Diversity of parasitic micro-eukaryotes in the urine of *Gadus morhua* L

Estelle Grønneberg

Master of Science Thesis

Department of Biosciences

18.12. 2015

UiO : University of Oslo

Year: 2015

Title: Diversity of parasitic micro-eukaryotes in the urine of *Gadus morhua* L

Author: Estelle Grønneberg

http://www.duo.uio.no

Printed: Reprosentralen, University of Oslo

Abstract

Micro-eukaryotes are a divers and abundant group of organisms, that hold a highly important ecological role in all aquatic ecosystems. They are responsible for a large portion of the primary production and are an important food source for many organisms. Of the micro-eukaryotes, many lineages function as parasites,which means that they exploit another organism (a host) for their own benefit.

In this present study, DNA from the urine of *Gadus morhua* was extracted, amplified with general V4 primers and investigated using Illumina Miseq sequencing approach, to see what kind of micro-eukaryotic parasites one might find in the urine bladder of *Gadus morhua*, with a special emphasize on Microsporidia and Myxozoa. These are parasitic groups commonly known to parasitize the gadoids. Micro-eukaryotic parasites can cause different malfunctions in the fish, and in worst cases death, and it is therefore important to address factors that can threaten the health and abundance of *Gadus morhua,* a fish that is of highly commercial importance in Norway and has a IUCN (International Union for conservation of Nature) status as vulnerable.

The urinary bladder/tract might be an important target organ for parasites in aquatic systems, as urinary excretion could be a transmission pathway for parasite spores, and can therefore be an important factor for spreading parasites.

56 % of the OTUs detected were micro-eukaryotic parasites. Among these over 50% of the OTUs were assigned to Myxozoa and Microsporidia. Within these groups a vast majority were assigned to the genera *Gadimyxa*, *Loma* and *Microsporidium*. Potential new species within the genus *Micrsporidium* and *Gadimyxa* were found, demonstrating that there are probably more micro-eukaryotic parasites likely remaining to be discovered. The parasitic findings in the urine of G. morhua, may be a first step in the right direction of resolving several unfamiliar transmission path way of parasitic species infecting fish, and thereby also get a better understanding of the life cycle of these parasites, in the sense of how the spores spreads from one host to another.

Acknowledgements

Endelig var master levert inn, skal si det føles godt på kropp og sinn. Som bølgene i havet har tiden rundt masteren gått opp og ned. Mange har støttet meg, og jeg vil gjerne takke dere for det.

På Universitetet I Oslo ved CEES og EVOGENE var laboratoriearbeid og bioinformatikk gjort. I perioden september 2012 til desember 2015, slikt arbeid blir ikke utført fort.

En stor takk til veiledere Dag Klaveness, Kamran Schalchian-Tabrizi, Jean Françoise Mangot og Haakon Hansen, for å ha gitt meg denne sjansen.

En sjanse til å få være med i et stort og spennende prosjekt, det har vært et meget lærerikt og meste parten av tiden veldig kjekt.

David Bass, Kevin Hopkins og Egil Karlsbakk, er verdige en stor takk. De har bidratt med referanser, kunnskap og Illumina sekvensering. Med et gledeshyl kom urinprøvene tilbake fra England sekvensert og uten kontaminering.

En stor takk til Anders Krabberød og Arthur Haraldsen for hjelp med bioinformatikk.

Koding og fylogeni gav meg lettere sagt nokså mye problematikk.

Grete og Hans Erik Karlsen sender jeg en stor takk, for et herlig opphold i idylliske Drøbak.

Til mine kjære veiledere Janina Fuss og Marit. F.M Bjorbækmo, hva skulle jeg gjort uten dere to? Dere var der som to klipper da jeg slet på mitt meste, dere fikk meg til å forstå at tingene jeg gikk gjennom kan skje den beste. I tillegg til å ha gitt meg mengder med kunnskap og kompetanse. Har dere rettet oppgavens innhold, språk og referanse. Jeg er evig takknemlig for all hjelp dere har gitt, alt fra personlige ting til hver enkelt parasitt.

Torbjørn Gylt, vi har virkelig hatt en moro tid sammen, på lab og bølgan blå. Du fisket store fisk og jeg fisket små. Vi har jobbet, slitt og ledd i fleng, og jaggu meg har vi ikke fått en master i samme sleng.

Mine herlige venner som har stått ved min side, alle tiders døgn, i tid og i u-tide.

Tusen takk for at dere har vært der. Dere er spektakulære alle og enhver.

Til mine kjære og nære.

Jeg er så inderlig takknemlig for at dere har stått ved min side og heiet meg frem. Takk for tørking av tårer, latter, mat, kyss, klapp og klem. Ingenting er som å kunne komme hjem til en som står der med åpne armer og tar meg i mot, når kroppen er sliten og hodet et rot. Dere har hørt mangt om fiske, parasitter og sekvensering. Hold ut, nå skjer det

LEVERING.

Table of Contents

Introduction

The famous Norwegian theologian and author Petter Pettersen Dass said in his poem Nordlands Trompet ”The fish in the water, it is our bread”. Fish has been a way of life and an important source of food since the ancient times. According to the Food and Agriculture Organisation (FAO), fish is the most important single source of high quality protein, providing 16.7% of the animal protein consumed by world´s population (Beveridge *et al.*, 2013). Fish does not only contain healthy proteins but is also a well-known contributor of essential nutrients, as vitamin A, D, B and magnesium (Beveridge *et al.*, 2013). In the past 50 years, the consumption of fish *per capita* has doubled from 9.9 kg in 1960 to 18.6 kg in 2010 (Beveridge *et al.*, 2013).

Fisheries have historically been, and still are, a significant export industry for Norway. Our country is the second largest seafood exporter considering financial figures in the world, including both, wild catch and aquaculture. Cod (*Gadus morhua*) is one of the most important commercial fish stocks (WWF-Norway 2008). In 2010 the production of wild and farmed cod where at a staggering 21 240 tons, contributing to the Norwegian economy with billions of kroner (NOK)(Heuch, 2001). As the cod are of such great value to Norway, it is important to address factors that can threaten its health and abundance.

Parasites are one of the known factors that can pose a major threat to fish populations. A parasite is an organism that lives in non-mutualistic symbiosis with a host, where only the parasite benefits from the relationship, at the expense of the host (Brooker *et al.*, 2008). In an earlier report by the Norwegian Research Council (NRC), they stress the importance of gaining knowledge about the transmission of parasites and the diseases that could potentially be transmitted from and between the Norwegian cods (Heuch, 2001). Even though several studies have investigated the parasitic fauna of cod (Timi & Mackenzie, 2015), knowledge about how parasites are spreading in fish, why they inhabit different organs, and lastly, how the parasites spread to the outside environment is still an understudied field.

In the present study, an investigation of the urinary bladder of *Gadus morhua* was done to examine what kind of parasites might be present there, with a special emphasize on micro-eukaryotic parasites like Microsporidia and Myxozoa. These are parasitic groups that are commonly known to parasitize *G. morhua* and might cause different malfunctions in the fish,

and in worst cases death (Håstein, 2011). Very little is known about the parasite fauna of the urinary bladder of cod, but it is thought to be an important target organ, as urinary excretion might be a transmission pathway for parasite spores, and can therefore be an important factor for spreading parasites. Using High Throughput Sequencing (HTC), We aimed at answering following questions that could provide better insight on potential parasites in Oslofjord cod population.

- Does urine have a large concentration of parasites?
- Are there any parasite taxa that are more abundant than others?
- Is there any undescribed parasitic diversity present in the urine?
- Is urine a pathway for excreting parasites?
- Is urine generally a good way to check the fish for parasites?

Gadus morhua

Gadus morhua (Linnaeus 1758), or commonly known as the Atlantic cod, is a widely recognized marine fish of the genus *Gadus*, and under the family Gadidae. It is easily recognized by its elongated body, big head, and a swollen belly. A pronounced beard line and a slightly overbite is a familiar trademark. The underside is silvery grey, and the side-lines white and curved, with a skin coloration that varies according to different habitats, for example red-green colour if the cod lives in areas with red and green algae. The cod can be 1.8 m long and 55.6 kg depending on its living condition (Jonsson, 2008). *G. morhua* is a cold-water fish and is generally found in the northern part of the Atlantic Ocean, from south of Biscaya up to the northern part of Spitsbergen, around Iceland, Newfoundland, south Greenland and the east coast of Cape Hatteras (Store Norske Leksikon, 2009). Depending on the feeding, spawning and changes of hydrographic conditions, one can find Atlantic cod at several depths, ranging from 600 meters at the continental shelf, to a couple of meters near the shores lines (Jørstad *et al.*, 2003).

In general, the cod is considered to be a demersal fish, but under certain conditions, for example whilst feeding, it may become pelagic (Wilmot, 2005). As a top predator and omnivore fish, the cod will hunt shoals of herring and capelin. Benthic animals like echinoderms, annelids and crustaceans are also an important part of their diet (Du Buit, 1995).

In Norwegian waters there are two main groups of cod, Norwegian arctic cod (*Skrei*), and the coastal cod (Pethon, 1998; Jonsson, 2006). The latter group is the object of study in this thesis. In contrast to the migrating Norwegian arctic cod, the coastal cod stocks typically inhabit shallow waters and are less mobile than their pelagic north-arctic counterparts that migrate long distances to spawn (Store Norske Leksikon, 2009). Spawning takes place from February through April within small fjords and bays. The eggs stay afloat in the upper segments of the water column and hatch after 2-3 weeks. Juvenile cod will after 8-10 weeks sink to the bottom and there they will continue their growth into mature adults (van der Meeren, 2015).

Studies show that the coastal cod have a tendency to reach maturity at early age (between two and three years of age) and small sizes of 35 cm (Olsen *et al.*, 2004; Jonsson, 2006), whilst the Norwegian arctic cod reaches maturity for the first time at 6 – 15 years of age (Store Norske Leksikon, 2009). Coastal cod in the Oslofjord are assumed to mature at the age of two years on average (Otterbech, 1953). According to Otterbech's findings, the cods used in this study were either preparing for first time spawning or were first time spawners at the time. In the same size group, older fish can also be found. Those do not fall into to the same maturity/developmental/life cycle? categories as the others, and the small size could indicate that these fish are challenged by some kind of health problems, like parasites, and are using more energy fighting the infection than increasing the body size.

Gadus morhua and its parasitic fauna

Due to its economic value, the Atlantic cod and its numerous diseases that may threaten its ability to survive is one of the most studied marine organism in the world, (Køie *et al.*, 2007). One of its biggest threats are parasites. Even though the cod is well studied, there has been little research on micro-parasites (invisible to the unaided eye) associated with cod, compared to studies on macro-parasites (immediately visible). One study showed that the Atlantic cod is the host to more than 120 parasite species, whereof only 13 were micro-parasites (Køie *et al.*, 2007). Another thorough study done by Heuch et al. (2011) showed as many as 107 parasites that parasitized different areas of the cods body. 11 of those were micro-parasites and 96 were macro-parasites. Gills, blood path, muscles, flesh, intestines and urinary bladder are some of the target organs that seem to be preferred as infection site (Hemmingsen & MacKenzie, 2001; Heuch *et al.*, 2011). The cod is only one of a few fish species that actually features a urine bladder (Teshima *et al.*, 1986). It has previously been considered to only function in the osmoregulation in the fish and as a storage chamber for urine. However, several studies have

shown that it does not only contain urine, but other organisms, for example protist parasites and viruses (Hendriks *et al.*, 1998; Heuch, 2001; Samuelsen *et al.*, 2006; Heuch *et al.*, 2011).

Traditionally, morphology and different microscopy techniques have been used as the main criteria for identification and taxonomic classification of parasites, such as microsporidia and myxozoa in fish (Schmahl *et al.*, 1990; Heuch *et al.*, 2011; Karlsbakk & Køie, 2012). Even though it is a fundamental and important way to confirm and identify infections in fish and other vertebrates, using it as a primary feature in systematics can lead to genera becoming polyphyletic or cryptic, and therefore new species of micro-organisms can be overlooked or grouped with species they do not belong to (Freeman & Kristmundsson, 2015). This entails the risk of non-suitable treatment and one can, by also including molecular methods to back-up parasite identity forestall these complications.

In the study of Hemmingsen and Macenzie (2001) only 6.5% of the affecting parasites were oioxenic parasite species[1], and 15.9 % were stenoxenic species[2]. The remaining majority were parasites with a broader host-range that could affect other fish and eukaryotes (Hemmingsen & MacKenzie, 2001). When the cod is malnourished and already infected, the chances for having greater fauna of parasites increases due to so called opportunistic infections. That means that bacterial, viral, fungal, or protozoan pathogens take advantage of a host with a weakened immune system. In addition to general factors like the overall health, environmental factors and food availability seem to be additional explanations for the high amount of infections with unspecific pathogens in *G.morhua*. Some of them can also be explained by its lifestyle being a benthic omnivore (Hemmingsen & MacKenzie, 2001), but also by the cod's special immune system (Star *et al.*, 2011).

Some gadoid fish, like the Atlantic cod have an unusual immune system, which distinguishes them from other teleost fish. For a long time it has been thought that all jawed vertebrates were born with the same immune system. The innate immunity is the immunity which you are born with, while the acquired immunity develops during encounters of bacteria, parasites and viruses that cause diseases (Brooker *et al.*, 2008; Vøllestad, 2009). The acquired immunity is regulated by the Major Histocompatibility Complex (MHC), which is divided into MHC class I (MHCI) and MHC class II (MHCII). The function of MHCI molecules is to target invading intracellular entities, while MCHII molecules function in the animals' extracellular space (Janeway *et al.*, 2001).

