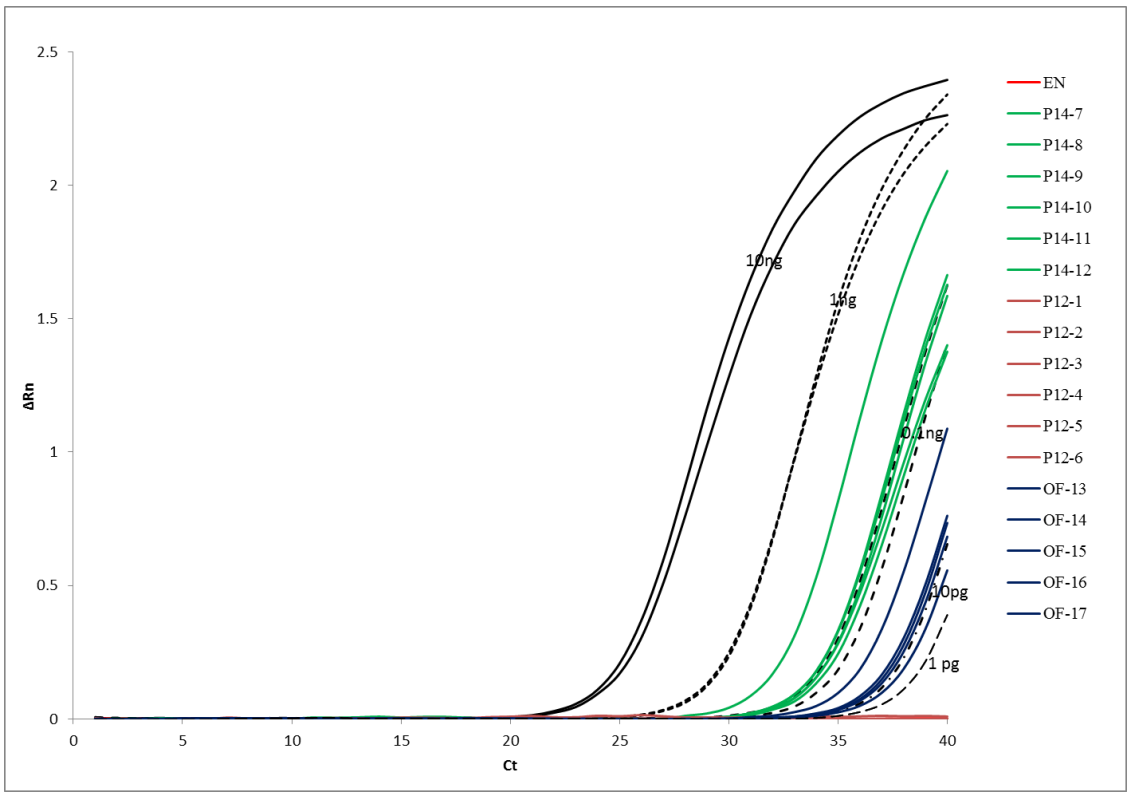


Figure 5.3. Amplification plots of DNA extractions from the decommissioned ornamental fish farm pre-*Pseudorasbora parva* eradication (2013). Change in fluorescence (ΔRn) is plotted against cycle threshold number (C_t-values). The *Sphaerothecum destruens* DNA standards (10 ng, 1 ng, 0.1 ng, 10 pg and 1 pg) are labelled on the graph (please note each DNA standard was duplicated). Extractions corresponding to positive samples from Pond 14 (green) and the overflow pond (blue) and negative samples from Pond 12 (red) are displayed.

B) Sampling of the decommissioned ornamental fish farm post *P. parva* eradication

S. destruens DNA was amplified from the upstream section of the Tadburn Lake stream (section 1d- Figure 5.2) with high C_t-values ranging from 34.3 to 36.5 that indicates *S. destruens* presence at low concentrations (Table 5.6, Figure 5.4).

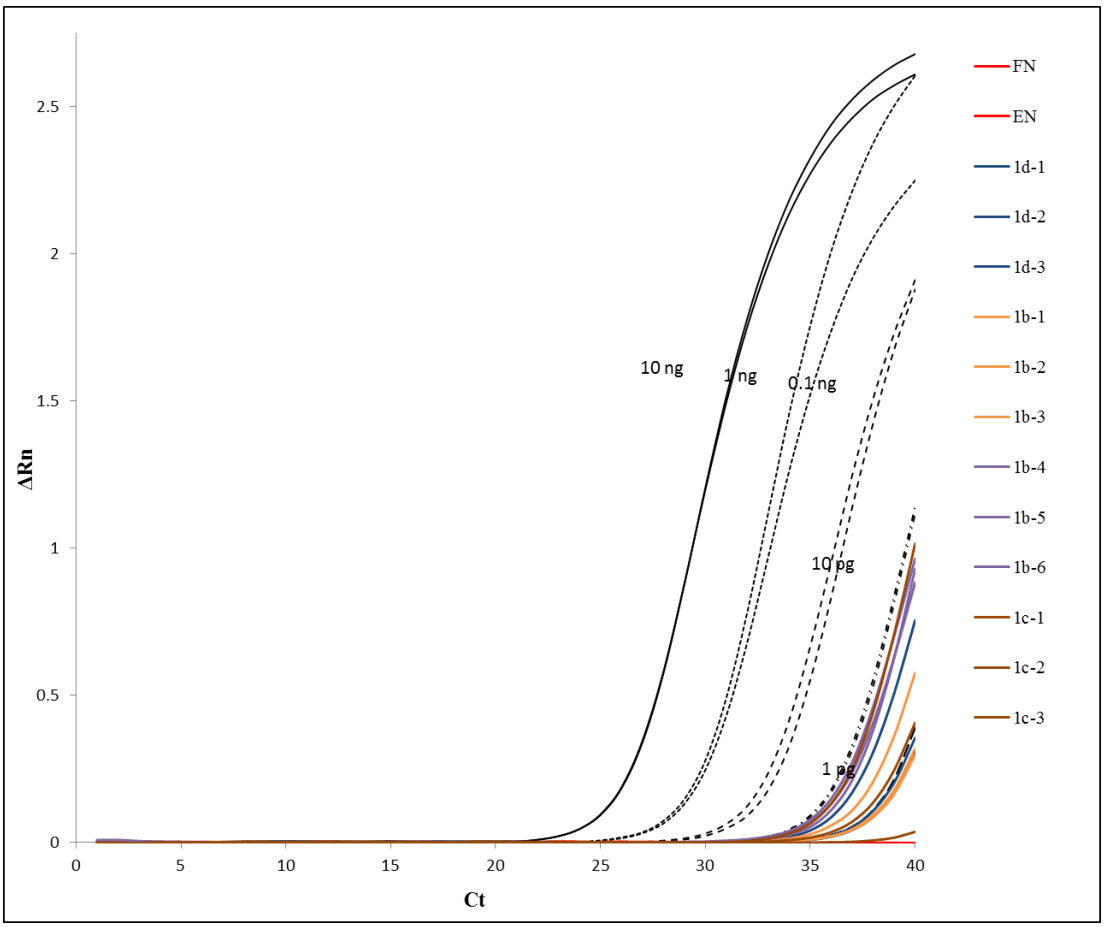


Figure 5.4. Amplification plots of DNA extractions from decommissioned ornamental fish farm post *Pseudorasbora parva* eradication (2016), Section 1a-1d. Change in fluorescence (ΔRn) is plotted against cycle threshold number (C_t-values). The *Sphaerothecum destruens* DNA standards (10 ng, 1ng, 0.1 ng, 10 pg and 1 pg) are labelled on the graph (please note each DNA standard was duplicated). The negative control (FN, filtration negative; EN, extraction negative)- red and extractions corresponding to water samples from section 1d (1d 1-3)- blue, section 1b (1b 1-3)- mustard, section 1b (1b 4-6)- purple and section 1c (1c 1-3)- brown are displayed.

The C_t-values for the samples obtained from outflow of the decommissioned ornamental fish farm into Tadburn Lake stream (section 1b-Figure 5.2) indicated the presence of *S. destruens* DNA. Specifically, C_t-values ranged from 35.3 to 36.7 for sampling points 1-3 (closest to the decommissioned fishery) and 33.11 to 33.4 at sampling points 4-6 (approximately 360 m away from the decommissioned fishery) (Table 5.6). *S. destruens* DNA was also detected in 2 out of 3 samples from section 1c (500 m away from the decommissioned Fishery-Figure 5.2) with C_t values 33.2 to 34.4.

C) Fish sampling from downstream of decommissioned ornamental fish farm.

None of the fish samples screened for *S. destruens* presence tested positive with nested Cyt-b PCR (Table 5.6).

Table 5.6. Validation of the eDNA technique for *Sphaerothecum destruens* using environmental samples.

Site	**Pond**	**Samples**	**Mean C_t-values (n=2)**	**RT-PCR *S. destruens* status**	***S. destruens* status Cyt-b**
Decommissioned ornamental fish farm (i) pre-eradication of *P. parva*	Section 1a Pond 12	P12-1 to 12-6	Undetected	-	
	Section 1a Pond 14	P14-7	31.21	✓	
		P14-8	31.01	✓	
		P14-9	31.37	✓	
		P14-10	31.52	✓	
		P14-11	29.12	✓	
		P14-12	31.03	✓	
	Section 1a Overflow Pond	OF-13	34.34	✓	
		OF-14	32.82	✓	
		OF-15	34.43	✓	
		OF-16	34.84	✓	
		OF-17	34.19	✓	
(ii) post-eradication of *P. parva*	Section 1a Pond 1	P1a	Undetected	-	
		P1b	35.8	✓	
	Section 1d	1d-1	35.95	✓	
		1d-2	34.3	✓	
		1d-3	36.2	✓	
	Section 1b	1b-1	36.71	-	
		1b-2	36.6	-	
		1b-3	35.3	✓	- (0/17)
		1b-4	33.4	✓	
		1b-5	33.5	✓	
		1b-6	33.11	✓	
	Section 1c	1c-1	34.4	✓	- (0/2)
		1c-2	38.46	-	
		1c-3	33.2	✓	
Bourne stream	Site 1	S1-1	Undetected	-	
		S1-2	37.15	-	
	Site 2	S2-1	37.25	-	
		S2-2	Undetected	-	
	Site 3	S3-1	37.39	-	
		S3-2	Undetected	-	
River Teme	Knightwick	K	Undetected	-	
	Powick below	PB	Undetected	-	

(ii) Bourne Stream and River Teme

No amplification for *S. destruens* DNA was detected in the Bourne stream and River Teme as the mean C_t-values were >36.5 (Table 5.6 and Figure 5.5).

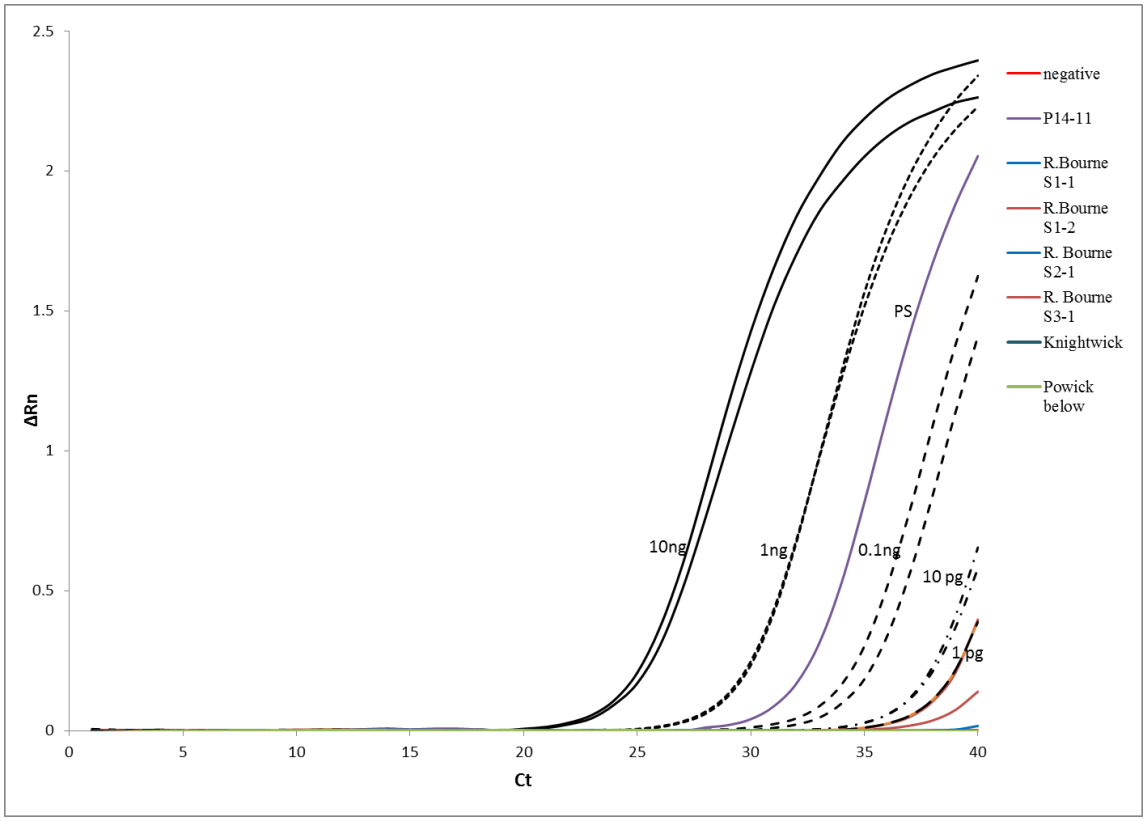


Figure 5.5. Amplification plots of DNA extractions from Bourne stream and River Teme. Change in fluorescence (ΔRn) is plotted against cycle threshold number (C_t-values). The *Sphaerothecum destruens* DNA standards (10 ng, 1 ng, 0.1 ng, 10 pg and 1 pg) are labelled, the negative control (NTC) and extractions corresponding to a positive sample (PS) for *S. destruens* from decommissioned ornamental fish farm (P14-11-purple) and negative samples (NS) from the Bourne stream (S1-1- S1-3) and River Teme (Knightwick- blue and Powick below-green). The multiple curves indicate duplicates for standards.

(iii) Negative controls

All the negative controls used during field sampling, filtration, DNA extraction and PCR were negative for *S. destruens* DNA (C_t-value = undetected), indicating that there was no *S. destruens* carry-over at any step of the process.

5.4 Discussion

In this study, an eDNA detection method for *S. destruens* was successfully developed and validated both in the laboratory and in the field. In the lab, *S. destruens* DNA could be detected using the real-time assay post inoculation both in the presence of sterile water and turbid water conditions across *S. destruens* concentrations from 1,500 spores/ml to 50 spores/ml. In the turbid water conditions, *S. destruens* DNA was not detected in all replicates as eDNA can get adsorbed to soil particles as it settles where it is subjected to slow degradation processes affecting its detectability (Turner et al. 2015). In addition ultraviolet irradiation levels and pH levels can influence the speed of DNA degradation (Pilliod et al. 2014, Strickler et al. 2015) reducing its detectability in natural systems.

The real-time PCR assay had a detection limit of 1 pg (equivalent to 500 spores, explained in Section 2.3.4, Table 2.2) which was slightly inferior to efficiencies reported for qPCR for chytrid fungus *Batrachochytrium dendrobatidis* (0.1 zoospore) and trematode *Ribeiroia ondatrae* (14 fg) (Walker et al. 2007, Huver et al. 2015). The assay was specific to *S. destruens* as it did not cross-react with any of the tested fishes and the closely related *D. salmonis*. In addition, water from water bodies where both *S. destruens* and its reservoir host *P. parva* have never been reported (e.g. the Bourne stream and River Teme) tested negative for *S. destruens* DNA. In contrast, *S. destruens* DNA was detected from the decommissioned ornamental fish farm where it was detected in *P. parva* tissues (Chapter 4 Table 4.1).

Water samples were collected from the decommissioned ornamental fish farm prior to the eradication of *P. parva* and 2 years post its eradication. Whilst *P. parva* was present, *S. destruens* DNA was detected from ponds with medium and high *P. parva* densities and was not detected in the pond with low *P. parva* density. The *S. destruens* prevalence in *P. parva* in the facility was approximately 3 % (referred to as UK site 1a in Chapter 4; Table 4.1). Transmission of *S. destruens* can be density dependent, with higher transmission rates and prevalence in denser populations (Al-Shorbaji et al. 2015), which can explain the detection and lower C_t-values in the high *P. parva* density ponds.

Sphaerothecum destruens DNA was detected in water samples from the Tadburn Lake stream both upstream and downstream of the decommissioned ornamental fish farm two years post *P. parva* eradication. Whilst the facility was active, water was diverted form upstream through the facility. Fish movement between the facility and the upstream was not possible due to the flow and the restriction of the fishery fishes in ponds. The presence of *S. destruens* in the upstream could thus be potentially due to otter's movement between the fishery and the stream for their food or movement of brown trout *Salmo trutta* from lower stretches of the Tadburn Lake stream. Interestingly, the *S. destruens* DNA was in higher concentration in water samples collected downstream of the fishery. In particular, *S. destruens* DNA concentration was highest in the samples further downstream (section 1b, samples 1b4- 1b6) versus the samples collected closer to the fishery (section 1b, samples 1b1- 1b3). This could be due to higher abundance of fish in downstream stretch (section 1b; 1b4- 1b6) and in section 1c (sample 1c; 1-3). Additionally, the native fishes *S. trutta, S. cephalus*, *R. rutilus*, and *L. leuciscus* sampled from section 1b in 2013 showed high prevalence (33-100 %) of *S. destruens* in the fish tissue (referred as UK site 1b in Chapter 4; Table 4.1). Sampling in 2016 was unsuccessful in capturing *S. cephalus*, *R. rutilus*, and *L. leuciscus* in the Tadburn Lake stream. Although present *S. trutta* could not be sampled due to permission restriction set by the Environment Agency. Additional potential host species sampled (Stone loach *N. barbatulus*, bullhead *C. gobio*, and stickleback *G. aculeatus*) were all negative when tested for the presence of *S destruens* using nested PCR. Therefore, future work should include a comprehensive survey of the stream, including invertebrates, in order to determine the potential reservoirs for *S. destruens*.

The detection of *S. destruens* DNA two years post *P. parva* eradication suggests that the parasite has established in the fish community of Tadburn stream, as it is highly unlikely that *S. destruens* spore DNA would persist for 2 years in the environment in the absence of its reservoir host *P. parva*. *Sphaerothecum destruens* spores zoosporulate in the presence of freshwater with a maximum survival of 7 days (Andreou et al. 2009). In natural field conditions, zoospores are more prone to cell disruption and ultimately DNA degradation (Strickler et al. 2015). In the experimental setup, the *S. destruens* eDNA persistence was up to 20 days which is within the established eDNA persistence rate of 14-60 days in freshwater ecosystems (Goldberg et al. 2015). It should be noted that these figures are for macroorganisms (Dejean et al. 2011, Pilliod et al. 2013) which

have higher abundances and can shed more DNA into waterbodies compared to microbes (Thomsen et al. 2012, Pilliod et al. 2013). The present results could support the theoretical prediction that the parasite can maintain transmission in other fish species even in the absence of its reservoir host, *P. parva* (Al-Shorbaji et al. 2016).

The majority of eDNA studies carried out involve aquatic vertebrates or macro organisms, with a little focus on microscopic parasites (Bass et al. 2015). The eDNA detection tool for *S. destruens* fills this gap and can be used to screen fish consignments both from within UK movements and fish imports by simply testing for the presence of *S. destruens* DNA in the water. Where *S. destruens* DNA is detected, further tests should include histopathological investigation of a subsample of fishes coupled with molecular detection in fish tissues (Huver et al. 2015). Species detection through eDNA is often verified with independent detection methods (Hyman and Collins 2012, Thomsen et al. 2012).

The eDNA tool can be combined with site occupancy models (Schmidt et al. 2013) to develop epidemiological maps for *S. destruens* across its suspected range. This is particularly important as the parasite can spread through water transfer including contaminated angling equipment. This raises the need of a wider survey for the parasite prevalence especially in the waters adjacent to the decommissioned ornamental fish farm and the waterbodies where *P. parva* is or has been present (Britton et al. 2010). To increase probability of detection, sampling should occur during the spring as *S. destruens* infections highest during the spring season (Ercan et al. 2015). It is important to note that the use of this tool is not limited to the UK, as Chapter 3 has shown that the parasite is present in at least 3 more European countries (Spain, Netherlands and Turkey). Thus, a cheap and quick detection method is now available for a pan-European survey for *S. destruens* which would inform the future management of this non-native parasite.

5.5 Summary

A new eDNA method was developed for the detection of *S. destruens* in the wild in the present study. The method successfully detected *S. destruens* DNA where expected (decomissioned ornamental fishfarm and its associated Tadburn Lake stream). Compared to the traditional survey methods, eDNA method is fast (*S. destruens* status can be determined in 2 days) and non-invasive (no fish killing involved). The eDNA method can be effectively employed in the development of epidemiological maps for *S. destruens* across its suspected range. This study has demonstrated that despite the eradication of source host *P. parva*, it is impossible to eradicate the environmentally transmitted propagule (*S. destruens* spores), once it has established in the community. This emphasizes that preventive measures against pathogen expansion should be implemented, as reactive measures such as eradication would not be effective. The relevant measures such as, early detection of *S. destruens* can be achieved using the new eDNA method to initiate rapid actions to prevent its further dispersal. The newly developed eDNA method will also serve as a rapid detection tool for a pan-European survey for *S. destruens* which would inform the future management of this non-native parasite. In case of positive detection with the eDNA method, it is highly recommended to couple it with fish dissection to reliably ascertain the parasite's distribution and its impact on fish hosts in the suspected waters.

Chapter 6

Sphaerothecum destruens taxonomy and mitochondrial genome organisation

6.1 Introduction

Mitochondria are double membrane organelles ubiquitous to eukaryotes with a few exceptions (Burger et al. 2003a). In addition to their primary role in energy production through the electron transport chain coupled with oxidative phosphorylation, mitochondria also play role in translation, transcription, RNA processing and protein import and maturation (Gray 2012). A broad spectrum of biochemical and phylogenetic studies have supported that the mitochondria originated from a single alpha-proteobacterial ancestor from within the order Rickettsiales (Wang and Wu 2015). Animal mtDNA are usually small, circular with invariable gene content and usually range in size from 13 to 19 kbp, are compactly arranged without intergenic regions (only a few bp in some cases) and are intron-less (with a few exceptions where mitochondrial group I introns were found in some Cnidaria and Placozoa (Lavrov 2007). Relatively large mtDNAs (20 to 43 kbp) have been found in some Cnidaria, Demospongiae and Placozoa (Lavrov 2007). In animals, the standard set of mtDNA genes encode for 12-13 proteins involved in the electron transport chain and oxidative phosphorylation and 24-25 structural RNAs, including small and large subunit rRNA and tRNAs (Anderson et al. 1981, Bibb et al. 1981, Anderson et al. 1982).

Mitochondrial DNA organization varies in the unicellular relatives of animals (Choanozoa) and provides insights into the mtDNA evolution. They have additional mtDNA proteins: the extra respiratory proteins and ribosomal proteins (Gray et al. 1999). A great diversity was observed in the mitochondrial genome size and topology of the Choanozoa, ranges from single-chromosome circular mitochondria in *Monosiga brevicollis*, to linear single-chromosome mitochondria in *Ministeria vibrans* and *Capsaspora owczarzaki* and linear multiple-chromosome mitochondria has been identified for *Amoebidium parasiticum* with variable genome sizes range 76 kbp to > 200 kbp (Burger et al. 2003a, Lavrov and Lang 2014).

Previous studies on the animal mtDNA and their comparisons with early diverging animals (sponges) and unicellular relatives (Choanoflagellate and Ichthyosporea) have highlighted that mtDNA has undergone tremendous transitions from large, spacious

mtDNA to compactly arrayed. The comparison of mtDNA genomes of bilateral and radially symmetrical animals has identified variations in the size and gene content. Usually, mitochondrial evolution is correlated with two main events in animal evolution i.e. the origin of multicellularity and the origin of Bilateria (Lavrov 2007).

The origin of multicellularity is usually linked with the loss of multiple genes and reduction of non-coding mtDNA, as evident from the large mt-genomes in the unicellular relatives of animal with either long repeat sequences or non-coding regions in Choanoflagellate, Filasterea and Ichthyosporea (Burger et al. 2003a, Lavrov and Lang 2014). The emergence of Bilateria is usually correlated with multiple changes in genetic code associated with loss of tRNA genes, increased rate of sequence evolution and emergence of several genetic novelties, such as highly modified structures of ribosomal and transfer RNAs and the presence of single noncoding "control region" in mtDNA (Wolstenholme 1992). However, it is not fully established whether these changes have happened simultaneously with the morphological transitions or if mtDNA evolved independently in different lineages (Lavrov 2007).

The findings from the mtDNA studies of *M. brevicollis* and *A. parasiticum* led to the hypothesis that the last common ancestor of Holozoa (multicellular animals and their closest unicellular relatives) had gene-rich mtDNA. Mitochondrial DNA in the Holozoa (animals and their unicellular relatives) has been hypothesized to evolve from its common ancestor, which is assumed to have possessed a large and non-compact mtDNA, along three main routes: the Ichthyosporea lineage (accumulation of repeat sequences), the Choanoflagellate lineage (amplified intergenic region) and the Metazoan lineage (extensive gene loss with size contraction) (Burger et al. 2003a, Signorovitch et al. 2007). However, this hypothesis is based on limited taxonomic sampling for mitochondrial analysis with one organism from both Ichthyosporea and Choanoflagellate lineages.

Sphaerothecum destruens is an obligate intracellular organism sitting at the animal-fungal boundary. A recent phylogenomic study based on the flagellar and chitin synthase characters has placed *S. destruens* in a group termed the "Teretospores" which

is comprised of the Ichthyosporea and *Corallochytrium limacisporum* and is designated as the earliest Holozoan divergence (Torruella et al. 2015). This group is interesting in that they are phylogenetically located where animals first diverged from the fungi and can provide important clues into the origin of higher organisms and mtDNA evolution. The extant organisms from this group are not always easily available and experimentally amenable.

To date, the *S. destruens* phylogeny has not been evaluated based on the mt-genome. *Sphaerothecum destruens* and *A. parasiticum* belong to the same Class Ichthyosporea, within Orders Dermocystida and Icthyophonida respectively. The presence of peculiar mt-genome architecture in *A. parasiticum*, which shows a rampant expansion and fragmentation of its mt-genome, raises the question whether a similar trend is also present for *S. destruens*. Thus, the aims of this study were (i) to investigate the mitochondrial genome organization and content of *S. destruens* through its mtDNA amplification by Long Range PCR and subsequent DNA sequencing (Primer walk); (ii) to better decipher its taxonomic position, reconstruction of the phylogenetic tree based on amino acid sequences derived from the mtDNA encoded genes and (iii) critically evaluate how the *S. destruens* mtDNA structure and organization contributes to knowledge about mitochondrial evolution in unicellular animals.

6.2 Material and methods

6.2.1 DNA extraction of *Sphaerothecum destruens* spores

Sphaerothecum destruens whole DNA was extracted from its spores using DNeasy Blood and tissue kit (Qiagen). All the steps were performed according to manufacturer's guidelines and DNA was eluted in 100 µl elution buffer and quantified using the Nanodrop (Thermofisher).

6.2.2 Amplification of 'anchor regions' of mitochondrial DNA (mtDNA)

A number of universal mtDNA primers for Metazoa and degenerate primers specific for Cnidarians were used to amplify short gene fragments of *S. destruens* mtDNA. The primers tested and their outputs are listed in Table 6.1.

6.2.3 Designing of *Sphaerothecum destruens*-specific long PCR primers

The sequences obtained from the mtDNA gene fragments cox1, cob and nad5 were used to manually design long PCR primers specific to *S. destruens*. Primers had a length of 26-30 nt, a GC content of 40 %-60 % and were checked for primer dimer and hairpin formation using OligoAnalyzer 3.1 (https://www.idtdna.com/calc/analyzer). Various combinations of long-PCR primers were tested as it was uncertain which primers face each other on the genome and if they were separated by an appropriate distance.

6.2.3.1 Long-PCR based genome sequencing protocol

The mitochondrial fragments spanning the cob-cox1 and cox1-nad5 were amplified by using two long PCR kits; Long range PCR kit (Thermofisher) and LA PCR kit (TAKARA, Clontech). The successful primer combinations for each fragment, PCR cycling conditions and their output are presented in Table 6.2.

6.2.3.2 Step-out long PCR

The peripheral regions of the mitochondrial genome were amplified with an alternative approach "modified step-out approach" (Burger et al. 2007) The step-out primers (Table 6.1) were coupled with species-specific primers LR-cob-R2 and LR-nad5-R4 to amplify the peripheral regions. The successful step-out primers that produced an amplified product with species-specific primer are listed in Table 6.4. The PCR cycling conditions were as follows: 94 °C-1 min, 1 x (94 °C for 20 s; 30 °C for 2 min; 68 °C for 8 min), Pause to add specie-specific primers, 16 x (94 °C for 20 s; 65 °C (decrement= 0.3 °C per cycle) for 20 s; 68 °C for 8 min), 19 x [94 °C for 20 s, 60 °C for 20 s, 68 °C for 8 min (increment= 15 s per cycle)] ,68 °C 12 min.

6.2.3.3 Gel electrophoresis and sequencing

The amplified products were run on 1 % agarose gel stained with SYBR safe DNA gel stain at 80 V for 1 hour. The products were sent off for direct sequencing for small fragments up to 1,500 bp. The long fragments which were 12,986 bp and 7,048 bp in length were sequenced by primer walk service (Beckman coulter genomics).

Primer walk is a sequencing technique to sequence long DNA fragments which cannot be sequenced in a single sequence read using the chain termination method. The set of primer pairs were initially used to amplify the PCR product, and were also used to sequence approximately 1,000 bp from the two extremities of the amplified fragment. This generated sequence was then used to design new set of primers pair (20 bp in length) and sequenced further into the amplified fragment. The process was continued until the sequences met in the middle. The short sequence fragments generated were then assembled to generate a consensus sequence of the long amplified fragment.

6.2.4 Gene annotation and phylogenetic analysis

Gene annotation of mitochondrial genome of *S. destruens* was performed using the automated annotation tool MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl), followed by manual inspection. The 22 tRNA genes were further scanned and secondary structures were generated with the tRNA scan-SE program (Lowe and Eddy 1997). The predicted secondary structures of small (rns) and large (rnl) subunits ribosomal RNA of *S. destruens* mtDNA were generated in the RNA fold program of Vienna RNA package (Gruber et al. 2008).

Table 6.1. Universal primers for animal and Cnidarians mitochondrial DNA used in this study. The primer sequences and their output for *Sphaerothecum destruens* are listed.

mt DNA genes and respective primers (5′→3′)		Amplification	Sequence	Species DNA amplified	Reference
cox1	HCO: TAAACTTCAGGGTGACCAAAAAATCA LCO: GGTCAACAAATCATAAAGATATTGG	✓	✓ (410 bp)	*S. destruens*	(Folmer et al. 1994)
cob	CobF424: GGWTAYGTWYTWCCWTGRGGWCARAT CobR876: GCRTAWGCRAAWARRAARTAYCAYTCWGG	✓	✓ (440 bp)	*S. destruens*	(Boore and Brown 2000)
rnl	16Sar: CGCCTGTTTATCAAAAACAT 16Sbr: CCGGTCTGAACTCAGATCACGT	✓	✓	Fish	(Palumbi 1996)
rns	12Sai: AAACTAGGATTAGATACCCTATTAT 12Sb: GAGGGTGACGGGCGGTGTGT	✓	✓	Fish	(Kocher et al. 1989)
cox2	Cox2F1: AAGCWAATWGGNCATCARTGRTATTG Cox2R1: CTCCRCATATTTCNGARCATTGNCC	✓	✓	Fish	(Burger et al. 2007)
cox3	Cox3F: TGGTGGCGAGATGTKKTNCGNGA Cox3R: ACWACGTCKACGAAGTGTCARTATCA	Multiple bands obtained	-	-	(Burger et al. 2007)
nad4	Nad4F: CCKAARGCYCAYGTKGARGCYCC Nad4R: GARGAWCAKAWWCCRTGAGCAATYAT	-			(Shao et al. 2006)
nad5	Nad5F: TWYTATTAGGKTGAGATGGKYTNGG Nad5R: TARAAKCCWGMTARAAAWGGKAWWCC	✓	✓	Mixed read contaminated	(Lavrov et al. 2004)
rnl	16S 1471: CCTGTTTANCAAAAACAT 16S1472: AGATAGAAACCAACCTGG	✓	✓	Fish	(Schubart et al. 1998)
Universal mt DNA primers for Cnidarians					
cox2	diplo-cox2-f1: AAGCWATWGGRCATCARTGRTATTG diplo-cox2-r1: CWATWGGCATAAANGARTGATTNGC				(Lavrov et al. 2008)
rnl	diplo-rnl-f1: TCGACTGTTTACCAAAAACATAGC diplo-rnl-r1: AATTCAACATCGAGGTSGGAAAC	✓	✓	Fish	(Lavrov et al. 2008)

nad1	spong-nad1-r1: AATGGTRCTCKATTNGTTTCNGC sponge-nad1-f1: CTATTMGRGCAGCRGCHCAAATG	Multiple bands which faint at after 52°C			(Lavrov et al. 2008)
nad2	sponge-nad2-f1: TGRGCNCCAGATGTNTATGADGG sponge-nad2-r1: TTACTTAAAAAYCCNGCTAARGG	Very faint multiple bands			(Lavrov et al. 2008)
nad4	diplo-nad4-f1: TATTTGARGGNRTATTRATHCCANTG diplo-nad4-r1: CCATRTGNGCCACHGAAGAATARGC	-			(Lavrov et al. 2008)
nad5	sponge-nad5-f1: TGGGAGGGWGTWGGNTTATGTTC sponge-nad5-r1: ACTGGTGTNGGNCCTTCCATWGC	✓	✓ (320 bp)	*S. destruens*	(Lavrov et al. 2008)

Table 6.2. Step-out primers employed in "modified step-out approach" to amplify the peripheral regions of the *Sphaerothecum destruens* mitochondrial genome.