[1] Parasites which infects a single species of host

[2] Parasites that target e.g. a certain family of hosts

Research done by (Star *et al.*, 2011; Malmstrøm *et al.*, 2013) demonstrated that the Atlantic cod has lost the normal function of the acquired immunity. Since cod doesn't possess MCHII, it compensates with a high copy number of MCHI loci and a rare composition of the highly conserved Toll-like Receptor (TLR) that play a major role in the innate immune response and the initial detection of pathogens (Persson *et al.*, 1999; Miller *et al.*, 2002). Although the cod might experience some problems getting rid of all infections at (Birkbeck *et al.*, 2011; Star *et al.*, 2011) the lack of MCHII does not seem to cause any sort of problem for the cod. As long as the infections are small scaled it seems like this unique immune system gives the adult cods a higher tolerance for infections, and prevents all kinds of diseases from being malicious for the cod (Star *et al.*, 2011). Although there are many parasites infecting cods, there are several that do not cause life-threatening conditions like copepods and *Cryptocotyle lingua* (Figure 8a), which will be discussed further down.

Protists

There are four traditional kingdoms in the Eukarya domain: Fungi, Animalia, Plantae and Protista. Of these four kingdoms one might say that Protista is the most diverse and difficult to study, as their morphological features can be hard to distinguish (Pawlowski et al 2012). The use of the term *Protista* as a definition of a kingdom has historical rather than biological reasons. The term *Protista* was first used by (Haeckel, 1866), and was largely based on superficial commonalities: a dominant single-celled life stage and the lack of a body plan (Hebert, 2008). The protists were sub-grouped according to their similarity to the three other kingdoms Fungi, Animalia and Plantae (Caron *et al.*, 2012). The "Algae" were the photosynthetic, plant-like protist, "Protozoa" were the animal-like protist, that comprised heterotrophic organisms that were mobile, while "Slime moulds" were fungus-like protists that resembled true fungi (Brooker *et al.*, 2008; Caron *et al.*, 2012). These phylogenetic definitions lasted through the latter part of the 1980s, even though increasing information that contradicted these characterizations became apparent. These older terms are in some cases still used as informal names to describe the morphology and ecology of various protists.

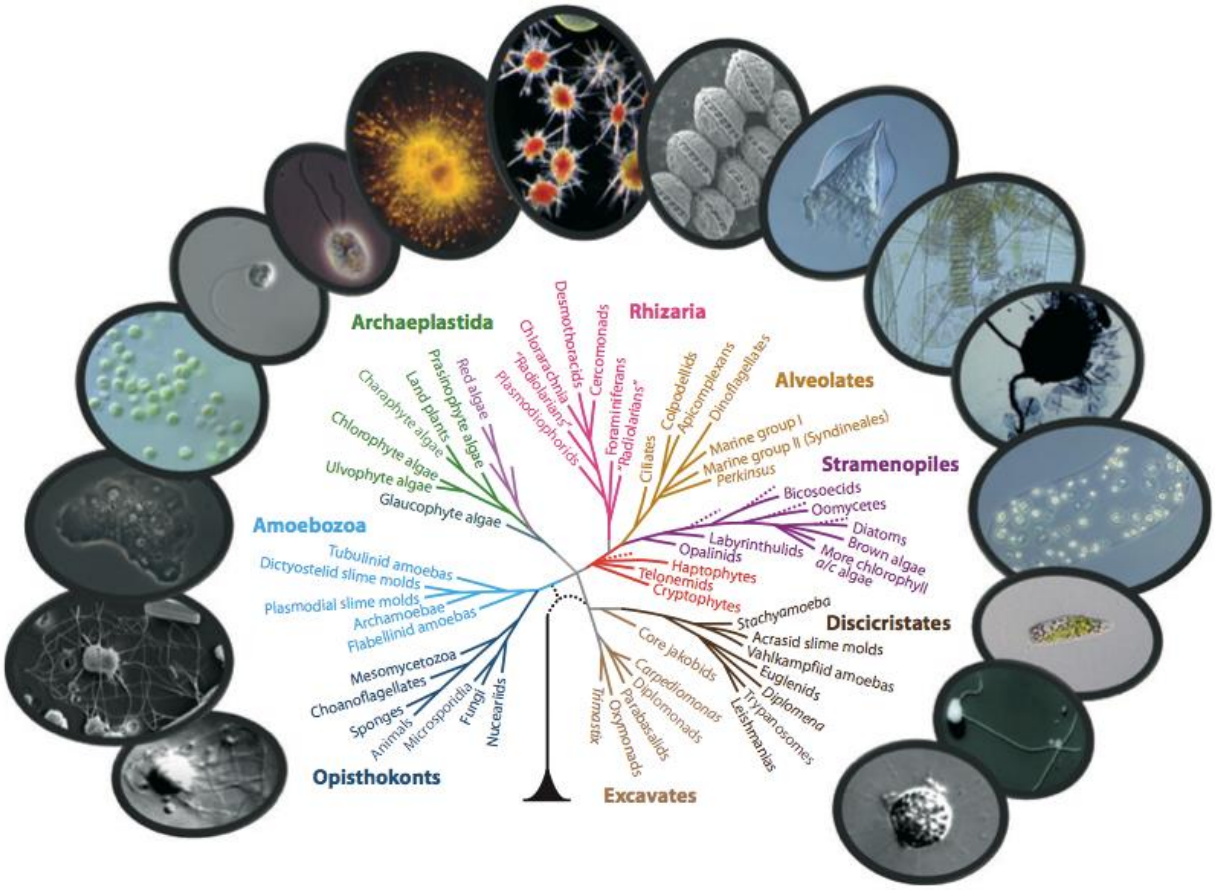


Figure 1: Phylogenetic tree showing the eight major groups of eukaryotes. The super groups are Opisthokonts, Amoebozoa, Archaeplastida, Rhizaria, Alveolates, Stramenophiles, Discicristates and Excavates (Caron *et al.*, 2012).

Instead of categorizing protists as one taxonomic group, it is today common to define the term "protist" in two ways. The phylogenetic way explains what protists are by defining them as a paraphyletic[3] group (Schlegel & Hülsmann, 2007; Brooker *et al.*, 2008). Protists are, unlike Metazoa plants (Viridiplantea) and true fungi (Fungi) paraphyletic and therefore they cannot be defined correctly as one, true taxonomic group (Schlegel & Hülsmann, 2007; Brooker *et al.*, 2008). Therefore, the phylogenetic definition implies that a protist is any eukaryote that is not metazoan, plant, or a true fungus (see figure 1).

The second definition is based on the function and biology of protists. Protists are in this particular definition: those eukaryotes that are never multicellular, that exist as independent cells, and if they occur in colonies, they should not show any differentiation into tissue (Adl *et al.*, 2005).

[3] A group that contains a common ancestor and some of its descendants

Both definitions are mainly overlapping, but exclude or include different groups. The phylogenetic definition excludes for example unicellular eukaryotes as Microsporidia because they are true fungi (Adl *et al.*, 2005). The functional definition includes Microsporidia, but excludes members of the otherwise protist "supergroup" Stramenopiles that are truly multicellular organisms like seaweeds and kelps in the Phaeophycea (Adl *et al.*, 2007). As increasing amounts of DNA sequence information and new molecular methods have become available, reconstructions for a better phylogeny and overview of protist diversity is a still ongoing process (Caron *et al.*, 2012). Whilst the work on protist taxonomy is proceeding it is common to classify the protists into the eukaryotic "supergroups": Amoebozoa, Ophistokonta, Archaeplastida, Rhizaria, Alveolata Strampenopiles, Discicristata and Excavata (Caron *et al.*, 2012).

Due to the definition problems for protists, I will use the term micro-eukaryotes to include all protists according to Adl et al. (2007), plus Myxozoa and Microsporidia in this thesis.

Wherever one may find a moist place, micro-eukaryotes ought to be found. They have been detected in all terrestrial and aquatic environments, and despite their small size they have a great impact on global ecology and human affairs (Brooker *et al.*, 2008). Micro-eukaryotes play an important role in connecting organisms at every trophic level in ecological food webs as primary producers and consumers (Brooker *et al.*, 2008).

Because of their large size range, rapid metabolic rate and their many nutritional modes, eukaryotic microorganisms can enter the food web at several trophic levels (Caron *et al.*, 2012). While a large fraction of the oceans' global primary production is meditated by the photosynthetic (autotrophic) micro-eukaryotes (Caron *et al.*, 2012), heterotrophic micro-eukaryotes ensure that energy (dissolved organic carbon) is returned to higher trophic levels by consuming heterotrophic bacteria that feed on dissolved organic matter (Azam *et al.*, 1983; Worden *et al.*, 2015). The protists play therefore a significant role in the highly important biomechanical and carbon cycle (Liu *et al.*, 2009).

Micro-eukaryote parasites

Parasitism is considered to be the most common consumer strategy in nature. It is argued that roughly half of the species on Earth are parasites (Windsor, D, 1998; de Meeus & Renaud, 2002; Skovgaard, 2014). Of the micro-eukaryotes, many lineages function as parasites, which means that they exploit another organism (a host) for their own benefit (Skovgaard, 2014).

The parasites embrace many groups of micro-eukaryotes, and a broad spectre of these organisms can have a substantial effect on human health or economy (Brooks *et al.*, 2015). Parasites infect organisms like plants, vertebrates, invertebrates and other protists, including other parasitic taxa (hyperparasitism) (Brooks *et al.*, 2015). The parasites can live as ectoparasites, which means that the outer most living cell layer of the host is parasitized (Hopla *et al.*, 1994; Bellay *et al.*, 2008). The X-cell, a disease causing agent that has only recently been identified as a protistan parasite, can form severe epidermal differentiations in a variety of fish species including gadoids (cod fish) from the Pacific and Atlantic Oceans (Freeman, 2009). The parasite can develop large epidermal tumours in skin, fins, head and the inner operculum region of the host. X-cell infections have been shown to impact fish health in varying degrees and can even lead to death in juveniles (Freeman, 2009; Ohtsuka *et al.*, 2015). The parasites of *Trichodina spp* .. are another example of ectoparasites on fish that infect the skin surface and fins, and feed on waterborne particles (Heuch *et al.*, 2011). The movement of the parasite irritates the fish skin making the fish increase the production of mucus, thus adding more food particles into the water (Hopla *et al.*, 1994; Heuch, 2001).

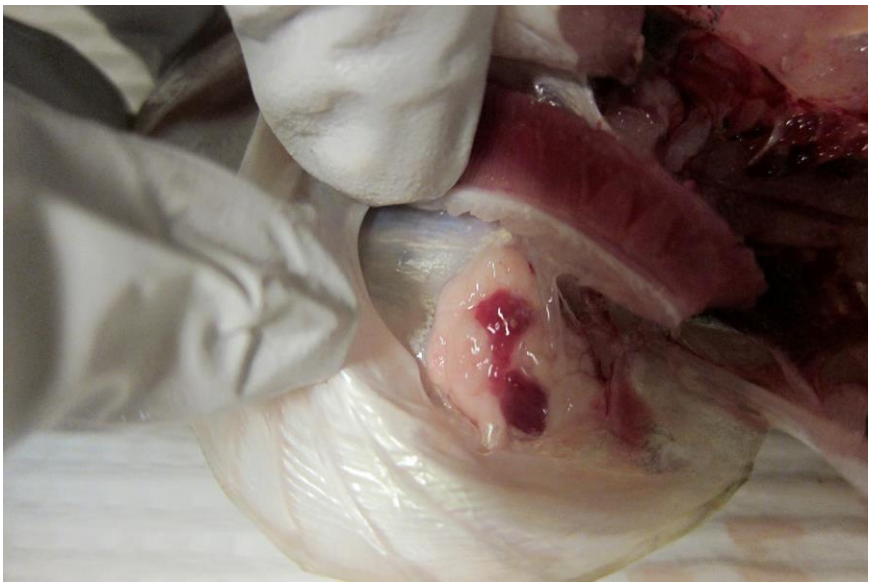


Figure 2: The white, granulate mass is the pseudobranch of a juvenile *Gadus morhua* infected with X-Cell parasites. Photo: Estelle Grønneberg

Micro-eukaryotic parasites can also live as endoparasites that live inside the host's body (Bellay *et al.*, 2008; Brooker *et al.*, 2008). After the parasite has been ingested or penetrated the host's skin, the spores of endoparasites can either live intracellular or intercellular, depending on their reproduction and life strategies. The spores travel around the host's circulatory system until they find their target organ or tissue, where they continue with their life cycle.

Micro-eukaryotic parasites' life cycles may involve one or several hosts, in which they can have different life stages (Brooker *et al.*, 2008). Parasites with indirect life cycles involving developmental stages with two or more host species can only persist in an area if all hosts necessary for the completion of their life cycle are present. Such parasites can multiply very fast and become a serious threat (Heuch *et al.*, 2011). The malaria causing alveolate *Plasmodium falciparvum* is a prominent example of a micro-eukaryotic parasite that has more than one host. This parasite's alternate hosts are mosquitos and humans. Humans are the final hosts and can get infected by the parasites, when bitten by the intermediate mosquito host (Brooker *et al.*, 2008).

Even though there are few studies that have addressed hyperparasitism (parasites infecting other parasites) within micro-eukaryote taxa, there are examples that indicate that they are more common than described. A review paper by (Skovgaard, 2014) mentions several cases of hyperparasitism including the microsporidian *Metchnikovellida* that infects gregarines, which themselves are parasites of polychaetes (Sokolova *et al.*, 2014). Other examples for hyperparasitism are: *Amoebophrya sp.* that parasitizes hosts like the dinoflagellate *Oodinium sp.* or the gregarine *Paraophioidina haeckeli* that infects the copepod *Sapphirina spp* ..

Many of these micro-eukaryotic parasites can cause major threats to fish in more or less enclosed areas, depending on the nature of their life cycle. If the parasites for example have a one-host life cycle with fish as its host, it could cause a tremendous threat in areas such as spawning-grounds, fish farms and feeding-grounds were fish occur in higher densities, and thereby infect each other (Burt & MacKinnon, 1997; Hemmingsen & MacKenzie, 2001; Bricknell *et al.*, 2006).

Examples of serious pathogens are Microsporidia and Myxozoa. They are known to have complex life strategies as those mentioned above (Sleigh, 1992). Through addition of molecular data, it has become more apparent that they contain cryptic as well as paraphyletic taxa, and are therefore difficult to study. Although there are many species that have been described morphologically and taxonomically (e.g., Morrison & Sprague, 1981; Heuch, 2001; Rodriguez-Tovar *et al.*, 2003; Køie *et al.*, 2007; Karlsbakk & Køie, 2012) and to a lesser degree, studies examining transmission pathways and pathology, environmental sequencing studies show that there is still a huge diversity of organisms that have not been described.

Microsporidia

The Phylum Microsporidia is a diverse group of intracellular eukaryotic parasites, within the super-group Ophistokonta. Over 1400 species in 187 genera have been described (Vávra & Lukeš, 2013; Didier *et al.*, 2014). Microsporidia are very small unicellular organisms, mostly 2-7 μm with a eukaryotic nucleus (Bylen, 1994). They are found in a vast variety of hosts, and in different types of tissue cells and organs, for example in the urinary bladder, muscle cells, cartilage, gastrointestinal tract, gallbladder, skin, gills, fins and blood (Nilsen, 2000). Microsporidia can cause diseases like nosematosis (*Nosema apis*) of the honey-bee (*Apis mellifera*) and microsporidosis in immune-deficient humans.

Before the introduction of the molecular phylogeny, microsporidia posed a difficult evolutionary problem for scientists. Historically, the description of genera and species were based on characteristics of spores and developmental forms that could be observed under light microscopy, as well as characteristics of infections and host- and tissue-specificity (Hylis *et al.*, 2005). However, the unique structure and biology of these micro-eukaryotic parasites made it difficult for scientists to classify them in correct phylogenetic position. The microsporidians lack several features considered to be universal to eukaryotes, including mitochondria, classically stacked Golgi-membranes, peroxisomes, 80S ribosomes, cilia and flagella. Eventually, the absence of these "eukaryotic" features in the Microsporidia, particularly mitochondria, led to the proposal that Microsporidia never had these features because they diverged from other eukaryotes prior to the origin of these features. They were collectively called "Archezoa" because of the obvious lack of mitochondria, and were proposed to be ancient basal lineages in the eukaryotic tree of life (Keeling, 1998; Didier *et al.*, 2014).

Today, accumulated data and sophisticated analyses confirm that the microsporidians belong within the super-group Opistokonta, within the basal fungi group, like zygomycetes and cryptomycetes (Adl *et al.*, 2005; Vávra & Lukeš, 2013). This is based on characteristics that most basal fungi possess, such as lacking flagella, a unique eleven amino acid-long insertion in the EF -1α gene, and having their THS-DHFR genes separated (that usually are occur together) (Stechmann & Cavalier-Smith, 2002; Vivares *et al.*, 2002). Although Microsporidia do not resemble fungi in terms of morphology, they have some similar characters. The deposition of cell wall material on the cell wall membranes, the formation of chitin containing spores and the features of meiosis are some morphologic character traits that are shared. Molecular evidence that strengthens the relationship between Microsporidia and fungi are the sequences of their protein-coding genes like α- and β- tubulin, the TATAbox binding protein an the vacuolar ATPase as well as sequences from the rRNA genes (Vávra & Lukeš, 2013; Didier *et al.*, 2014).

Microsporidia are strictly intracellular protist parasites, and have an interesting, but not fully understood mechanism of cell invasion (Slamovits, 2004). The infective stage of the Microsporidia is the dormant spore, during which it has a specialized mode of infecting their host's cell (Slamovits, 2004; Vávra & Lukeš, 2013). They possess a unique feature which is a long, coiled, polar filament (also referred to as injection tube, polar tube or invasion tube) present in the spore, that they use to inject sporoplasm (spore content) into the host's cell upon spore germination (Keeling, 1998; Vossbrinck & Debrunner-Vossbrinck, 2005a; Vávra & Lukeš, 2013). It is hypothesized that changes in the extracellular environment of the host is the plausible cause of initiation of injection of the polar tube into the host's cell (Slamovits, 2004). Another hypothesis is that the infection may be initiated by a molecular anchoring of the spore to the host cell surface glycosaminoglycans prior to germination (Southern *et al.*, 2006). An interaction between host and parasite prior to infection may provide an understanding of the mechanisms behind host specificity of the different microsporidians (Slamovits, 2004).

Once the spore content is in the host cell, different life cycles will happen, depending on which host species the Microsporidium has infected. In general, there is a phase of meront[4] cell division, then a stage of sporogony (where the spore and the injection tube is made),

[4] A stage in the life cycle of sporozoans, in which multiple asexual fission occur, resulting in production of merozoites (Stedmans 2006)

followed by maturation and release of spores, that is presumed to be responsible for spreading the infection inside the host (Bylen, 1994; Slamovits, 2004). It is still unknown how the parasite is transported from the site of infection to the area of reproduction. One hypothesis suggests that microsporidians probably are transported passively by the haemolymph to all organs,and as soon as they encounter suitable conditions they start to develop (Canning, 1976).

As the Microsporidia lack mitochondria, they are obligated to attach themselves to the host's mitochondria, which are their only source for energy and which they are totally dependent on for survival (Bylen, 1994). Microsporidia have been recorded from a wide range of host taxa, some with commercial significance like farm animals, silkworms, honey bees and fish (Didier *et al.*, 2014). Fish seem to be particularly prone to microsporidian infection and are the definitive host for a large number of known vertebrate-infecting microsporidia (Rodriguez-Tovar *et al.*, 2003). At least 18 genera have been recorded in fish, and the majority appears to be host specific (Lom & Nilsen, 2003; Didier *et al.*, 2014).Those host-parasite pairs can be found in a wide variety of aquatic habitats, many of which have only recently been identified (Brown *et al.*, 2010).

They can transmit directly from fish to fish, or with an intermediate host (Nilsen, 2002). *Pleistophorea gadi* may cause a fatal muscular dysfunctions in juvenile cod (Nilsen, 2002). The parasites of the genera *Glugea* and *Pleistophora* can also infect tissue of fishes. The microsporidia invade the host is mesenchymatic cells (cells that lack polarity and are surrounded by a large extracellular matrix) and alter the gene expression in the nucleus, which makes the host cell divide uncontrolled, like tumour tissue. The infections look like lumps in the tissue, and are filled up with spores and other development stages (Bylen, 1994; Nilsen, 2002).

The spores' ability of being extracellular and tolerating environmental stress, in addition to microsporidias wide host range, makes the microsporidians capable of spreading fast. If these parasites can be transmitted from one fish host to another fish host, they will cause serious damage in areas where the hosts are close to each other (Heuch *et al.*, 2011). This would especially be critical in aquaculture, were it is hard to control the load and spreading of Microsporidia, as well as other fish parasites like the Myxozoa.

Myxozoa

Myxozoans are highly specialized endoparasites that constitute a vast assemblage of more than 2200 parasite species (Holzer *et al.*, 2010; Rocha *et al.*, 2011; Hartikainen *et al.*, 2014b). For a long time, there have been discussions about the taxonomical placement of these enigmatic organisms. Myxozoa were for a long time classified as protists because of their extremely simple morphologies (Canning & Okamura, 2003; Holland *et al.*, 2011; Hartikainen *et al.*, 2014b; Foox & Siddall, 2015). It is only a few decades ago that they were confirmed as being metazoans due to their differentiated cells and desmosomes[5], which are both metazoan apomorphies (Smothers *et al.*, 1994). Phylogenomic analyses based on 128 protein coding genes of *Myxobolus cerebralis*, as well as sequencing and characterization of a homologue minicollagen [6] (designated Tb-Ncol-1) in the myxozoan *Tetracapsuloides bryosalmonae*, with an addition of 50 protein coding genes confirmed the placement of Myxozoa in the Metazoa, more specifically within the phylum Cnidaria (Holland *et al.*, 2011; Hartikainen *et al.*, 2014b). Although myxozoans are now attributed to the status of phylum Cnidaria, many texts have continued to classify them as Protozoa (Canning & Okamura, 2003).

The myxozoa are characterized as extracellular endoparasites, which can be either coelozoic[7] or histozoic[8] (Canning & Okamura, 2003). They have a two host life cycle, usually alternating between marine- or freshwater fish and invertebrates (Canning & Okamura, 2003; Jimenez-Guri *et al.*, 2007). The group Myxozoa are constituted by two classes: The Myxosporea, which alternate between fish and annelid hosts, and the Malacosporea that parasitize mainly bryozoans with fish as the only intermediate host (Canning & Okamura, 2003). Few host species exert symptoms that one might think of as problematic, although seldom, some members of the Myxozoa can be severe pathogens of teleosts. Diseases caused by the parasite *Myxobolus cerebralis* is considered as one of the most devastating diseases among salmonid populations (Canning & Okamura, 2003). *Tetracapsula brosalmonae* and *Sphaerospora renicola* are two important pathogens in cultured fish that can cause proliferative kidney disease (PKD) of salmonids and swimbladder inflammation of carp, respectively (Canning & Okamura, 2003; Kallert, 2006).

[5] A plaque like site on a cell surface that functions in maintaning cohesion with an adjacant cell.

[6] A unique constituent of nematocysts only found in the Phylum Cnidaria

[7] The parasite inhabits the intestinal canal of the hosts body.

[8] The parasite lives in the tissue of the host.

It was previously thought that there were two groups of Myxozoa: the Myxosporea and the Actinosporea. It is now known that these were in fact two stages of the myxozoan life cycle. Once again this shows the potential pitfalls of using morphology studies as the only way of classifying species. Because life cycle and morphology studies are difficult to conduct and confirm with molecular data only about 25 life cycles have been elucidated today (Hervio *et al.*, 1997; Kallert, 2006). The complete chronological development of a Myxozoa species has only been carried out with *Myxobolus cerebralis* (Kallert, 2006). Therefore the following description of the myxozoan life cycle will be described in the light of *M. cerebralis* with some examples of other myxozoans.

Actinospores (the alternate stage of the myxozoan parasite) are released from oligochaetes to the environment and infect the teleost hosts though its skin or gills. *Myxobolus cerebralis* uses the triactinomyxontype actinospore, which are an infective spore with three projections that anchor in the epidermis, or to the intestinal tract if ingested during feeding. In the centre of the triactinomyxospore are the sporoplasm containing germcells, and polar capsules containing polarfilaments (Kallert *et al.*, 2005). The polarfilament shoots into the host's tissue, creating a pathway for the sporoplasm, from the actinospore to the tissue of the host. In the target tissue, the sporoplasmodia undergo mitosis to produce amoeboid cells that contain an enveloping cell and an inner cell (figure 3, 1-16). *M. cerebralis* amoeboid cells can then migrate through several tissues to reach their target area, for example teleost cartilage (Kent *et al.*, 2001). When the amoeboid cells have reached the target area, the cells undergo sporogeny (formation of new spores), resulting in multicellular myxospores. *Zschokella hilda, Gadimyxa atlantica,* and *Gadimyxa arctica* spores have been detected in collecting ducts, ureters of the kidney, gall bladder and urine of *Gadus morhua* (Køie *et al.*, 2007). The myxospores are composed of at least six cells (two polar capsules, two binucleate sporoplasms and two form protective valves). These myxospores are released from the intermediate vertebrate host to its surroundings, and are ingested by invertebrates, where they infect the intestine of the new host individuals. In the finite invertebrate host (17- 30 in figure 3) the myxospores undergo merogony and later sporogenesis (sexual stage), which results in the triactionmyxon stage spores that are again released into the intermediate host's environment (Hedrick *et al.*, 2004) (see figure 3).