Primer	**Sequence (5′→3′)**
Step-out 1	GTCAGTCAGANNNNAGA
Step-out 2	TCAGGAACGATCGTNNNNTCA
Step-out 3	AACAAGCCCACCAAAATTTNNNATA
Step-out 4	TTGTTCGGGTGGTTTTAAANNNTAT

Table 6.3. *Sphaerothecum destruens*-specific long range Polymerase Chain Reaction (PCR) primers, PCR conditions and the size of their amplified fragment.

Gene fragment	**Primer combination (5′→3′)**	**Cycling conditions**	**Amplified product (bp)**
cob-cox1	**LR-COB-F** ATG AGG AGG GTT TAG TGT GGA TAA TGC **LR-COX1-R** GCT CCA GCC AAC AGG TAA GGA TAA TAA C	94 °C-2 min, 10 x (94 °C for 20 s, 58 °C for 30 s, 68 °C for 7 min), 25 x (94 °C for 20 s, 58 °C for 30 s, 68 °C for 7 min (increment 5 s/cycle) 68 °C for 10 min	1,200
cox1-nad5	**LR-COX1-R3** GTT ATT ATC CTT ACC TGT GTT GGC TGG AGC **LR-NAD5-R1** CCA TTG CAT CTG GCA ATC AGG TAT GC	94 °C-1 min, 16 x (94 °C for 20 s, 60 °C for 20 s, 68 °C for 8 min) 19 x (94 °C for 20 s, 60 °C, for 20 s, 68 °C for 8 min) 68 °C for 12 min	12,986

Table 6.4. Modified Step-out approach for the amplification of peripheral regions of *Sphaerothecum destruens* mitochondrial DNA.

Gene fragment	**Primer combination (5′→3′)**	**Amplified product (bp)**
cob-ccmF	LR-COB-R2 TCA ACA TGC CCT AAC ATA TTC GGA AC Step-out3: AAC AAG CCC ACC AAA ATT TNN NAT A	7,048
nad5-trnL	LR-nad5-R4: TGG GGC AAG ATC CTC ATT TGT Step-out 3 AAC AAG CCC ACC AAA ATT TNN NAT A	3,128

6.2.5 Phylogenetic analysis

Due to *S. destruens* being an earliest off-shoot from Holozoa, a higher species representation was chosen of basal organisms from Holozoa (animals and their unicellular relatives) and Holomycota (Fungi and their unicellular relatives). For the phylogenetic analysis, 23 species were used in order to have representatives from Proteobacteria, Jakobid, Fungi, Nucleariidae, Stramenopiles, Rhodophyta, Choanoflagellate, Filasterea, Ichthyosporea and Metazoa (Table 6.5). A bacterial outgroup *Rickettsia* from α-Proteobacteria that represents the closest ancestor to mitochondria. The eight most conserved protein coding genes cox1-3, cob, nad3, nad4l, nad5 and atp6 were used in the alignment. The genes nad1, 2, 4 and 6 were excluded as they have not been completely sequenced for *A. parasiticum* and are not available online (Burger et al. 2003a).

Nucleotide sequences for each of eight protein coding genes of *S. destruens* were translated into amino acids using the Mold, Protozoan mitochondrial translation genetic code. The resulting amino acid sequences were then aligned for each gene using Clustal W with default options (Gap open cost: 15 and Gap extend cost: 6.66). Due to the wide taxonomic range of species, the Clustal W multiple alignment programs cannot guarantee uniform results due to sequence length variation, sequence divergence and rate variation among lineages. To avoid these problems, the most conserved sequence regions of 8 aligned genes were selected for phylogenetic analysis using the web-based Gblocks program with default options of "medium" stringent selection (Castresana 2000). The final dataset consisted of 1875 aligned amino acid positions from eight conserved proteins.

Phylogenetic analysis of the concatenated dataset of conserved amino acid sequences were performed using Bayesian inference approach with LG+I+G+F model for amino acid substitution. The best model for our data was calculated by ProtTest v 3.4.2 (Darriba et al. 2011). The Bayesian analysis was run in Mr Bayes (Ronquist et al. 2012) and posterior probabilities were obtained after 2,500,000 generations with a burn-in of 25 %.

Table 6.5. The species, taxonomy and GenBank accession numbers for the species used in the phylogenetic analysis.

Species	**Taxonomic group**	**GenBank accession no.**
Rickettsia prowazekii	Proteobacteria	NC_000963
Reclinomonas americana	Jakobida	AF007261
Sphaerothecum destruens	Ichthyosporea	To be deposited
Amoebidium parasiticum	Ichthyosporea	AF538043-45, AF538047-49, AF538051-52
Ministeria vibrans	Ministeria/Filasterea	KC573040
Capsaspora owczarzaki	Filasterea	KC573038
Monosiga brevicollis	Choanoflagellida	AF538053
Tethya actinia	Porifera	NC_006991
Oscarella carmela	Porifera	EF081250
Geodia neptuni	Porifera	NC_006990
Iphiteon panicea	Porifera	EF537576
Sympagella nux	Porifera	EF537577
Trichoplax adhaerens	Placozoa	NC_008151
Sarcophyton glaucum	Cnidaria	AF064823, AF063191
Metridium senile	Cnidaria	NC_000933
Porphyra purpurea	Rhodophyta	NC_002007
Chondrus crispus	Rhodophyta	NC_001677
Chrysodidymus synuroideus	Stramenopiles	NC_002174
Allomyces macrogynus	Fungi	NC_001715
Podospora anserina	Fungi	NC_001329
Schizophyllum commune	Fungi	NC_003049
Rhizopus oryzae	Fungi	NC_006836
Nuclearia simplex	Nucleariidae	NC_020369

6.3 Results

6.3.1 Amplification of *Sphaerothecum destruens* mitochondrial DNA

The three universal primers for cox1, cob and nad5 genes successfully amplified *S. destruens* mitochondrial gene fragments (referred as anchor regions). The sequenced

fragments of cox1, cob and nad5 gene were 410 bp, 440 bp and 320 bp in length respectively. These anchor sequences served for the design of long-PCR primers specific to *S. destruens*. The whole mitochondrion was amplified in four overlapping fragments spanning ccmF-cob (7,048 bp), cob-cox1 (1,200 bp), cox1-nad5 (12,986 bp) and nad5-trnL (3,127 bp) Table 6.3 and 6.4.

6.3.2 Gene content and organization

The mitochondrial genome of *S. destruens* was 23,939 bp in size, circular, with an overall A+T content of 71.2 %, starting with the ccmF gene. A list of gene order, gene length, and intergenic spacer regions of *S. destruens* mtDNA is given in Figure 6.1 Table 6.6. The nucleotide composition of the entire *S. destruens* mtDNA sequences is 40.8 % Thymine, 31 % Adenine, 19.7 % Guanine and 8.5 % Cytosine (detailed nucleotide composition is listed in Table 6.7). It consisted of a total of 47 genes including protein-coding genes (21), rRNA (2) and tRNA (22) and two unidentified Open Reading Frames (ORFs), with all genes encoded by the same strand in the same transcriptional orientation (Figure 6.1).

The standard proteins encoded by mitochondria include 13 energy pathway proteins, including subunits 6, 8 and 9 of ATP synthase (atp6, 8 and 9), three subunits of cytochrome oxidase (cox1-3), NADH dehydrogenase subunits 1-6 and 4L (nad1-6 and 4L), apocytochrome b (cob), small and large subunit rRNAs (rns and rnl). The *S. destruens* mtDNA included genes that are usually absent from standard animal and fungal mtDNAs, such as four ribosomal proteins (small subunit rps13 and 14; large subunit rpl2 and 16), tatC (twin-arginine translocase component C), ccmC and ccmF (cytochrome c maturation protein CcmC and heme lyase). The mitochondrial genome of *S. destruens* was intron-less and compact with a few intergenic regions, a maximum of 357 bp between tatC and nad2 and several neighbouring genes overlapping by 1-31 nucleotides (Table 6.6, Figure 6.1). Table 6.8 provides an overview of the *S. destruens* mitochondrial genes and their biological function.

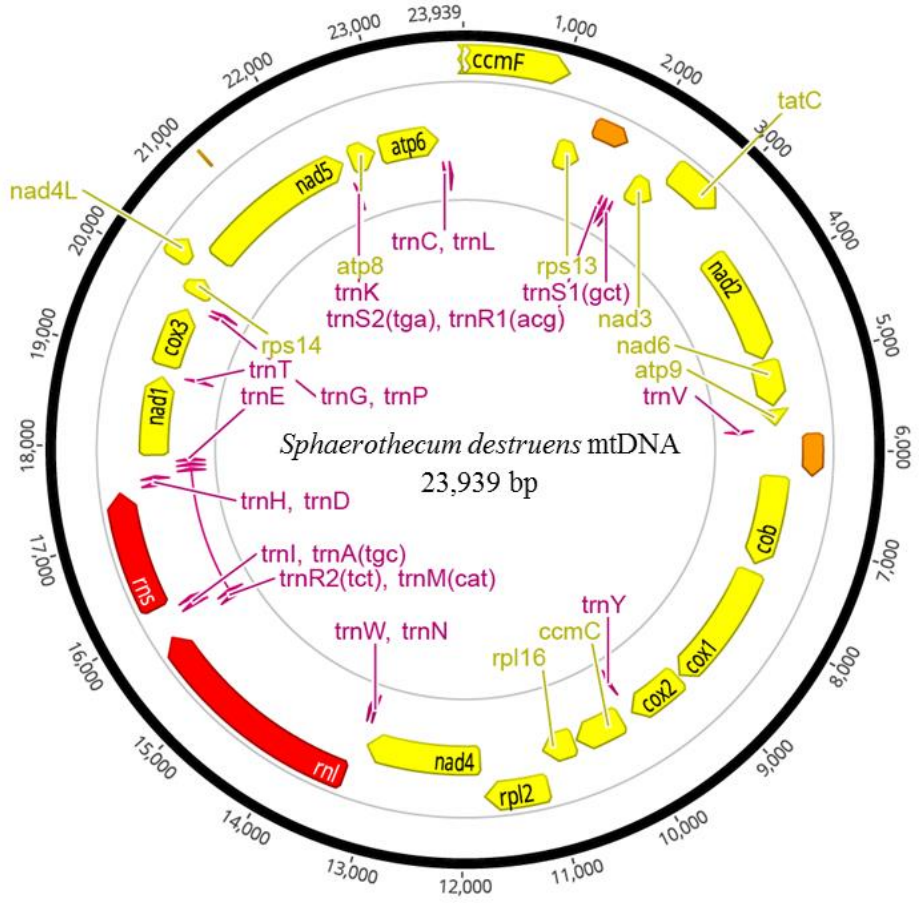


Figure 6.1. The complete mitochondrial genome for *Sphaerothecum destruens*. All genes are encoded in the same transcriptional orientation. 22 tRNA genes (pink), 2 rRNA genes (red), 19 protein coding genes (yellow) and 2 open reading frames (ORFs) (orange) are labelled. 22 transfer RNA genes are designated with single-letter amino acid code; A-alanine, C-cysteine, D-aspartic acid, E-glutamic acid, G-Glycine, H-histidine, I-isoleucine, K-Lysine, L-leucine, M-methionine, N-asparagine, P-proline, R-Arginine, S-Serine, T-threonine, V-valine, W-tryptophan and Y-tyrosine. Three methionine (M) and two serine (S) and arginine (R) tRNA genes are labelled along with their anticodon sequence.

Table 6.6. The mitochondrial genome organization of *Sphaerothecum destruens*.

Gene	Position		Size		Codons		Intergenic sequence (bp)
	Start	Finish	No. of nt	No. of aa	Initiation	Termination	
ccmF	1	1,080	1,080	359	GTG	TAG	55
rps13	1,136	1,459	324	107	GTG	TAA	3
orf144	1,463	1,897	435				4
trnS2	1,902	1,974	73	-	-	-	1
trnR1	1,976	2,046	71	-	-	-	0
trnS1	2,047	2,126	80	-	-	-	6
nad3	2,133	2,486	354	117	ATG	TAG	-31
tatC	2,456	3,115	660	219	GTG	TAG	357
nad2	3,473	4,909	1,437	478	ATG	TAG	0
nad6	4,910	5,500	591	196	GTG	TAA	13
atp9	5,514	5,738	225	74	ATG	TAA	7
trnV	5,746	5,817	72	-	-	-	3
orf167	5,821	6,324	504				-1
cob	6,324	7,466	1,143	380	ATG	TAG	60
cox1	7,527	9,119	1,593	530	ATG	TAA	1
cox2	9,121	9,870	750	249	ATG	TTA	-1
trnY	9,870	9,944	75	-	-	-	45
ccmC	9,990	10,622	633	210	ATG	TAA	4
rpl16	10,627	11,067	441	146	ATG	TAG	-11
rpl2	11,057	11,806	750	249	TTG	TAA	-1
nad4	11,806	13,236	1,431	476	ATG	TAG	0
trnW	13,237	13,308	72	-	-	-	2
trnN	13,311	13,382	72	-	-	-	-46
rnl	13,337	15,828	2,317	-	-	-	-4
trnR2	15,825	15,897	73	-	-	-	1
trnM3	15,899	15,969	71	-	-	-	28
trnL	15,998	16,069	72	-	-	-	1
trnA	16,071	16,142	72	-	-	-	25
rns	16,168	17,536	1,222	-	-	-	-4
trnH	17,533	17,606	74	-	-	-	0
trnD	17,607	17,679	73	-	-	-	3
trnM2	17,683	17,754	71	-	-	-	0
trnM	17,754	17,824	71	-	-	-	1
trnE	17,826	17,898	73	-	-	-	6
nad1	17,905	18,912	1,008	335	TTG	TAG	3
trnT	18,916	18,987	72	-	-	-	22
cox3	19,010	19,801	792	264	ATG	TAA	2
trnG	19,804	19,877	74	-	-	-	7
trnP	19,885	19,956	72	-	-	-	1
rps14	19,958	20,200	243	80	ATG	TAA	-7
nad4L	20,194	20,493	300	99	ATG	TAA	0

nad5	20,494	22,458	1,965	654	GTG	TAG	-1
trnK	22,458	22,530	73	-	-	-	1
atp8	22,532	22,867	336	111	ATG	TAA	45
atp6	22,913	23,659	747	248	ATG	TAA	6
trnC	23,666	23,738	73	-	-	-	12
trnL	23,751	23,822	72	-	-	-	117

*****Stop codon not included in AA sequence, nt = nucleotides, aa = amino acid.

Table 6.7. Nucleotide composition of mitochondrial genome of *Sphaerothecum destruens*.

Nucleotide	**Length (bp)**	**A (%)**	**C (%)**	**T (%)**	**G (%)**	**A+T (%)**	**G+C (%)**
Entire sequence	23,939	31	8.5	40.8	19.7	71.8	28.2
Protein coding sequences	17,691	28.8	8.0	43.2	20	72	28
rRNA genes sequences	3,539	37.9	9.9	33.2	19.0	71.1	28.9
Transfer RNA gene sequences	1,601	33.4	11.3	36.2	19.1	69.5	30.5
Non-coding regions	964	38.3	7.3	36.2	18.2	74.5	25.5
NCR 1	357	35.9	11.7	30.8	21.6	66.7	33.3
NCR 2	117	33.3	8.5	35.1	23.1	68.4	31.6

6.3.3 Protein coding genes and codon usage

A total of 21 protein coding genes were identified in the *S. destruens* mt-genome. Fourteen of these genes (nad1-6, 4L, cox1-3, cob, atp6, 8, and 9) encode for proteins involved in respiration and oxidative phosphorylation. Four ribosomal genes encode for small and large ribosomal subunits (rps13, rps14, rpl2 and 16). The longest gene fragment encodes for nad5 which was comparable to the size of nad5 genes of *A. parasiticum*, *M. vibrans* and *C. owczarzaki*. Three unusual proteins to Opisthokonts were encoded in *S. destruens* mt-genome: tatC, ccmF and ccmC.

The tatC gene (also known as mttB and ymf16) is present in *M. brevicollis* (Choanoflagellate) and also reported in only one other animal mt genome that of *Oscarella carmela* (Homoscleromorph) (Burger et al. 2003a, Wang and Lavrov 2007). This protein, a component of twin-arginine translocase (Tat) pathway, is involved in the

transport of fully folded proteins and enzyme complexes across lipid membrane bilayers and is usually present in prokaryotes, chloroplasts and some mitochondria (Lee et al. 2006). The tatC gene in *S. destruens* is 660 bp in length and utilizes GTG as its initiation codon. The derived amino acid sequence of *S. destruens* TatC is most similar to *M. brevicollis* TatC- 21.4 % (Choanoflagellate) followed by *Reclinomonas americana*- 19.2 % (Jakobid) and *O. carmela*-15.9 % (Porifera).

The CcmF protein also known as yejR is involved in Heme c maturation (protein maturation) and CcmC (also known as yejU) plays role in heme delivery (protein import). The genes for both proteins have been reported in all Jakobids studied to date (Burger et al. 2013), only *R. americana* is compared here, *C. owczarzaki* and *Rickettsia prowazekii*. Among the reported organisms the ccmF in *S. destruens* was shortest (1,080 bp) compared to *C. owczarzaki* (2,547 bp), *R. prowazekii* (2,013 bp) and *R. americana* (1,914 bp). The derived amino acid sequence of *S. destruens* ccmF gene was 14.9 %, 16.1 % and 18.7 %, identical with those of *C. owczarzaki, R. prowazekii,* and *R. americana* respectively. The ccmC gene fragment was of comparable length with those of the aforementioned species. The inferred amino acid sequence of ccmC of *S. destruens* was 25.1 %, 26.9 % and 22.9 % identical to those of *C. owczarzaki, R. prowazekii* and *R. americana* respectively.

Among 21 protein coding gene, 14 genes (atp6, 8, 9, cob, cox1-3, nad2-4, nad4l, rps14, rpl16 and ccmC) were inferred to use ATG as initiation codon, 5 genes (nad5-6, ccmF, tatC and rps13) used GTG as a start codon and the remaining rpl2 was initiated with TTG. Eleven proteins were terminated with the stop codon TAA (atp6, 8, 9, cox1, cox3, nad6, ccmC, rps 13-14), nine genes used the stop codon TAG (nad1-5, cob, tatC, ccmF and rpl16) and TTA was used in the termination of cox2.

Table 6.8. Genes in *Sphaerothecum destruens* mitochondrial DNA and their function

Sr No.	Functions	Genes
1	**Electron transport and oxidative phosphorylation**	
	Complex I (NADH: ubiquinone oxidoreductase)	nad1-6, 4L
	Complex III (ubiquinone: cytochrome c oxidoreductase)	cob
	Complex IV (cytochrome c:02 reductase)	cox1-3
	Complex V (F1F0 ATP synthase)	atp6, 8, 9
2	**Translation**	
	Ribosomal RNAs	rnl, rns
	Ribosomal proteins	
	Small subunit (SSU)	rps13, 14
	Large subunit (LSU)	rpl2, 16
	Transfer RNAs	trnA, C-E, G-I, K-N, P, R1, R2, S1, S2, T, V, W, Y
3	**Protein import**	
	Heme delivery	ccmC
	Sec-independent transporter	tatC
4	**Protein maturation**	
	Heme c maturation	ccmF

6.3.4 Ribosomal RNA and transfer RNA genes

Genes for the small and large subunits for mitochondrial rRNAs (rns and rnl, respectively) were present. Both genes were separated by four tRNA genes (rnl-trnR2-trnM-trnI-trnA-rns). The rns and rnl (1,369 and 2,449 bp) had sizes approximately similar to those in *M. brevicollis* (1,596 and 2,878 bp) and *A. parasiticum* (1,385 and 3,053 bp). These sizes were comparable to their eubacterial homologs (1,542 and 2,904 bp in *Escherichia coli*). The predicted secondary structures of small (rns) and large (rnl) subunits ribosomal RNA of *S. destruens* mtDNA are presented in Figure 6.2.

Twenty-two tRNA genes, ranging in size from 71-80 bp, were identified in *S. destruens* mtDNA and their predicted secondary structures had a cloverleaf-like shape (Figure 6.3). Three copies of tRNA M (Methionine CAT) of the same length (71 bp) with overall 6-nt differences were identified. The tRNA M1 was at a distance of 1,713 bp from tRNA M2, whereas tRNA M2 and M3 were adjacent (Figure 6.1). Additionally, there were duplicated genes for tRNA S (Serine) and tRNA R (Arginine). The 2 copies, tRNA S1 (gct) and tRNA S2 (tga) were of 80 bp and 73 bp in length respectively and

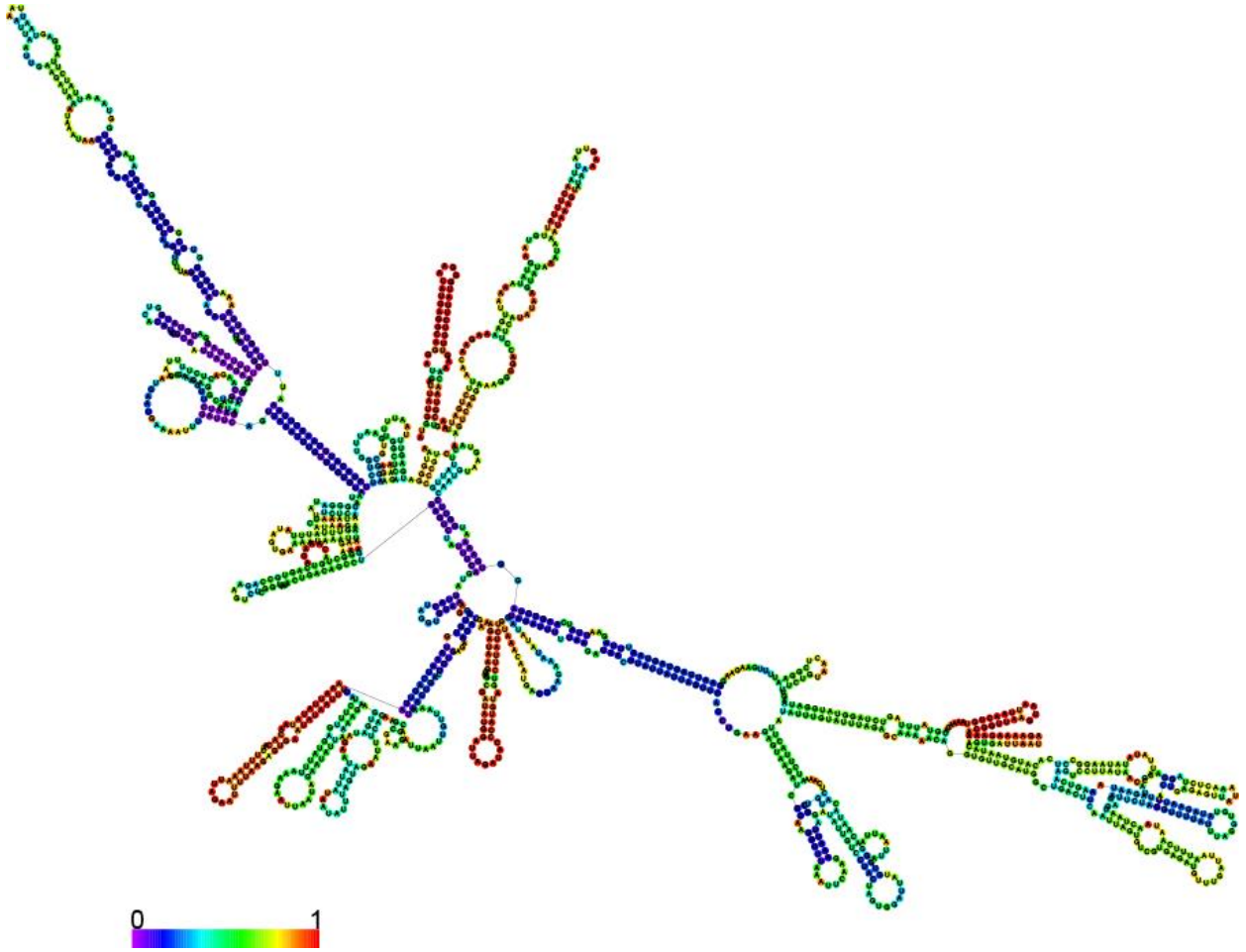
were 73 % identical. The tRNA R1 (acg) and tRNA R2 (tct) were approximately same length with 60 % similarity. The tRNA gene duplication, tRNAM, tRNAS and tRNAR, is also present in *A. parasiticum*, *M. brevicollis*, *C. owczarzaki* and *M. vibrans* except for only two copies instead of 3 for tRNA M in *M. vibrans*.

All the tRNAs secondary structures had a dihydrouridine (DHU) arm, a pseudouridin (TΨC) arm and an anticodon stem except for tRNA S1(gct) that had an additional short variable loop. The TΨC and D-loop is comprised of 7 and 7-10 nucleotides respectively. Similar to *M. brevicollis* and *A. parasiticum* none of the mitochondrial tRNA in *S. destruens* had a truncated D or T loop structure, a feature that is widespread in animal mitochondrial tRNAs.

6.3.5 Non-coding region

The total length of the non-coding region was 842 bp and was comprised of 32 intergenic sequences ranging in size from 1-357 bp. Only two intergenic regions had lengths greater than 100 bp: a) the non-coding region 1 NCR 1- 357 bp which was between the tatC and nad2 genes and b) the non-coding region 2 NCR 2- 117 bp which was found towards the periphery between the trnL and ccmF genes (Figure 6.1).

A



B

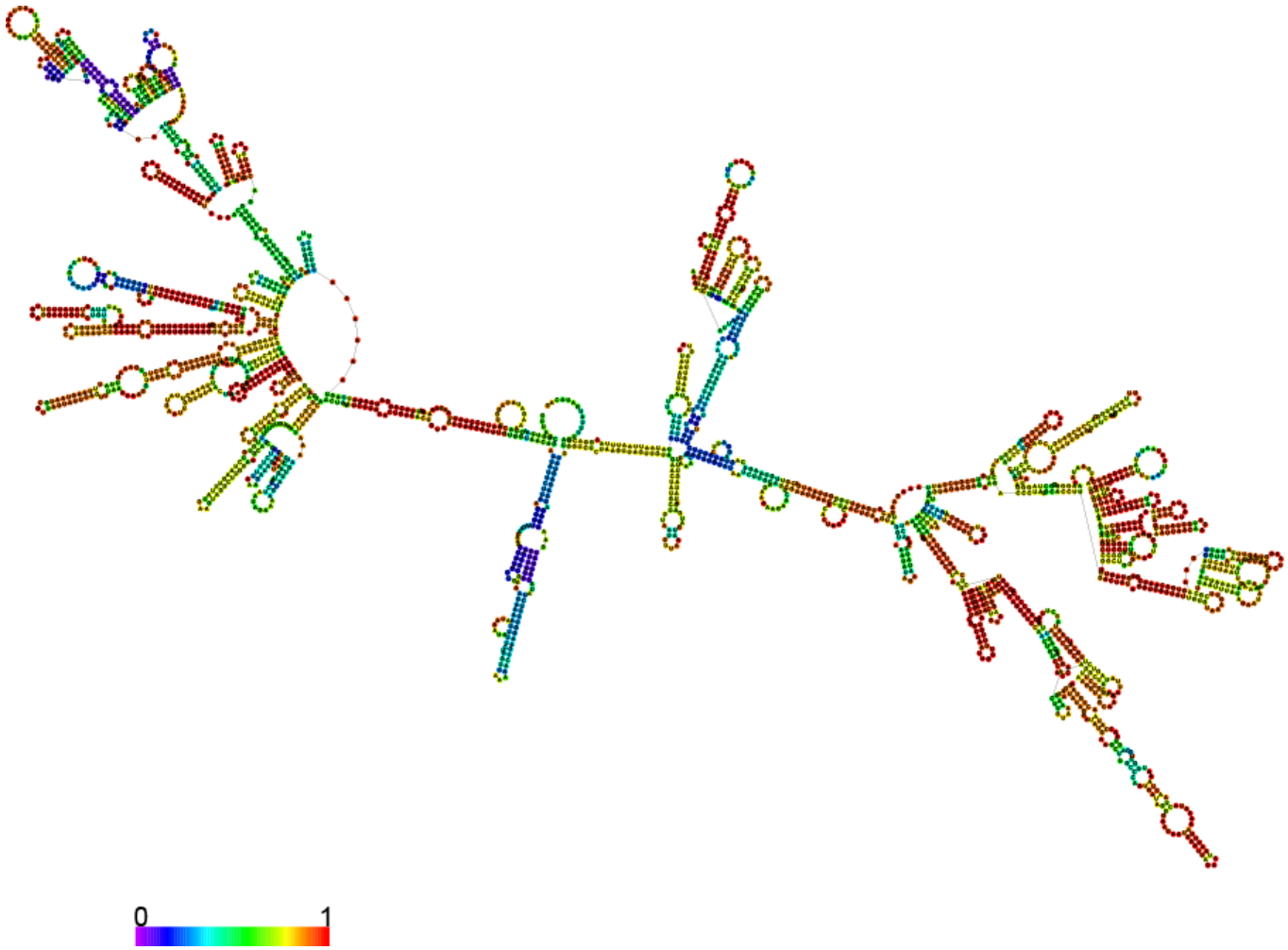
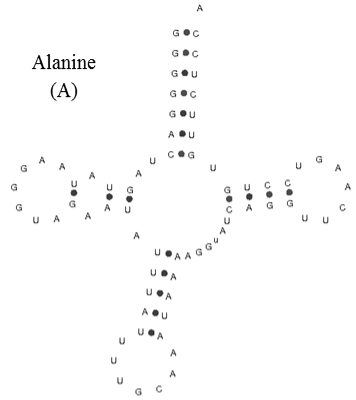
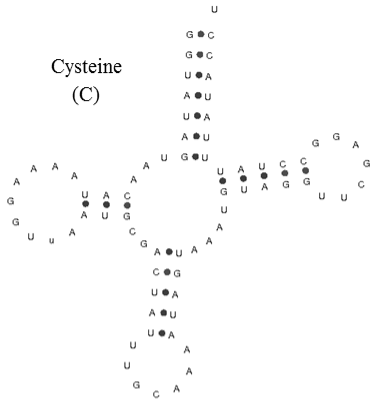
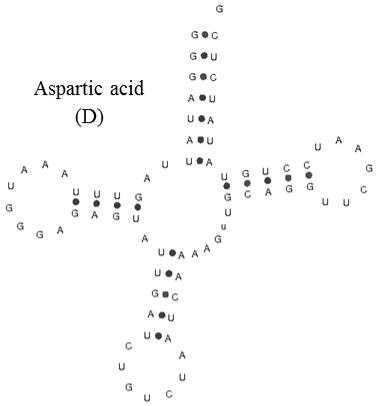
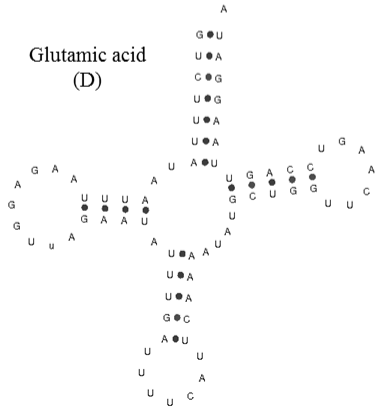


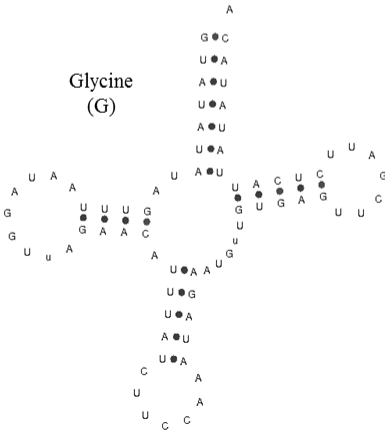
Figure 6.2. Predicted secondary structures of mitochondrial small and large subunit ribosomal RNA (a) rns and (b) rnl of *Sphaerothecum destruens*. The secondary structures were predicted using the online tool RNA fold in the Vienna RNA package (Gruber et al. 2008).

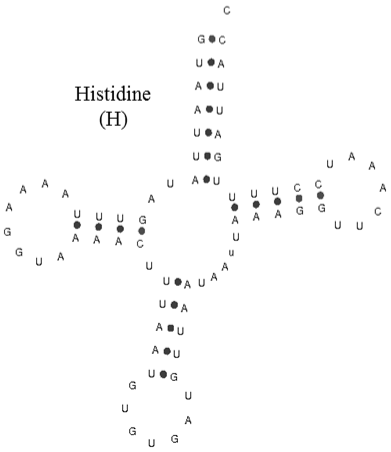

Alanine
(A)

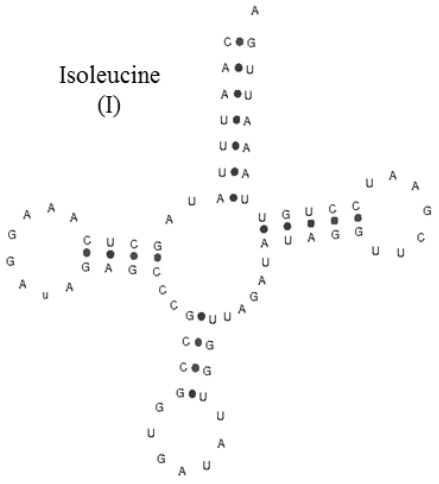

Cysteine
(C)

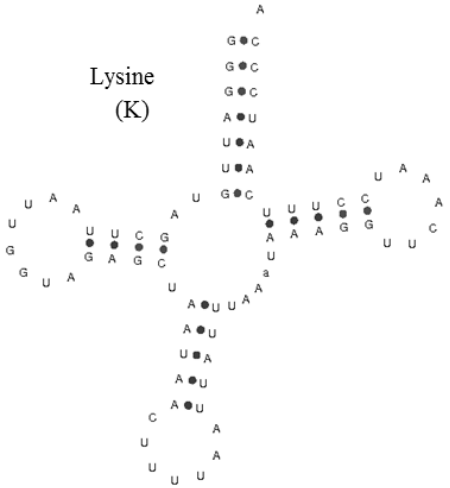

Aspartic acid
(D)

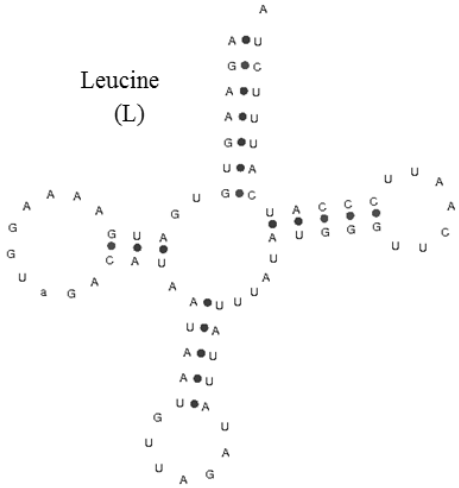

Glutamic acid
(D)

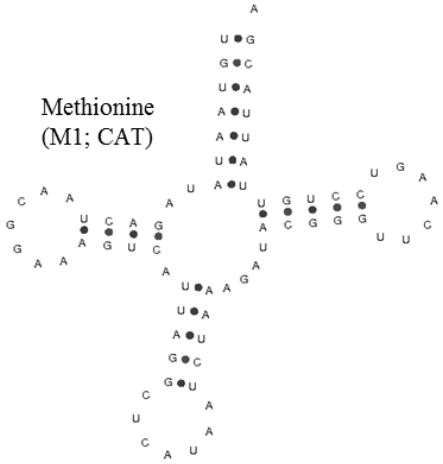

Glycine
(G)

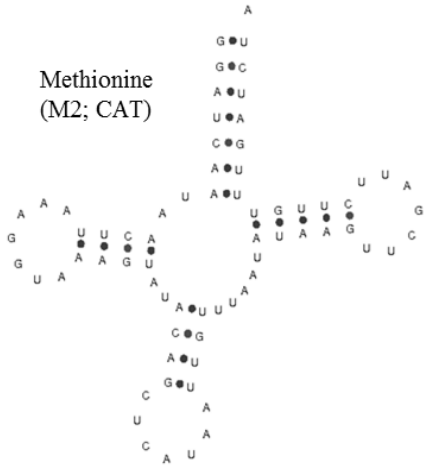

Histidine
(H)

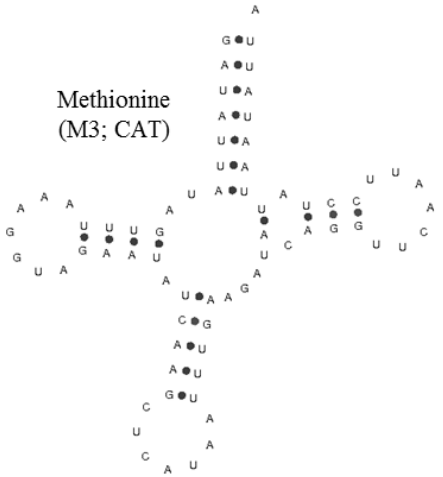

Isoleucine
(I)

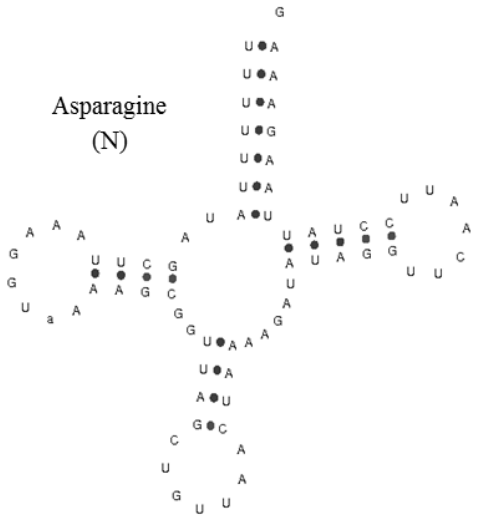

Lysine
(K)

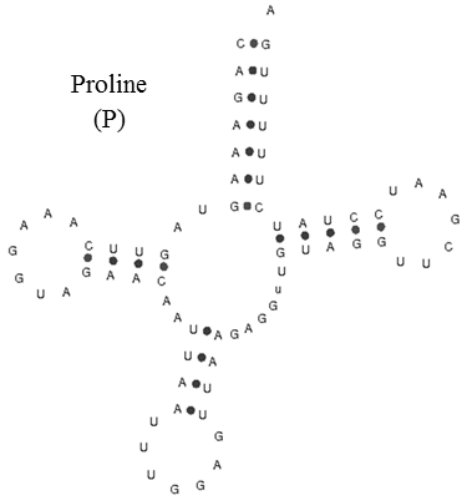

Leucine
(L)

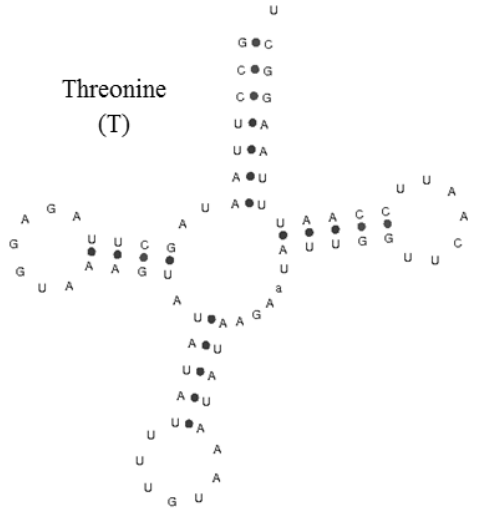

Methionine
(M1; CAT)

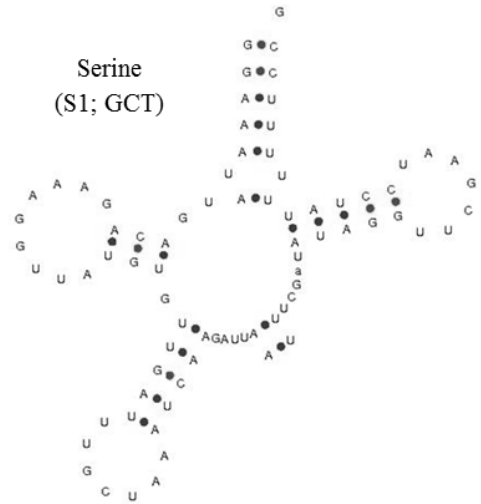

Methionine
(M2; CAT)

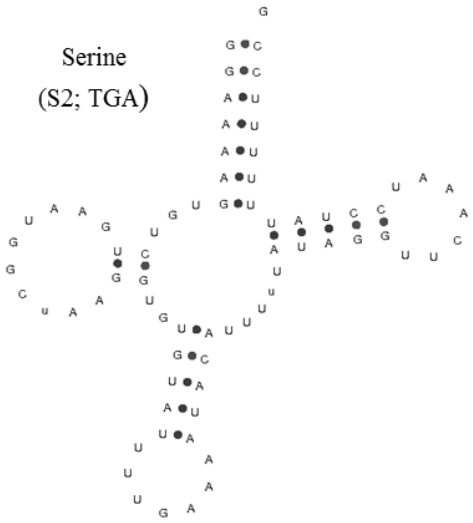

Methionine
(M3; CAT)

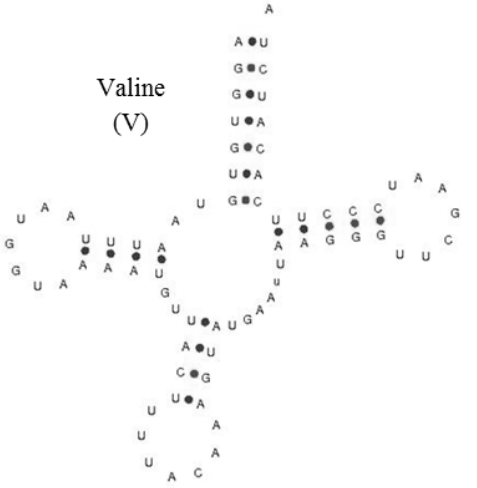

Asparagine
(N)


Proline
(P)


Threonine
(T)


Serine
(S1; GCT)


Serine
(S2; TGA)


Valine
(V)

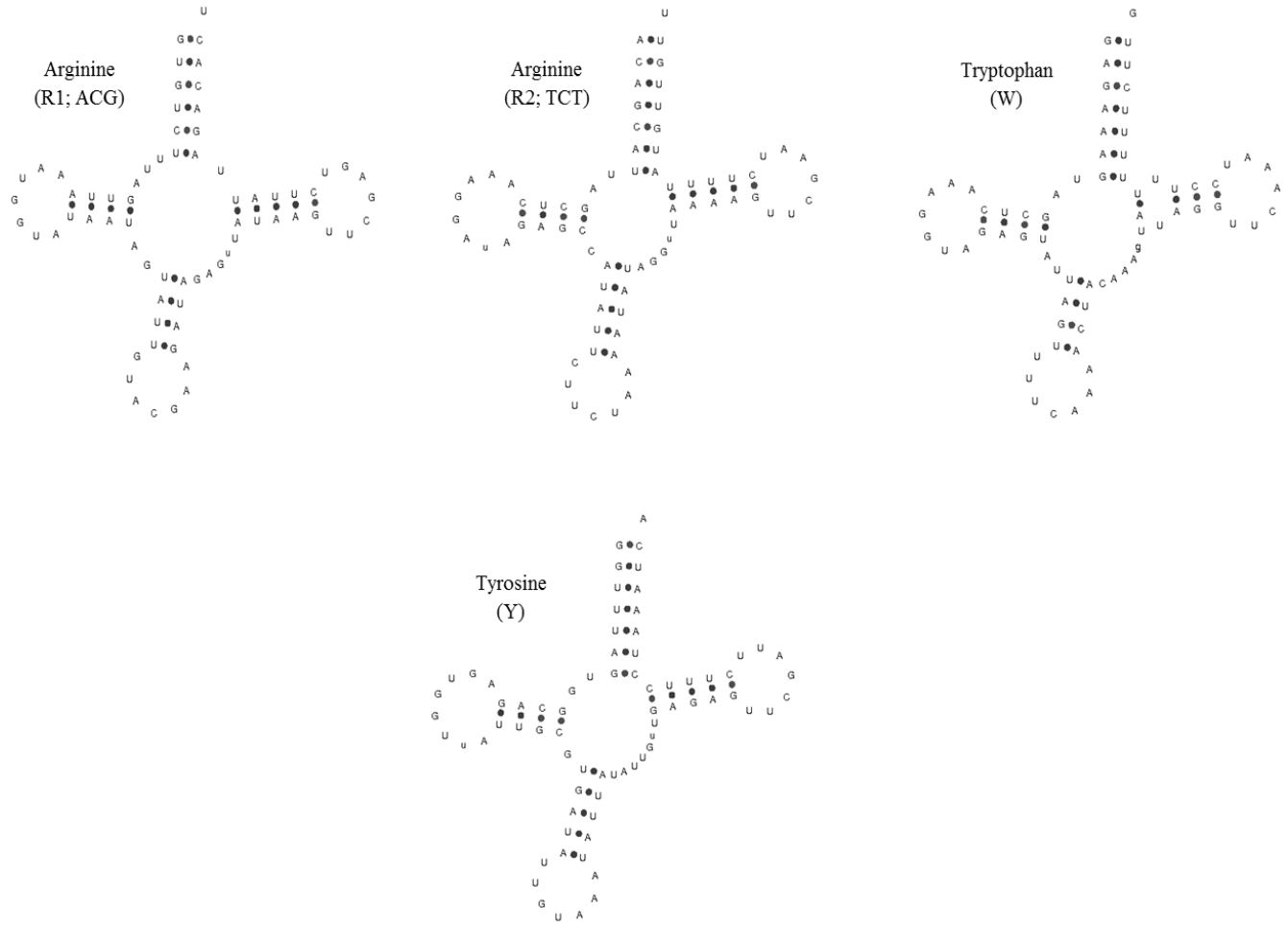


Figure 6.3. The predicted secondary structures of 22 tRNAs of *Sphaerothecum destruens* mitochondrial DNA generated in tRNAscan-SE (Lowe and Eddy 1997).

6.3.6 Phylogenetic position of *Sphaerothecum destruens* based on mitochondrial protein sequences

Phylogenetic analysis using the concatenated amino acid sequences from eight protein coding genes revealed a conventional tree for eukaryotic relationships especially for Holomycota and Metazoa with strong support for almost all inferred clades (Figure 6.4). The tree supports the sister group association of Choanoflagellate (*M. brevicollis*) with the Metazoa. In the tree *C. owczarzaki* and *M. vibrans* strongly grouped together within the Filasterea. Interestingly, *S. destruens* has come up as a sister group to the Choanoflagellate and the Metazoa group along with the Filasterea with strong bootstrap posterior probability (BPP; 1.00). *S. destruens* and *A. parasiticum* have not grouped together in the same class Ichthyosporea, with *A. parasiticum* being placed as an earliest Holozoan divergence with strong branch support (BPP; 1).

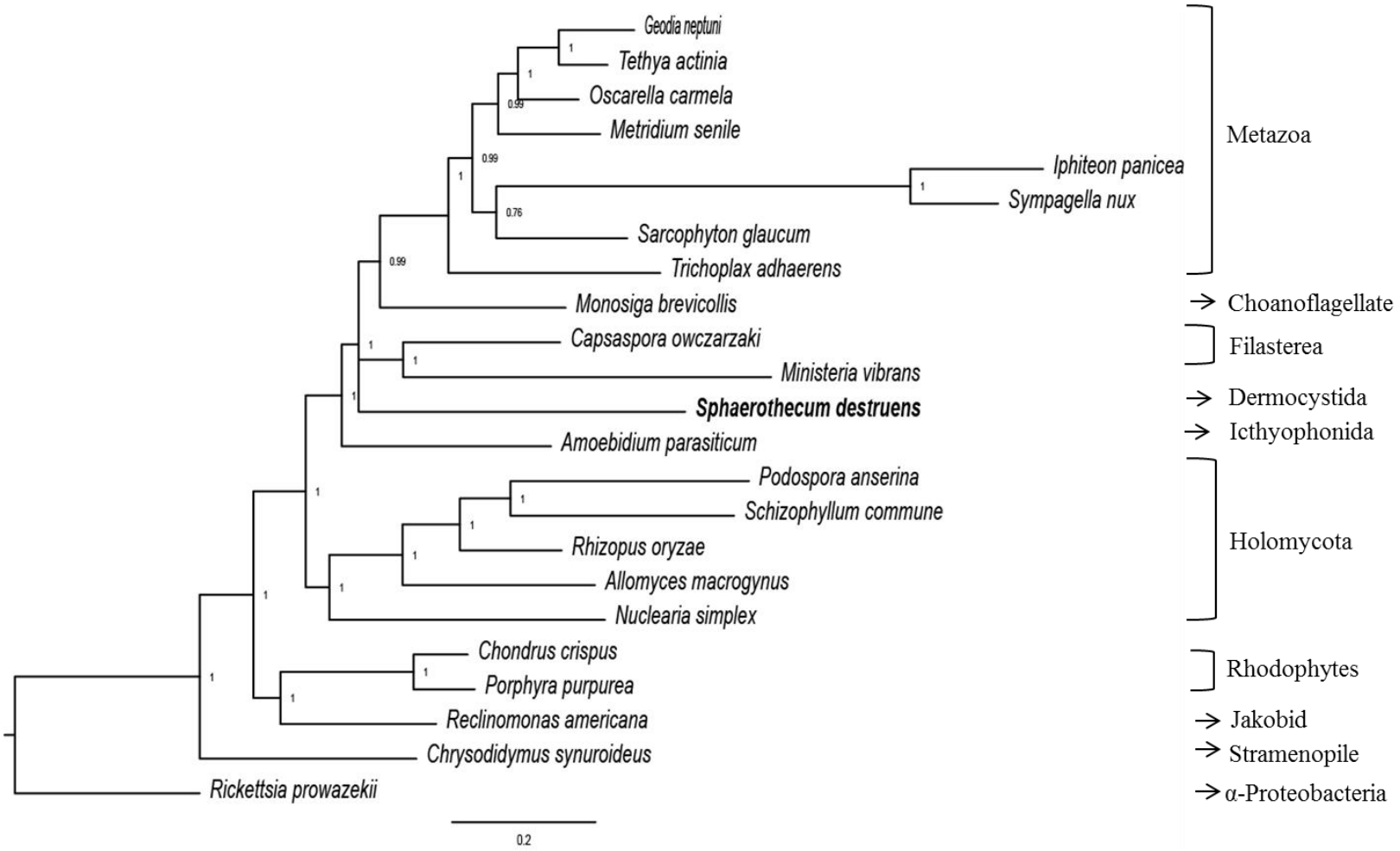


Figure 6.4. Phylogeny of *Sphaerothecum destruens* inferred from 1,875 aligned amino acid positions from 8 concatenated mitochondrial genes (cob, cox1-3, nad3, nad4l, nad5 and atp6) using LG+I+G model of protein evolution in Mr Bayes (Ronquist et al. 2012). The branch support values for each node are shown as Bayesian posterior probability (BPP). *Rickettsia prowazekii* was used as an out-group.

6.3.7 Comparison of mitochondrial genomes of two members of the Class Ichthyosporea

S. destruens and *A. parasiticum* are members of the order Dermocystida and Icthyophonida respectively within the Class Ichthyosporea. The mitochondrial genome of *A. parasiticum* was investigated by Burger et al. (2003a). *A. parasiticum* mt-genome is very peculiar compared to its close relatives, as it is larger than >200 kbp and consists of several hundred linear chromosomes. To date, 65 % of the mt-genome has been sequenced. Compared to *A. parasiticum*, *S. destruens* mt-genome is 8 times smaller (23,939 bp) with all the genes encoded by a single circular strand in the same transcriptional orientation.

There is a remarkable difference in the coding portion of the genomes of both species with only 20 % of the *A. parasiticum* mitochondrial genome coding for proteins whereas 93 % of the *S. destruens* mitochondrial genome codes for proteins. *A. parasiticum* genome is very spacious, gene-rich and is comprised of long regions of repeat sequences whereas in *S. destruens* mt-genome is very compact, with very little intergenic regions, no repeat sequences and even gene-overlap is evident. The *S. destruens* mt-genome is made up of 47 intron-less genes (including two ORFs) while the *A. parasiticum* genome is gene rich comprising of altogether 87 genes out of which 44 genes have been identified (includes approximately 24 ORFs) and is very intron-rich Burger et al. (2003a).

The above comparison illustrates that mitochondrial genome architecture and content is very different in both closely related species which could be a possible explanation as to why the two species did not group together in the phylogenetic tree (Figure 6.4). However, a few features shared by their mtDNA confirm their close phylogenetic relationship as evident from the phylogenetic tree constructed from their derived amino acid sequences (Figure 6.4). Both organism use mitochondrial UGA (“stop”) codon to specify tryptophan. Several tRNA genes are duplicated with a maximum of four copies as for tRNAM in *A. parasiticum* mt-genome. Unusual structural similarity in tRNAS of *A. parasiticum* and *M. brevicollis* was observed. Similar to that, in *S. destruens* nucleotide 8 that connects aminoacyl and D stems of tRNAS is missing, and in position

26 there is pyrimidine (Uracil) instead of purine. However, unlike *A. parasiticum* the second nucleotide in D-loop of *S. destruens* is Adenine (A) instead of Uracil.

6.4 Discussion

Due to unique taxonomical position of *S. destruens*, its mt-genome was investigated to better decipher its phylogeny and investigate if *S. destruens* shares same mt-genome features with other unicellular relatives of animals (Choanozoa) which will help in better understating of mtDNA evolution in this important group of eukaryotes. The overall coding portion (including two ORFs) of *S. destruens* mitochondrial genome is 96.4 % and represents the highest amongst the unicellular protists (*M. brevicollis* (46.9 %), *M. vibrans* (80.0 %), *C. owczarzaki* (28.6 %) and *A. parasiticum* (20 %) and even greater than the protozoan *Jakoba bahamiensis* (93.0 %) that has had the most primitive mitochondrial genome and highest gene content reported to date (Burger et al. 2003a, Burger et al. 2013, Lavrov and Lang 2014).

S. destruens mtDNA is considerably compact with overlapping genes of up to 31 nt (between nad3 and tatC), a feature (distinct compaction) that is usually ascribed to more evolved mtDNA (Metazoan) and to the emergence of a multicellular body plan (Lavrov 2007). The *S. destruens* mitochondrial genome is three times smaller than *M. brevicollis* (76 kbp), two times smaller than *M. vibrans* (55.9 kbp), and approximately eight times smaller than its closely related organism *A. parasiticum* (> 200 kbp) from the sister group Icthyophonida and *C. owczarzaki* (200 kbp). *Sphaerothecum destruens* had extensive gene loss especially for ribosomal proteins compared to Filasterea (*M. vibrans* and *C. owczarzaki*) and Choanoflagellate (*M. brevicollis*), with only four ribosomal genes left in its mitochondrial genome. *Amoebidium parasiticum* cannot be compared at this moment as its whole mitochondria have not been sequenced yet (Burger et al. 2003a). The number of tRNAs in *S. destruens* is reduced to 22 which is the lowest among the unicellular protists studied so far. The reduced number of 22 tRNAs as evident in bilaterian animals is usually associated with genetic code changes to utilize a fewer number of mitochondrial tRNAs compared to the translation under the standard genetic code (Marck and Grosjean 2002).

The remarkable differences in mitochondrial genome of *S. destruens* from its related unicellular protists could also be attributed to their different lifestyles. For example, parasite genes tend to evolve faster compared to their free-living relatives (Baldauf et al.

2013). The mt-genomes of organisms studied so far from the Choanozoa include *M. brevicollis* (Choanoflagellate); a unicellular free-living heterotroph, *A. parasiticum* (Icthyophonida); a symbiont present in external exoskeleton of insects, *C. owczarzaki* – an endosymbiont of tropical freshwater snail where it acts as a parasite of *Schistosoma mansoni* and *M. vibrans*- a free-living protist (Baldauf et al. 2013). Mt-genomes studies of parasites have shown development of great diversity of gene content, organization and expression machineries (Feagin 2000). The parasitic lifestyle *of S. destruens* could have contributed to the observed accelerated mtDNA evolution but accelerated growth rates have also been proposed as a possible explanation for highly derived mtDNAs (Burger et al. 2003b). *S. destruens* divides within cells asexually and can reach high numbers in less than 20 days in cell cultures within the lab. Therefore, these accelerated growth rates could also have contributed to the observed accelerated mtDNA evolution.

The phylogenetic analysis revealed conventional relationships for the Metazoa and the Holomycota, however it revealed new interrelationships of early branching Metazoans. The phylogeny of *S. destruens* based on mitochondrial data revealed the interesting grouping of *S. destruens* with the Filasterea and the Metazoa in one clade with a strong BPP value. However, this is in contradiction to a previous multi-gene study that has identified S. destruens as one of the earliest Holozoan divergence after *Corallochytrium limacisporum* (Torruella et al. 2015). This contradiction could be due to the different set of genes studied which can be under different evolutionary pressures and could result in more evolved genes for mitochondria compared to flagellar and chitin synthase genes (which were used by (Torruella et al. 2015). Moreover, the highly derived mitochondrial genome content and the architecture of *S. destruens* mtDNA compared to *A. parasiticum* could be a possible explanation for the non-grouping of both organisms into a monophyletic group and the association of *S. destruens* with the Filasterea, the Choanoflagellate and the Metazoa group. It is also noteworthy to point out that most of the multi-gene phylogenetic studies (except for Torruella et al. 2015), did not include *S. destruens* as a representative of the Ichthyosporea (Lang et al. 2002, Shalchian-Tabrizi et al. 2008, Torruella et al. 2012).

Based on the mt-genome expansion trends in the unicellular relatives of animals (*M. brevicollis* and *A. parasiticum*), due to the accumulation of repeat sequences and

increased size of intergenic regions, it was proposed that mtDNA reorganization occurred recently in the Metazoan evolutionary history and was linked with the emergence of multicellular body plans (Burger et al. 2003a, Signorovitch et al. 2007). *Sphaerothecum destruens* mtDNA in comparison with its unicellular relatives *A. parasiticum, M. vibrans, C. owczarzaki* and *M. brevicollis* is very compact and without any repeat sequences or large non-coding regions. Although a few unusual genes (ccmC, ccmF and tatC) are present, it has a comparatively low gene-content; in particular, there is evidence of loss of larger number of ribosomal genes compared to other member of the Choanozoa. The *S. destruens* mtDNA indicates that the distinctive mtDNA compaction has already started in the unicellular organisms as opposed to the findings in *M. brevicollis* and *A. parasiticum* (Burger et al. 2003a). This highlights the need to increase the quantity of mt-genome of unicellular organisms sequenced, particularly those of the closest relatives of animals (Holozoa) to better understand the evolution of mitochondrial DNA in this group of eukaryotes.

In conclusion, the study of *S. destruens* mtDNA has challenged the conventional assumption about the presence of big mt-genomes in the unicellular relatives of animals (Choanozoa). The absence of long intergenic regions and even a few genes overlaps indicates that the mtDNA compaction feature is not strongly linked with the multicellularity of animal lineage as generally assumed. The presence of comparatively evolved mtDNA in one of the earliest Holozoan organism supports the hypothesis that evolution of mtDNA occurs in parallel trajectories.

6.5 Summary

The present study has provided novel data on mtDNA evolution by providing first evidence that distinctive mtDNA compaction has already occurred before the emergence of a multicellular body plan in the animal lineage as previously hypothesized. The presence of a compact mt-genome in the Ichthyosporea supports the parallel evolution hypothesis for mtDNA. The compact, intron-less and reduced gene-content of the *S. destruens* mt-genome is very unique for the species phylogenetic position. Based on mitochondrial data, *S. destruens* has a sister group relationship to the Filasterea, the Choanoflagellate and the Metazoa group.

Chapter 7

Discussion and Conclusion

7.1 Synthesis of principal results

Sphaerothecum destruens was identified as a potential novel parasite in 2005 after its discovery in the invasive fish species *Pseudorasbora parva* (Gozlan et al. 2005). Since its first record in Europe, the parasite's presence has been confirmed in *P. parva* populations from the Netherlands, Turkey and France (Spikmans et al. 2013, Ercan et al. 2015, Charrier et al. 2016), with declines in endemic freshwater species in the later (Ercan et al. 2015). Due to the management implications associated with this parasite's status (i.e. native or non-native parasite) this work aimed at determining *S. destruens* origin and distribution across native and non-native *P. parva* populations (Chapter 3); investigating its distribution and potential impact in the UK (Chapter 4) whilst also developing eDNA detection methods in order to assess the efficacy of *P. parva* eradication as a viable control measure for *S. destruens* (Chapter 5). In order to achieve the aims in Chapters 3 and 4 a new phylogenetic marker, the mitochondrial Cyt-b, was developed and a nuclear ITS marker was optimized (Chapter 2). The global distribution of *S. destruens* (Chapter 3) was determined through the examination of its reservoir host *P. parva*. The work expanded the confirmed range of *S. destruens* to more locations in Europe and is the first study to determine its presence in China (Figure 7.1). This study provided the first evidence to support the novel pathogen hypothesis i.e. that *S. destruens* had been introduced into Europe via the accidental introduction of its reservoir host *P. parva*. Therefore, *S. destruens* is a non-native parasite to the UK and continental Europe and should be managed as such.

In addition to this parasite potentially being an important pathogen for fishes, it also has a unique taxonomical position, where animals first diverged from fungi in the tree of life. Most studies which look at the deep roots of the tree are often limited by the inability to culture these organisms and thus do not have good quality DNA to investigate these relationships using genetic markers. *Sphaerothecum destruens* is the only member of the Dermocystida that has been successfully cultured in the lab and thus this represented a strong opportunity to both understand its mitochondrial evolution and better decipher its taxonomic position in the tree of life. In Chapter 6, the complete mitochondrial genome for *S. destruens* was sequenced and used to reconstruct phylogenetic tree based on mitochondrial derived protein sequences which revealed an interesting position for *S. destruens* as a sister group to the Filasterea, the

Choanoflagellate and the Metazoa group (which is contrast to other studies - (Torruella et al. 2015). The sequencing of the mitochondrial DNA has however yielded novel results in terms of the species' mitochondrial organisation which showed extreme compaction and relatively low gene content of it mitochondrial DNA compared to its only sequenced close relative, *Amoebidium parasiticum.*

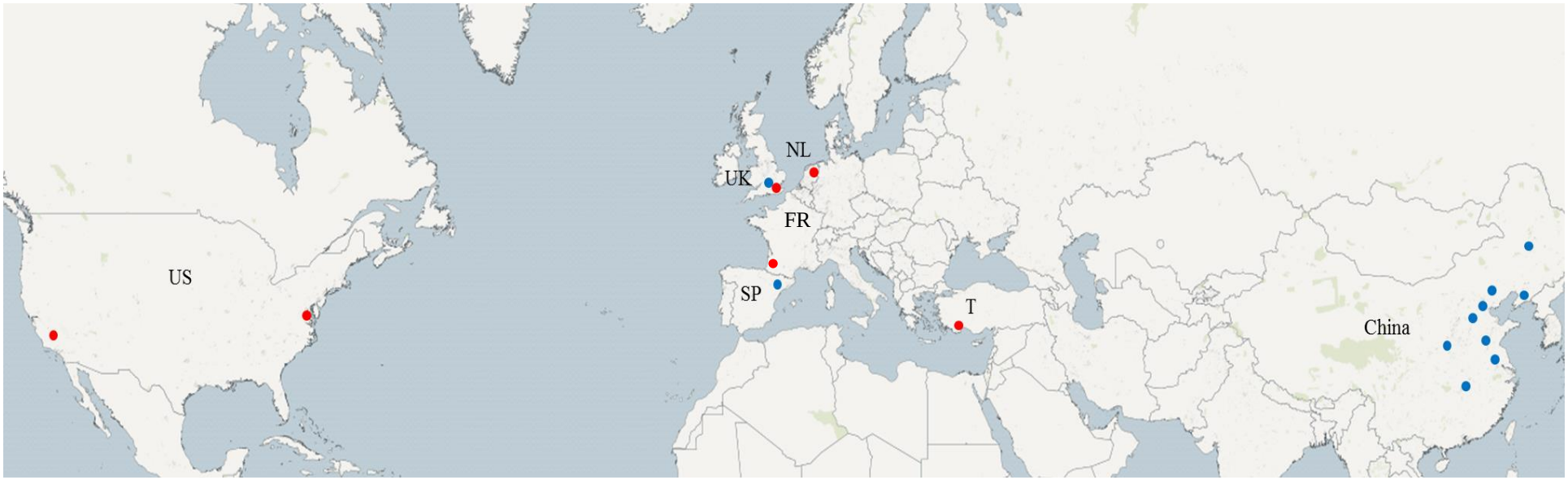


Figure 7.1. Global distribution of *Sphaerothecum destruens*. The sites labelled red are from previous reported studies (Harrell et al. 1986, Hedrick et al. 1989, Arkush et al. 1998, Spikmans et al. 2013, Ercan et al. 2015, Charrier et al. 2016) and blue are this thesis's (Chapter 3) contribution. The sites detected positive for *S. destruens* in its reservoir host Pseudorasbora parva in this study are multiple sites across China, from Spain and the UK. (Abbreviations: US-United states of America, UK-United Kingdom, SP-Spain, NL-the Netherlands, FR- France and T-Turkey).

In the UK, the risk of all non-native parasites is assessed and their inclusion on the Environment Agency's Category 2 and the non-native parasite list is determined and reviewed by a panel of experts using a number of criteria including the evidence for histopathological impacts on the infected fish. It was thus important to investigate the distribution and potential threat of *S. destruens* through histopathology in the UK (Chapter 4). The epidemiological map (Figure 4.1) developed for *S. destruens'* prevalence in the UK indicated that *S. destruens* was present in 50 % of the sampled sites. Histopathology of native fishes which were detected positive for *S. destruens* through molecular analysis (i.e. PCR) revealed that chub, dace, roach and brown trout displayed no signs of *S. destruens* related histopathology (Chapter 4). However, the study confirmed the increased range of *S. destruens* including new potential hosts in the wild. These fish hosts with low *S. destruens* prevalence can become reservoirs of infection themselves and can result in the emergence of the disease.

The phylogeographical analysis using the mt Cyt-b indicated the presence of two unique *S. destruens* haplotypes in the UK; one of which was similar to Chinese haplotype (Chapter 3) and the other was unique to UK. The presence of a unique *S. destruens* haplotype in UK could indicate that *P. parva* sampling in China was not exhaustive and that not all possible sources of *P. parva* introduction into Europe were captured. This is also confirmed by the presence of *P. parva* Cyt-b haplotype that is unique to Europe and has not been captured in the sampled Chinese native populations. Despite the lack of histopathology, it is highly recommended that *S. destruens* should be considered for listing on the Environment Agency's Category 2 and novel parasites list due to its association with chronic host mortalities (Gozlan et al. 2005, Andreou et al. 2011, Ercan et al. 2015).