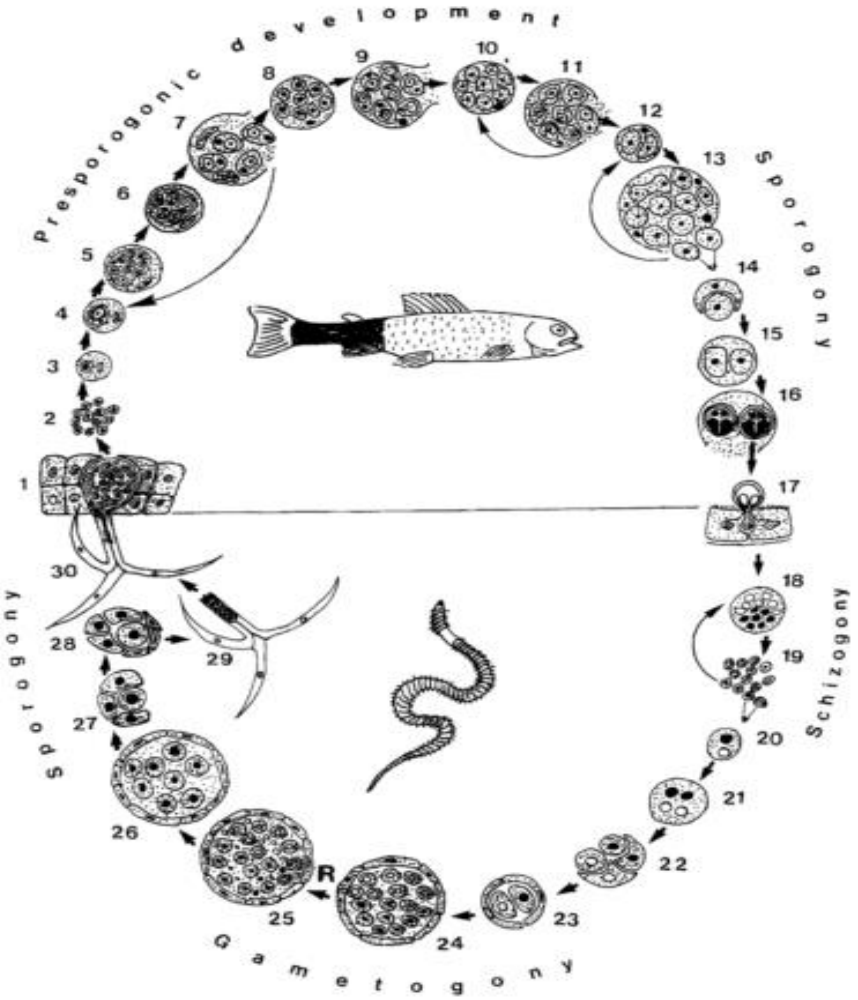


Figure 3: The life cycle and development of Myxozoa, based largely on the species *Myxobolus cerebralis.* Details can be found in the text, 1-16 is the development stage in the infected fish, and 17- 30, the development of actinospore in the invertebrate host (Kent *et al.*, 2001).

The existing knowledge on transmission pathways of parasites is limited. Myxozoa are thought to be released from fish hosts by death of the host, but there are results that indicate that they may be released from live fish as well (Kallert, 2006). As the spores are known to be enduring in the enviroment they may threatens to the fish stocks and aquaculture, even if there hasn't been an outbreak for a long time in the region (Kent *et al.*, 2001).

G. morhua has been described as one of the hosts of Microsporidia and Myxozoa. Both parasites have been found in the urinary bladder, but the information about transmission pathway, diversity of parasites and their life cycles is limited. With a broader knowledge about the transmission pathways of parasites, one might gain a better insight into how the cod gets infected. One can for example avoid building fish farms where several of the parasite's intermediate hosts resident. One can also predict the impact newly introduced species may

have on fish if they are potential intermediate host for parasites at the same time is is important to know which parasites might come with intermediate vertebrates in an area, where one might be doing some fishing or fish farming. It is also important to have a general knowledge about how the fish stocks will be in the future; if the areas with potential parasites are also areas where vulnerable juvenile cods spend their first years, it can cause a drop in the abundance of the cods. Furthermore a general knowledge of what kind of parasites infect cod is of general interest as the cod do migrate and therefore transports their parasites from one area to another they and can hence be a potential secondary threat to other organism in their surroundings. Such questions can be answered by increasing knowledge through the use of molecular methods, like the high sensitivity offered by high throughput sequencing.

High throughput sequencing

The development of new High Throughput Sequencing (HTS) methods has had an immense impact on the studies of microbial diversity in the last decade. It has given a more reliable annotation of species, in connection with morphological studies. Several studies (e.g., Karlsbakk & Køie, 2009; Freeman *et al.*, 2011b; Bittner *et al.*, 2013; Stentiford *et al.*, 2014), show that morphology studies have to be connected to molecular work to have more reliable results. Studies investigating microbial communities in fish have predominantly been based on Sanger sequencing (Star *et al.*, 2013). Previous research has, because of limitations with sequencing technology, overlooked a substantial amount of lineages, due to the species' low abundance in the environment (Pawlowski *et al.*, 2012; Bittner *et al.*, 2013; Lepère *et al.*, 2013; Logares *et al.*, 2014a). It is now known that this unseen diversity contains numerous species with unknown biological functions, yet to be identified (Logares *et al.*, 2014b). The increase in sequencing depth offered by HTS compared to Sanger sequencing allows researchers to detect new lineages that are less abundant (Guillou *et al.*, 2013; Lepère *et al.*, 2013; Logares *et al.*, 2014a).

The new HTS technologies provides millions of reads, instead of previously only hundreds, with the Illumina platform providing the highest sequencing depth per run (Kozich *et al.*, 2013). The percentage sequenced of the total diversity in an environmental sample depends highly on the sample type and its species diversity. Detecting the total diversity in an environmental sample with high diversity when using HTS in an Illumina Miseq is approximately 64% (Logares *et al.*, 2014a), and was therefore chosen as a tool for this thesis.

Illumina MiSeq sequencing enables read lengths long enough to cover the entire length of common genetic barcode regions targeting micro-eukaryotes (Pawlowski *et al.*, 2012). The

application of this method can improve the picture of the micro-eukaryotic community, identify rare taxa, and get a decent taxonomic assignment and phylogenetic placement. With amplicon sequencing[9] specific genetic regions called barcodes can be chosen to target specific taxonomic groups. This means that one can target only wanted gene regions and /or relevant organisms (Pawlowski *et al.*, 2012; Hadziavdic *et al.*, 2014). Optimal genetic regions for amplification are conserved in all organisms of interest, but do vary enough to distinguish species from each other. The V4 region of the 18S ribosomal subunit is the largest variable region in eukaryotes, but shorter in prokaryotes. There are some groups that are very different within the species or even the individuals in V4, like the radiolarians or very conserved like the haptophytes, where even one change in V4 can make a difference in species (Egge *et al.*, 2013; Decelle *et al.*, 2014). Several studies have applied this region in assessments of the composition of microbial communities, and concluded that it is the best suited region for biodiversity studies (Dunthorn *et al.*, 2014; Hadziavdic *et al.*, 2014). Incomplete species information and unclear microbial species definitions do limit next generation sequencing data annotation, when trying to link environmental sequences to species information (Boenigk *et al.*, 2012; Zinger *et al.*, 2012). Therefore, microbial signatures from environmental data are rather divided into operational taxonomic units (OTUs) based on sequence similarity than into actual species. Once the degree of sequence-similarity and different cut-off values has been chosen, the sequences that are similar are then grouped together and form one OTU (Zinger *et al.*, 2012).

[9] Sequencing of Polymer Chain Reaction (PCR) products referred to as amplicons

Material and methods

Fieldwork and sampling

A total of fifty juvenile cods, *Gadus morhua* were collected for parasitological examination during the summer of 2013 (July 5th to August 21th). The sampling took place at four different localities in the Oslofjord (see figure 4): Bergholmen (59°40`33.8”N 10°35`00.2”E), Hallangspollen (59°41`48.1”N 10°37`27.5”E), Kaldekota (59°40`07.4”N 10°35`41.1”E) and the Drøbak strait (59°40`02.0”N 10°37`17.8”E). The dissections were conducted in the laboratory facility at Tollboden, Biological research station, UIO in Drøbak.

Figure 4 : Map displaying the four sampling locations: Bergholmen, Drøbak strait ,Hallangspollen ,and Kaldekota.

The cods caught were under the Norwegian legal size, therefore a permit from The Norwegian Food Safety Authority was ensured before the fieldwork was performed (https://lovdata.no/dokument/SF/forskrift/2004-12-22-1878/KAPITTEL_10#KAPITTEL_10)

All cods were caught using fishing rods, line and bait - in our case frozen shrimp and herring. Fish were kept alive in containers with seawater and freezer packs during transportation, and whilst the fishing took place. Water temperature, depth and salinity was measured with an immersible temperature/salinity probe (YSI 30 TSC), to gain an impression on the environmental factors in the sampling sites. The measurements were taken at one-meter intervals from the sea surface to the sea floor, or to 15 m, that was the maximum cable length of the instrument.

Shortly after capturing, cods from different localities was transported to shore, and kept in tanks filled with seawater before dissection. The fish were killed by a blow to the head, at the last minute before dissection. Length and weight measurement were done instantly after sacrifice.

Fish dissection and tissue collection

Each fish was dissected in a dissection tray covered with Versi drylab soakers, to prevent any cross-contamination. The dissection tray was cleaned with 70% ethanol (EtOH) and Chlorine each time a new fish was to be dissected. The skin was then cleaned with Dermica disposable baby cloths dipped in 96% ethanol, preventing DNA from the outside of the fish entering the inside while opening the fish. This was done with the blunt end of surgical scissors to prevent the internal organs from bursting. A vertical cut towards the pelvic fins was done on each side of the fish, and from there a horizontal cut towards the anus and urino-genital opening was made. The internal organs were moved to the side, and a soft push to the rectum area was done, so that the urine bladder could appear easily. Furthermore, by clamping the anal canal with surgical clamps, faecal leakage in to the fish cavity was forestalled. Sex determination of the fish was done before the intestinal system was removed using scissors.

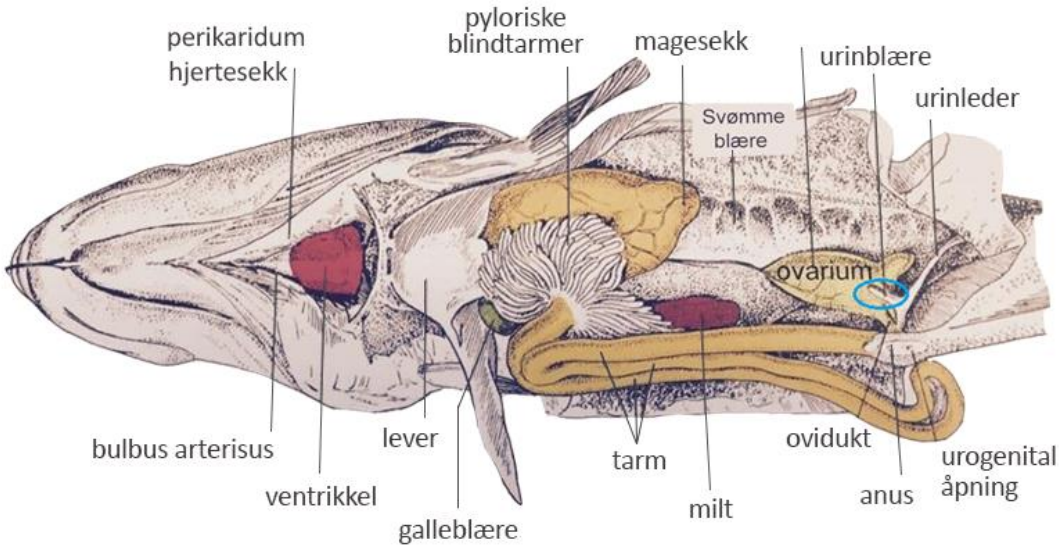


Figure 5: Anatomy of inner organs The blue ring in the picture shows the area of the urine bladder (Karlsen, 2012).

Urine was collected from the urine bladder using a BD microlance syringe with a 0,4 x 13 mm needle. The urine samples were transferred to cryotubes and immediately flash-frozen in liquid nitrogen. All samples were stored at -80°C until DNA extraction.

A general body examination (tail, skin, fins, mouth, gills and pseudobranchs) was done on every cod specimen to have an overall view on the cod's state of infections. If tumours were suspected or skin lesions observed, a picture was taken and a skin sample was collected from that area for further potential molecular examination. Macro-parasites were sampled as well, following the same procedure.

Finally, otoliths from each fish were collected for age determination. To get a clear view of the head cavity, a piece of the skull was removed with a sharp knife. The otoliths were removed with an acute tweezer and collected individually in Eppendorf tubes with 96% EtOH and stored at 4 °C. The age determination was done by counting the number of paired summer/winter growth zones. To clearly discern the growth zones, the otoliths were broken in two and slightly burnt before exanimating them under the kitmicroscope (Brothers *et al.*, 1976).

Molecular methods

Extraction of DNA

DNA was purified individually from the 50 urine samples. It was extracted and purified with the DNeasy® Blood & Tissue (Qiagen kit). The procedures were followed according to the manufacturer's recommendations. Due to the large differences in the amount of urine collected from the different individuals, (Appendix S1), the first step of the DNA purification procedure was carefully modified. Urine samples with a volume above 100 µL were treated as cultured cells and were centrifuged for 5 min at 300 x g. To have a visible pellet, some of the samples had to be centrifuged 2 minutes longer than the protocol recommended. Samples with a volume under 100 µL were extracted directly, as suggested by the supplier, for blood samples (ideal for samples of low quantity). The samples were only eluted once at the final step with 200 µL Buffer AE.

Gel electrophoresis and quantification

Gel electrophoresis was done using Agarose (1 %). 3 µL GelRed was added for every 50 mL of Agarose. 1 µL Gel loading dye blue mixed with 4 µL DNA sample was used per well. 4 µL Middle range (FastRuler^{TM}) was used as a DNA ladder. The gels ran 20-50 minutes at 80 volt.

All of the 50 purified DNA samples quantified using the Qubit ® dsDNA BR Assay kit (Invitrogen) and their purity was assessed by considering the A260/A280 ratio measured with Nanodrop ND-1000 (Thermo Scientific).

Amplification and sequencing

Polymerase chain reaction

The region of interest to be amplified using Polymerase Chain Reaction (PCR) in this study was the hypervariable region V4 of eukaryotic 18S rDNA. The PCR amplification was accomplished using general eukaryote forward primer 3NDF and the reverse primer V4_euk_ R2 (Bråte *et al.*, 2010a). PCR reactions were carried out using Phusion High-Fidelity Master Mix (Thermo Scientific) following the recommendations of the manufacturer.