The eDNA detection method developed in Chapter 5 could be used to inform health checks if the parasite is listed on the Environment Agency's Category 2 non-native parasite list. The technique has been successfully used to confirm the presence of *S. destruens'* eDNA in the decommissioned ornamental fish farm and Tadburn lake stream which is a tributary of River Test and has confirmed that despite the eradication of reservoir host *P. parva*, it is practically impossible to eradicate the pathogen *S. destruens* once it has established in the community (Al-Shorbaji et al. 2016). The

Tadburn Lake stream which runs into River Test, increases the risk of *S. destruens* expansion. This raises the need of effective measures to be taken to minimize *S. destruens* further spread. These relevant measures include its monitoring in suspected waters and its detection during fish health checks undertaken before the fish movements are carried out (Figure 7.2).

In England and NRW, fish movements and introductions are regulated by the Environment Agency. During fish health checks, the Environment Agency requires a minimum of 30 fish to be examined. This sample size is considered sufficient for the detection of most of the pathogens. The sampled fishes are dissected and are examined for Category 2 non-native parasites. The fish movements are carried out depending on the absence/ presence of the non-native parasites. In order to monitor *S. destruens* presence it is worthwhile to sample water from the inspected site and check for *S. destruens* through the eDNA method. Positive water samples can be further confirmed through the detection of the parasite in fish tissues (Figure 7.2).

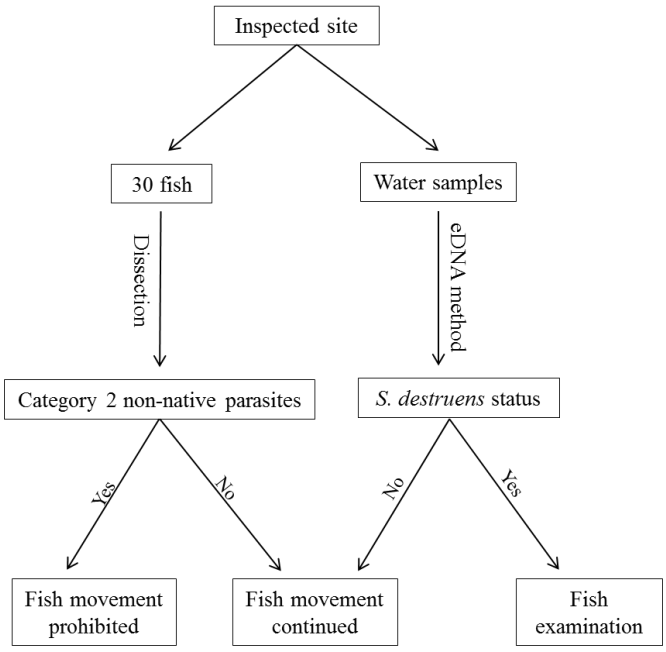


Figure 7.2. A proposed lay-out for the inclusion of *Sphaerothecum destruens* screening during fish health checks.

The mitochondrial DNA study of *S. destruens* revealed unexpected features of its mt-genome when compared to other organisms that are phylogenetically related. Compared to its close relatives (*Amoebidium parasiticum*, *Ministeria vibrans*, *Capsaspora owczarzaki* and *Monosiga brevicollis*), *S. destruens* mtDNA is the smallest. This small-size of its genome is due to the scarcity of intergenic nucleotides which accounts for only 3.5 % of its genome, loss of greater number of ribosomal protein genes and complete absence of introns from its mt-genome. The *S. destruens* mtDNA encodes the same set of proteins involved in oxidative phosphorylation, as other members of Choanozoa. The extra genes carried by its genome are present in other members such as tatC in *M. brevicollis*, ccmC and ccmF in *C. owczarzaki*. *S. destruens* mtDNA has the lowest number of ribosomal protein genes and tRNA genes compared to other holozoan protists and is the most derived among the mt-genomes from the Choanozoa reported to date.

This is the first study where *S. destruens* phylogenetic position was investigated based on the mt-genome. The phylogenetic analysis based on the 8 most conserved mtDNA protein sequences (cox1-3, cob, nad3, nad4l, nad5 and atp6) revealed an interesting position for *S. destruens* as a sister group to Filasterea and Choanoflagellate and Metazoa group. This was in contrast to the previous studies where *S. destruens* was designated as one of the earliest holozoan divergence (Torruella et al. 2015).

Table 7.1. Risk assessment to determine the hazard risk associated with *Sphaerothecum destruens* based on findings from previous studies and this thesis. The risk assessment follows the guidelines by Williams et al. (2013).

Risk query	Score	Rationale	
		Knowledge from previous reports	Thesis contribution
A. Value/susceptibility of native resources			
1. What is the economic value of susceptible host(s) to freshwater fisheries?	4	*Sphaerothecum destruens* is a generalist parasite with a broad range of hosts of high economic and ecologic value (Arkush et al. 2003, Andreou et al. 2011, Ercan et al. 2015).	New potential hosts of *S. destruens* identified in the wild are dace *leuciscus* l*euciscus*, chub *Squalius cephalus*, roach *Rutilus rutilus* and brown trout *Salmo trutta* (Chapter 4).
2. What is the ecological value of susceptible host(s) to freshwater fisheries?	3		
3. Does the parasite infect a host that is endangered, vulnerable or threatened (yes/no)?	1	None of these fish species are threatened or endangered in England and Wales. It is scored 1 based on the high susceptibility of *L. delineatus* which is listed as endangered species in Europe	
B. Colonisation potential			
4. Based upon climatic conditions of source and recipient localities (including those expected through climate change), what is the likelihood that the parasite will become established?	3	The colonisation potential for *S. destruens* is high based on its wide range of hosts and direct life cycle. Broad range of host- *S. salar*, *O. tshawytscha*, *S. trutta*, *O. mykiss* (Hedrick et al. 1989, Arkush et al. 1998), *A. brama*, *C. carpio*, *L. delineatus* (Andreou et al. 2012). The presence of life stages (spores, zoospore) with wider temperature tolerance and long environmental persistence (Andreou et al. 2009) increases its	Detection of *S. destruens* in *P. parva* in decommissioned ornamental fish farm and in native fish species in its adjacent waters "Tadburn Lake stream" in 2013 (Chapter 4) and later in 2016 through eDNA method (Chapter 5) indicated the colonization success for the parasite and its spread through its environmentally transmitted infectious spores.
5. Based upon life-cycle development and host specificity of the parasite, what is the likelihood of successful colonisation and spread?	4		

		colonisation potential.
6. How many legal fish movements take place annually within risk assessment area comprising susceptible hosts? (0–10 = v. low, 10–50 = low, 50–250 = medium, 250–500 = high, >500 v. high)	3	Fish movements are considered high risk for the spread of *S. destruens*. There are approximately 450 movements for *A. brama* recorded annually in England and Wales (Williams et al. 2013).
C. Potential disease risk		
7. What is the likely pathogenicity of the parasite to fish populations based on disease occurrence in other geographical regions?	3	Fish losses have been reported in *O. tshawytscha* and *S. salar* (Harrell et al. 1986, Hedrick et al. 1989). Population declines have been reported in centrarchid species in Turkey (Ercan et al. 2015).
8. What is the likely pathogenic importance of the parasite to fisheries based on pathological descriptions and host level changes	1	Previous studies have reported pathology in both salmonids and cyprinids (Arkush et al. 1998, Andreou et al. 2011, Ercan et al. 2015). However, it is scored 1 due to no signs of disease in the histopathology of fish samples detected positive for *S. destruens* through PCR. (Chapter 4).
9. What is the potential disease risk based on the pathogenicity of congeners of the parasite?	3	*Amphibiocystidium ranae*, a close relative that has caused high mortalities in frogs (Pascolini et al. 2003).
Total	**25**	

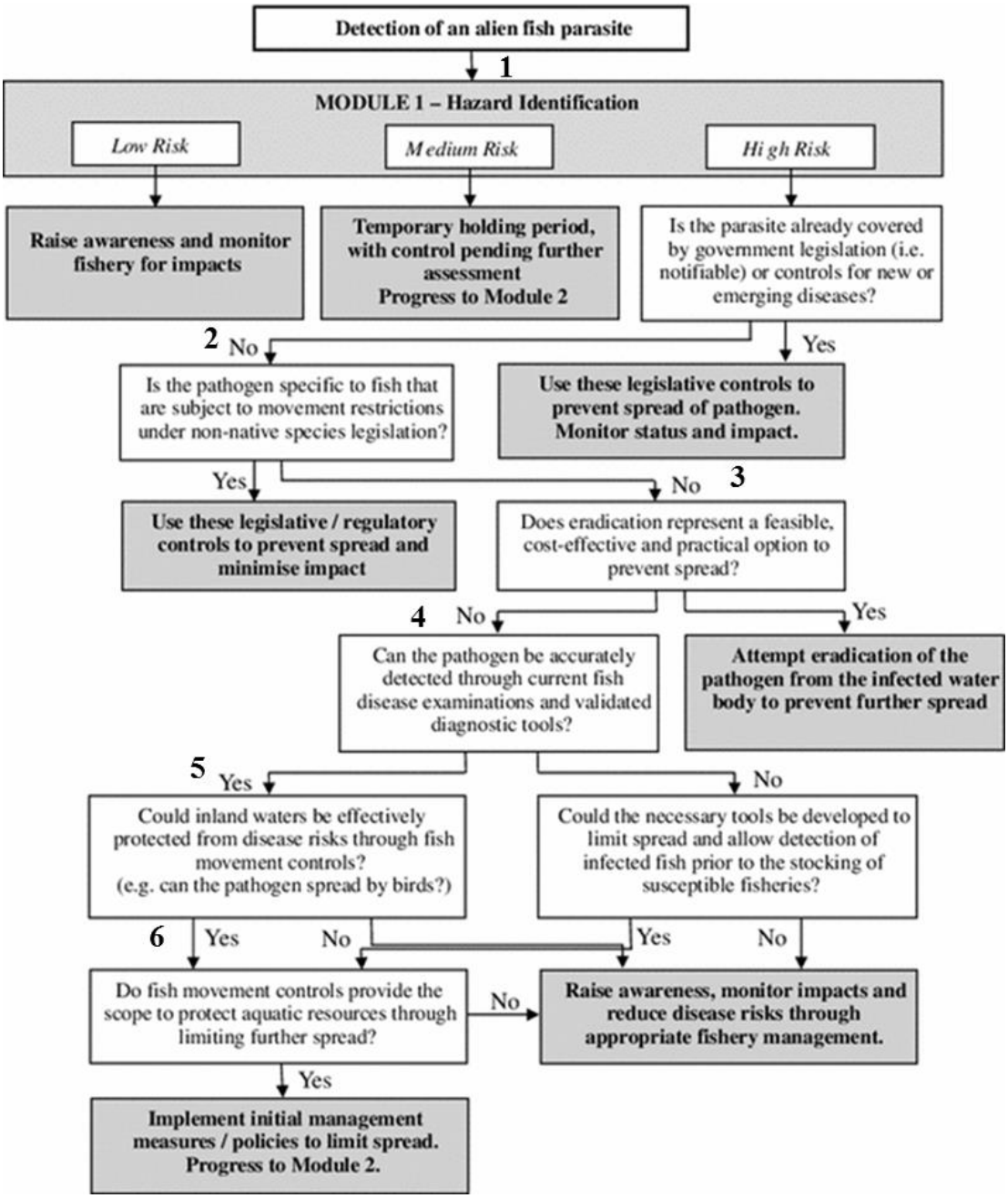


Figure 7.3. Assessment of the management options to control the spread of *Sphaerothecum destruens* based on hazard identification (Table 7.1). Rationale supporting the decisions made at each step. 1. *S. destruens* has been identified as non-native parasite to Europe (Chapter 3) and as high risk based on Hazard score 25. 2. *S. destruens* is not covered by any legislation in the UK. 3. *S. destruens* can infect *Salmo salar*, *Abramis brama* and *Cyprinus carpio*. The movement of these species is not restricted under the national exotic fish legislation (Andreou and Gozlan 2016). 4. The *P. parva* eradication was not found effective in terms of control of *S. destruens* (Chapter 5). 5. The parasite can be detected through histology and molecular methods (Andreou et al. 2011). eDNA method was found effective in *S. destruens* detection outside its host (Chapter 5). 6. Fish movement restriction could be the only effective method to control the further spread of *S. destruens* to wild freshwater habitats. The decision diagram has been adapted from Williams et al. (2013).

7.2 Future work

The present work demonstrated the introduction of *S. destruens* to continental Europe and the UK along with its reservoir host *P. parva*. The epidemiology of *S. destruens* in the continental Europe needs to be further investigated in countries especially where *P. parva* has spread and colonised river catchments (Gozlan et al. 2010a). Due to low infection levels in the reservoir host *P. parva*, the detection is usually difficult. Therefore, the native fish communities living in the vicinity of *P. parva* populations must also be examined. The eDNA method developed in the present work (Chapter 5) can be used for the initial survey of *S. destruens* presence in continental Europe. In case of positive detection with eDNA method, it should be coupled with fish examination. The mt Cyt-b marker (Chapter 2) can be effectively employed on *S. destruens* isolates from Europe to further confirm the invasive status of *S. destruens*.

Sphaerothecum destruens was detected in native fish species at high prevalence (57 %) living in the Tadburn Lake stream, which receives effluents from a decommissioned ornamental fish farm which carried *P. parva*, after one year of *P. parva* eradication from the facility. The fact that *S. destruens* has persisted in the stream in absence of its source host exhibits the ineffectiveness of the eradication programs in terms of control and spread of the environmentally transmitted propagules i.e. *S. destruens'* spores and zoospores. This emphasizes that all the *P. parva* eradicated waterbodies across the UK with direct connections to the freshwater habitats should be monitored for *S. destruens* presence. This necessitates that further work should be carried out to examine all the native fish species that have been in proximity of *P. parva* populations across the UK. A minimum of 30 fish should be examined through molecular means (i.e. PCR) and in case of positive detection histopathology should be carried out.

The eDNA method can be employed as an initial survey tool for *S. destruens* presence in the suspected waters. The outputs of *S. destruens* detection surveys can be largely dependent on the timing of the year they are carried out. The empirical studies have identified *S. destruens* infections to be highest during spring season (Ercan et al. 2015), that could be the ideal time for water sampling with maximal number of zoospores

released in inhabiting environment. But it is also important to note that zoospore have higher survival rates at lower temperatures (Andreou et al. 2009) .

The detection limit of real-time PCR assay was 1 pg/µl for *S. destruens* DNA which was slightly inferior to the detection assays for other organisms (Walker et al. 2007, Huver et al. 2015). The detection limit can be improved further by targeting the region of mitochondrial DNA of *S. destruens*, due to presence of multiple copies per cell (Avise 2000) and also its whole mtDNA sequence is available now (Chapter 6).

Comparison of the mitochondrial genomes of the unicellular organisms from the Choanozoa group revealed a rich diversity of mtDNA organization in comparison to a relatively small number of organisms with complete mt-genomes (n=5). In order to better understand the mitochondrial evolution in the unicellular relatives of animals, additional mitochondrial genomes should be obtained from this important group of eukaryotes.

7.3 Conclusions

This thesis has provided new information on *S. destruens* invasive status in Europe and its epidemiology in the UK. Novel findings include the introduction of *S. destruens* to Europe from China along with its reservoir host *P. parva* which has important implications in term of its management as a non-native parasite in the UK (Chapter 3). Identification of new potential hosts of *S. destruens* in the wild and the development of epidemiological map for *S. destruens* in the UK (Chapter 4) can play significant role in its monitoring and to minimize its further spread to adjacent waters with fisheries of high economic and conservation values. The development of eDNA method served as a fast detection tool for *S. destruens* presence in the suspected waterbodies and for monitoring the effectiveness of *P. parva* eradication programs in terms of *S. destruens* control (Chapter 5). The mitochondrial genome study of *S. destruens*, led to the development of a new phylogenetic informative marker Cyt-b (Chapter 2), and revealed a new phylogenetic position of *S. destruens* in the tree of life (Chapter 6). The features of *S. destruens* mtDNA- compaction, intron-less genes, gene-loss particularly of

ribosomal protein genes had supported the parallel evolutionary hypothesis of mtDNA evolution (Chapter 6). Lastly, the risk of *S. destruens* (Table 7.1, Figure 7.3) in the UK was re-evaluated based on the findings from Chapter 4 and 5.

References

Adl, S. M., Simpson, A. G., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., Brugerolle, G., Fensome, R. A. and Fredericq, S. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology, 52*(5), pp. 399-451.

Al-Shorbaji, F., Roche, B., Gozlan, R., Britton, R. and Andreou, D. (2016) The consequences of reservoir host eradication on disease epidemiology in animal communities. *Emerging Microbes & Infections, 5*(5), pp. e46.

Al-Shorbaji, F. N., Gozlan, R. E., Roche, B., Britton, J. R. and Andreou, D. (2015) The alternate role of direct and environmental transmission in fungal infectious disease in wildlife: threats for biodiversity conservation. *Scientific reports, 5*.

Anderson, Staden, R. and Young, I. (1981) Sequence and organization of the human mitochondrial genome. *Nature, 290*(5806), pp. 457-465.

Anderson, S., Bankier, A. T., Barrell, B. G., Debruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A. and Sanger, F. (1982) Comparison of the human and bovine mitochondrial genomes. *Cold Spring Harbor Monograph Archive, 12*, pp. 5-43.

Andree, K. B., MacConnell, E. and Hedrick, R. P. (1998) A nested polymerase chain reaction for the detection of genomic DNA of Myxobolus cerebralis in rainbow trout Oncorhynchus mykiss. *Diseases of aquatic organisms, 34*(2), pp. 145-154.

Andreou, D. (2010) *Sphaerothecum destruens:Life history traits and host range*. Unpublished PhD, Cardiff University.

Andreou, D., Arkush, K. D., Guégan, J.-F. and Gozlan, R. E. (2012) Introduced pathogens and native freshwater biodiversity: a case study of Sphaerothecum destruens. *PloS one, 7*(5), pp. e36998.

Andreou, D., Gozlan, R. and Paley, R. (2009) Temperature influence on production and longevity of Sphaerothecum destruens zoospores. *Journal of Parasitology, 95*(6), pp. 1539-1541.

Andreou, D. and Gozlan, R. E. (2016) Associated disease risk from the introduced generalist pathogen Sphaerothecum destruens: Management and policy implications. *Parasitology*.

Andreou, D., Gozlan, R. E., Stone, D., Martin, P., Bateman, K. and Feist, S. W. (2011) Sphaerothecum destruens pathology in cyprinids. *Diseases of aquatic organisms, 95*(2), pp. 145.

Annis, S. L., Dastoor, F. P., Ziel, H., Daszak, P. and Longcore, J. E. (2004) A DNA-based assay identifies Batrachochytrium dendrobatidis in amphibians. *Journal of Wildlife Diseases, 40*(3), pp. 420-428.

Arkush, K., Frasca Jr, S. and Hedrick, R. (1998) Pathology associated with the rosette agent, a systemic protist infecting salmonid fishes. *Journal of Aquatic Animal Health, 10*(1), pp. 1-11.

Arkush, K. D., Mendoza, L., Adkison, M. A. and Hedrick, R. P. (2003) Observations on the life stages of Sphaerothecum destruens ng, n. sp., a mesomycetozoean fish pathogen formally referred to as the rosette agent. *Journal of Eukaryotic Microbiology, 50*(6), pp. 430-438.

Avise, J. C. (2000) *Phylogeography: the history and formation of species*, Harvard university press.

Baldauf, S., Romeralo, M. and Carr, M. (2013) The Evolutionary Origin of Animals and Fungi. in *Evolution from the Galapagos*: Springer. pp. 73-106.

Bancroft, J., Stevens, A. and Turner, D. (1996) Theory and practice of histological techniques: Churchill Livingstone New York. *the text*, pp. 766.

Bandelt, H.-J., Forster, P. and Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution, 16*(1), pp. 37-48.

Bass, D., Stentiford, G. D., Littlewood, D. and Hartikainen, H. (2015) Diverse applications of environmental DNA methods in parasitology. *Trends in parasitology, 31*(10), pp. 499-513.

Beja-Pereira, A., Oliveira, R., Alves, P. C., Schwartz, M. K. and Luikart, G. (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources, 9*(5), pp. 1279-1301.

Bell, A. S., Yokoyama, H., Aoki, T., Takahashi, M. and Maruyama, K. (1999) Single and nested polymerase chain reaction assays for the detection of Microsporidium seriolae (Microspora), the causative agent of\'Beko\'disease in yellowtail Seriola quinqueradiata. *Diseases of aquatic organisms, 37*(2), pp. 127-134.

Berger, L., Speare, R., Daszak, P., Green, D. E., Cunningham, A. A., Goggin, C. L., Slocombe, R., Ragan, M. A., Hyatt, A. D. and McDonald, K. R. (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences, 95*(15), pp. 9031-9036.

Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. and Clayton, D. A. (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell, 26*(2), pp. 167-180.

Blaustein, A. R. and Kiesecker, J. M. (2002) Complexity in conservation: lessons from the global decline of amphibian populations. *Ecology letters, 5*(4), pp. 597-608.

Blehert, D. S., Hicks, A. C., Behr, M., Meteyer, C. U., Berlowski-Zier, B. M., Buckles, E. L., Coleman, J. T., Darling, S. R., Gargas, A. and Niver, R. (2009) Bat white-nose syndrome: an emerging fungal pathogen? *Science, 323*(5911), pp. 227-227.

Boore, J. L. and Brown, W. M. (2000) Mitochondrial genomes of Galathealinum, Helobdella, and Platynereis: sequence and gene arrangement comparisons indicate that Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. *Molecular biology and evolution, 17*(1), pp. 87-106.

Briolay, J., Galtier, N., Brito, R. M. and Bouvet, Y. (1998) Molecular phylogeny of cyprinidae inferred fromcytochrome bDNA Sequences. *Molecular Phylogenetics and Evolution, 9*(1), pp. 100-108.

Britton, J. and Brazier, M. (2006) Eradicating the invasive topmouth gudgeon, Pseudorasbora parva, from a recreational fishery in northern England. *Fisheries Management and Ecology, 13*(5), pp. 329-335.

Britton, J., Brazier, M., Davies, G. and Chare, S. (2008a) Case studies on eradicating the Asiatic cyprinid Pseudorasbora parva from fishing lakes in England to prevent their riverine dispersal. *Aquatic Conservation: Marine and Freshwater Ecosystems, 18*(6), pp. 867-876.

Britton, J., Davies, G. and Brazier, M. (2008b) Contrasting life history traits of invasive topmouth gudgeon (Pseudorasbora parva) in adjacent ponds in England. *Journal of Applied Ichthyology, 24*(6), pp. 694-698.

Britton, J., Davies, G., Brazier, M. and Pinder, A. (2007) A case study on the population ecology of a topmouth gudgeon (Pseudorasbora parva) population in the UK and the implications for native fish communities. *Aquatic Conservation: Marine and Freshwater Ecosystems, 17*(7), pp. 749-759.

Britton, J. R., Davies, G. D. and Brazier, M. (2010) Towards the successful control of the invasive Pseudorasbora parva in the UK. *Biological Invasions, 12*(1), pp. 125-131.

Brown, W. M., George, M. and Wilson, A. C. (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences, 76*(4), pp. 1967-1971.

Burger, G., Forget, L., Zhu, Y., Gray, M. W. and Lang, B. F. (2003a) Unique mitochondrial genome architecture in unicellular relatives of animals. *Proceedings of the National Academy of Sciences, 100*(3), pp. 892-897.

Burger, G., Gray, M. W., Forget, L. and Lang, B. F. (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome biology and evolution, 5*(2), pp. 418-438.

Burger, G., Gray, M. W. and Lang, B. F. (2003b) Mitochondrial genomes: anything goes. *TRENDS in Genetics, 19*(12), pp. 709-716.

Burger, G., Lavrov, D. V., Forget, L. and Lang, B. F. (2007) Sequencing complete mitochondrial and plastid genomes. *Nature protocols, 2*(3), pp. 603-614.

Burns, M. and Valdivia, H. (2008) Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology, 226*(6), pp. 1513-1524.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution, 17*(4), pp. 540-552.

Cavalier-Smith, T., Coombs, G., Vickerman, K., Sleigh, M. and Warren, A. (1998) Neomonada and the origin of animals and fungi. *Evolutionary relationships among protozoa.*, pp. 375-407.

Charrier, A., Peudpiece, M., Lesne, M. and Daniel, P. (2016) First report of the intracellular fish parasite Sphaerothecum destruens associated with the invasive topmouth gudgeon (Pseudorasbora parva) in France. *Knowl. Manag. Aquat. Ecosyst.*, (417), pp. 4.

Costa-Dias, S., Dias, E., Lobon-Cervia, J., Antunes, C. and Coimbra, J. (2010) Infection by Anguillicoloides crassus in a riverine stock of European eel, Anguilla anguilla. *Fisheries Management and Ecology, 17*(6), pp. 485-492.

Crowl, T. A., Crist, T. O., Parmenter, R. R., Belovsky, G. and Lugo, A. E. (2008) The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment, 6*(5), pp. 238-246.

Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics, 27*(8), pp. 1164-1165.

Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature methods, 9*(8), pp. 772-772.

Daszak, P., Cunningham, A. and Hyatt, A. (2001) Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta tropica, 78*(2), pp. 103-116.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P. and Miaud, C. (2011) Persistence of environmental DNA in freshwater ecosystems. *PloS one, 6*(8), pp. e23398.

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E. and Miaud, C. (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog Lithobates catesbeianus. *Journal of Applied Ecology, 49*(4), pp. 953-959.

Domaniewski, J. and Wheeler, A. (1996) The topmouth gudgeon has arrived. *Fish, 43*, pp. 40.

Domrös, M. and Gongbing, P. (1988) Climate Zones of China. in *The Climate of China*,Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 258-280.

Dunn, A. M. (2009) Parasites and biological invasions. *Advances in parasitology, 68*, pp. 161-184.

Dupuis, J. R., Roe, A. D. and Sperling, F. A. (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology, 21*(18), pp. 4422-4436.

Dyall, S. D., Brown, M. T. and Johnson, P. J. (2004) Ancient invasions: from endosymbionts to organelles. *Science, 304*(5668), pp. 253-257.

Elliott, M. (2003) Biological pollutants and biological pollution—an increasing cause for concern. *Marine Pollution Bulletin, 46*(3), pp. 275-280.

Ercan, D., Andreou, D., Sana, S., Öntaş, C., Baba, E., Top, N., Karakuş, U., Tarkan, A. S. and Gozlan, R. E. (2015) Evidence of threat to European economy and biodiversity following the introduction of an alien pathogen on the fungal–animal boundary. *Emerging Microbes & Infections, 4*(9), pp. e52.

Eskew, E. A. and Todd, B. D. (2013) Parallels in amphibian and bat declines from pathogenic fungi. *Emerging infectious diseases, 19*(3), pp. 379.

Excoffier, L. and Lischer, H. E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources, 10*(3), pp. 564-567.

Feagin, J. E. (2000) Mitochondrial genome diversity in parasites. *International Journal for Parasitology, 30*(4), pp. 371-390.

Fenton, A., Streicker, D. G., Petchey, O. L. and Pedersen, A. B. (2015) Are all Hosts created equal? Partitioning host species contributions to parasite persistence in multihost communities. *The American Naturalist, 186*(5), pp. 610-622.

Ficetola, G. F., Coïc, C., Detaint, M., Berroneau, M., Lorvelec, O. and Miaud, C. (2007a) Pattern of distribution of the American bullfrog Rana catesbeiana in Europe. *Biological Invasions, 9*(7), pp. 767-772.

Ficetola, G. F., Miaud, C., Pompanon, F. and Taberlet, P. (2008) Species detection using environmental DNA from water samples. *Biology letters, 4*(4), pp. 423-425.

Ficetola, G. F., Thuiller, W. and Miaud, C. (2007b) Prediction and validation of the potential global distribution of a problematic alien invasive species—the American bullfrog. *Diversity and Distributions, 13*(4), pp. 476-485.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L. and Gurr, S. J. (2012) Emerging fungal threats to animal, plant and ecosystem health. *Nature, 484*(7393), pp. 186-194.

Folmer, O., Black, M., Hoeh, W., Lutz, R. and Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase Subunit 1 from diverse metazoan invertebrates. *Molecular marine biology and biotechnology,* (5).

Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*(2), pp. 915-925.

Galtier, N., Nabholz, B., Glémin, S. and Hurst, G. (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology, 18*(22), pp. 4541-4550.

Garner, T. W. J., Walker, S., Bosch, J., Hyatt, A. D., Cunningham, A. A. and Fisher, M. C. (2005) Chytrid fungus in Europe. *Emerging infectious diseases, 11*(10), pp. 1639-1641.

GBNNSS (2015) The Great Britain Invasive Non-native Species Strategy. Last accessed August 2016.

Ghosh, K. and Weiss, L. M. (2009) Molecular diagnostic tests for microsporidia. *Interdisciplinary perspectives on infectious diseases, 2009*.

Goldberg, C. S., Pilliod, D. S., Arkle, R. S. and Waits, L. P. (2011) Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. *PloS one, 6*(7), pp. e22746.

Goldberg, C. S., Sepulveda, A., Ray, A., Baumgardt, J. and Waits, L. P. (2013) Environmental DNA as a new method for early detection of New Zealand mudsnails (Potamopyrgus antipodarum).

Goldberg, C. S., Strickler, K. M. and Pilliod, D. S. (2015) Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. *Biological Conservation, 183*, pp. 1-3.

Gonzalez, S. F., Osorio, C. R. and Santos, Y. (2003) Development of a PCR-based method for the detection of Listonella anguillarum in fish tissues and blood samples. *Diseases of aquatic organisms, 55*(2), pp. 109-115.

Gozlan, R. (2012) Pathogens: Monitoring fungal infections in fish. *Nature, 485*(7399), pp. 446-446.

Gozlan, R., Andreou, D., Asaeda, T., Beyer, K., Bouhadad, R., Burnard, D., Caiola, N., Cakic, P., Djikanovic, V. and Esmaeili, H. R. (2010a) Pan-continental invasion of Pseudorasbora parva: towards a better understanding of freshwater fish invasions. *Fish and Fisheries, 11*(4), pp. 315-340.

Gozlan, R., Britton, J., Cowx, I. and Copp, G. (2010b) Current knowledge on non-native freshwater fish introductions. *Journal of Fish Biology, 76*(4), pp. 751-786.

Gozlan, R., Pinder, A. and Shelley, J. (2002) Occurrence of the Asiatic cyprinid Pseudorasbora parva in England. *Journal of Fish Biology, 61*(1), pp. 298-300.

Gozlan, R. E. (2008) Introduction of non-native freshwater fish: is it all bad? *Fish and Fisheries, 9*(1), pp. 106-115.

Gozlan, R. E., St-Hilaire, S., Feist, S. W., Martin, P. and Kent, M. L. (2005) Biodiversity: disease threat to European fish. *Nature, 435*(7045), pp. 1046-1046.

Gozlan, R. E., Whipps, C. M., Andreou, D. and Arkush, K. D. (2009) Identification of a rosette-like agent as Sphaerothecum destruens, a multi-host fish pathogen. *International Journal for Parasitology, 39*(10), pp. 1055-1058.

Gray, M. W. (2012) Mitochondrial evolution. *Cold Spring Harbor perspectives in biology, 4*(9), pp. a011403.

Gray, M. W., Burger, G. and Lang, B. F. (1999) Mitochondrial evolution. *Science, 283*(5407), pp. 1476-1481.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. and Hofacker, I. L. (2008) The vienna RNA websuite. *Nucleic acids research, 36*(suppl 2), pp. W70-W74.

Gurnell, J., Wauters, L. A., Lurz, P. W. and Tosi, G. (2004) Alien species and interspecific competition: effects of introduced eastern grey squirrels on red squirrel population dynamics. *Journal of Animal Ecology, 73*(1), pp. 26-35.

Hajibabaei, M., Singer, G. A., Hebert, P. D. and Hickey, D. A. (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *TRENDS in Genetics, 23*(4), pp. 167-172.

Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. in *Nucleic acids symposium series*. pp. 95-98.

Harpending, H. C., Sherry, S. T., Rogers, A. R. and Stoneking, M. (1993) The genetic structure of ancient human populations. *Current Anthropology, 34*(4), pp. 483-496.

Harrell, L. W., Elston, R. A., Scott, T. M. and Wilkinson, M. T. (1986) A significant new systemic disease of net-pen reared chinook salmon (Oncorhynchus tshawytscha) brood stock. *Aquaculture, 55*(4), pp. 249-262.

Hasegawa, M., Kishino, H. and Yano, T.-a. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution, 22*(2), pp. 160-174.

Hedrick, R., FRIEDMAN, C. and MODIN, J. (1989) Systemic infection in Atlantic salmon Salmo salar with a Dermocystidium-like species. *Diseases of aquatic organisms, 7*(3), pp. 171-177.

Herr, R. A., Ajello, L., Taylor, J. W., Arseculeratne, S. N. and Mendoza, L. (1999) Phylogenetic analysis of Rhinosporidium seeberi's 18S small-subunit ribosomal DNA groups this pathogen among members of the protoctistan Mesomycetozoa clade. *Journal of Clinical Microbiology, 37*(9), pp. 2750-2754.

Hudson, P. J., Rizzoli, A., Grenfell, B. T., Heesterbeek, H. and Dobson, A. P. (2002) *The ecology of wildlife diseases*, Oxford University Press Oxford.

Hulme, P. E. (2006) Beyond control: wider implications for the management of biological invasions. *Journal of Applied Ecology, 43*(5), pp. 835-847.

Huver, J., Koprivnikar, J., Johnson, P. and Whyard, S. (2015) Development and application of an eDNA method to detect and quantify a pathogenic parasite in aquatic ecosystems. *Ecological Applications, 25*(4), pp. 991-1002.

Hyman, O. J. and Collins, J. P. (2012) Evaluation of a filtration-based method for detecting Batrachochytrium dendrobatidis in natural bodies of water. *Diseases of aquatic organisms, 97*(3), pp. 185-195.

Jerde, C. L., Chadderton, W. L., Mahon, A. R., Renshaw, M. A., Corush, J., Budny, M. L., Mysorekar, S. and Lodge, D. M. (2013) Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. *Canadian Journal of Fisheries and Aquatic Sciences, 70*(4), pp. 522-526.

Jerde, C. L., Mahon, A. R., Chadderton, W. L. and Lodge, D. M. (2011) "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters, 4*(2), pp. 150-157.

Johnson, M. L. and Speare, R. (2005) Possible modes of dissemination of the amphibian chytrid Batrachochytrium dendrobatidis in the environment. *Diseases of aquatic organisms, 65*, pp. 181-186.

Kats, L. B. and Ferrer, R. P. (2003) Alien predators and amphibian declines: review of two decades of science and the transition to conservation. *Diversity and Distributions, 9*(2), pp. 99-110.

Kennedy, C. (2007) The pathogenic helminth parasites of eels. *Journal of fish diseases, 30*(6), pp. 319-334.

Kimmerer, W. J., Gartside, E. and Orsi, J. J. (1994) Predation by an introduced clam as the likely cause of substantial declines in zooplankton of San Francisco Bay. *Marine ecology progress series. Oldendorf, 113*(1), pp. 81-93.

Kirk, R. (2003) The impact of Anguillicola crassus on European eels. *Fisheries Management and Ecology, 10*(6), pp. 385-394.

Kirshtein, J. D., Anderson, C. W., Wood, J. S., Longcore, J. E. and Voytek, M. A. (2007) Quantitative PCR detection of Batrachochytrium dendrobatidis DNA from sediments and water. *Diseases of aquatic organisms, 77*(1), pp. 11-15.

Knapp, R. A. and Matthews, K. R. (2000) Non-Native Fish Introductions and the Decline of the Mountain Yellow-Legged Frog from within Protected Areas. *Conservation Biology, 14*(2), pp. 428-438.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X. and Wilson, A. C. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences, 86*(16), pp. 6196-6200.

Koops, H. and Hartmann, F. (1989) Anguillicola-infestations in Germany and in German eel imports. *Journal of Applied Ichthyology, 5*(1), pp. 41-45.

Krasnov, B. R., Khokhlova, I. S., Shenbrot, G. I. and Poulin, R. (2008) How are the host spectra of hematophagous parasites shaped over evolutionary time? Random choice vs selection of a phylogenetic lineage. *Parasitology research, 102*(6), pp. 1157-1164.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, pp. msw054.

Kwok, S. a. and Higuchi, R. (1989) Avoiding false positives with PCR. *Nature, 339*, pp. 237-238.

Lang, B., O'kelly, C., Nerad, T., Gray, M. and Burger, G. (2002) The closest unicellular relatives of animals. *Current Biology, 12*(20), pp. 1773-1778.

Lavrov, D. V. (2007) Key transitions in animal evolution: a mitochondrial DNA perspective. *Integrative and Comparative Biology, 47*(5), pp. 734-743.

Lavrov, D. V., Brown, W. M. and Boore, J. L. (2004) Phylogenetic position of the Pentastomida and (pan) crustacean relationships. *Proceedings of the Royal Society of London B: Biological Sciences, 271*(1538), pp. 537-544.

Lavrov, D. V. and Lang, B. F. (2014) Mitochondrial Genomes in Unicellular Relatives of Animals. *Molecular Life Sciences, 9091*.

Lavrov, D. V. and Pett, W. (2016) Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. *Genome biology and evolution, 8*(9), pp. 2896-2913.

Lavrov, D. V., Wang, X. and Kelly, M. (2008) Reconstructing ordinal relationships in the Demospongiae using mitochondrial genomic data. *Molecular Phylogenetics and Evolution, 49*(1), pp. 111-124.

Lee, P. A., Tullman-Ercek, D. and Georgiou, G. (2006) The bacterial twin-arginine translocation pathway. *Annual review of microbiology, 60*, pp. 373.

Librado, P. and Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*(11), pp. 1451-1452.

Lorch, J. M., Gargas, A., Meteyer, C. U., Berlowski-Zier, B. M., Green, D. E., Shearn-Bochsler, V., Thomas, N. J. and Blehert, D. S. (2010) Rapid polymerase chain reaction diagnosis of white-nose syndrome in bats. *Journal of Veterinary Diagnostic Investigation, 22*(2), pp. 224-230.

Lowe, T. M. and Eddy, S. R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic acids research, 25*(5), pp. 955-964.

Lyubetsky, V., Piel, W. H. and Quandt, D. (2014) Current advances in molecular phylogenetics. *BioMed research international, 2014*.

Mahon, A. R., Jerde, C. L., Galaska, M., Bergner, J. L., Chadderton, W. L., Lodge, D. M., Hunter, M. E. and Nico, L. G. (2013) Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. *PloS one, 8*(3), pp. e58316.

Maltagliati, F., Di Giuseppe, G., Barbieri, M., Castelli, A. and Dini, F. (2010) Phylogeography and genetic structure of the edible sea urchin Paracentrotus lividus (Echinodermata: Echinoidea) inferred from the mitochondrial cytochrome b gene. *Biological Journal of the Linnean Society, 100*(4), pp. 910-923.

Marck, C. and Grosjean, H. (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *Rna, 8*(10), pp. 1189-1232.

Martel, A., Blooi, M., Adriaensen, C., Van Rooij, P., Beukema, W., Fisher, M., Farrer, R., Schmidt, B., Tobler, U. and Goka, K. (2014) Recent introduction of a chytrid fungus endangers Western Palearctic salamanders. *Science*, *346*(6209), pp. 630-631.

McKee, A. M., Spear, S. F. and Pierson, T. W. (2015) The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biological Conservation, 183*, pp. 70-76.

Mendonca, H. L. and Arkush, K. D. (2004) Development of PCR-based methods for detection of Sphaerothecum destruens in fish tissues. *Diseases of aquatic organisms, 61*(3), pp. 187-197.

Mendoza, L., Ajello, L. and Taylor, J. W. (2001) The taxonomic status of Lacazia loboi and Rhinosporidium seeberi has been finally resolved with the use of molecular tools. *Revista Iberoamericana de Micologia, 18*(3), pp. 95-98.

Michel, C., Faivre, B. and De Kinkelin, P. (1986) A clinical case of enteric red-mouth in minnows (Pimephales promelas) imported into Europe as bait-fish. *Bulletin of the European Association of Fish Pathologists, 6*, pp. 97-99.

Minamoto, T., Pu, X., Xie, J., Dong, Y., Wu, D., Kong, H., Yang, X., Takahara, T., Honjo, M. N. and Yamanaka, H. (2015) Monitoring fish pathogenic viruses in natural lakes in Yunnan, China. *Limnology, 16*(1), pp. 69-77.

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M. N. and Kawabata, Z. i. (2012) Surveillance of fish species composition using environmental DNA. *Limnology, 13*(2), pp. 193-197.

Mohammed, S., Birhan, G., Admassu, B., Shite, A. and Yeneneh, H. (2015) Review on Polymerase Chain Reaction and its Diagnostic Merit Over Conventional Techniques in Animal Disease. *African Journal of Basic & Applied Sciences, 7*(5), pp. 262-281.

Mordecai, E. A. (2013) Consequences of pathogen spillover for cheatgrass-invaded grasslands: coexistence, competitive exclusion, or priority effects. *The American Naturalist, 181*(6), pp. 737-747.

Murray, A. G. and Peeler, E. J. (2005) A framework for understanding the potential for emerging diseases in aquaculture. *Preventive veterinary medicine, 67*(2), pp. 223-235.

Navidi, W., Arnheim, N. and Waterman, M. (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American journal of human genetics, 50*(2), pp. 347.

Paley, R., Andreou, D., Bateman, K. and Feist, S. (2012) Isolation and culture of Sphaerothecum destruens from Sunbleak (Leucaspius delineatus) in the UK and pathogenicity experiments in Atlantic salmon (Salmo salar). *Parasitology, 139*(7), pp. 904.

Palumbi, S. R. (1996) Nucleic acids II: the polymerase chain reaction. *Molecular systematics, 2*, pp. 205-247.

Pascolini, R., Daszak, P., Cunningham, A., Tei, S., Vagnetti, D., Bucci, S., Fagotti, A. and Di Rosa, I. (2003) Parasitism by Dermocystidium ranae in a population of Rana esculenta complex in Central Italy and description of Amphibiocystidium n. gen. *Diseases of aquatic organisms, 56*(1), pp. 65-74.

Patwardhan, A., Ray, S. and Roy, A. (2014) Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology, 2014*.

Peeler, E. J., Oidtmann, B. C., Midtlyng, P. J., Miossec, L. and Gozlan, R. E. (2011) Non-native aquatic animals introductions have driven disease emergence in Europe. *Biological Invasions, 13*(6), pp. 1291-1303.

Pelicice, F. M., Vitule, J. R. S., Lima Junior, D. P., Orsi, M. L. and Agostinho, A. A. (2014) A serious new threat to Brazilian freshwater ecosystems: the naturalization of nonnative fish by decree. *Conservation Letters, 7*(1), pp. 55-60.

Pereira, C. N., Di Rosa, I., Fagotti, A., Simoncelli, F., Pascolini, R. and Mendoza, L. (2005) The pathogen of frogs Amphibiocystidium ranae is a member of the order Dermocystida in the class Mesomycetozoea. *Journal of Clinical Microbiology, 43*(1), pp. 192-198.

Perkins, S. E., Altizer, S., Bjornstad, O., Burdon, J. J., Clay, K., Gómez-Aparicio, L., Jeschke, J. M., Johnson, P. T., Lafferty, K. D. and Malmstrom, C. M. (2008) *Invasion biology and parasitic infections,* Princeton, New Jersey: Princeton University Press.

Peters, G. and Hartmann, F. (1986) Anguillicola, a parasitic nematode of the swim bladder spreading among eel populations in Europe. *Diseases of aquatic organisms, 1*(3), pp. 229-230.

Pichler, F. B. (2002) Genetic assessment of population boundaries and gene exchange in Hector's dolphin.

Pilliod, D. S., Goldberg, C. S., Arkle, R. S. and Waits, L. P. (2013) Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences, 70*(8), pp. 1123-1130.

Pilliod, D. S., Goldberg, C. S., Arkle, R. S. and Waits, L. P. (2014) Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources, 14*(1), pp. 109-116.

Pimentel, D., Lach, L., Zuniga, R. and Morrison, D. (2000) Environmental and economic costs of nonindigenous species in the United States. *BioScience, 50*(1), pp. 53-65.

Pinder, A., Gozlan, R. and Britton, J. (2005) Dispersal of the invasive topmouth gudgeon, Pseudorasbora parva in the UK: a vector for an emergent infectious disease. *Fisheries Management and Ecology, 12*(6), pp. 411-414.

Pinder, A. C. and Gozlan, R. E. (2003) Sunbleak and topmouth gudgeon: two new additions to Britain's freshwater fishes. *British Wildlife, 15*(2), pp. 77-83.

Poulin, R., Paterson, R., Townsend, C., Tompkins, D. and Kelly, D. (2011) Biological invasions and the dynamics of endemic diseases in freshwater ecosystems. *Freshwater Biology, 56*(4), pp. 676-688.

Prenter, J., MacNeil, C., Dick, J. T. and Dunn, A. M. (2004) Roles of parasites in animal invasions. *Trends in ecology & evolution, 19*(7), pp. 385-390.

Puechmaille, S. J., Wibbelt, G., Korn, V., Fuller, H., Forget, F., Mühldorfer, K., Kurth, A., Bogdanowicz, W., Borel, C. and Bosch, T. (2011) Pan-European distribution of white-nose syndrome fungus (Geomyces destructans) not associated with mass mortality. *PloS one, 6*(4), pp. e19167.

Rachowicz, L. J., HERO, J., Alford, R. A., Taylor, J. W., Morgan, J. A., Vredenburg, V. T., Collins, J. P. and Briggs, C. J. (2005) The novel and endemic pathogen hypotheses: competing explanations for the origin of emerging infectious diseases of wildlife. *Conservation Biology, 19*(5), pp. 1441-1448.

Ragan, M. A., Goggin, C. L., Cawthorn, R. J., Cerenius, L., Jamieson, A., Plourde, S. M., Rand, T. G., Söderhäll, K. and Gutell, R. R. (1996) A novel clade of protistan parasites near the animal-fungal divergence. *Proceedings of the National Academy of Sciences, 93*(21), pp. 11907-11912.

Ramos-Onsins, S. E. and Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular biology and evolution, 19*(12), pp. 2092-2100.

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. and Gough, K. C. (2014) REVIEW: The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology, 51*(5), pp. 1450-1459.

Riley, S. (2005) Invasive alien species and the protection of biodiversity: the role of quarantine laws in resolving inadequacies in the international legal regime. *Journal of Environmental Law, 17*(3), pp. 323-359.

Rogers, A. R. and Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution, 9*(3), pp. 552-569.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A. and Huelsenbeck, J. P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology, 61*(3), pp. 539-542.

Ruiz-Trillo, I., Roger, A. J., Burger, G., Gray, M. W. and Lang, B. F. (2008) A phylogenomic investigation into the origin of metazoa. *Molecular biology and evolution, 25*(4), pp. 664-672.

Ruiz, G. M., Carlton, J. T., Grosholz, E. D. and Hines, A. H. (1997) Global invasions of marine and estuarine habitats by non-indigenous species: mechanisms, extent, and consequences. *American Zoologist, 37*(6), pp. 621-632.

Schmidt, B. R., Kery, M., Ursenbacher, S., Hyman, O. J. and Collins, J. P. (2013) Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution, 4*(7), pp. 646-653.

Schneider, S. and Excoffier, L. (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics, 152*(3), pp. 1079-1089.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., Bolchacova, E., Voigt, K. and Crous, P. W. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences, 109*(16), pp. 6241-6246.

Schubart, C. D., Reimer, J. and Diesel, R. (1998) Morphological and molecular evidence for a new endemic freshwater crab, Sesarma ayatum sp. n.,(Grapsidae, Sesarminae) from eastern Jamaica. *Zoologica Scripta, 27*(4), pp. 373-380.

Shalchian-Tabrizi, K., Minge, M. A., Espelund, M., Orr, R., Ruden, T., Jakobsen, K. S. and Cavalier-Smith, T. (2008) Multigene phylogeny of choanozoa and the origin of animals. *PloS one, 3*(5), pp. e2098.

Shao, Z., Graf, S., Chaga, O. Y. and Lavrov, D. V. (2006) Mitochondrial genome of the moon jelly Aurelia aurita (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. *Gene, 381*, pp. 92-101.

Signorovitch, A. Y., Buss, L. W. and Dellaporta, S. L. (2007) Comparative genomics of large mitochondria in placozoans. *PLoS Genet, 3*(1), pp. e13.

Simon, A., Britton, R., Gozlan, R., Van Oosterhout, C., Volckaert, F. A. and Hänfling, B. (2011) Invasive cyprinid fish in Europe originate from the single introduction of an admixed source population followed by a complex pattern of spread. *PloS one, 6*(6), pp. e18560.

Simon, A., Gozlan, R. E., Britton, J. R., Van Oosterhout, C. and Hänfling, B. (2015) Human induced stepping-stone colonisation of an admixed founder population: the spread of topmouth gudgeon (Pseudorasbora parva) in Europe. *Aquatic Sciences, 77*(1), pp. 17-25.

Skerratt, L. F., Berger, L., Speare, R., Cashins, S., McDonald, K. R., Phillott, A. D., Hines, H. B. and Kenyon, N. (2007) Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth, 4*(2), pp. 125-134.

Spikmans, F., van Tongeren, T., van Alen, T. A., van der Velde, G. and den Camp, H. J. O. (2013) High prevalence of the parasite Sphaerothecum destruens in the invasive topmouth gudgeon Pseudorasbora parva in the Netherlands, a potential threat to native freshwater fish. *Aquatic Invasions, 8*(3), pp. 355-360.

Strand, D. A., Jussila, J., Johnsen, S. I., Viljamaa-Dirks, S., Edsman, L., Wiik-Nielsen, J., Viljugrein, H., Engdahl, F. and Vrålstad, T. (2014) Detection of crayfish plague spores in large freshwater systems. *Journal of Applied Ecology, 51*(2), pp. 544-553.

Strayer, D. L. and Smith, L. C. (1996) Relationships between zebra mussels (Dreissena polymorpha) and unionid clams during the early stages of the zebra mussel invasion of the Hudson River. *Freshwater Biology, 36*(3), pp. 771-780.

Strickler, K. M., Fremier, A. K. and Goldberg, C. S. (2015) Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation, 183*, pp. 85-92.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P. and Bouvet, J. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic acids research, 24*(16), pp. 3189-3194.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), pp. 585-595.

Takahara, T., Minamoto, T. and Doi, H. (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PloS one, 8*(2), pp. e56584.

Takahara, T., Minamoto, T. and Doi, H. (2015) Effects of sample processing on the detection rate of environmental DNA from the Common Carp (Cyprinus carpio). *Biological Conservation, 183*, pp. 64-69.

Thomsen, P., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L. and Willerslev, E. (2012) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology, 21*(11), pp. 2565-2573.

Thomsen, P. F. and Willerslev, E. (2015) Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183*, pp. 4-18.

Tompkins, D. M., Dunn, A. M., Smith, M. J. and Telfer, S. (2011) Wildlife diseases: from individuals to ecosystems. *Journal of Animal Ecology, 80*(1), pp. 19-38.

Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M. A., del Campo, J., Eme, L., Pérez-Cordón, G., Whipps, C. M. and Nichols, K. M. (2015) Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Current Biology, 25*(18), pp. 2404-2410.

Torruella, G., Derelle, R., Paps, J., Lang, B. F., Roger, A. J., Shalchian-Tabrizi, K. and Ruiz-Trillo, I. (2012) Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Molecular biology and evolution, 29*(2), pp. 531-544.

Tsui, C. K., Woodhall, J., Chen, W., Lévesque, C. A., Lau, A., Schoen, C. D., Baschien, C., Najafzadeh, M. J. and de Hoog, G. S. (2011) Molecular techniques for pathogen identification and fungus detection in the environment. *IMA Fungus: The Global Mycological Journal, 2*(2), pp. 177.

Turner, C. R., Uy, K. L. and Everhart, R. C. (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation, 183*, pp. 93-102.

Turner, G. G., Reeder, D. and Coleman, J. T. (2011) A Five-year Assessment of Mortality and Geographic Spread of White-Nose Syndrome in North American Bats, with a Look at the Future. Update of White-Nose Syndrome in Bats. *Bat Research News*, pp. 13.

Van Zon, J. (1977) Grass carp (Ctenopharyngodon idella Val.) in Europe. *Aquatic Botany, 3*, pp. 143-155.

Voyles, J., Kilpatrick, A. M., Collins, J. P., Fisher, M. C., Frick, W. F., McCallum, H., Willis, C. K., Blehert, D. S., Murray, K. A. and Puschendorf, R. (2014) Moving beyond too little, too late: Managing emerging infectious diseases in wild populations requires international policy and partnerships. *EcoHealth, 2014*, pp. 1-4.

Vredenburg, V. T. (2004) Reversing introduced species effects: experimental removal of introduced fish leads to rapid recovery of a declining frog. *Proceedings of the National Academy of Sciences of the United States of America, 101*(20), pp. 7646-7650.

Walker, S. F., Salas, M. B., Jenkins, D., Garner, T. W., Cunningham, A. A., Hyatt, A. D., Bosch, J. and Fisher, M. C. (2007) Environmental detection of Batrachochytrium dendrobatidis in a temperate climate. *Diseases of aquatic organisms, 77*(2), pp. 105-112.

Wang, X. and Lavrov, D. V. (2007) Mitochondrial genome of the homoscleromorph Oscarella carmela (Porifera, Demospongiae) reveals unexpected complexity in the common ancestor of sponges and other animals. *Molecular biology and evolution, 24*(2), pp. 363-373.

Wang, Z. and Wu, M. (2015) An integrated phylogenomic approach toward pinpointing the origin of mitochondria. *Scientific reports, 5*, pp. 7949.

Weiss, J. B. (1995) DNA probes and PCR for diagnosis of parasitic infections. *Clinical Microbiology Reviews, 8*(1), pp. 113-130.

Welcomme, R. (1992) A history of international introductions of inland aquatic species. in *ICES Marine Science Symposium*. pp. 3-14.

Whipps, C. M., Burton, T., Watral, V. G., St-Hilaire, S. and Kent, M. L. (2006) Assessing the accuracy of a polymerase chain reaction test for Ichthyophonus

hoferi in Yukon River Chinook salmon Oncorhynchus tshawytscha. *Diseases of aquatic organisms, 68*(2), pp. 141-147.

White, T. J., Bruns, T., Lee, S. and Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications, 18*(1), pp. 315-322.

Whittington, R. and Chong, R. (2007) Global trade in ornamental fish from an Australian perspective: the case for revised import risk analysis and management strategies. *Preventive veterinary medicine, 81*(1), pp. 92-116.

Wibbelt, G., Kurth, A., Hellmann, D., Weishaar, M., Barlow, A., Veith, M., Prüger, J., Görföl, T., Grosche, L. and Bontadina, F. (2010) White-nose syndrome fungus (Geomyces destructans) in bats, Europe. *Emerging infectious diseases, 16*(8), pp. 1237-1243.

Williams, C. F., Britton, J. R. and Turnbull, J. F. (2013) A risk assessment for managing non-native parasites. *Biological Invasions, 15*(6), pp. 1273-1286.

Wolstenholme, D. R. (1992) Animal mitochondrial DNA: structure and evolution. *International review of cytology, 141*, pp. 173-216.

Yamamoto, Y. (2002) PCR in diagnosis of infection: detection of bacteria in cerebrospinal fluids. *Clinical and diagnostic laboratory immunology, 9*(3), pp. 508-514.

Zuccarello, G., Burger, G., West, J. and King, R. (1999) A mitochondrial marker for red algal intraspecific relationships. *Molecular Ecology, 8*(9), pp. 1443-1447.

Appendices

Appendix 1

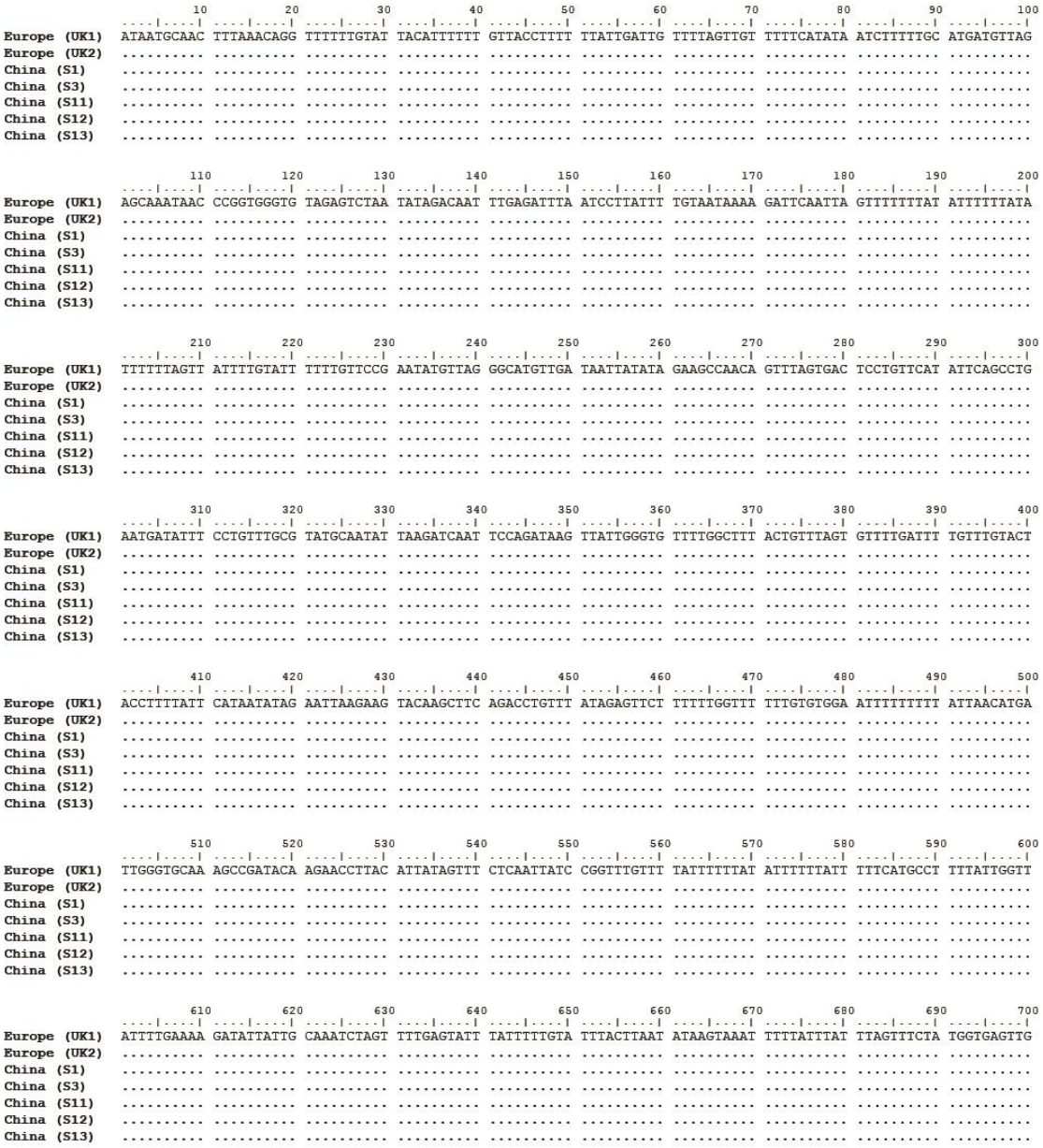


Figure 1: Cytochrome b sequences of *Sphaerothecum destruens* isolates from two European samples (UK 1: FN996945 and UK-2: RA isolate from this study) and 5 samples across China (sites S1, S3, S11, S12 and S13). The sequences showed no DNA polymorphism.

Appendix 2

Minimum detection threshold for *Sphaerothecum destruens* using nested PCR

Day6

S. destruens DNA was detected in all spore concentrations in distilled water. In turbid water, all three replicates produced amplification at High spore concentration and two

replicates were detected positive at medium and low spore concentration (Figure 2.1 Table 2.1).

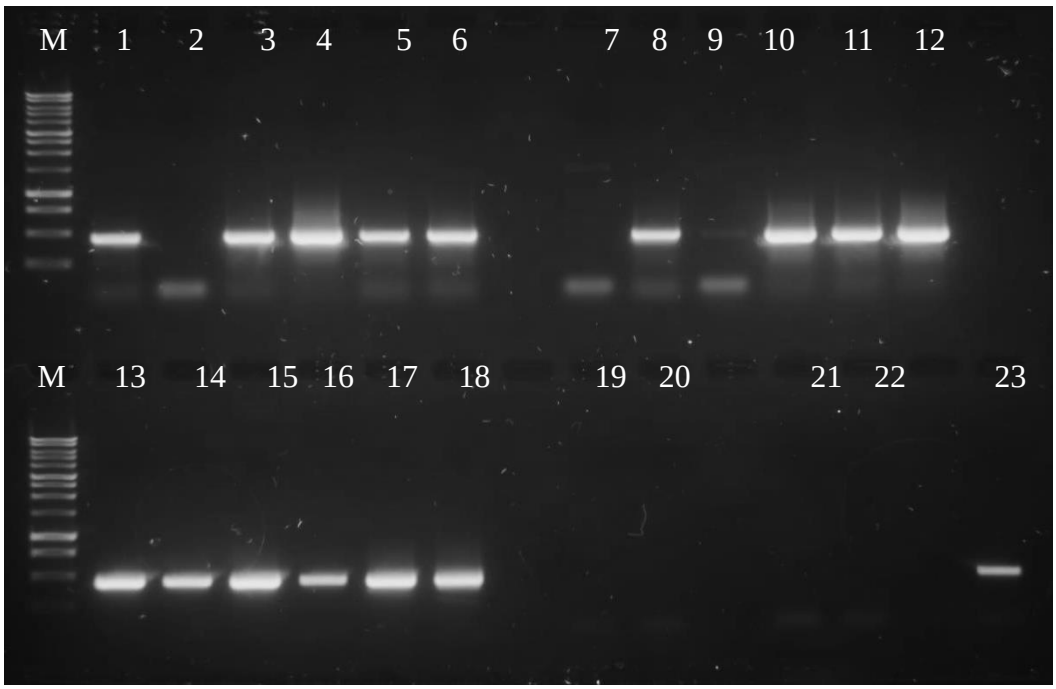


Figure 2.1: Detection limit of conventional nested PCR for *Sphaerothecum destruens* spores in turbid and natural water at Day 6 incubation period. Lane 1-3, 7-9, 13-15: second round PCR products from turbid water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 4-6, 10-12, 16-18: PCR products from natural water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 19, 20, 21, 22: negative controls. Lane 23: PCR +ve control, Lane M: 1Kb DNA ladder.

Day13

At Day 13, *S. destruens*-specific DNA was detected in presence of distilled water in all 3 replicates at High spore concentration and 2/3 for medium spore concentration and 3/3 for low spore concentration of which amplification was very low for one replicate. In turbid water *S. destruens*-specific DNA was detected in only High spore concentration in 2/3 replicates (Figure 2.2 Table 2.1).

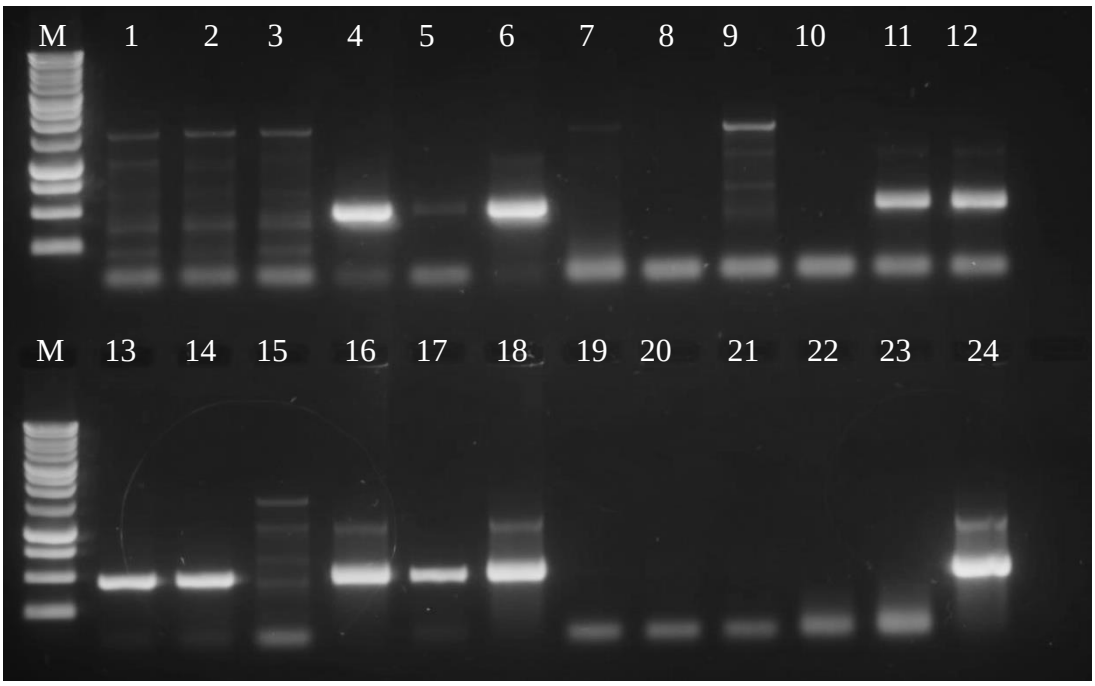


Figure 2.2: Detection limit of conventional nested PCR for *Sphaerothecum destruens* spores in turbid and natural water at Day 13 incubation period. Lane M: 1Kb DNA ladder, Lane 1-3, 7-9, 13-15: second round PCR products from turbid water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 4-6, 10-12, 16-18: PCR products from natural water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 19-23: negative controls. Lane 24: PCR +ve control.

Day20

For distilled water, all the replicates at all the spore concentrations were detected positive for *S. destruens*-specific DNA. In the turbidity conditions two replicates were positive at high spore concentration and one replicate at medium spore concentration. No *S. destruens* DNA was detected at low spore concentration in any replicate Figure 5.3. An overview of *S. destruens* DNA detection by conventional PCR over the course of 20 days is summarised in Table 5.1. No positive amplifications were obtained from extractions of filter papers used in filtration equipment post-disinfection, indicating that there was no carry-over of *S. destruens* DNA.