Concentration of primers: 10 micro molar original/ working solution

The PCR was run in 20 µL reactions containing, 0.5 µM of each forward and reverse primer, 1X Phusion Master Mix and 1,6 µL template DNA and 1,6 µL BSA. To increase PCR yields and prevent adhesions of enzymes, Bovine serum albumin (BSA) was added to the PCR reactions. DNA template from brown algae was used as positive control. To ensure that enough DNA template was acquired for sequencing and to reduce amplification biases, PCRs were set up in three replicates. The conditions for the amplification procedure were as followed for 3NDF-V4_euk_ R2 primers: an initial denaturation for 30 sec at 98°C, followed by 25 amplification cycles of 10 sec denaturation at 98°C, 30 sec at 59°C (annealing), 30sec at 72°C (extension), followed by the final extension step for 10 min at 72°C in a thermal cycler (Mastercycler ep gradient S,Eppendorf, Hamburg, Germany). Each of the 150 PCR products was checked with agarose (1%) gel-electrophoresis for correct amplification size as described for the DNA. The quantity of the resulting amplicons were determined using Qubit dsDNA BR Assay kit (Invitrogen, Waltham, USA) and the purity was assessed with Nanodrop.

Primers used in the study:
Universal eukaryotic forward 3NDF (5'- GGAAGTCTGGTGCCAG-3') (Cavalier-Smith *et al.,* 2009)

Reverse primer V4_euk_ R2 (5'- ACGGTATCTRATCRTCTTCG-3') (Bråte *et al.*, 2010b).

Preparation for sequencing

The replicates were pooled according to the individual fish. Since one of the main aims of this Master thesis was to get a general overview of the fauna in the cod's urine, focusing on the cod in the Oslofjord, the samples were equimolarly pooled into one tube (Appendix S3). ChargeSwitch® (Invitro) was used to clean the pooled PCR products. The clean-up was done according to the manufacturers protocol, with some minor modifications. The incubation steps were prolonged to three minutes, and the transfer of the supernatant in the elution step was repeated to ensure that no magnetic beads were left behind in the samples. Concentration and impurity measurement was done with Qubit and Nanodrop ND-1000 (Thermo Scientific) before sending the sample for library preparation and sequencing at the Natural History Museum (NHM) in London.

The total of submitted DNA was 1720 ng.

Sequencing

Illumina MiSeq sequencing was the optimal sequencing platform for this study, due to its long read length, high output and low error rate, in comparison to other HTS platforms. Illumina Miseq offers a read length of 2x 300 base pairs (bp) and a total output between 13.2 -15 Gigabases with the V3 chemistry introduced in the beginning of 2013 (or end of 2012) (Loman *et al.*, 2012). The Miseq system leverages Illumina sequences by Sequencing by synthesis technology (SBS), where the base order in the DNA sample is determined when fluorescently labelled nucleotides attach to the DNA templates and emit light after laser excitation (Illumina, Miseq system).

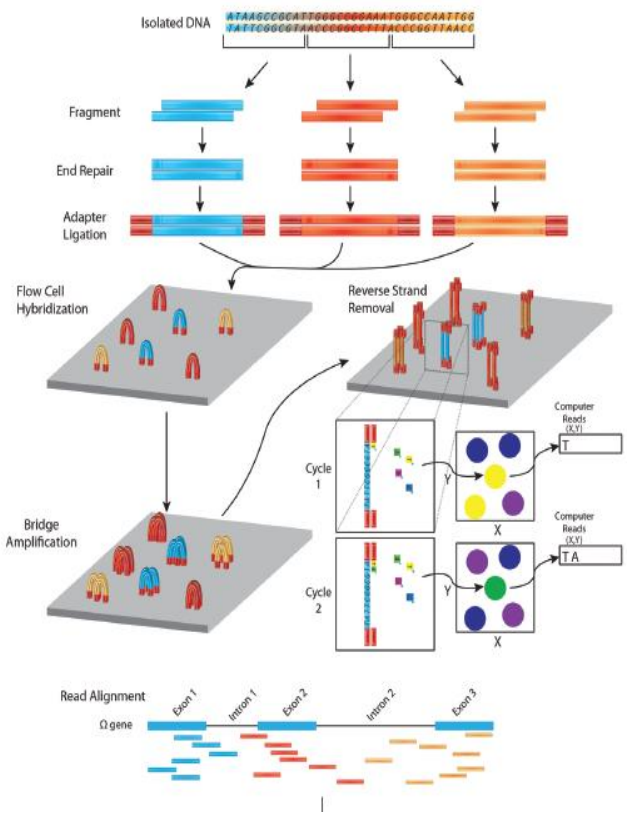


Figure 6 : Overview of DNA sequencing, using the Illumina platform (Churko *et al.*, 2013).

Before SBS can be performed, oligonucleotide adaptors, indexes and seuquencing primer-binding sites have to be ligated to both ends of the amplicon (DNA fragment). The adaptors attach to its complementary oligonucleotides in the flow cell (Churko *et al.*, 2013). On the surface of the flow cells, a polymerase forms a reverse complimentary strand, to the original amplicon. The fragments are then denaturized, and the original strand is removed. The DNA fragments on the flow cell undergo a hybridization step by bridge amplification, and then form clusters of identical sequences. The reverse strand is removed, and the forward strand is left to be sequenced (Churko *et al.*, 2013).

During the actual sequencing, fluorescently labelled nucleotides guanine (G), cytosine (C), adenine (A) and thymine (T) are added to the flow cell, and pass each cluster of sequences. One at a time, the nucleotides bind to the complementary strand and emit specific light signal upon excitation for each of the nucleotides. The light is detected and used to determine the

base order of the complementary DNA strand. After 300 cycles, the read products are removed and the procedure is repeated with the other strand in paired end sequencing approaches done. This way, both ends of the DNA fragment can be sequenced and detections of rearrangements such as insertions, deletions and inversions is possible as well as an assemble of longer fragments, due to overlapping sequences that cover the whole V4 region.

The DNA sample was prepared for sequencing with Regular TruSeqTM adapter ligation, pooled with 7 samples, and a radiolarian transcriptome was added to increase sequence diversity and loaded onto one run with Illumina MiSeq 300bp paired-end. Library preparations, sequencing and demultiplexing of the data set was done at NHM in London.

Bioinformatics

Bioinformatics processing of data was done using the Abel computer cluster, which is owned by the University of Oslo and the Norwegian meta-centre, (NOTUR) for computational science and operated by the Department of Research Computing (USIT). The dataset was mainly processed using Qiime 1.8.0 pipeline (Caporaso *et al.*, 2011).

The sequencing centre in London had sequence the samples twice, after low output the first time, which in return gave two datasetss. The raw sequence data (1 557 179 sequences) consisted of both forward and reverse reads, each about 300 bp in length.

Quality check

To have a quick visual impression of raw sequence data from the high throughput sequencing pipeline, a quality control was done with FastQC. It provide a set of analysis that can give you an overview of the quality of your data and awareness of whether your sequencing run encounteredany problems or not (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

The command was called like exemplified for file R1 of the first run:

```
fastqc /usit/abel/u1/estelleg/Urine-Gen-Euk_S8_L001_R1_001A.fastq -o
/usit/abel/u1/estelleg/Urine/miseq-qualities-fastqc/
```

The linux command `cat`, was then used to combine the two forward reads files and the two reverse reads files separately.

```
bash-4.1$ cat Urine-Gen-Euk_S8_L001_R1_001A.fastq Urine-Gen-
Euk_S8_L001_R1_001B.fastq > ~/Urine/combined_reads/R1comb.fastq

-bash-4.1$ cat Urine-Gen-Euk_S8_L001_R2_001A.fastq Urine-Gen-
Euk_S8_L001_R2_001B.fastq > ~/Urine/combined_reads/R2com.fastq
```

Assembly of the forward and reverse reads was done using the pair–end read merger PEAR (version 0.9.6-64bit) (Zhang *et al.*, 2014).

A minimum limit of overlap (-v) was set to 40. As the sequence quality scores and the per base sequence quality were good according to FastQC (mean score was over 30 over the whole run), the parameter –q (quality score) to minimum 20, meaning that sequences that have a right end quality of lower than 20 were trimmed until the phred score was consistently higher thatn 20. With such a high error rate, corresponding to an error rate of 1%, pairing can

not be ensured. The minimum length after trimming (-t) was set to 220, all sequences shorter than that were excluded. The assembly was done in a job script using these command lines:

```
/usit/abel/u1/estelleg/pear-0.9.0-bin-64/pear-0.9.6-bin-64/pear-0.9.6-bin-64      -f
/usit/abel/u1/estelleg/Urine/combined_reads/R1comb.fastq                           -r
/usit/abel/u1/estelleg/Urine/combined_reads/R2com.fastq                            -o
/usit/abel/u1/estelleg/Urine/pear -v 40 -q 20 -t 220
```

The total read length after assembly was approximately 500 bp. This ensured a complete coverage of the V4-region of the 18S rRNA gene, that is about 400 bp (Decelle *et al.*, 2014).

Quality filtering

Quality filtering and trimming was performed with Prinseq (version lite-0.20.4) (Schmieder & Edwards, 2011). The sequences were trimmed from the 3'ends to exclude all ambiguous bases and nucleotides. From the FastQC results it seemed like there was a grouping at the sequence length between 260-300, in the assembled 500bp reads, the minimum length parameter was set to 220. All sequences ahorter than 220 was rejected, as well as all sequences with minimum quality mean scores under 30. Sequences with a Phred score at the 3' end under 25 were trimmed until the quality was high enough to be retained A phred quality score of 25 refers to a 99.5% base call accuracy. The command used was:

```
perl /usit/abel/u1/estelleg/prinseq-lite-0.20.4/prinseq-lite.pl -fastq
/usit/abel/u1/estelleg/Urine/Pear/pear.assembled.fastq -ns_max_n 0 -min_len 220 -
trim_qual_right 25 -min_qual_mean 30
```

Removal of unwanted sequences and clustering

To get an overview of the diversity in the samples, one have to cluster similar sequences together at a certain sequence similarity threshold to be used as a species proxy. This grouping of sequences is called Operational Taxonomic Unit (OTU). De novo clustering of the complete dataset was done with USEARCH (version 7.0.1001) within the qiime pipeline. Using USEARCH one can exclude unwanted sequences from for example host and larger metazoan organisms, and also check for chimeras (Edgar, 2010).

Clustering was done using the script `pick_otus.py`. To find the right clustering similarity for the dataset, clustering similarity of 99, 97, 95, 93 and 91 were tested separately on the dataset.

These four clustering similarities were run twice to check if there were big differences in the dataset with changing parameter g (that controls how many sequences there must be in a cluster to keep it in the analysis) as either five or two. Two removes only where only singletons, while five means that the smallest OTU contains five sequences. In the first run parameter g five was used, and in the second round parameter g two was used. The word-length in all runs was the default 64. The command also checked for chimeras and were:

```
pick_otus.py -i ~/Urine/Fasta_Urine/pear_prinseq_good_PKOu.fna -m usearch -o
~/Urine/Qiime_usearch/qiime_usearch_99/ --word_length 64 -s 0.95 -x -F intersection
-g 5
```

resultater:

The choice of 95% similarity threshold and g 5 as parameter was then made to not overestimate the diversity in the samples, but not underestimate either (Kunin *et al.*, 2010; Logares *et al.*, 2014a). Chimeras constituted only 3% of the dataset, giving the final OTU count at 123. The script `pick_rep_set.py` was used to build a dataset of representative sequences from each OTU The parameter –m (mode for picking the representative sequence) was set to most abundant, meaning that the most abundante sequence with an OTU chosen to represent the OTU.

```
pick_rep_set.py -i
~/Urine/Qiime_usearch/qiime_usearch_95/pear_prinseq_good_PKOu_otus.txt -f
~/Urine/Fasta_Urine/pear_prinseq_good_PKOu.fna -o ~/Urine/Pick_rep_set -m
most_abundant
```

Taxonomic assignment with SILVA

Assignment of taxonomy and preparation of OTU tables was done within Qiime using blast (-m), setting an e- value cutoff (-e) of 0.01 against the comprehensive Silvadata base (Quast *et al.*, 2013). This was used to get an overview of the diversity in the samples.

```
assign_taxonomy.py -i ~/Urine/Pick_rep_set/rep_set95.fna -o
~/Urine/Taxonomic_ass_Silva/rep_set95_blast_tax_assignments.txt -t
~/Urine/Database/Silva_108_database_curated/97_taxa_map_Silva_108.txt -r
~/Urine/Database/Silva_108_database_curated/97_rep_set_Silva_108.fasta -e 0.01 -m
blast

make_otu_table.py -i
~/Urine/Qiime_usearch/qiime_usearch_95/pear_prinseq_good_PKOu_otus.txt -o
~/Urine/Taxonomic_ass_Silva/otu_silva95_g5.t
able.biom -t
~/Urine/Taxonomic_ass_Silva/rep_set95_blast_tax_assignments.txt/rep_set95_tax_assig
nments.txt


biom convert -i ~/Urine/Taxonomic_ass_Silva/otu_silva95_g5.table.biom -o
~/Urine/Taxonomic_ass_Silva/silva_table95_g5.from_biom.txt -b --header-key taxonomy
```

Taxonomic assignment with NCBI

Due to limited taxonomical information of blast hits for certain groups in SILVA, reads annotated as Microsporidia and Myxozoa, were extracted into a new fasta file containing only OTU numbers and sequences from these phyla and were blasted against the NCBI nucleotide database using Blastn (http://www.blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul *et al.*, 1997).