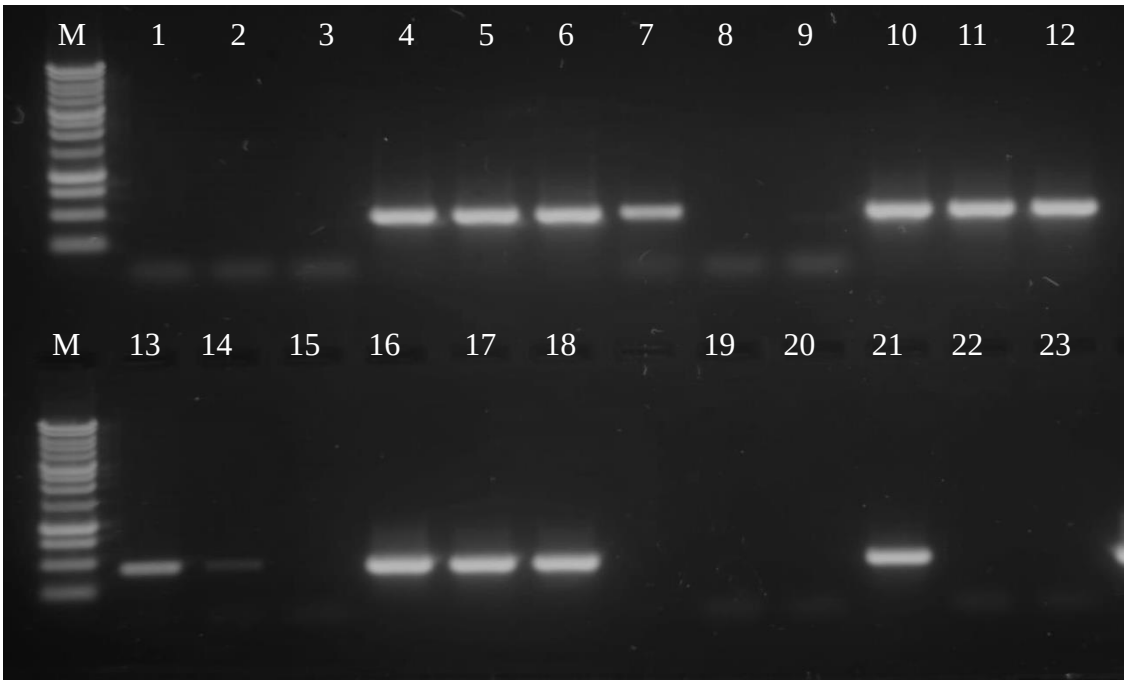


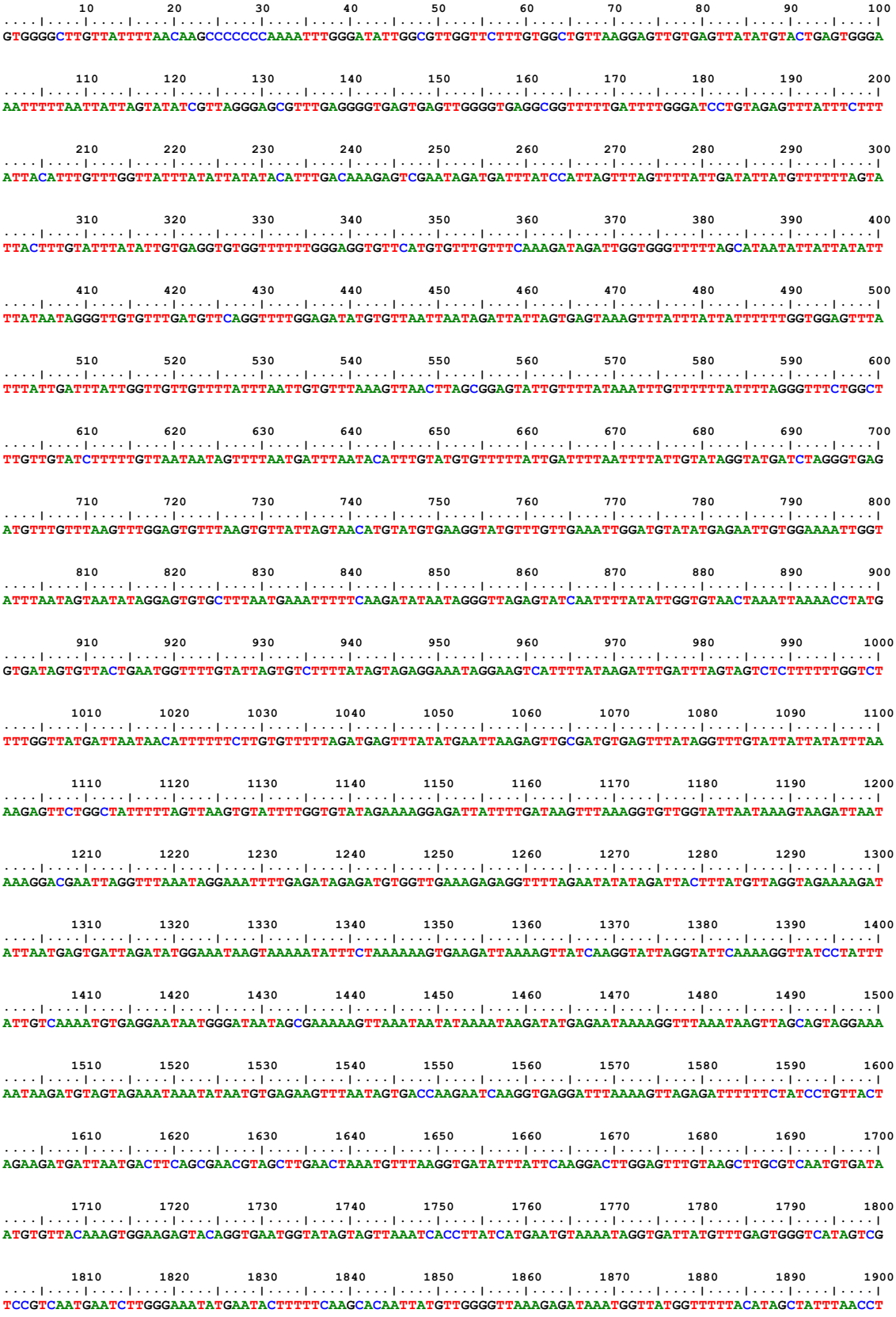
Figure 2.3: Detection limit of conventional nested PCR for *Sphaerothecum destruens* spores in turbid and natural water at Day 20 incubation period. Lane 1-3, 7-9, 13-15: second round PCR products from turbid water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 4-6, 10-12, 16-18: PCR products from natural water spiked with 100K spores/2L, 300K spores/2L and 3M spores/2L. Lane 19, 20, 22, 23: negative controls. Lane 21: PCR +ve control, Lane M: 1Kb DNA ladder.

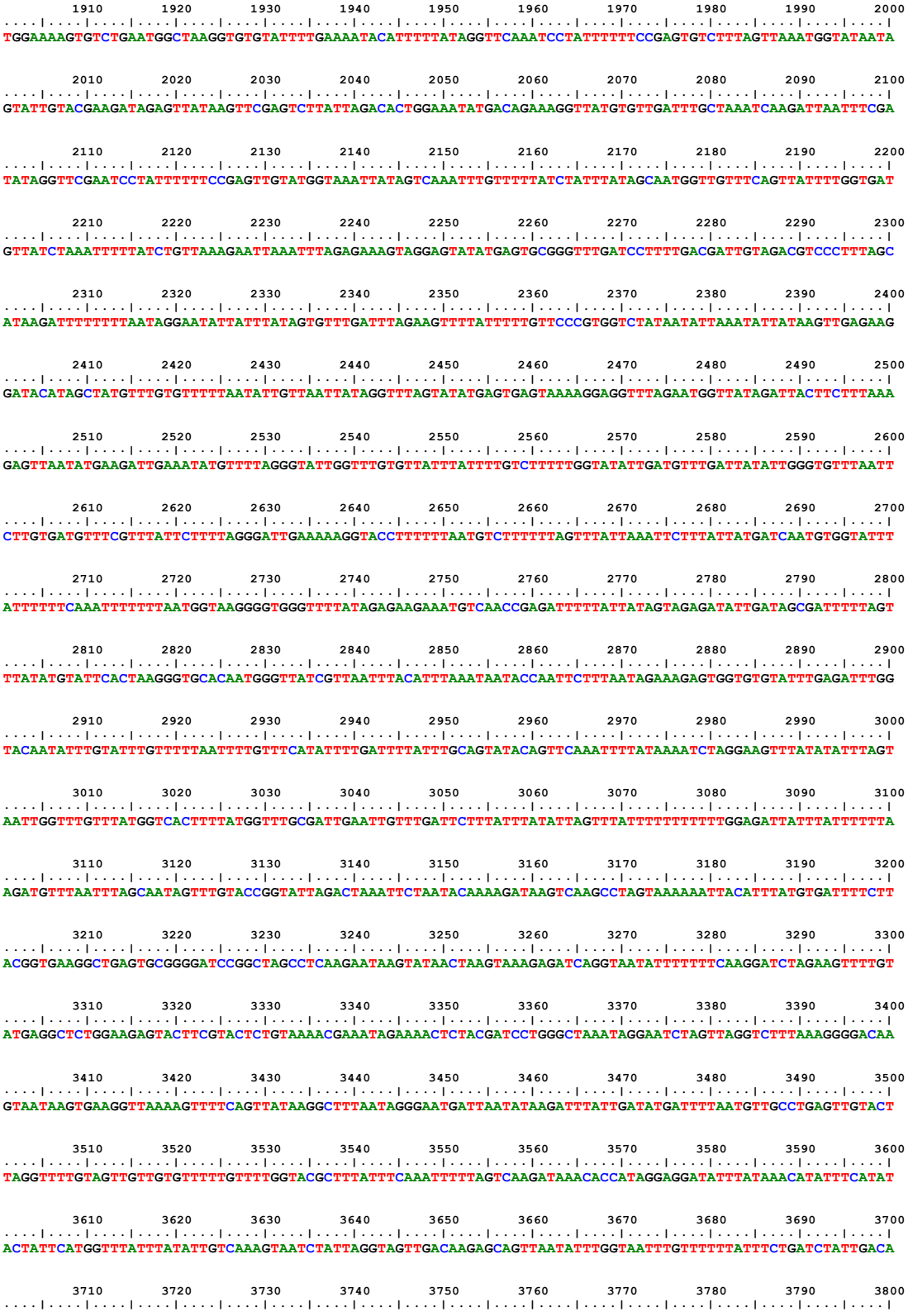
Table 2.1: Detection limit of *Sphaerothecum destruens* DNA in distilled and turbid water over the course of 20 days by conventional PCR. Results are presented as the number of positive replicate/total number of replicates for each treatment.

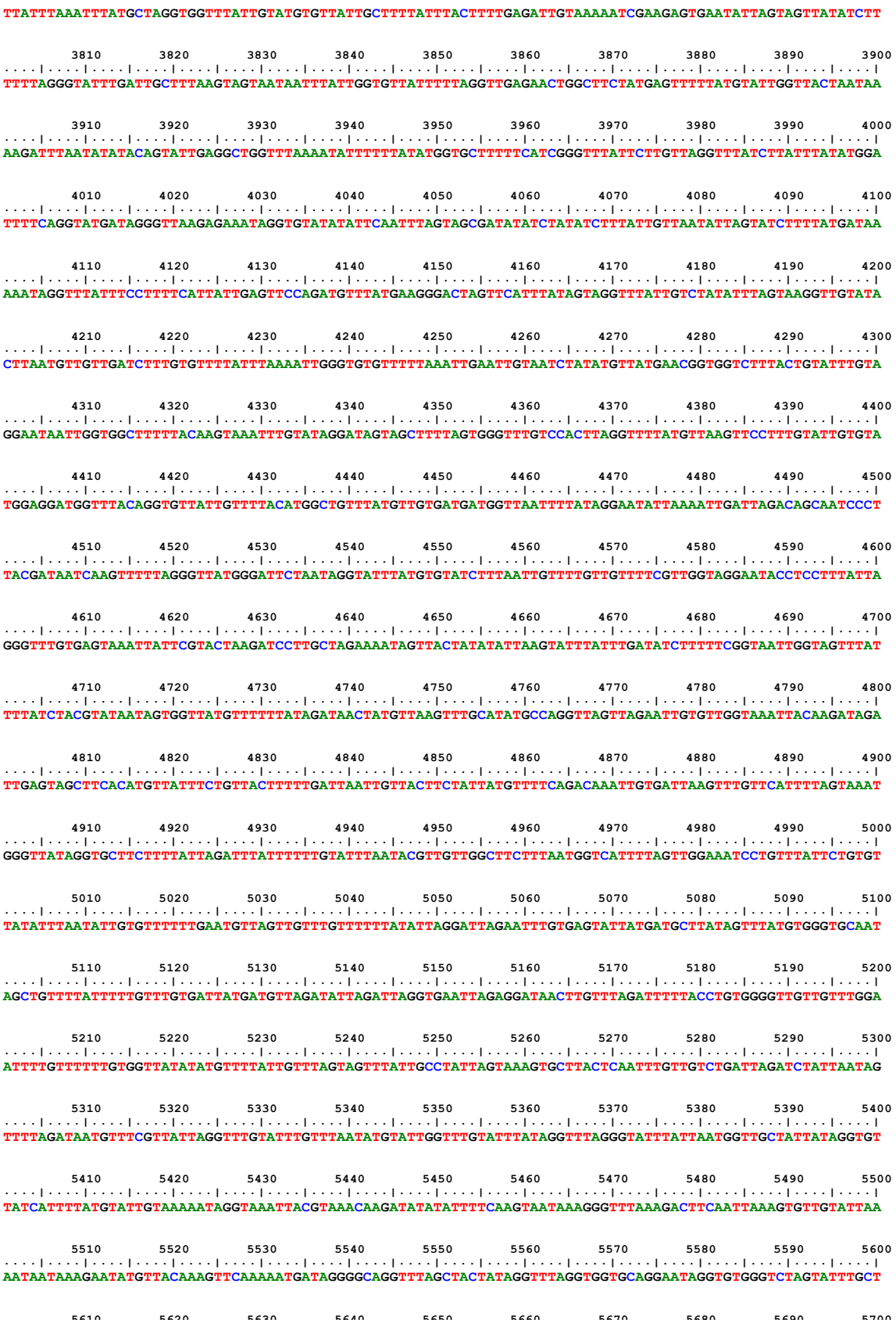
Spore count	**1,500,000/L (High)**		**150,000/L (Medium)**		**50,000/L (Low)**	
Estimated spores filtered	(150,000 spores)		(15,000 spores)		(5000 spores)	
Incubation point	**Sterile water**	**Turbidity**	**Sterile water**	**Turbidity**	**Sterile water**	**Turbidity**
Day 6	3/3	3/3	3/3	3/3	3/3	2/3
Day 13	3/3	2/3	2/3	0/3	3/3	0/3
Day 20	3/3	2/3	3/3	1/3	3/3	0/3

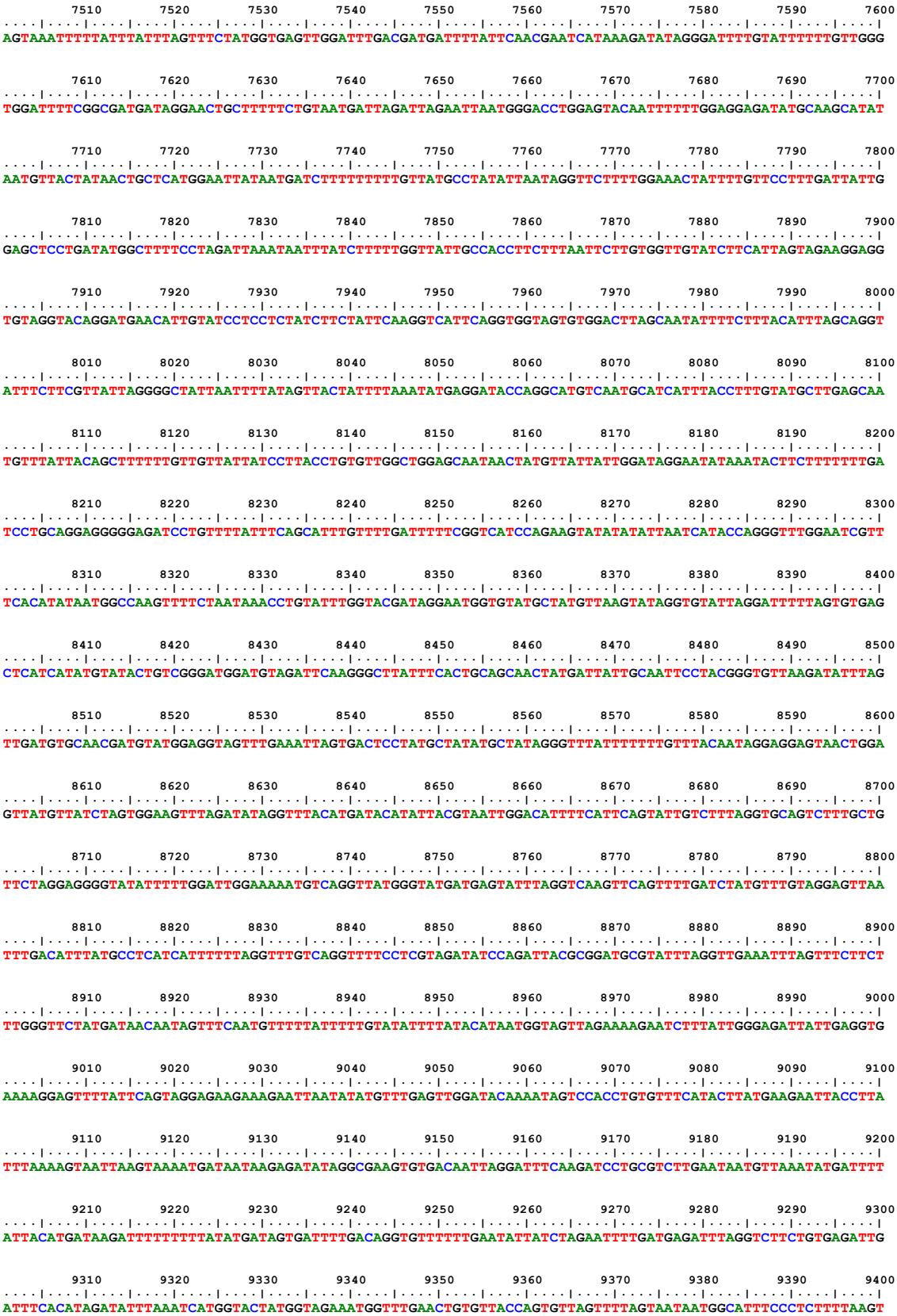
Appendix 3

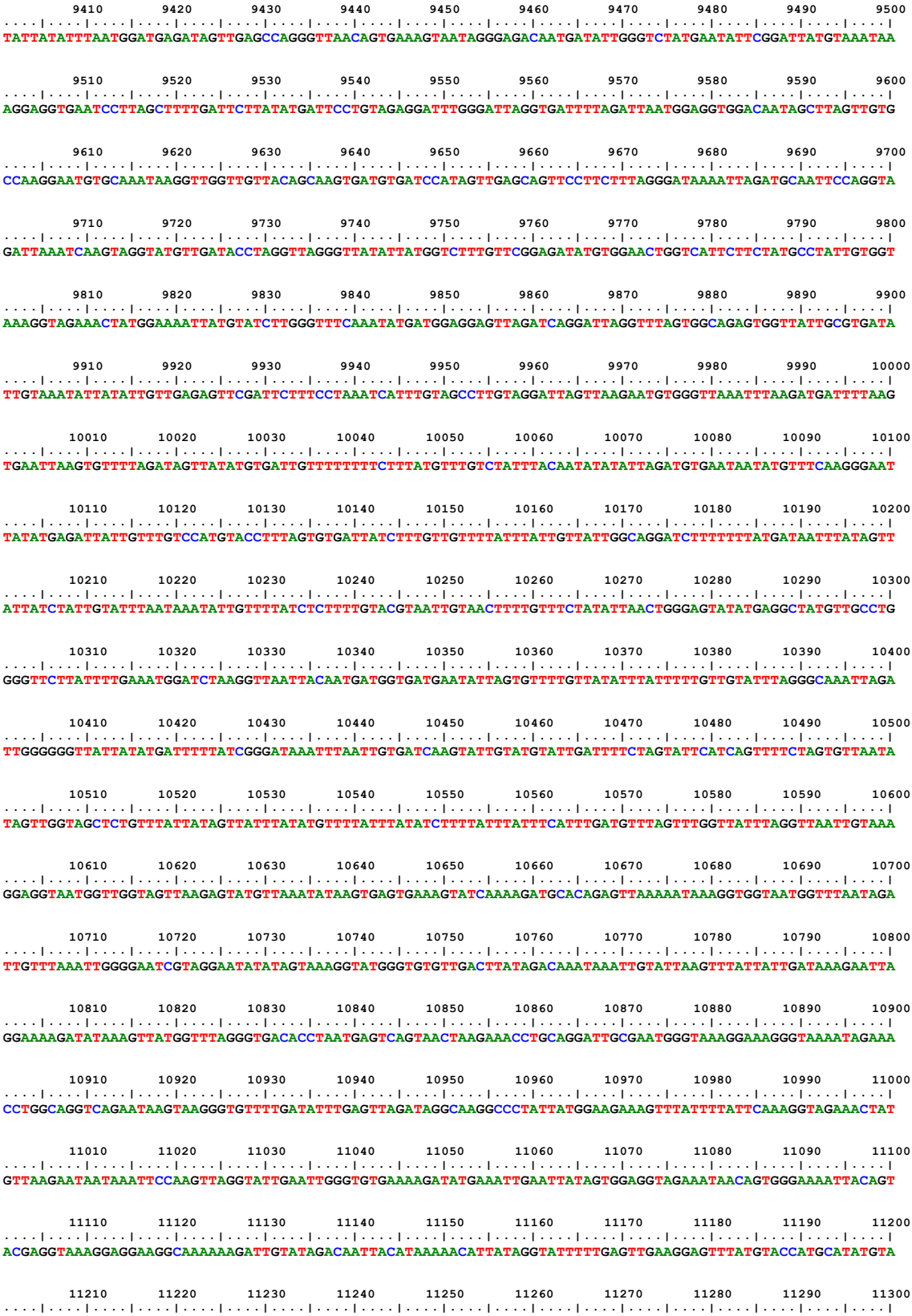
Sphaerothecum destruens mtDNA sequence

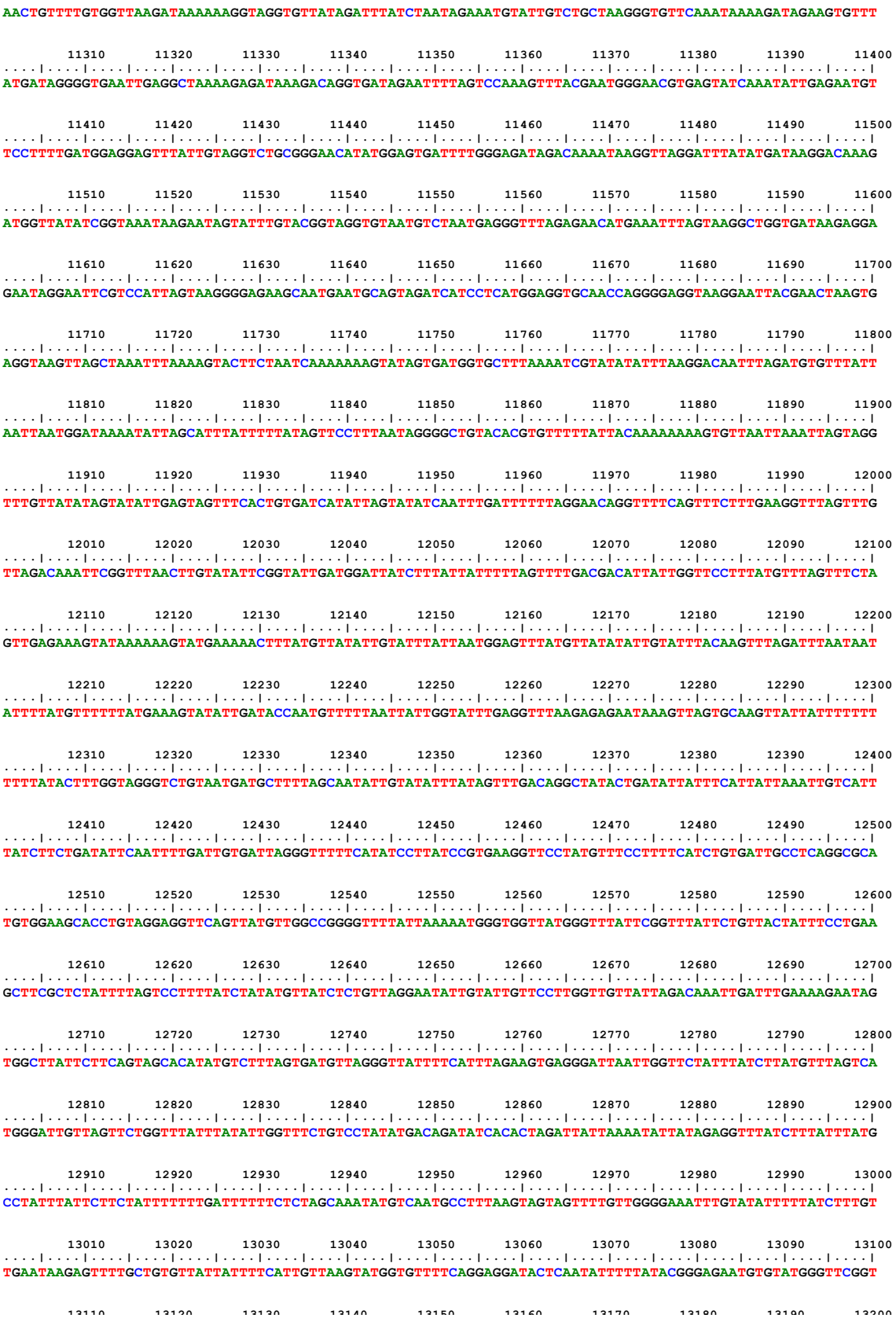


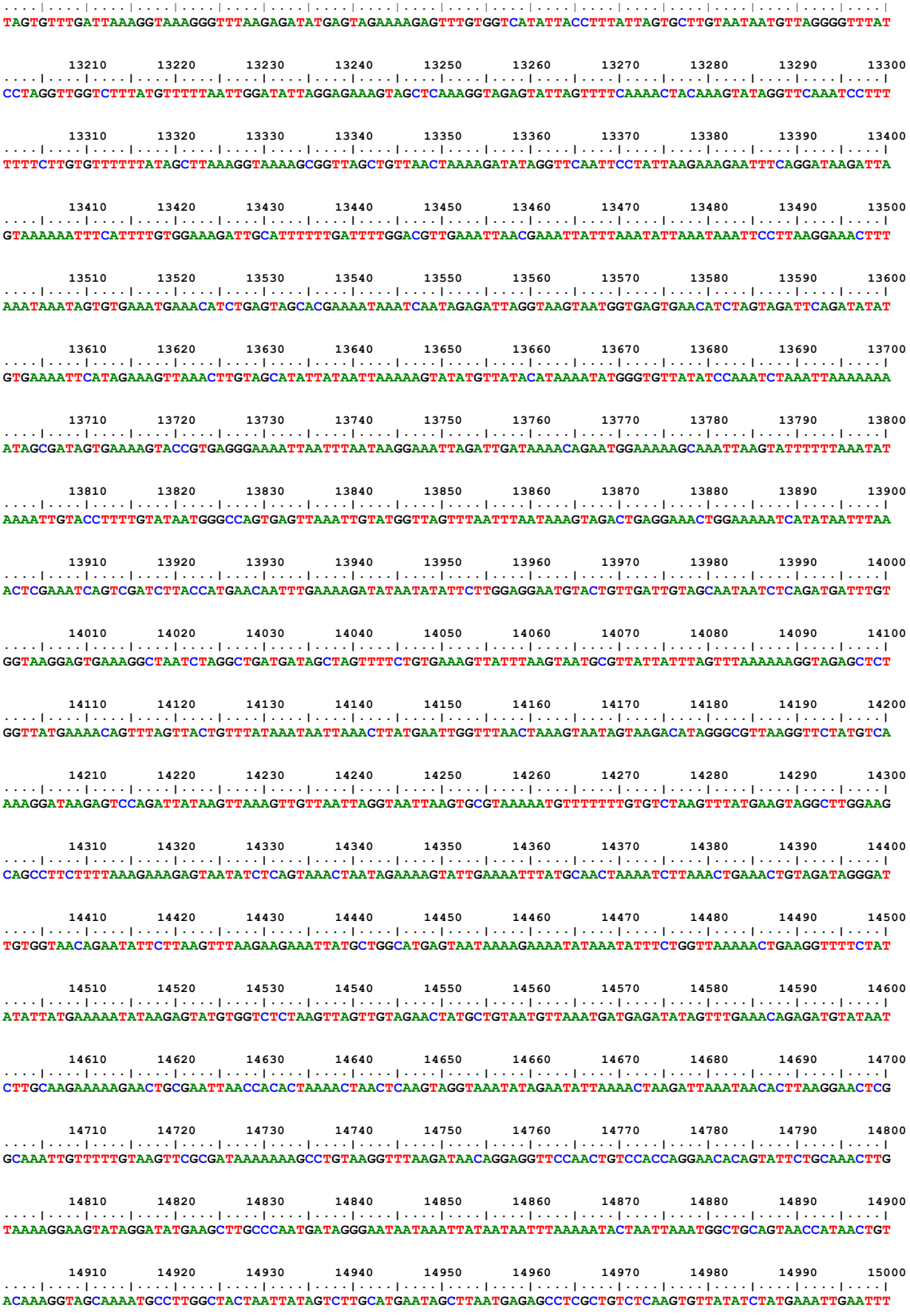

1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
TGGAAAAGTGTCTGAATGGCTAAGGTGTGTATTTTGAAAATACATTTTTATAGGTTCAAATCCTATTTTTTCCGAGTGTCTTTAGTTAAATGGTATAATA
2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
GTATTGTACGAAGATAGAGTTATAAGTTCGAGTCTTATTAGACACTGGAAATATGACAGAAAGGTTATGTGTTGATTTGCTAAATCAAGATTAATTTCGA
2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
TATAGGTTCGAATCCTATTTTTTCCGAGTTGTATGGTAAATTATAGTCAAATTTGTTTTTATCTATTTATAGCAATGGTTGTTTCAGTTATTTTGGTGAT
2210 2220 2230 2240 2250 2260 2270 2280 2290 2300
GTTATCTAAATTTTTATCTGTTAAAGAATTAAATTTAGAGAAAGTAGGAGTATATGAGTGCGGGTTTGATCCTTTTGACGATTGTAGACGTCCCTTTAGC
2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
ATAAGATTTTTTTTAATAGGAATATTATTTATAGTGTTTGATTTAGAAGTTTTATTTTTGTTCCCGTGGTCTATAATATTAAATATTATAAGTTGAGAAG
2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
GATACATAGCTATGTTTGTGTTTTTAATATTGTTAATTATAGGTTTAGTATATGAGTGAGTAAAAGGAGGTTTAGAATGGTTATAGATTACTTCTTTAAA
2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
GAGTTAATATGAAGATTGAAATATGTTTTAGGGTATTGGTTTGTGTTATTTATTTTGTCTTTTTGGTATATTGATGTTTGATTATATTGGGTGTTTAATT
2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
CTTGTGATGTTTCGTTTATTCTTTTAGGGATTGAAAAAGGTACCTTTTTTAATGTCTTTTTTAGTTTATTAAATTCTTTATTATGATCAATGTGGTATTT
2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
ATTTTTTCAAATTTTTTTAATGGTAAGGGGTGGGTTTTATAGAGAAGAAATGTCAACCGAGATTTTTATTATAGTAGAGATATTGATAGCGATTTTTAGT
2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
TTATATGTATTCACTAAGGGTGCACAATGGGTTATCGTTAATTTACATTTAAATAATACCAATTCTTTAATAGAAAGAGTGGTGTGTATTTGAGATTTGG
2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
TACAATATTTGTATTTGTTTTTAATTTTGTTTCATATTTTGATTTTATTTGCAGTATACAGTTCAAATTTTATAAAATCTAGGAAGTTTATATATTTAGT
3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
AATTGGTTTGTTTATGGTCACTTTTATGGTTTGCGATTGAATTGTTTGATTCTTTATTTATATTAGTTTATTTTTTTTTTTGGAGATTATTTATTTTTTA
3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
AGATGTTTAATTTAGCAATAGTTTGTACCGGTATTAGACTAAATTCTAATACAAAAGATAAGTCAAGCCTAGTAAAAAATTACATTTATGTGATTTTCTT
3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
ACGGTGAAGGCTGAGTGCGGGGATCCGGCTAGCCTCAAGAATAAGTATAACTAAGTAAAGAGATCAGGTAATATTTTTTTCAAGGATCTAGAAGTTTTGT
3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
ATGAGGCTCTGGAAGAGTACTTCGTACTCTGTAAAACGAAATAGAAAACTCTACGATCCTGGGCTAAATAGGAATCTAGTTAGGTCTTTAAAGGGGACAA
3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
GTAATAAGTGAAGGTTAAAAGTTTTCAGTTATAAGGCTTTAATAGGGAATGATTAATATAAGATTTATTGATATGATTTTAATGTTGCCTGAGTTGTACT
3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
TAGGTTTTGTAGTTGTTGTGTTTTTGTTTTGGTACGCTTTATTTCAAATTTTTAGTCAAGATAAACACCATAGGAGGATATTTATAAACATATTTCATAT
3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
ACTATTCATGGTTTATTTATATTGTCAAAGTAATCTATTAGGTAGTTGACAAGAGCAGTTAATATTTGGTAATTTGTTTTTTATTTCTGATCTATTGACA
3710 3720 3730 3740 3750 3760 3770 3780 3790 3800