Rarefaction plot

To determine whether all species in the samples had been detected, a α-rarefaction plot was made in Microsoft Excel (Office 2013) based on the calculation of observed species for randomly re-sampling the pool of sequences multiple times. Then the plotting the numbers of OTUs observed per sub-sample were plotted in Excel against the number of sequences in the subsamples that increased in size until all sequences were included.

Phylogenetic analyses

For the Microsporidia and Myxozoa a sequence alignment were constructed in Geneious (version 7.1.4), against reference alignement extracted from literature and NCBI. The reference alignement for the microsporidians was taken from the article Vossbrinck & Debrunner-Vossbrinck, (2005) with additional unpublished microsporidia sequences received from Egil Karlsbakk. Sequences for the Myxozoa reference tree was taken from the article Kodádková *et al.* (2014), with additional Myxozoa sequences received from Egil Karlsbakk. The accession numbers of the Myxozoa and a separate for Microsporidia species were used to extract the respective sequences from NCBI (http://www.ncbi.nlm.gov/sites/batchentrez). The sequences were then compiled into a fasta-file for Microsporidia and Myxozoa, respectively, and imported to Geneious. A fasta-file containing the representative sequences from both groups, was then imported and combined with the reference 18S sequences. For alignment the MAFFT-algorithm (Katoh *et al.*, 2002) was used. Gblocks was used to remove ambiguous sections with little phylogenetic value (http://molvol-cmima.csic.es/castresana/Gblocks_server.html). Following the suggestion of Dunthorn et al. (2014), the parameters were set to the least stringent alternative. A Fast Tree for each of the alignments was conducted to see where the species grouped.

J-Model Test 2.1.7 was used to compute which substitution model that could fit the dataset. J-model test computes the likelihood scores, AIC and BIC scores, and gives the substitution model that fits the input sequences best.

```
/cygdrive/c/Documents and Settings/estelleg/Downloads/jmodeltest-2.1.7-win.tar/jmodeltest-2.1.7-win/jmodeltest-2.1.7

$ java -jar jModelTest.jar
```

The finaltree building was done on freebee using RAxML and MR Bayes (version 3.2.1). General time reversal (GTR) + gamma +I was used for RAxML and GTR + gamma was used for MR Bayes. In MR Bayes the Markov Chain Monte Carlo (MCMC) algorithm used four heated chains and was set to a chain length of 1 100 000 generations for Myxozoa and 2 200 000 for Microsporidia, sampling at every 200 generation , and a burn-in at 10 %.

```
Myxozoa

RAxMl
nohup raxmlHPC-PTHREADS-SSE3 -T 8 -m GTRGAMMA -b 12345 -# 100 -s MyxozoaSeqEK2usegblock.phy -p 1000 -n bootstrap &

nohup raxmlHPC-PTHREADS-SSE3 -T 8 -m GTRGAMMA -# 100 -s MyxozoaSeqEK2usegblock.phy -n topology -p 1000 &


nohup raxmlHPC -f b -m GTRGAMMA -s MyxozoaSeqEK2usegblock.phy -z RAxML_bootstrap.bootstrap -t RAxML_bestTree.topology -n Beste_Tree_With_bootstrapMyxofinal


nohup mrbayes MyxozoaSeqEK2usegblock.nex &

Nst=6 rates=gamma ngammacat=4;
prset brlenspr=unconstrained:exponential(100) shapepr=exponential(100);
mcmc ngen=1100000 samplefreq=200 printfreq=1000 nchains=4 temp=02 savebrlens=yes starttree=random;
set seed=2752;
sumt burnin=500



Microsporidia

nohup raxmlHPC-PTHREADS-SSE3 -T 8 -m GTRGAMMAI -b 12345 -# 100 -s Microsporidiafinalgblock2use.phy -p 1000 -n bootstrap &

nohup raxmlHPC-PTHREADS-SSE3 -T 8 -m GTRGAMMAI -b 12345 -# 100 -s Microsporidiafinalgblock2use.phy -p 1000 -n bootstrap &

nohup raxmlHPC -f b -m GTRGAMMAI -s Microsporidiafinalgblock2use.phy -z RAxML_bootstrap.bootstrap -t RAxML_bestTree.topology -n Best_tree_with_bootstrapMicro2 &

nohup mrbayes Microsporidia2use137seqgblock.nex &


Nst=6 rates=gamma ngammacat=4;
prset brlenspr=unconstrained:exponential(10.0) shapepr=exponential(10.0);
mcmc ngen=2200000 samplefreq=200 printfreq=1000 nchains=4 temp=0.2 savebrlens=yes starttree=random;
set seed=14114;
sumt burnin=500;
```

Finished trees were further analyse using Figtree (version 1.4.29) and Adobe illustrator CC.

Graphic representations

Microsoft Excel (office 2010-13) was used to create graphs. Adobe Illustrator CC was used to improve picture quality.

Results

The purpose of this study was to investigate the diversity of micro-eukaryotic parasites in the urine of *Gadus morhua*, focusing on Microsporidia and Myxozoa, which are known to be resident the urinary system of *G.morhua*.

The V4 region of 18S was investigated using amplicon sequencing and general eukaryotic primers. In addition, metadata like weight, age, length and health status from the each collected fish was used to obtain a general overview of the individual fishes that were collected. This may be used later for example to further examine the individual cods.

This is the first study investigating the content of urine of *Gadus morhua*, with a special focus on parasitic micro-eukaryotes, using HTS. The general diversity was in fact quite low, but on the contrary there were surprisingly high presence of some species.

Sampling locations

All the sampling areas were located in the Oslofjord, near the research station (laboratory facilities) in Drøbak. Hallangspollen was the locality furthest away (4 km in linear distance). The other locations were approximately 1-2 km in linear distance from the research station. Hallangspollen is an enclosed area, where little wind enters, so the water is rather still. Drøbak strait on the other hand is located right in the middle of the fjord. As it is the narrowest part of the fjord strong currents influences this areas, especially in high and low tide. Bergholmen and Kaldekota are intermediate between Hallangspollen and Drøbak strait, as the water stillness depends on the tide and wind directions. Wind from the northwest will stir up the water in Bergholmen and Kaldekota. The depth where the fish were caught varied between 4 – 25 m.(Appendix S7).

Fish morphology

The average age, weight and length at the different locations is shown in the table below (Table 1). The cods caught in Drøbak strait were the biggest fish, with an average length of 33.3 cm and weight 390 gram. Cods from Bergholmen were intermediary with an average length of 28.9 cm and 240 gram. This was similar to Kaldekota but only two fish were caught there.Although Drøbak strait had on average the biggest cods, it was at Bergholmen the biggest cod was caught (length: 44 cm and weight 820 gram). Cod from the location Hallangspollen were the smallest ones regarding length and weight (length: 23 cm and weight 129 gram).

Average	Bergholmen	Drøbak strait	Hallangspollen	Kaldekota
Age (year)	1.86	1.9	1.29	2
Weight (kg)	0.24	0.39	0.16	0.22
Length (cm)	28.9	33.3	23	29

Table 1: Average age, weight and length of fish being caught in the respective sampling areas. Hallangspollen had on average the smallest fish. Drøbak strait had on average the biggest fish and Bergholmen is an intermediate of the two localities. Kaldekota shall not be taken under consideration of this study as it only contains two fish.

The age of the cods varied from one to four years. Within the different age groups weight and length are for the most part proportional, although there are some fish that differ from the trend (Figure 7). The one-year-old fish that have roughly the same length also have a somewhat similar weight. There is one fish in this age group that sticks out, a fish from Hallangspollen is 18 cm and only 18 g, which could be an indication of sickness or little food in the stomach at the moment the fish was caught. The two-year-old fish have a linear relationship between weight and length. Most of the fish caught cluster around the length of 30 cm and 300 g, but it is also the age group with extreme outliers, where the smallest fish is 26 cm and 60 g, and the largest fish 44 cm and 820 g. Both of those fish were caught at Bergholmen. In context of this thesis two fish are of special interest. They are of smaller size then expected in regard to (three and four respectively; figure 7, arrows), which can indicate poor health. There was only one fish caught in the four-year-old age group, as they usually are significantly bigger than the size that was targeted in the sampling approach.

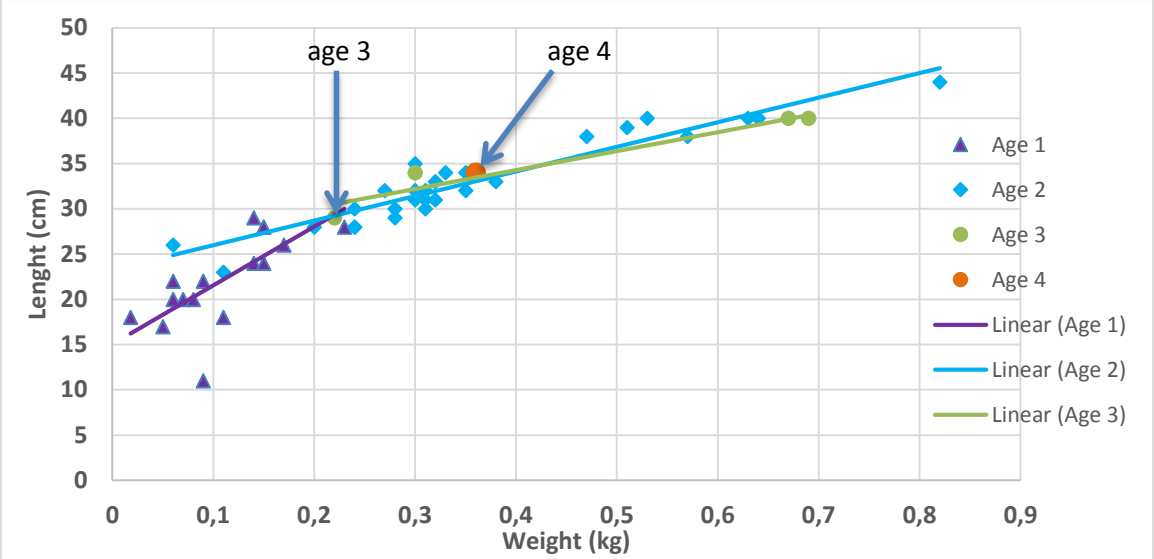


Figure 7: Correlation between age, length (cm) and weight (kg) of *G.morhua*. The different shapes represents each individual fish and their respective length and weight. The shape and colour represents the age group. A trend line has been added for each age group besides age 4, due to small sample size. Marked with arrows are the two fish that were significantly smaller that the others according to their age, which can be a sign of poor health.

Infections and macro parasites- a visual overview

In total 70 % (35 of 50) of the fish sampled had parasites and other disease like structures (tumors, fungus, skin lesions) that were visible to the naked eye. In Table 2 below one can see a tendency for prevalence of specific infections at the different locations. But it has to be taken into consideration that the fish metadata is based on a relative small sample size, and should only be regarded as an indication of the condition of cods in these specific areas, not as statistically significant in a broader sense. Data from Kaldekota should not be given any statistical weight as it is only based on two samples

In the two locations where most fish were caught (Bergholmen and Drøbak strait), Bergholmen (B.H) had on average a higher infestation of copepods 43 % (see Figure 8c), compared to 38 % in D.S and 29% in H.P (table 2). The amount of copepods found on the individual fishes and areas varied quite significantly. Some of the cods had only one copepod, and others had as much as 11 (D.S). In three of the four locations (B.H, D.S and H.P), over 50% of the fishes had the so-called "Black spot disease". "Black spot disease" is caused by the parasite larvae *Cryptocotyle lingua*. The parasite larva protects itself by forming resistant cysts. As a response to this the host encapsulate the parasite and encloses the cysts with black pigments (Figure 8a) (Karlsbakk *et al.*, 2009).

”White spot infection” was less abundant at all localities, with the highest percentage of 14 % infected fish in D.S. It is an infection with the Myxozoa parasite genus Myxobolus (Figure 8b). Myxobolus feeds of the fish cartilage and infects the eyeball (Baldwin *et al.*, 2000; Kaur & Singh, 2010). The table is only to be viewed as an indication of the general health of the cods in these particular localities, and not as a comprehensive analysis of the health of *G. morhua* in the Oslofjord.

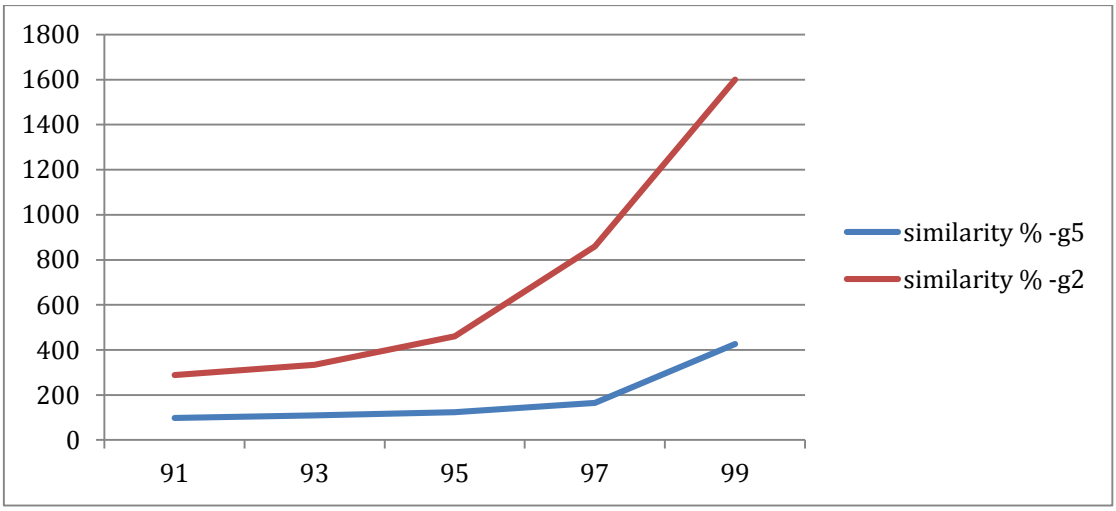
Diseases	**B.H**	**D.S**	**H.P**	**K.K**
Black spots	13 of 21 (62 %)	11 of 21 (52%)	4 of 7 (57 %)	0 of 2 (0 %)
Copepods	9 of 21 (43 %)	8 of 21 (38%)	2 of 7 (29 %)	2 of 2 (100 %)
White spots	0 of 21 (0 %)	3 of 21 (14%)	0 of 7 (0%)	0 of 2 (0 %)
Total infections	71.4%	61%	85%	100%

Table 2: Table giving an overview of the visual infections on G.morhua across the sampling locations. B.H = Bergholmen, D.S = Drøbak Strait, H.P =Hallangspollen and K.K = Kaldekota.