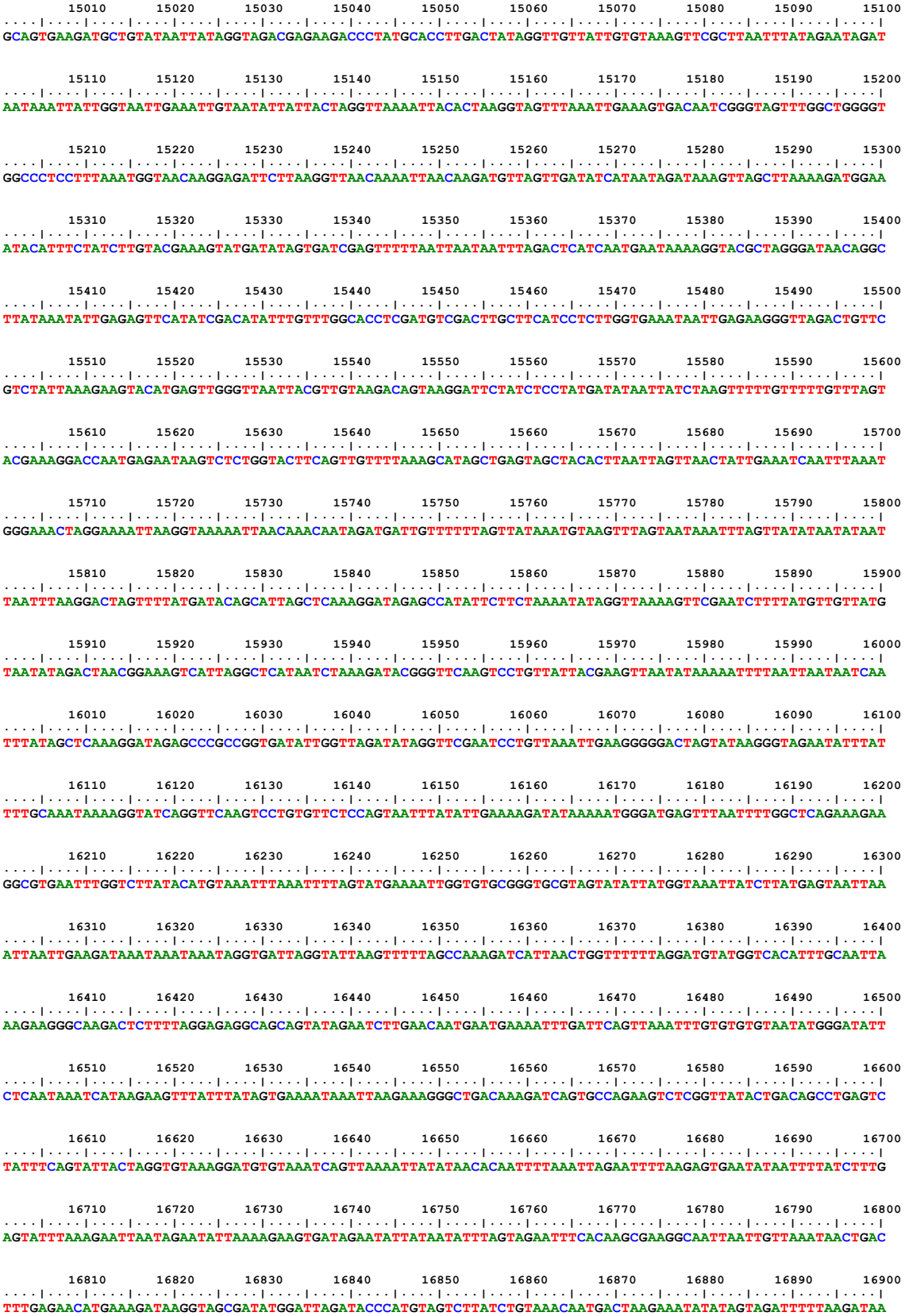

TTATTTAAATTTATGCTAGGTGGTTTATTGTATGTGTTATTGCTTTTATTTACTTTTGAGATTGTAAAAATCGAAGAGTGAATATTAGTAGTTATATCTT
3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
TTTTAGGGTATTTGATTGCTTTAAGTAGTAATAATTTATTGGTGTTATTTTTAGGTTGAGAACTGGCTTCTATGAGTTTTTATGTATTGGTTACTAATAA
3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
AAGATTTAATATATACAGTATTGAGGCTGGTTTAAAATATTTTTTTATATGGTGCTTTTTCATCGGGTTTATTCTTGTTAGGTTTATCTTATTTATATGGA
4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
TTTTCAGGTATGATAGGGTTAAGAGAAATAGGTGTATATATTCAATTTAGTAGCGATATATCTATATCTTTATTGTTAATATTAGTATCTTTTATGATAA
4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
AAATAGGTTTATTTCCTTTTCATTATTGAGTTCCAGATGTTTATGAAGGGACTAGTTCATTTATAGTAGGTTTATTGTCTATATTTAGTAAGGTTGTATA
4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
CTTAATGTTGTTGATCTTTGTGTTTTATTTAAAATTGGGTGTGTTTTTAAATTGAATTGTAATCTATATGTTATGAACGGTGGTCTTTACTGTATTTGTA
4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
GGAATAATTGGTGGCTTTTTACAAGTAAATTTGTATAGGATAGTAGCTTTTAGTGGGTTTGTCCACTTAGGTTTTATGTTAAGTTCCTTTGTATTGTGTA
4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
TGGAGGATGGTTTACAGGTGTTATTGTTTTACATGGCTGTTTATGTTGTGATGATGGTTAATTTTATAGGAATATTAAAATTGATTAGACAGCAATCCCT
4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
TACGATAATCAAGTTTTTAGGGTTATGGGATTCTAATAGGTATTTATGTGTATCTTTAATTGTTTTGTTGTTTTCGTTGGTAGGAATACCTCCTTTATTA
4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
GGGTTTGTGAGTAAATTATTCGTACTAAGATCCTTGCTAGAAAATAGTTACTATATATTAAGTATTTATTTGATATCTTTTTCGGTAATTGGTAGTTTAT
4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
TTTATCTACGTATAATAGTGGTTATGTTTTTTATAGATAACTATGTTAAGTTTGCATATGCCAGGTTAGTTAGAATTGTGTTGGTAAATTACAAGATAGA
4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
TTGAGTAGCTTCACATGTTATTTCTGTTACTTTTTGATTAATTGTTACTTCTATTATGTTTTCAGACAAATTGTGATTAAGTTTGTTCATTTTAGTAAAT
4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
GGGTTATAGGTGCTTCTTTTATTAGATTTATTTTTTGTATTTAATACGTTGTTGGCTTCTTTAATGGTCATTTTAGTTGGAAATCCTGTTTATTCTGTGT
5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
TATATTTAATATTGTGTTTTTTGAATGTTAGTTGTTTGTTTTTTATATTAGGATTAGAATTTGTGAGTATTATGATGCTTATAGTTTATGTGGGTGCAAT
5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
AGCTGTTTTATTTTTGTTTGTGATTATGATGTTAGATATTAGATTAGGTGAATTAGAGGATAACTTGTTTAGATTTTTACCTGTGGGGTTGTTGTTTGGA
5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
ATTTTGTTTTTTGTGGTTATATATGTTTTATTGTTTAGTAGTTTATTGCCTATTAGTAAAGTGCTTACTCAATTTGTTGTCTGATTAGATCTATTAATAG
5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
TTTTAGATAATGTTTCGTTATTAGGTTTGTATTTGTTTAATATGTATTGGTTTGTATTTATAGGTTTAGGGTATTTATTAATGGTTGCTATTATAGGTGT
5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
TATCATTTTATGTATTGTAAAAATAGGTAAATTACGTAAACAAGATATATATTTTCAAGTAATAAAGGGTTTAAAGACTTCAATTAAAGTGTTGTATTAA
5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
AATAATAAAGAATATGTTACAAAGTTCAAAAATGATAGGGGCAGGTTTAGCTACTATAGGTTTAGGTGGTGCAGGAATAGGTGTGGGTCTAGTATTTGCT

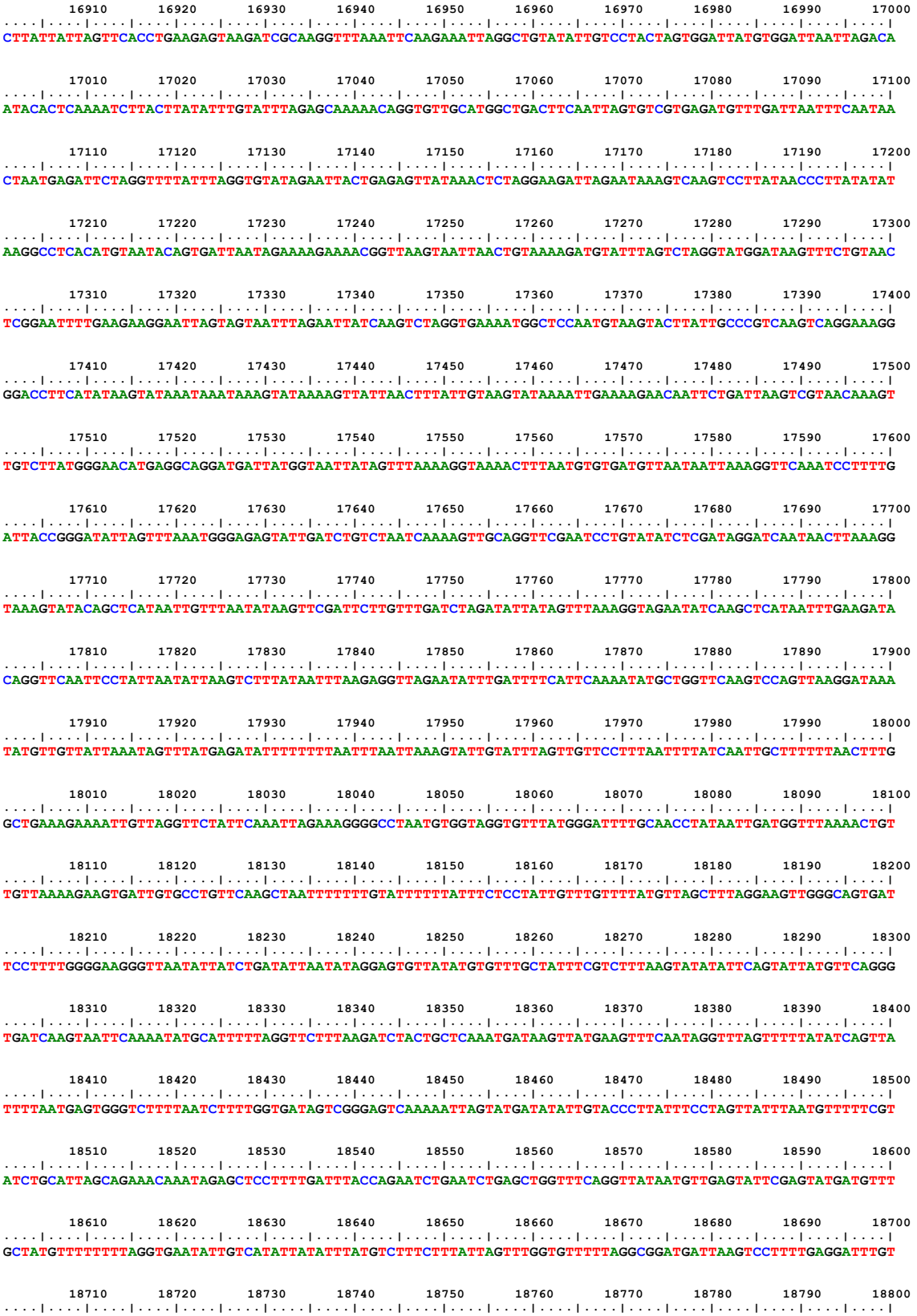

7510 7520 7530 7540 7550 7560 7570 7580 7590 7600
AGTAAATTTTTATTTATTTAGTTTCTATGGTGAGTTGGATTTGACGATGATTTTATTCAACGAATCATAAAGATATAGGGATTTTGTATTTTTTGTTGGG
7610 7620 7630 7640 7650 7660 7670 7680 7690 7700
TGGATTTTCGGCGATGATAGGAACTGCTTTTTCTGTAATGATTAGATTAGAATTAATGGGACCTGGAGTACAATTTTTTGGAGGAGATATGCAAGCATAT
7710 7720 7730 7740 7750 7760 7770 7780 7790 7800
AATGTTACTATAACTGCTCATGGAATTATAATGATCTTTTTTTTTGTTATGCCTATATTAATAGGTTCTTTTGGAAACTATTTTGTTCCTTTGATTATTG
7810 7820 7830 7840 7850 7860 7870 7880 7890 7900
GAGCTCCTGATATGGCTTTTCCTAGATTAAATAATTTATCTTTTTGGTTATTGCCACCTTCTTTAATTCTTGTGGTTGTATCTTCATTAGTAGAAGGAGG
7910 7920 7930 7940 7950 7960 7970 7980 7990 8000
TGTAGGTACAGGATGAACATTGTATCCTCCTCTATCTTCTATTCAAGGTCATTCAGGTGGTAGTGTGGACTTAGCAATATTTTCTTTACATTTAGCAGGT
8010 8020 8030 8040 8050 8060 8070 8080 8090 8100
ATTTCTTCGTTATTAGGGGCTATTAATTTTATAGTTACTATTTTAAATATGAGGATACCAGGCATGTCAATGCATCATTTACCTTTGTATGCTTGAGCAA
8110 8120 8130 8140 8150 8160 8170 8180 8190 8200
TGTTTATTACAGCTTTTTTGTTGTTATTATCCTTACCTGTGTTGGCTGGAGCAATAACTATGTTATTATTGGATAGGAATATAAATACTTCTTTTTTTGA
8210 8220 8230 8240 8250 8260 8270 8280 8290 8300
TCCTGCAGGAGGGGGAGATCCTGTTTTATTTCAGCATTTGTTTTGATTTTTCGGTCATCCAGAAGTATATATATTAATCATACCAGGGTTTGGAATCGTT
8310 8320 8330 8340 8350 8360 8370 8380 8390 8400
TCACATATAATGGCCAAGTTTTCTAATAAACCTGTATTTGGTACGATAGGAATGGTGTATGCTATGTTAAGTATAGGTGTATTAGGATTTTTAGTGTGAG
8410 8420 8430 8440 8450 8460 8470 8480 8490 8500
CTCATCATATGTATACTGTCGGGATGGATGTAGATTCAAGGGCTTATTTCACTGCAGCAACTATGATTATTGCAATTCCTACGGGTGTTAAGATATTTAG
8510 8520 8530 8540 8550 8560 8570 8580 8590 8600
TTGATGTGCAACGATGTATGGAGGTAGTTTGAAATTAGTGACTCCTATGCTATATGCTATAGGGTTTATTTTTTTGTTTACAATAGGAGGAGTAACTGGA
8610 8620 8630 8640 8650 8660 8670 8680 8690 8700
GTTATGTTATCTAGTGGAAGTTTAGATATAGGTTTACATGATACATATTACGTAATTGGACATTTTCATTCAGTATTGTCTTTAGGTGCAGTCTTTGCTG
8710 8720 8730 8740 8750 8760 8770 8780 8790 8800
TTCTAGGAGGGGTATATTTTTGGATTGGAAAAATGTCAGGTTATGGGTATGATGAGTATTTAGGTCAAGTTCAGTTTTGATCTATGTTTGTAGGAGTTAA
8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
TTTGACATTTATGCCTCATCATTTTTTAGGTTTGTCAGGTTTTCCTCGTAGATATCCAGATTACGCGGATGCGTATTTAGGTTGAAATTTAGTTTCTTCT
8910 8920 8930 8940 8950 8960 8970 8980 8990 9000
TTGGGTTCTATGATAACAATAGTTTCAATGTTTTTATTTTTGTATATTTTATACATAATGGTAGTTAGAAAAGAATCTTTATTGGGAGATTATTGAGGTG
9010 9020 9030 9040 9050 9060 9070 9080 9090 9100
AAAAGGAGTTTTATTCAGTAGGAGAAGAAAGAATTAATATATGTTTGAGTTGGATACAAAATAGTCCACCTGTGTTTCATACTTATGAAGAATTACCTTA
9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
TTTAAAAGTAATTAAGTAAAATGATAATAAGAGATATAGGCGAAGTGTGACAATTAGGATTTCAAGATCCTGCGTCTTGAATAATGTTAAATATGATTTT
9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
ATTACATGATAAGATTTTTTTTTATATGATAGTGATTTTGACAGGTGTTTTTTGAATATTATCTAGAATTTTGATGAGATTTAGGTCTTCTGTGAGATTG
9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
ATTTCACATAGATATTTAAATCATGGTACTATGGTAGAAATGGTTTGAACTGTGTTACCAGTGTTAGTTTTAGTAATAATGGCATTTCCCTCTTTTAAGT

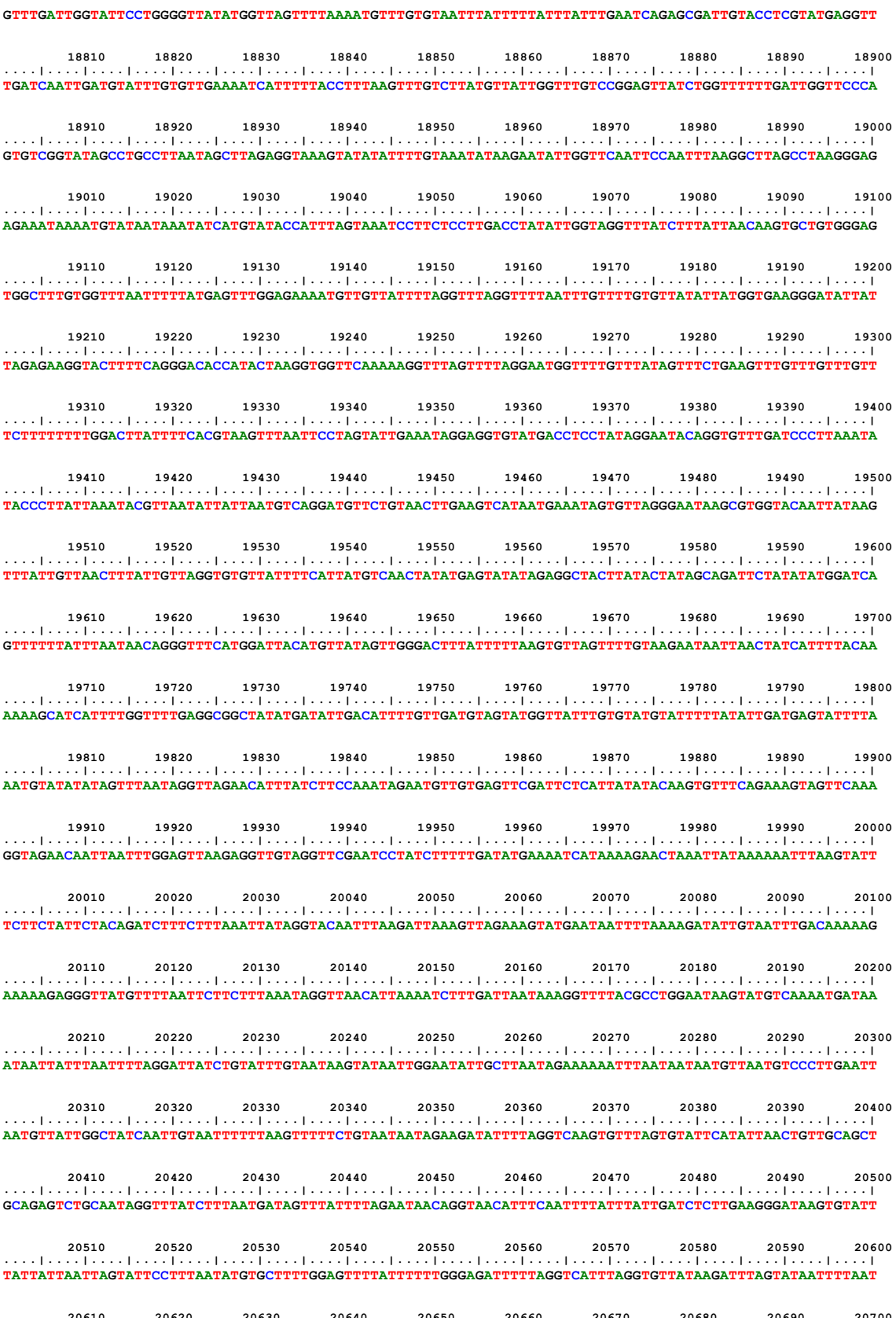

9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
TATTATATTTAATGGATGAGATAGTTGAGCCAGGGTTAACAGTGAAAGTAATAGGGAGACAATGATATTGGGTCTATGAATATTCGGATTATGTAAATAA
9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
AGGAGGTGAATCCTTAGCTTTTGATTCTTATATGATTCCTGTAGAGGATTTGGGATTAGGTGATTTTAGATTAATGGAGGTGGACAATAGCTTAGTTGTG
9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
CCAAGGAATGTGCAAATAAGGTTGGTTGTTACAGCAAGTGATGTGATCCATAGTTGAGCAGTTCCTTCTTTAGGGATAAAATTAGATGCAATTCCAGGTA
9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
GATTAAATCAAGTAGGTATGTTGATACCTAGGTTAGGGTTATATTATGGTCTTTGTTCGGAGATATGTGGAACTGGTCATTCTTCTATGCCTATTGTGGT
9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
AAAGGTAGAAACTATGGAAAATTATGTATCTTGGGTTTCAAATATGATGGAGGAGTTAGATCAGGATTAGGTTTAGTGGCAGAGTGGTTATTGCGTGATA
9910 9920 9930 9940 9950 9960 9970 9980 9990 10000
TTGTAAATATTATATTGTTGAGAGTTCGATTCTTTCCTAAATCATTTGTAGCCTTGTAGGATTAGTTAAGAATGTGGGTTAAATTTAAGATGATTTTAAG
10010 10020 10030 10040 10050 10060 10070 10080 10090 10100
TGAATTAAGTGTTTTAGATAGTTATATGTGATTGTTTTTTTTCTTTATGTTTGTCTATTTACAATATATATTAGATGTGAATAATATGTTTCAAGGGAAT
10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
TATATGAGATTATTGTTTGTCCATGTACCTTTAGTGTGATTATCTTTGTTGTTTTATTTATTGTTATTGGCAGGATCTTTTTTTTATGATAATTTATAGTT
10210 10220 10230 10240 10250 10260 10270 10280 10290 10300
ATTATCTATTGTATTTAATAAATATTGTTTTATCTCTTTTGTACGTAATTGTAACTTTTGTTTCTATATTAACTGGGAGTATATGAGGCTATGTTGCCTG
10310 10320 10330 10340 10350 10360 10370 10380 10390 10400
GGGTTCTTATTTTGAAATGGATCTAAGGTTAATTACAATGATGGTGATGAATATTAGTGTTTTGTTATATTTATTTTTGTTGTATTTAGGGCAAATTAGA
10410 10420 10430 10440 10450 10460 10470 10480 10490 10500
TTGGGGGGTTATTATATGATTTTTATCGGGATAAATTTAATTGTGATCAAGTATTGTATGTATTGATTTTCTAGTATTCATCAGTTTTCTAGTGTTAATA
10510 10520 10530 10540 10550 10560 10570 10580 10590 10600
TAGTTGGTAGCTCTGTTTATTATAGTTATTTATATGTTTTATTTATATCTTTTATTTATTTCATTTGATGTTTAGTTTGGTTATTTAGGTTAATTGTAAA
10610 10620 10630 10640 10650 10660 10670 10680 10690 10700
GGAGGTAATGGTTGGTAGTTAAGAGTATGTTAAATATAAGTGAGTGAAAGTATCAAAAGATGCACAGAGTTAAAAATAAAGGTGGTAATGGTTTAATAGA
10710 10720 10730 10740 10750 10760 10770 10780 10790 10800
TTGTTTAAATTGGGGAATCGTAGGAATATATAGTAAAGGTATGGGTGTGTTGACTTATAGACAAATAAATTGTATTAAGTTTATTATTGATAAAGAATTA
10810 10820 10830 10840 10850 10860 10870 10880 10890 10900
GGAAAAGATATAAAGTTATGGTTTAGGGTGACACCTAATGAGTCAGTAACTAAGAAACCTGCAGGATTGCGAATGGGTAAAGGAAAGGGTAAAATAGAAA
10910 10920 10930 10940 10950 10960 10970 10980 10990 11000
CCTGGCAGGTCAGAATAAGTAAGGGTGTTTTGATATTTGAGTTAGATAGGCAAGGCCCTATTATGGAAGAAAGTTTATTTTATTCAAAGGTAGAAACTAT
11010 11020 11030 11040 11050 11060 11070 11080 11090 11100
GTTAAGAATAATAAATTCCAAGTTAGGTATTGAATTGGGTGTGAAAAGATATGAAATTGAATTATAGTGGAGGTAGAAATAACAGTGGGAAAATTACAGT
11110 11120 11130 11140 11150 11160 11170 11180 11190 11200
ACGAGGTAAAGGAGGAAGGCAAAAAAGATTGTATAGACAATTACATAAAAACATTATAGGTATTTTTGAGTTGAAGGAGTTTATGTACCATGCATATGTA
11210 11220 11230 11240 11250 11260 11270 11280 11290 11300

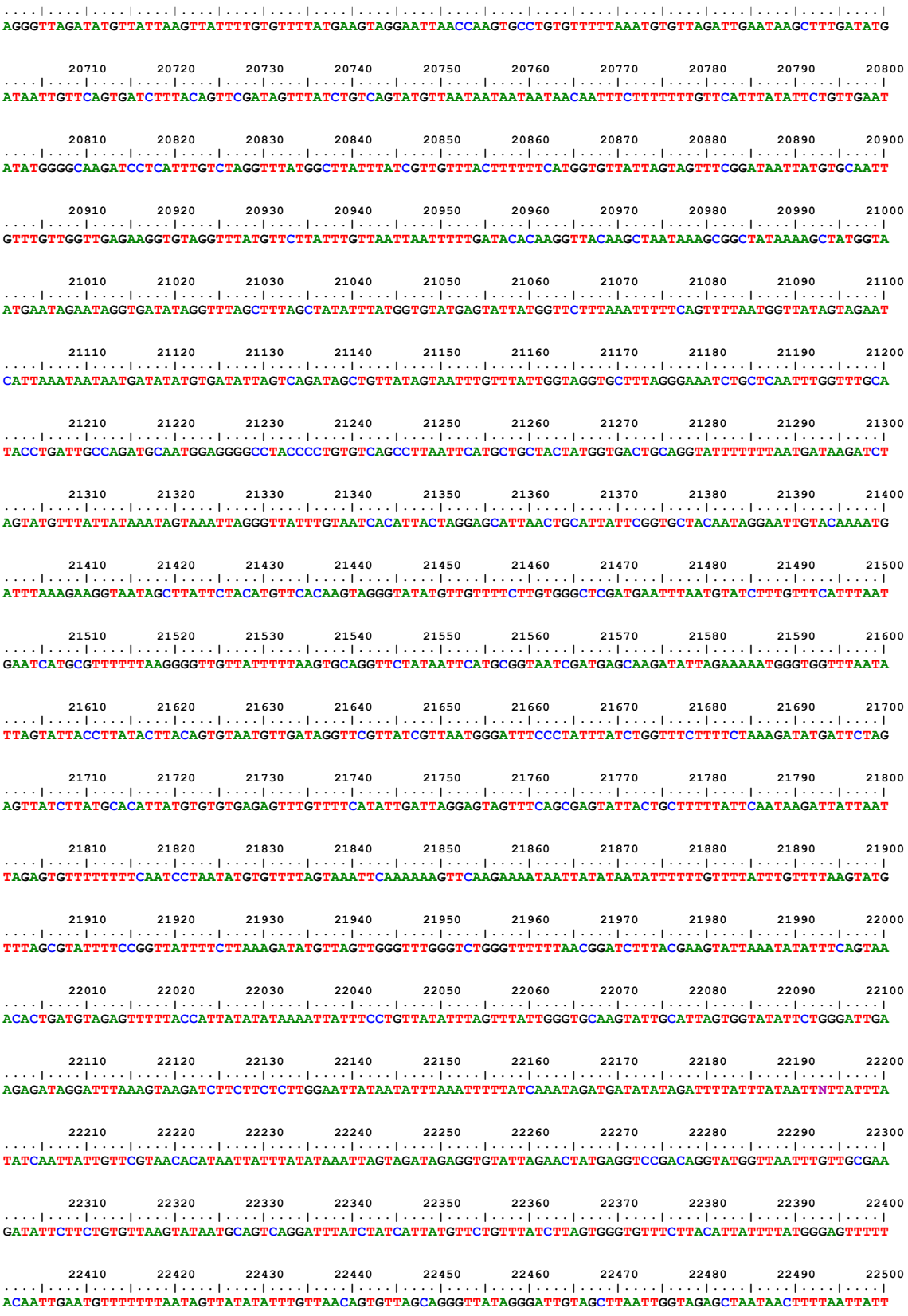

AACTGTTTTGTGGTTAAGATAAAAAAGGTAGGTGTTATAGATTTATCTAATAGAAATGTATTGTCTGCTAAGGGTGTTCAAATAAAAGATAGAAGTGTTT
11310 11320 11330 11340 11350 11360 11370 11380 11390 11400
ATGATAGGGGTGAATTGAGGCTAAAAGAGATAAAGACAGGTGATAGAATTTTAGTCCAAAGTTTACGAATGGGAACGTGAGTATCAAATATTGAGAATGT
11410 11420 11430 11440 11450 11460 11470 11480 11490 11500
TCCTTTTGATGGAGGAGTTTATTGTAGGTCTGCGGGAACATATGGAGTGATTTTGGGAGATAGACAAAATAAGGTTAGGATTTATATGATAAGGACAAAG
11510 11520 11530 11540 11550 11560 11570 11580 11590 11600
ATGGTTATATCGGTAAATAAGAATAGTATTTGTACGGTAGGTGTAATGTCTAATGAGGGTTTAGAGAACATGAAATTTAGTAAGGCTGGTGATAAGAGGA
11610 11620 11630 11640 11650 11660 11670 11680 11690 11700
GAATAGGAATTCGTCCATTAGTAAGGGGAGAAGCAATGAATGCAGTAGATCATCCTCATGGAGGTGCAACCAGGGGAGGTAAGGAATTACGAACTAAGTG
11710 11720 11730 11740 11750 11760 11770 11780 11790 11800
AGGTAAGTTAGCTAAATTTAAAAGTACTTCTAATCAAAAAAAGTATAGTGATGGTGCTTTAAAATCGTATATATTTAAGGACAATTTAGATGTGTTTATT
11810 11820 11830 11840 11850 11860 11870 11880 11890 11900
AATTAATGGATAAAATATTAGCATTTATTTTTATAGTTCCTTTAATAGGGGCTGTACACGTGTTTTTATTACAAAAAAAAGTGTTAATTAAATTAGTAGG
11910 11920 11930 11940 11950 11960 11970 11980 11990 12000
TTTGTTATATAGTATATTGAGTAGTTTCACTGTGATCATATTAGTATATCAATTTGATTTTTTAGGAACAGGTTTTCAGTTTCTTTGAAGGTTTAGTTTG
12010 12020 12030 12040 12050 12060 12070 12080 12090 12100
TTAGACAAATTCGGTTTAACTTGTATATTCGGTATTGATGGATTATCTTTATTATTTTTAGTTTTGACGACATTATTGGTTCCTTTATGTTTAGTTTCTA
12110 12120 12130 12140 12150 12160 12170 12180 12190 12200
GTTGAGAAAGTATAAAAAAGTATGAAAAACTTTATGTTATATTGTATTTATTAATGGAGTTTATGTTATATATTGTATTTACAAGTTTAGATTTAATAAT
12210 12220 12230 12240 12250 12260 12270 12280 12290 12300
ATTTTATGTTTTTTATGAAAGTATATTGATACCAATGTTTTTAATTATTGGTATTTGAGGTTTAAGAGAGAATAAAGTTAGTGCAAGTTATTATTTTTTT
12310 12320 12330 12340 12350 12360 12370 12380 12390 12400
TTTTATACTTTGGTAGGGTCTGTAATGATGCTTTTAGCAATATTGTATATTTATAGTTTGACAGGCTATACTGATATTATTTCATTATTAAATTGTCATT
12410 12420 12430 12440 12450 12460 12470 12480 12490 12500
TATCTTCTGATATTCAATTTTGATTGTGATTAGGGTTTTTCATATCCTTATCCGTGAAGGTTCCTATGTTTCCTTTTCATCTGTGATTGCCTCAGGCGCA
12510 12520 12530 12540 12550 12560 12570 12580 12590 12600
TGTGGAAGCACCTGTAGGAGGTTCAGTTATGTTGGCCGGGGTTTTATTAAAAATGGGTGGTTATGGGTTTATTCGGTTTATTCTGTTACTATTTCCTGAA
12610 12620 12630 12640 12650 12660 12670 12680 12690 12700
GCTTCGCTCTATTTTAGTCCTTTTATCTATATGTTATCTCTGTTAGGAATATTGTATTGTTCCTTGGTTGTTATTAGACAAATTGATTTGAAAAGAATAG
12710 12720 12730 12740 12750 12760 12770 12780 12790 12800
TGGCTTATTCTTCAGTAGCACATATGTCTTTAGTGATGTTAGGGTTATTTTCATTTAGAAGTGAGGGATTAATTGGTTCTATTTATCTTATGTTTAGTCA
12810 12820 12830 12840 12850 12860 12870 12880 12890 12900
TGGGATTGTTAGTTCTGGTTTATTTATATTGGTTTCTGTCCTATATGACAGATATCACACTAGATTATTAAAATATTATAGAGGTTTATCTTTATTTATG
12910 12920 12930 12940 12950 12960 12970 12980 12990 13000
CCTATTTATTCTTCTATTTTTTTGATTTTTTCTCTAGCAAATATGTCAATGCCTTTAAGTAGTAGTTTTGTTGGGGAAATTTGTATATTTTTATCTTTGT
13010 13020 13030 13040 13050 13060 13070 13080 13090 13100
TGAATAAGAGTTTTGCTGTGTTATTATTTTCATTGTTAAGTATGGTGTTTTCAGGAGGATACTCAATATTTTTATACGGGAGAATGTGTATGGGTTCGGT