Figure 8. Pictures of the infections found on *G.morhua*. The picture **A** shows the head of *G.morhua* infected with *Cryptocotyle lingua*, picture **B** shows an eye of *G.morhua* infected with Myxobolus. In picture **C** is a *G.morhua* fin infected with copepods.

Clustering of OTUs

For describing the diversity in the sample, the sequences were clustered into OTU's based on sequence similarity. The range of cluster similarity that was tested varied from 91-99% similarity. The parameters g = 5 and g = 2 determine the lower limit for OTU-size for further analysis. G = 5 are all OTUs containing five or more sequences, g =2 are OTUs containing more than two sequences, meaning than only singletons were removed. With the value of parameter g as two the OTU number increased exponentially at 95 % similarity with the value of g at 5 the OTU number increased exponentially at 97 %. The least strict similarity setting of 93 % gave 109 OTU's with g = 5 and 334 OTU's with g = 2, with 95% similarity the OTU number was 123 with g = 5, and 461 with g =2, at 97% similarity the otu number was 165 with g =5, and 860 OTU's with g =2. The strictest one at 99 % similarity the OTU number was 426 with g 5 and 1601 with g =2. Based on these results, a further analysis with 95 % with g = 5 was done, as 95% similarity threshold and g 5 parameter do not overestimate the diversity in the samples, but probably not underestimate either (Kunin *et al.*, 2010; Logares *et al.*, 2014a).



Figur 9. Clustering of OTUs. The pictures describes clustering similarities from 91 -99 % with the parameters g =5 and g = 2 to determine the lower limit for further analysis.95 % similarity and the parameter g = 5 was chosen to further analysis as it does not underestimate or overestimate the diversity in the sample.

Rarefaction

A rarefaction curve was made to see if the sequencing depth adequately covers the diversity of the population in the sample. The rarefaction curve was created by randomly re-sampling the pool of samples multiple times and then plotting the average number of species (OTU) found in each sample. As more sequences were added at each round, more and more OTUs were found. This continued until it reached a plateau (were the curve flattens out), which was at approximately 1 200 000 sequences of the 1.5 million in total. The plateau indicates a complete coverage of the diversity has been achieved, and that adding more sequences will not give any more OTUs.

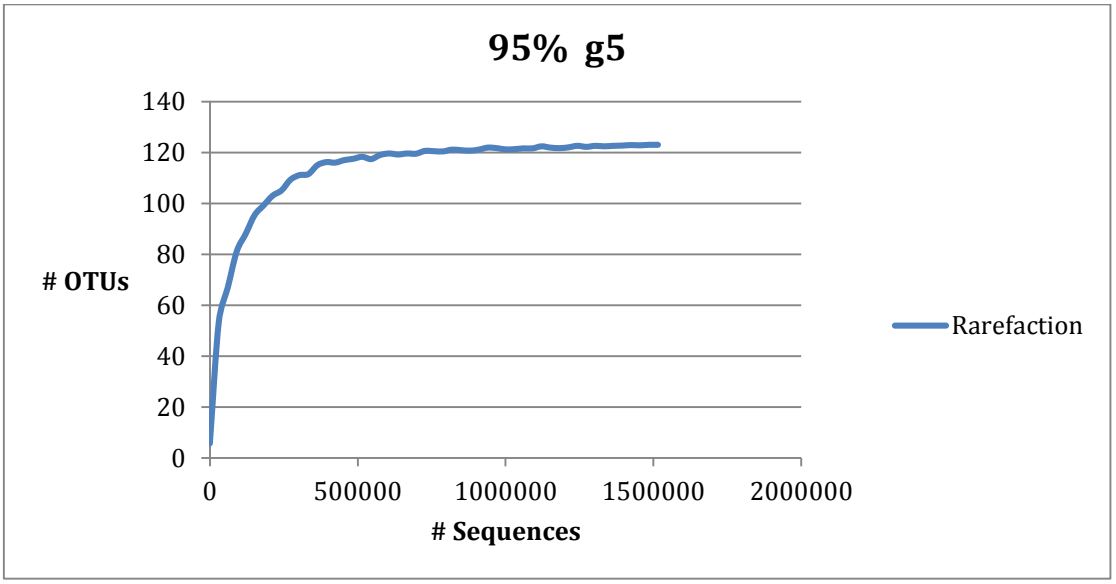


Figure 10: Rarefaction curve. The curve's plateau gives information of how many sequences that has to be gathered to find the complete coverage of the diversity in the sample.

Abundance of major taxa

After quality-filtering the total of number of sequences from the dataset was 1 515 114, and they were clustered into 123 OTUs, showing a rather low diversity. 70.4% of the sequences were assigned to Metazoa and 25.3% of the sequences were assigned to Fungi using the SILVA reference database. The rest (5%) of the sequences were divided amongst Stramenopiles, Alveolata, Streptophyta, Rhizaria and "no hits". The Metazoa and Fungi contain relatively few OTUs compared to the amount of sequences found. The Metazoa contain 28 OTUs (23% of the total OTU's) and the Fungi 21 OTUs (17% of the total OTUs) (Figure11). Sequences that showed "no hits" (sequences and OTUs with no taxonomic assignment) were blasted one more time against (NCBI), to see if there were any species that were not represented in the SILVA reference database. All the "no hits" OTUs comprised sequences that originated from PhiX Control v3, a virus genome that is used as a quality control for Illumina MiSeq sequencing runs. By removing the "no hits" from the analysis the Metazoa made up approximately 47% of the total OTUs.

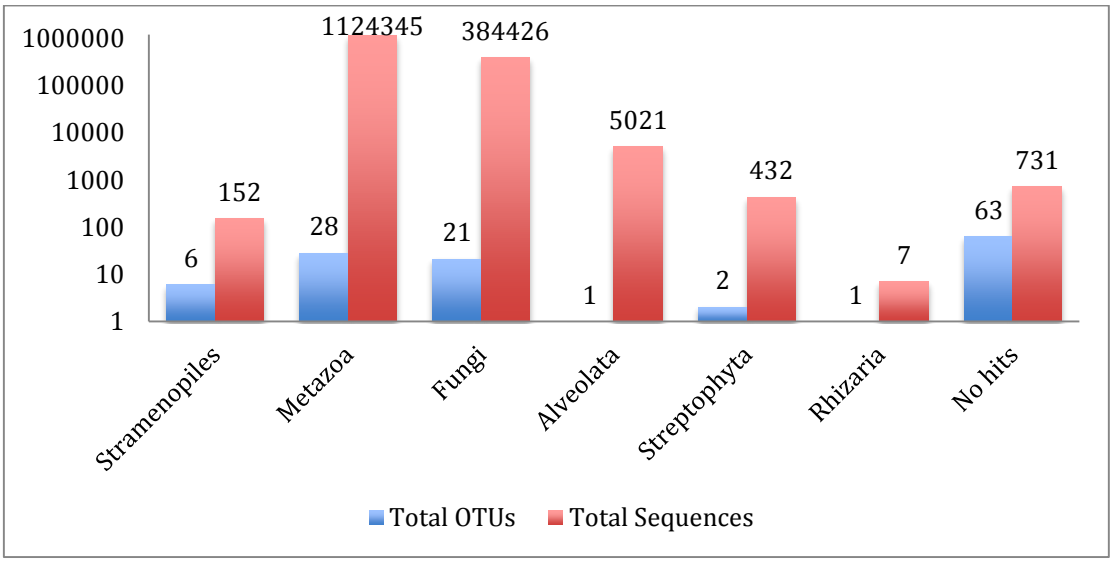


Figure 11: Graph showing the abundance of the major taxa found in the sample. The X-axis shows the major taxa, with their corresponding OTUs and number of sequences. The Y-axis is log- transformed to correspond to both number of sequences and OTUs.

Parasite overview of OTUs

Figure 12 shows that of the groups containing most OTUs (Metazoa and Fungi) the majority taxa was parasite with 79 % of the Metazoan OTUs and 51 % of the fungal OTUs. OTUs of the taxa Stramenophiles, Rhizaria and Streptophyta did not comprise any parasite taxa. The single OTUs of Alveolatea was also assigned to parasite. In total, approximately 56 % of the total OTUs recovered belong to parasite taxa.

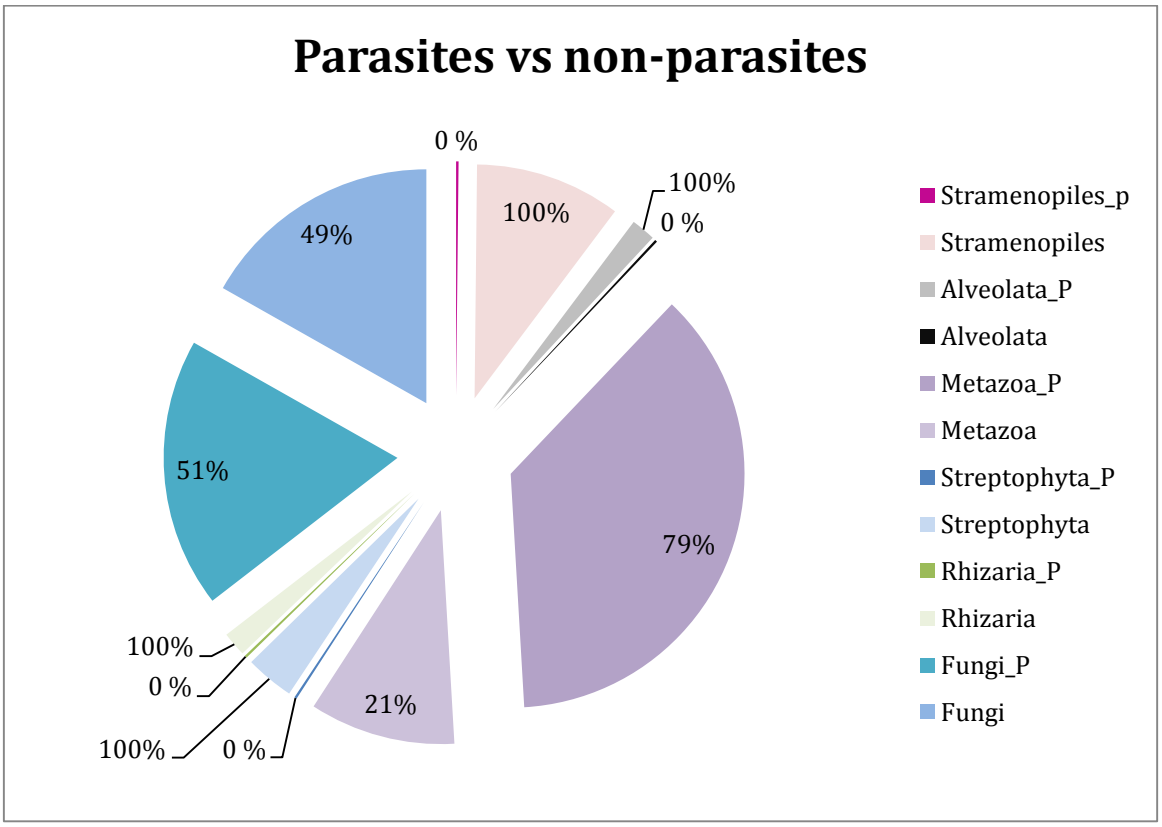


Figure 12: A proportional overview of which OTUs were parasites and non-parasites in the major taxa. Parasitic OTUs are labeled with taxonomic group and underscore P (e.g. Alveolata_P). The most striking of this results, where that 79% of the Metazoan OTUs were parasites and 50% of the Fungi OTUs were parasites. Overall 56% of the total OTUs were parasites.

Myxozoa

In figure 13 the diversity of Metazoa is shown. Of the 28 OTUs that comprise the Kingdom Metazoa, the phylum Myxozoa represents 78.6 % (22 OTus) of the total diversity. Craniata is the second largest phylum comprising 3 OTUs, all of them representing the host. The rest, which all only comprise 1 OTU, are Myriapoda, Phoroniformea and Eleutherozoa. The second bar (Figure 13 b) chart shows that the Myxozoa contains surprisingly low taxonomic diversity. Hundred percent of the 22 OTU's are of the genus Gadimyxa, divided into only two species in the urine sample of *G.morhua*: *Gadimyxa atlantica* comprises 14 OTus (64% of the total Myxozoa OTU's) and *Gadimyxa_sp_EK-2009* comprises 8 OTUs (36% of the total Myxozoa OTUs) (Figure 14). This seemed very suspicious, as previous studies have shown that cod urine had higher diversity of myxozoans (Køie *et al.*, 2007; Karlsbakk & Køie, 2009; Kodádková *et al.*, 2014). As suspected, the SILVA reference database did not contain a high diversity of Gadimyxa reference sequences. Familiar species known to be found in *G.morhua*'s urine, like for example *Gadimyxa arctica* and *Gadimyxa_sp_EK-2007* (Køie *et al.*, 2007) were not present in the database. The representative sequences of the 22 OTU's were therefore blasted manually against NCBI, and confirmed that there was a higher diversity of species detected by taxonomy assignment. This is illustrated in the phylogenetic tree ant the represented OTUs shown with in the tree (Figure 15).