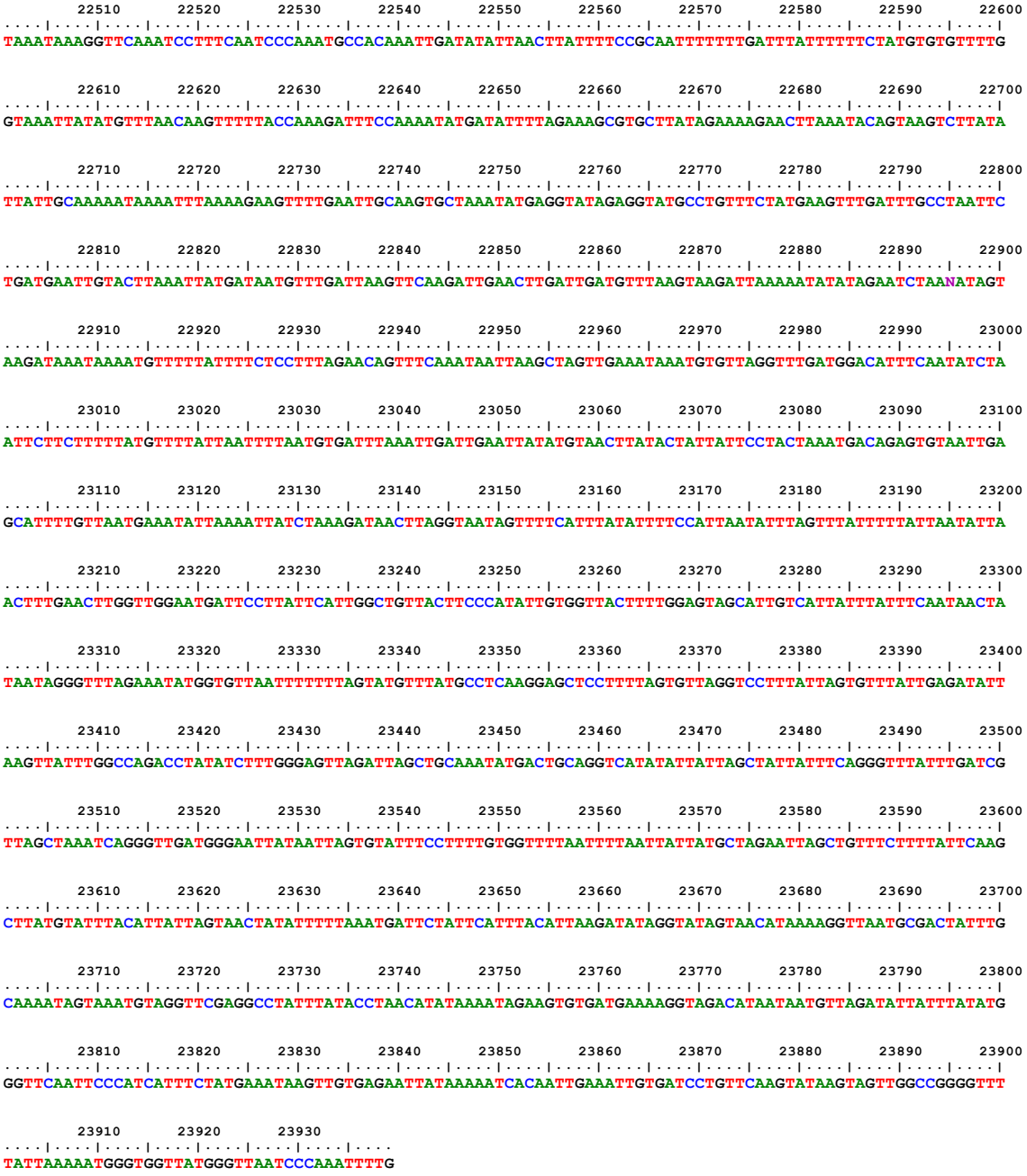

TAGTGTTTGATTAAAGGTAAAGGGTTTAAGAGATATGAGTAGAAAAGAGTTTGTGGTCATATTACCTTTATTAGTGCTTGTAATAATGTTAGGGGTTTAT
13210 13220 13230 13240 13250 13260 13270 13280 13290 13300
CCTAGGTTGGTCTTTATGTTTTTAATTGGATATTAGGAGAAAGTAGCTCAAAGGTAGAGTATTAGTTTTCAAAACTACAAAGTATAGGTTCAAATCCTTT
13310 13320 13330 13340 13350 13360 13370 13380 13390 13400
TTTTCTTGTGTTTTTTATAGCTTAAAGGTAAAAGCGGTTAGCTGTTAACTAAAAGATATAGGTTCAATTCCTATTAAGAAAGAATTTCAGGATAAGATTA
13410 13420 13430 13440 13450 13460 13470 13480 13490 13500
GTAAAAAATTTCATTTTGTGGAAAGATTGCATTTTTTGATTTTGGACGTTGAAATTAACGAAATTATTTAAATATTAAATAAATTCCTTAAGGAAACTTT
13510 13520 13530 13540 13550 13560 13570 13580 13590 13600
AAATAAATAGTGTGAAATGAAACATCTGAGTAGCACGAAAATAAATCAATAGAGATTAGGTAAGTAATGGTGAGTGAACATCTAGTAGATTCAGATATAT
13610 13620 13630 13640 13650 13660 13670 13680 13690 13700
GTGAAAATTCATAGAAAGTTAAACTTGTAGCATATTATAATTAAAAAGTATATGTTATACATAAAATATGGGTGTTATATCCAAATCTAAATTAAAAAAA
13710 13720 13730 13740 13750 13760 13770 13780 13790 13800
ATAGCGATAGTGAAAAGTACCGTGAGGGAAAATTAATTTAATAAGGAAATTAGATTGATAAAACAGAATGGAAAAAGCAAATTAAGTATTTTTTAAATAT
13810 13820 13830 13840 13850 13860 13870 13880 13890 13900
AAAATTGTACCTTTTGTATAATGGGCCAGTGAGTTAAATTGTATGGTTAGTTTAATTTAATAAAGTAGACTGAGGAAACTGGAAAAATCATATAATTTAA
13910 13920 13930 13940 13950 13960 13970 13980 13990 14000
ACTCGAAATCAGTCGATCTTACCATGAACAATTTGAAAAGATATAATATATTCTTGGAGGAATGTACTGTTGATTGTAGCAATAATCTCAGATGATTTGT
14010 14020 14030 14040 14050 14060 14070 14080 14090 14100
GGTAAGGAGTGAAAGGCTAATCTAGGCTGATGATAGCTAGTTTTCTGTGAAAGTTATTTAAGTAATGCGTTATTATTTAGTTTAAAAAAGGTAGAGCTCT
14110 14120 14130 14140 14150 14160 14170 14180 14190 14200
GGTTATGAAAACAGTTTAGTTACTGTTTATAAATAATTAAACTTATGAATTGGTTTAACTAAAGTAATAGTAAGACATAGGGCGTTAAGGTTCTATGTCA
14210 14220 14230 14240 14250 14260 14270 14280 14290 14300
AAAGGATAAGAGTCCAGATTATAAGTTAAAGTTGTTAATTAGGTAATTAAGTGCGTAAAAATGTTTTTTTGTGTCTAAGTTTATGAAGTAGGCTTGGAAG
14310 14320 14330 14340 14350 14360 14370 14380 14390 14400
CAGCCTTCTTTTAAAGAAAGAGTAATATCTCAGTAAACTAATAGAAAAGTATTGAAAATTTATGCAACTAAAATCTTAAACTGAAACTGTAGATAGGGAT
14410 14420 14430 14440 14450 14460 14470 14480 14490 14500
TGTGGTAACAGAATATTCTTAAGTTTAAGAAGAAATTATGCTGGCATGAGTAATAAAAGAAAATATAAATATTTCTGGTTAAAAACTGAAGGTTTTCTAT
14510 14520 14530 14540 14550 14560 14570 14580 14590 14600
ATATTATGAAAAATATAAGAGTATGTGGTCTCTAAGTTAGTTGTAGAACTATGCTGTAATGTTAAATGATGAGATATAGTTTGAAACAGAGATGTATAAT
14610 14620 14630 14640 14650 14660 14670 14680 14690 14700
CTTGCAAGAAAAAGAACTGCGAATTAACCACACTAAAACTAACTCAAGTAGGTAAATATAGAATATTAAAACTAAGATTAAATAACACTTAAGGAACTCG
14710 14720 14730 14740 14750 14760 14770 14780 14790 14800
GCAAATTGTTTTTGTAAGTTCGCGATAAAAAAAGCCTGTAAGGTTTAAGATAACAGGAGGTTCCAACTGTCCACCAGGAACACAGTATTCTGCAAACTTG
14810 14820 14830 14840 14850 14860 14870 14880 14890 14900
TAAAAGGAAGTATAGGATATGAAGCTTGCCCAATGATAGGGAATAATAAATTATAATAATTTAAAAATACTAATTAAATGGCTGCAGTAACCATAACTGT
14910 14920 14930 14940 14950 14960 14970 14980 14990 15000
ACAAAGGTAGCAAAATGCCTTGGCTACTAATTATAGTCTTGCATGAATAGCTTAATGAGAGCCTCGCTGTCTCAAGTGTTATATCTATGAAATTGAATTT


15010 15020 15030 15040 15050 15060 15070 15080 15090 15100
GCAGTGAAGATGCTGTATAATTATAGGTAGACGAGAAGACCCTATGCACCTTGACTATAGGTTGTTATTGTGTAAAGTTCGCTTAATTTATAGAATAGAT
15110 15120 15130 15140 15150 15160 15170 15180 15190 15200
AATAAATTATTGGTAATTGAAATTGTAATATTATTACTAGGTTAAAATTACACTAAGGTAGTTTAAATTGAAAGTGACAATCGGGTAGTTTGGCTGGGGT
15210 15220 15230 15240 15250 15260 15270 15280 15290 15300
GGCCCTCCTTTAAATGGTAACAAGGAGATTCTTAAGGTTAACAAAATTAACAAGATGTTAGTTGATATCATAATAGATAAAGTTAGCTTAAAAGATGGAA
15310 15320 15330 15340 15350 15360 15370 15380 15390 15400
ATACATTTCTATCTTGTACGAAAGTATGATATAGTGATCGAGTTTTTAATTAATAATTTAGACTCATCAATGAATAAAAGGTACGCTAGGGATAACAGGC
15410 15420 15430 15440 15450 15460 15470 15480 15490 15500
TTATAAATATTGAGAGTTCATATCGACATATTTGTTTGGCACCTCGATGTCGACTTGCTTCATCCTCTTGGTGAAATAATTGAGAAGGGTTAGACTGTTC
15510 15520 15530 15540 15550 15560 15570 15580 15590 15600
GTCTATTAAAGAAGTACATGAGTTGGGTTAATTACGTTGTAAGACAGTAAGGATTCTATCTCCTATGATATAATTATCTAAGTTTTTGTTTTTGTTTAGT
15610 15620 15630 15640 15650 15660 15670 15680 15690 15700
ACGAAAGGACCAATGAGAATAAGTCTCTGGTACTTCAGTTGTTTTAAAGCATAGCTGAGTAGCTACACTTAATTAGTTAACTATTGAAATCAATTTAAAT
15710 15720 15730 15740 15750 15760 15770 15780 15790 15800
GGGAAACTAGGAAAATTAAGGTAAAAATTAACAAACAATAGATGATTGTTTTTTAGTTATAAATGTAAGTTTAGTAATAAATTTAGTTATATAATATAAT
15810 15820 15830 15840 15850 15860 15870 15880 15890 15900
TAATTTAAGGACTAGTTTTATGATACAGCATTAGCTCAAAGGATAGAGCCATATTCTTCTAAAATATAGGTTAAAAGTTCGAATCTTTTATGTTGTTATG
15910 15920 15930 15940 15950 15960 15970 15980 15990 16000
TAATATAGACTAACGGAAAGTCATTAGGCTCATAATCTAAAGATACGGGTTCAAGTCCTGTTATTACGAAGTTAATATAAAAATTTTAATTAATAATCAA
16010 16020 16030 16040 16050 16060 16070 16080 16090 16100
TTTATAGCTCAAAGGATAGAGCCCGCCGGTGATATTGGTTAGATATAGGTTCGAATCCTGTTAAATTGAAGGGGGACTAGTATAAGGGTAGAATATTTAT
16110 16120 16130 16140 16150 16160 16170 16180 16190 16200
TTTGCAAATAAAAGGTATCAGGTTCAAGTCCTGTGTTCTCCAGTAATTTATATTGAAAAGATATAAAAATGGGATGAGTTTAATTTTGGCTCAGAAAGAA
16210 16220 16230 16240 16250 16260 16270 16280 16290 16300
GGCGTGAATTTGGTCTTATACATGTAAATTTAAATTTTAGTATGAAAATTGGTGTGCGGGTGCGTAGTATATTATGGTAAATTATCTTATGAGTAATTAA
16310 16320 16330 16340 16350 16360 16370 16380 16390 16400
ATTAATTGAAGATAAATAAATAAATAGGTGATTAGGTATTAAGTTTTTAGCCAAAGATCATTAACTGGTTTTTTAGGATGTATGGTCACATTTGCAATTA
16410 16420 16430 16440 16450 16460 16470 16480 16490 16500
AAGAAGGGCAAGACTCTTTTAGGAGAGGCAGCAGTATAGAATCTTGAACAATGAATGAAAATTTGATTCAGTTAAATTTGTGTGTGTAATATGGGATATT
16510 16520 16530 16540 16550 16560 16570 16580 16590 16600
CTCAATAAATCATAAGAAGTTTATTTATAGTGAAAATAAATTAAGAAAGGGCTGACAAAGATCAGTGCCAGAAGTCTCGGTTATACTGACAGCCTGAGTC
16610 16620 16630 16640 16650 16660 16670 16680 16690 16700
TATTTCAGTATTACTAGGTGTAAAGGATGTGTAAATCAGTTAAAATTATATAACACAATTTTAAATTAGAATTTTAAGAGTGAATATAATTTTATCTTTG
16710 16720 16730 16740 16750 16760 16770 16780 16790 16800
AGTATTTAAAGAATTAATAGAATATTAAAAGAAGTGATAGAATATTATAATATTTAGTAGAATTTCACAAGCGAAGGCAATTAATTGTTAAATAACTGAC
16810 16820 16830 16840 16850 16860 16870 16880 16890 16900
TTTGAGAACATGAAAGATAAGGTAGCGATATGGATTAGATACCCATGTAGTCTTATCTGTAAACAATGACTAAGAAATATATAGTAGATTTTTAAGATAA


16910 16920 16930 16940 16950 16960 16970 16980 16990 17000
CTTATTATTAGTTCACCTGAAGAGTAAGATCGCAAGGTTTAAATTCAAGAAATTAGGCTGTATATTGTCCTACTAGTGGATTATGTGGATTAATTAGACA
17010 17020 17030 17040 17050 17060 17070 17080 17090 17100
ATACACTCAAAATCTTACTTATATTTGTATTTAGAGCAAAAACAGGTGTTGCATGGCTGACTTCAATTAGTGTCGTGAGATGTTTGATTAATTTCAATAA
17110 17120 17130 17140 17150 17160 17170 17180 17190 17200
CTAATGAGATTCTAGGTTTTATTTAGGTGTATAGAATTACTGAGAGTTATAAACTCTAGGAAGATTAGAATAAAGTCAAGTCCTTATAACCCTTATATAT
17210 17220 17230 17240 17250 17260 17270 17280 17290 17300
AAGGCCTCACATGTAATACAGTGATTAATAGAAAAGAAAACGGTTAAGTAATTAACTGTAAAAGATGTATTTAGTCTAGGTATGGATAAGTTTCTGTAAC
17310 17320 17330 17340 17350 17360 17370 17380 17390 17400
TCGGAATTTTGAAGAAGGAATTAGTAGTAATTTAGAATTATCAAGTCTAGGTGAAAATGGCTCCAATGTAAGTACTTATTGCCCGTCAAGTCAGGAAAGG
17410 17420 17430 17440 17450 17460 17470 17480 17490 17500
GGACCTTCATATAAGTATAAATAAATAAAGTATAAAAGTTATTAACTTTATTGTAAGTATAAAATTGAAAAGAACAATTCTGATTAAGTCGTAACAAAGT
17510 17520 17530 17540 17550 17560 17570 17580 17590 17600
TGTCTTATGGGAACATGAGGCAGGATGATTATGGTAATTATAGTTTAAAAGGTAAAACTTTAATGTGTGATGTTAATAATTAAAGGTTCAAATCCTTTTG
17610 17620 17630 17640 17650 17660 17670 17680 17690 17700
ATTACCGGGATATTAGTTTAAATGGGAGAGTATTGATCTGTCTAATCAAAAGTTGCAGGTTCGAATCCTGTATATCTCGATAGGATCAATAACTTAAAGG
17710 17720 17730 17740 17750 17760 17770 17780 17790 17800
TAAAGTATACAGCTCATAATTGTTTAATATAAGTTCGATTCTTGTTTGATCTAGATATTATAGTTTAAAGGTAGAATATCAAGCTCATAATTTGAAGATA
17810 17820 17830 17840 17850 17860 17870 17880 17890 17900
CAGGTTCAATTCCTATTAATATTAAGTCTTTATAATTTAAGAGGTTAGAATATTTGATTTTCATTCAAAATATGCTGGTTCAAGTCCAGTTAAGGATAAA
17910 17920 17930 17940 17950 17960 17970 17980 17990 18000
TATGTTGTTATTAAATAGTTTATGAGATATTTTTTTTAATTTAATTAAAGTATTGTATTTAGTTGTTCCTTTAATTTTATCAATTGCTTTTTTAACTTTG
18010 18020 18030 18040 18050 18060 18070 18080 18090 18100
GCTGAAAGAAAATTGTTAGGTTCTATTCAAATTAGAAAGGGGCCTAATGTGGTAGGTGTTTATGGGATTTTGCAACCTATAATTGATGGTTTAAAACTGT
18110 18120 18130 18140 18150 18160 18170 18180 18190 18200
TGTTAAAAGAAGTGATTGTGCCTGTTCAAGCTAATTTTTTTTGTATTTTTTATTTCTCCTATTGTTTGTTTTATGTTAGCTTTAGGAAGTTGGGCAGTGAT
18210 18220 18230 18240 18250 18260 18270 18280 18290 18300
TCCTTTTGGGGAAGGGTTAATATTATCTGATATTAATATAGGAGTGTTATATGTGTTTGCTATTTCGTCTTTAAGTATATATTCAGTATTATGTTCAGGG
18310 18320 18330 18340 18350 18360 18370 18380 18390 18400
TGATCAAGTAATTCAAAATATGCATTTTTAGGTTCTTTAAGATCTACTGCTCAAATGATAAGTTATGAAGTTTCAATAGGTTTAGTTTTTATATCAGTTA
18410 18420 18430 18440 18450 18460 18470 18480 18490 18500
TTTTAATGAGTGGGTCTTTTAATCTTTTGGTGATAGTCGGGAGTCAAAAATTAGTATGATATATTGTACCCTTATTTCCTAGTTATTTAATGTTTTTCGT
18510 18520 18530 18540 18550 18560 18570 18580 18590 18600
ATCTGCATTAGCAGAAACAAATAGAGCTCCTTTTGATTTACCAGAATCTGAATCTGAGCTGGTTTCAGGTTATAATGTTGAGTATTCGAGTATGATGTTT
18610 18620 18630 18640 18650 18660 18670 18680 18690 18700
GCTATGTTTTTTTTAGGTGAATATTGTCATATTATATTTATGTCTTTCTTTATTAGTTTGGTGTTTTTAGGCGGATGATTAAGTCCTTTTGAGGATTTGT
18710 18720 18730 18740 18750 18760 18770 18780 18790 18800


GTTTGATTGGTATTCCTGGGGTTATATGGTTAGTTTTAAAATGTTTGTGTAATTTATTTTTATTTATTTGAATCAGAGCGATTGTACCTCGTATGAGGTT
18810 18820 18830 18840 18850 18860 18870 18880 18890 18900
TGATCAATTGATGTATTTGTGTTGAAAATCATTTTTACCTTTAAGTTTGTCTTATGTTATTGGTTTGTCCGGAGTTATCTGGTTTTTTGATTGGTTCCCA
18910 18920 18930 18940 18950 18960 18970 18980 18990 19000
GTGTCGGTATAGCCTGCCTTAATAGCTTAGAGGTAAAGTATATATTTTGTAAATATAAGAATATTGGTTCAATTCCAATTTAAGGCTTAGCCTAAGGGAG
19010 19020 19030 19040 19050 19060 19070 19080 19090 19100
AGAAATAAAATGTATAATAAATATCATGTATACCATTTAGTAAATCCTTCTCCTTGACCTATATTGGTAGGTTTATCTTTATTAACAAGTGCTGTGGGAG
19110 19120 19130 19140 19150 19160 19170 19180 19190 19200
TGGCTTTGTGGTTTAATTTTTATGAGTTTGGAGAAAATGTTGTTATTTTAGGTTTAGGTTTTAATTTGTTTTGTGTTATATTATGGTGAAGGGATATTAT
19210 19220 19230 19240 19250 19260 19270 19280 19290 19300
TAGAGAAGGTACTTTTCAGGGACACCATACTAAGGTGGTTCAAAAAGGTTTAGTTTTAGGAATGGTTTTGTTTATAGTTTCTGAAGTTTGTTTGTTTGTT
19310 19320 19330 19340 19350 19360 19370 19380 19390 19400
TCTTTTTTTTGGACTTATTTTCACGTAAGTTTAATTCCTAGTATTGAAATAGGAGGTGTATGACCTCCTATAGGAATACAGGTGTTTGATCCCTTAAATA
19410 19420 19430 19440 19450 19460 19470 19480 19490 19500
TACCCTTATTAAATACGTTAATATTATTAATGTCAGGATGTTCTGTAACTTGAAGTCATAATGAAATAGTGTTAGGGAATAAGCGTGGTACAATTATAAG
19510 19520 19530 19540 19550 19560 19570 19580 19590 19600
TTTATTGTTAACTTTATTGTTAGGTGTGTTATTTTCATTATGTCAACTATATGAGTATATAGAGGCTACTTATACTATAGCAGATTCTATATATGGATCA
19610 19620 19630 19640 19650 19660 19670 19680 19690 19700
GTTTTTTATTTAATAACAGGGTTTCATGGATTACATGTTATAGTTGGGACTTTATTTTTAAGTGTTAGTTTTGTAAGAATAATTAACTATCATTTTACAA
19710 19720 19730 19740 19750 19760 19770 19780 19790 19800
AAAAGCATCATTTTGGTTTTGAGGCGGCTATATGATATTGACATTTTGTTGATGTAGTATGGTTATTTGTGTATGTATTTTTATATTGATGAGTATTTTA
19810 19820 19830 19840 19850 19860 19870 19880 19890 19900
AATGTATATATAGTTTAATAGGTTAGAACATTTATCTTCCAAATAGAATGTTGTGAGTTCGATTCTCATTATATACAAGTGTTTCAGAAAGTAGTTCAAA
19910 19920 19930 19940 19950 19960 19970 19980 19990 20000
GGTAGAACAATTAATTTGGAGTTAAGAGGTTGTAGGTTCGAATCCTATCTTTTTGATATGAAAATCATAAAAGAACTAAATTATAAAAAATTTAAGTATT
20010 20020 20030 20040 20050 20060 20070 20080 20090 20100
TCTTCTATTCTACAGATCTTTCTTTAAATTATAGGTACAATTTAAGATTAAAGTTAGAAAGTATGAATAATTTTAAAAGATATTGTAATTTGACAAAAAG
20110 20120 20130 20140 20150 20160 20170 20180 20190 20200
AAAAAGAGGGTTATGTTTTAATTCTTCTTTAAATAGGTTAACATTAAAATCTTTGATTAATAAAGGTTTTACGCCTGGAATAAGTATGTCAAAATGATAA
20210 20220 20230 20240 20250 20260 20270 20280 20290 20300
ATAATTATTTAATTTTAGGATTATCTGTATTTGTAATAAGTATAATTGGAATATTGCTTAATAGAAAAAATTTAATAATAATGTTAATGTCCCTTGAATT
20310 20320 20330 20340 20350 20360 20370 20380 20390 20400
AATGTTATTGGCTATCAATTGTAATTTTTTAAGTTTTTCTGTAATAATAGAAGATATTTTAGGTCAAGTGTTTAGTGTATTCATATTAACTGTTGCAGCT
20410 20420 20430 20440 20450 20460 20470 20480 20490 20500
GCAGAGTCTGCAATAGGTTTATCTTTAATGATAGTTTATTTTAGAATAACAGGTAACATTTCAATTTTATTTATTGATCTCTTGAAGGGATAAGTGTATT
20510 20520 20530 20540 20550 20560 20570 20580 20590 20600
TATTATTAATTAGTATTCCTTTAATATGTGCTTTTGGAGTTTTATTTTTTGGGAGATTTTTAGGTCATTTAGGTGTTATAAGATTTAGTATAATTTTAAT
20610 20620 20630 20640 20650 20660 20670 20680 20690 20700


AGGGTTAGATATGTTATTAAGTTATTTTGTGTTTTATGAAGTAGGAATTAACCAAGTGCCTGTGTTTTTAAATGTGTTAGATTGAATAAGCTTTGATATG
20710 20720 20730 20740 20750 20760 20770 20780 20790 20800
ATAATTGTTCAGTGATCTTTACAGTTCGATAGTTTATCTGTCAGTATGTTAATAATAATAATAACAATTTCTTTTTTTGTTCATTTATATTCTGTTGAAT
20810 20820 20830 20840 20850 20860 20870 20880 20890 20900
ATATGGGGCAAGATCCTCATTTGTCTAGGTTTATGGCTTATTTATCGTTGTTTACTTTTTTCATGGTGTTATTAGTAGTTTCGGATAATTATGTGCAATT
20910 20920 20930 20940 20950 20960 20970 20980 20990 21000
GTTTGTTGGTTGAGAAGGTGTAGGTTTATGTTCTTATTTGTTAATTAATTTTTGATACACAAGGTTACAAGCTAATAAAGCGGCTATAAAAGCTATGGTA
21010 21020 21030 21040 21050 21060 21070 21080 21090 21100
ATGAATAGAATAGGTGATATAGGTTTAGCTTTAGCTATATTTATGGTGTATGAGTATTATGGTTCTTTAAATTTTTCAGTTTTAATGGTTATAGTAGAAT
21110 21120 21130 21140 21150 21160 21170 21180 21190 21200
CATTAAATAATAATGATATATGTGATATTAGTCAGATAGCTGTTATAGTAATTTGTTTATTGGTAGGTGCTTTAGGGAAATCTGCTCAATTTGGTTTGCA
21210 21220 21230 21240 21250 21260 21270 21280 21290 21300
TACCTGATTGCCAGATGCAATGGAGGGGCCTACCCCTGTGTCAGCCTTAATTCATGCTGCTACTATGGTGACTGCAGGTATTTTTTTAATGATAAGATCT
21310 21320 21330 21340 21350 21360 21370 21380 21390 21400
AGTATGTTTATTATAAATAGTAAATTAGGGTTATTTGTAATCACATTACTAGGAGCATTAACTGCATTATTCGGTGCTACAATAGGAATTGTACAAAATG
21410 21420 21430 21440 21450 21460 21470 21480 21490 21500
ATTTAAAGAAGGTAATAGCTTATTCTACATGTTCACAAGTAGGGTATATGTTGTTTTCTTGTGGGCTCGATGAATTTAATGTATCTTTGTTTCATTTAAT
21510 21520 21530 21540 21550 21560 21570 21580 21590 21600
GAATCATGCGTTTTTTAAGGGGTTGTTATTTTTAAGTGCAGGTTCTATAATTCATGCGGTAATCGATGAGCAAGATATTAGAAAAATGGGTGGTTTAATA
21610 21620 21630 21640 21650 21660 21670 21680 21690 21700
TTAGTATTACCTTATACTTACAGTGTAATGTTGATAGGTTCGTTATCGTTAATGGGATTTCCCTATTTATCTGGTTTCTTTTCTAAAGATATGATTCTAG
21710 21720 21730 21740 21750 21760 21770 21780 21790 21800
AGTTATCTTATGCACATTATGTGTGTGAGAGTTTGTTTTCATATTGATTAGGAGTAGTTTCAGCGAGTATTACTGCTTTTTATTCAATAAGATTATTAAT
21810 21820 21830 21840 21850 21860 21870 21880 21890 21900
TAGAGTGTTTTTTTTCAATCCTAATATGTGTTTTAGTAAATTCAAAAAAGTTCAAGAAAATAATTATATAATATTTTTTGTTTTATTTGTTTTAAGTATG
21910 21920 21930 21940 21950 21960 21970 21980 21990 22000
TTTAGCGTATTTTCCGGTTATTTTCTTAAAGATATGTTAGTTGGGTTTGGGTCTGGGTTTTTTAACGGATCTTTACGAAGTATTAAATATATTTCAGTAA
22010 22020 22030 22040 22050 22060 22070 22080 22090 22100
ACACTGATGTAGAGTTTTTACCATTATATATAAAATTATTTCCTGTTATATTTAGTTTATTGGGTGCAAGTATTGCATTAGTGGTATATTCTGGGATTGA
22110 22120 22130 22140 22150 22160 22170 22180 22190 22200
AGAGATAGGATTTAAAGTAAGATCTTCTTCTCTTGGAATTATAATATTTAAATTTTTATCAAATAGATGATATATAGATTTTATTTATAATTNTTATTTA
22210 22220 22230 22240 22250 22260 22270 22280 22290 22300
TATCAATTATTGTTCGTAACACATAATTATTTATATAAATTAGTAGATAGAGGTGTATTAGAACTATGAGGTCCGACAGGTATGGTTAATTTGTTGCGAA
22310 22320 22330 22340 22350 22360 22370 22380 22390 22400
GATATTCTTCTGTGTTAAGTATAATGCAGTCAGGATTTATCTATCATTATGTTCTGTTTATCTTAGTGGGTGTTTCTTACATTATTTTATGGGAGTTTTT
22410 22420 22430 22440 22450 22460 22470 22480 22490 22500
ACAATTGAATGTTTTTTTAATAGTTATATATTTGTTAACAGTGTTAGCAGGGTTATAGGGATTGTAGCTTAATTGGTAGAGCTAATAACTTTTAATTATT


22510 22520 22530 22540 22550 22560 22570 22580 22590 22600
TAAATAAAGGTTCAAATCCTTTCAATCCCAAATGCCACAAATTGATATATTAACTTATTTTCCGCAATTTTTTTGATTTATTTTTTCTATGTGTGTTTTG
22610 22620 22630 22640 22650 22660 22670 22680 22690 22700
GTAAATTATATGTTTAACAAGTTTTTACCAAAGATTTCCAAAATATGATATTTTAGAAAGCGTGCTTATAGAAAAGAACTTAAATACAGTAAGTCTTATA
22710 22720 22730 22740 22750 22760 22770 22780 22790 22800
TTATTGCAAAAATAAAATTTAAAAGAAGTTTTGAATTGCAAGTGCTAAATATGAGGTATAGAGGTATGCCTGTTTCTATGAAGTTTGATTTGCCTAATTC
22810 22820 22830 22840 22850 22860 22870 22880 22890 22900
TGATGAATTGTACTTAAATTATGATAATGTTTGATTAAGTTCAAGATTGAACTTGATTGATGTTTAAGTAAGATTAAAAATATATAGAATCTAANATAGT
22910 22920 22930 22940 22950 22960 22970 22980 22990 23000
AAGATAAATAAAATGTTTTTATTTTCTCCTTTAGAACAGTTTCAAATAATTAAGCTAGTTGAAATAAATGTGTTAGGTTTGATGGACATTTCAATATCTA
23010 23020 23030 23040 23050 23060 23070 23080 23090 23100
ATTCTTCTTTTTATGTTTTATTAATTTTAATGTGATTTAAATTGATTGAATTATATGTAACTTATACTATTATTCCTACTAAATGACAGAGTGTAATTGA
23110 23120 23130 23140 23150 23160 23170 23180 23190 23200
GCATTTTGTTAATGAAATATTAAAATTATCTAAAGATAACTTAGGTAATAGTTTTCATTTATATTTTCCATTAATATTTAGTTTATTTTTATTAATATTA
23210 23220 23230 23240 23250 23260 23270 23280 23290 23300
ACTTTGAACTTGGTTGGAATGATTCCTTATTCATTGGCTGTTACTTCCCATATTGTGGTTACTTTTGGAGTAGCATTGTCATTATTTATTTCAATAACTA
23310 23320 23330 23340 23350 23360 23370 23380 23390 23400
TAATAGGGTTTAGAAATATGGTGTTAATTTTTTTAGTATGTTTATGCCTCAAGGAGCTCCTTTTAGTGTTAGGTCCTTTATTAGTGTTTATTGAGATATT
23410 23420 23430 23440 23450 23460 23470 23480 23490 23500
AAGTTATTTGGCCAGACCTATATCTTTGGGAGTTAGATTAGCTGCAAATATGACTGCAGGTCATATATTATTAGCTATTATTTCAGGGTTTATTTGATCG
23510 23520 23530 23540 23550 23560 23570 23580 23590 23600
TTAGCTAAATCAGGGTTGATGGGAATTATAATTAGTGTATTTCCTTTTGTGGTTTTAATTTTAATTATTATGCTAGAATTAGCTGTTTCTTTTATTCAAG
23610 23620 23630 23640 23650 23660 23670 23680 23690 23700
CTTATGTATTTACATTATTAGTAACTATATTTTTAAATGATTCTATTCATTTACATTAAGATATAGGTATAGTAACATAAAAGGTTAATGCGACTATTTG
23710 23720 23730 23740 23750 23760 23770 23780 23790 23800
CAAAATAGTAAATGTAGGTTCGAGGCCTATTTATACCTAACATATAAAATAGAAGTGTGATGAAAAGGTAGACATAATAATGTTAGATATTATTTATATG
23810 23820 23830 23840 23850 23860 23870 23880 23890 23900
GGTTCAATTCCCATCATTTCTATGAAATAAGTTGTGAGAATTATAAAAATCACAATTGAAATTGTGATCCTGTTCAAGTATAAGTAGTTGGCCGGGGTTT
23910 23920 23930
TATTAAAAAATGGGTGGTTATGGGTTAATCCCAAATTTTG