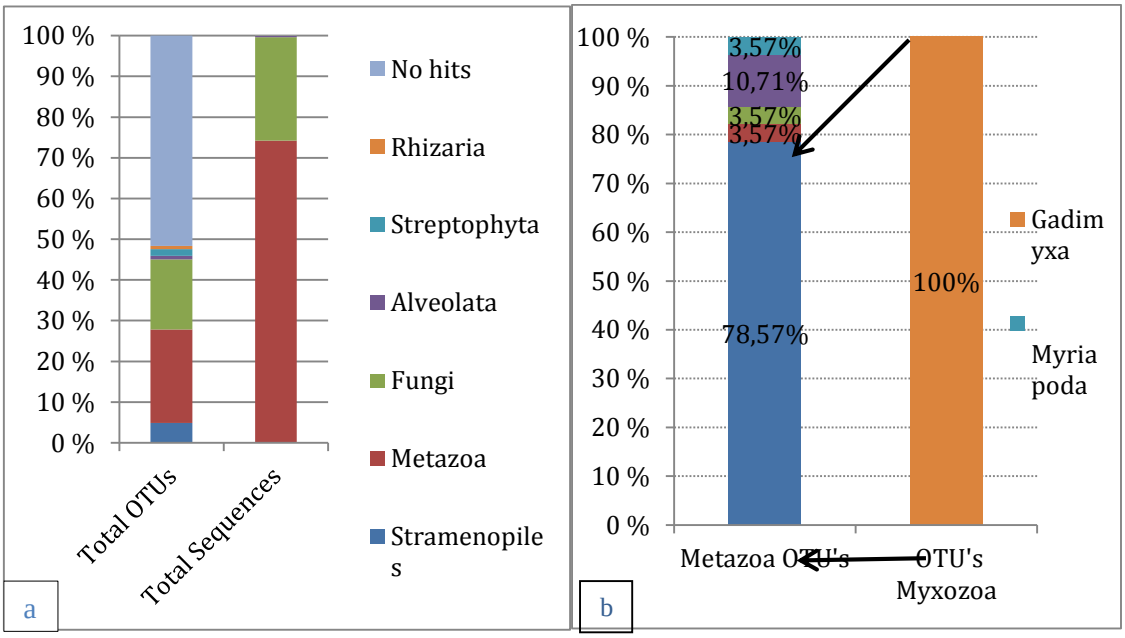


Figure 13: Figure 13 Diversity in urine : The Fungi contains 25.3% of the total sequences and 17% of the total OTUs. The Metazoa contains a 70.4% of the total sequences, but only 23% of the total OTUs. Myxozoa are the dominant phyla of the Metazoa kingdom. 100% of the Myxozoa phylum is the genus *Gadimyxa*.

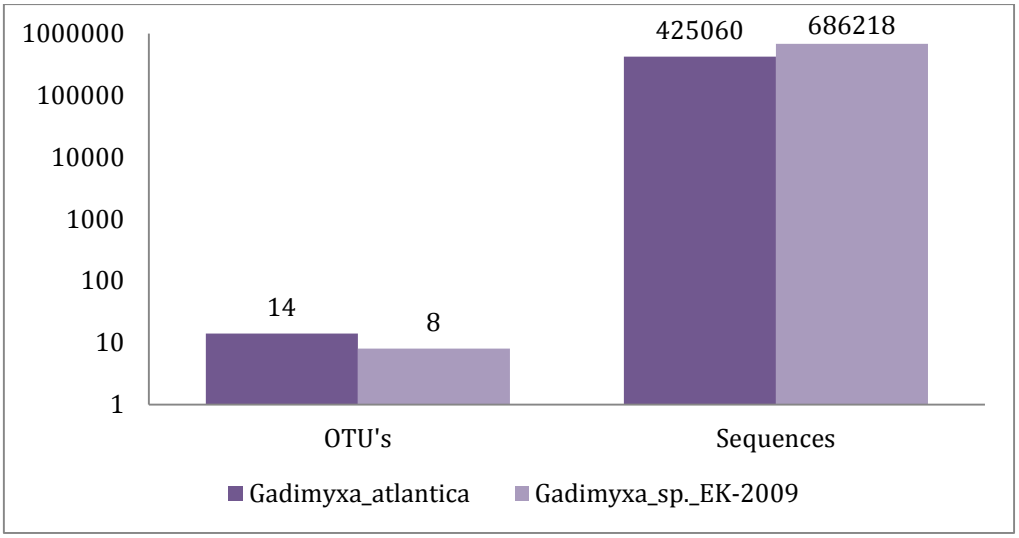


Figure 14: Taxonomic assignment of the genus Gadimyxa from SILVA database. Only two species were found using the SILVA database. *Gadimyxa atlantica* (14 OTUs) and *Gadimyxa sp EK-2009* (8 OTUs).

Phylogeny of Myxozoa OTUs.

To confirm the assignments of the OTUs found in the dataset, a phylogeny was made for the Myxozoa OTUs (figure 15). The reference sequences included in the alignment for the phylogenetic analyses of Myxozoa OTUs were based on a reference tree extracted from the article (Kodádková *et al.*, 2014). In addition sequences from Egil Karlsbakk that were not yet available in NCBI were included in the reference alignment (Appendix S7). All 22 sequences from the Myxozoa dataset cluster in a monophyletic group with a posterior probability of 99.39 Parvicapsula is basal to the Gadimyxa group with a posterior probability of 96.82. The Myxozoa group is again divided into two groups, where one group are clustered around the reference sequence Parvicapsula-indet-Sek6, and the other cluster as a subgroup around the remaining reference sequences of *Gadimyxa spp* .. The OTUs from the dataset do cluster to some extent with the respective sequence they were assigned to in the taxonomy assignment based on similarity by blast against NCBI. There are some OTUs that do not, though. For example 33_Gadimyxa_atlantica_100, which had a 100% similarity score with *Gadimyxa atlantica* (EU 163421) in NCBI, clustered more closely together with another OTU from the dataset (45_Gadimyxa_atlantica_99). However the posterior probability support is pretty low so one cannot rely on this information. Another example is the OTU 14_Gadimyxa_sp_EK-2007_95, which in NCBI had a 95% similarity to the reference sequence Gadimyxa_sp_EK-

2007 (FJ830379), how ever as seen in this tree it clusters closer to the reference sequence Parvicapsula-indet-Sek6.

In general the branches that includes the OTUs from the dataset are short. However, there are two exceptions, the OTU 123_Gadimyxa_sp_EK-2007_93, (OTU 123) that clusters with 80_Gadimyxa_sp_EK-2009_93 and the OTU 129_Gadimyxa_atlantica_91 (OTU129), that clusters with 107_Gadimyxa_atlantica_100. OTU 123 and OTU 129 branches has been cut down to 1/3 of its size, to ensure readability. Several of the clades from the dataset relatedness had a low support in both Bayesian posterior probability and maximum likelihood (Appendix S4).

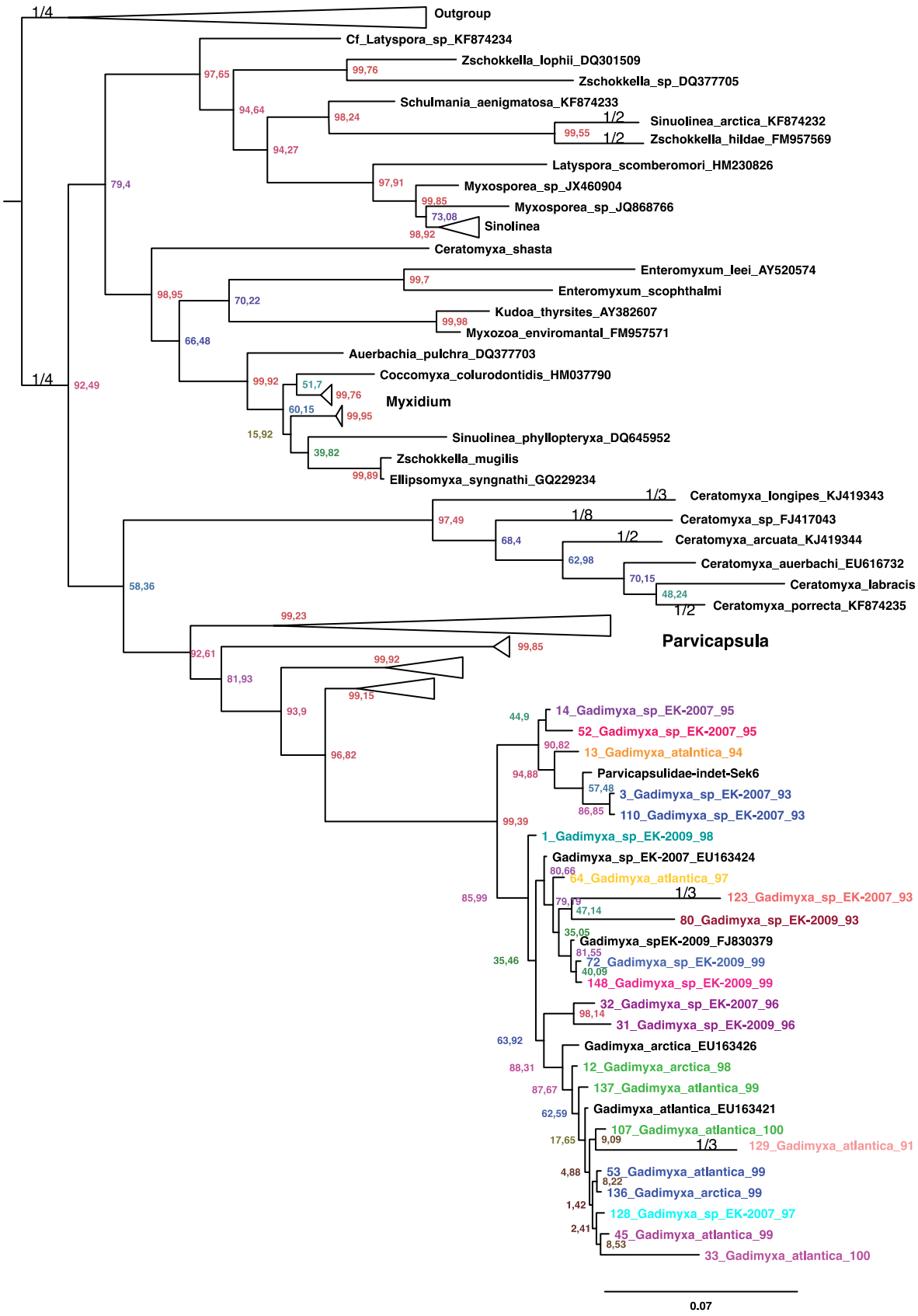


Figure 15: Phylogeny of Myxozoa constructed with Mr.Bayes, giving posterior probability of each cluster. This phylogeny is based on a reference tree from (Kodádková *et al.*, 2014) and supplementary sequences from Egil Karlsbakk. Each OTU has in this phylogeny been named firstly with an OTU number and then with the hit name and similarity percentage NCBI assigned to that OTU. For example : 33_Gadimyxa_atlantica_100, 33 = OTU number,

Gadimyxa atlantica = name assigned by blasting the sequence to NCBI and 100 = 100% similarity to *Gadimyxa atlantica* sequence that is present in the NCBI database. The names in black are reference sequences and accession numbers Collapsed branches contain all sequences belonging to the respective genus name. Numbers on reduced branches (e.g "1/3") is the length that is left after shortening the branches .OTUs that are coloured in the tree, have the same colouring in both MR bayes and RAxML trees. All the clades to the left of Parvicapsula, are collapsed groups of parvicapsula species.

Microsporidia

The kingdom Fungi comprised 23% (384 426 sequences) of the total sequences from the dataset, with 21 OTUs. Of these OTUs, 52 % were assigned to the phylum Microsporidia, giving a total of 11 OTUs of Microsporidia. Dikarya was the second most abundant subkingdom that contained 38% (8 OTUs) of the total fungal OTUs. The rest of the OTUs are divided between Fungi incertae sedis (1 OTU) and "uncultured" (1 OTU). Of the 11 Microsporidia OTUs 100% were assigned to the species *Trachipleistophora hominis* using the SILVA database (Figure 16). Again this seemed suspect, as *G.morhua* is proven to host more Microsporidia species than assigned with SILVA (Lom & Nilsen, 2003; Rodriguez-Tovar *et al.*, 2003). The sequences were therefore also blasted manually against NCBI. All OTUs were assigned to four sequences of the phylum Microsporidia. 4 OTUs were taxonomically assigned to the specie *Loma sp* (130 809 sequences), 5 OTUs (251 578 sequences) were assigned to *Microsporidium cerebralis* , 1 OTU (12 sequences) was assigned to *Kabatana sp* and 1 OTU (1327 sequences) was assigned to *Loma wallae*, (Figure 16 b). The results show that there is little diversity in respect of the total sequences, as only 4 species are found in 384 426 sequences, with the dominant species being *Microsporidium cerebralis* and *Loma sp*.

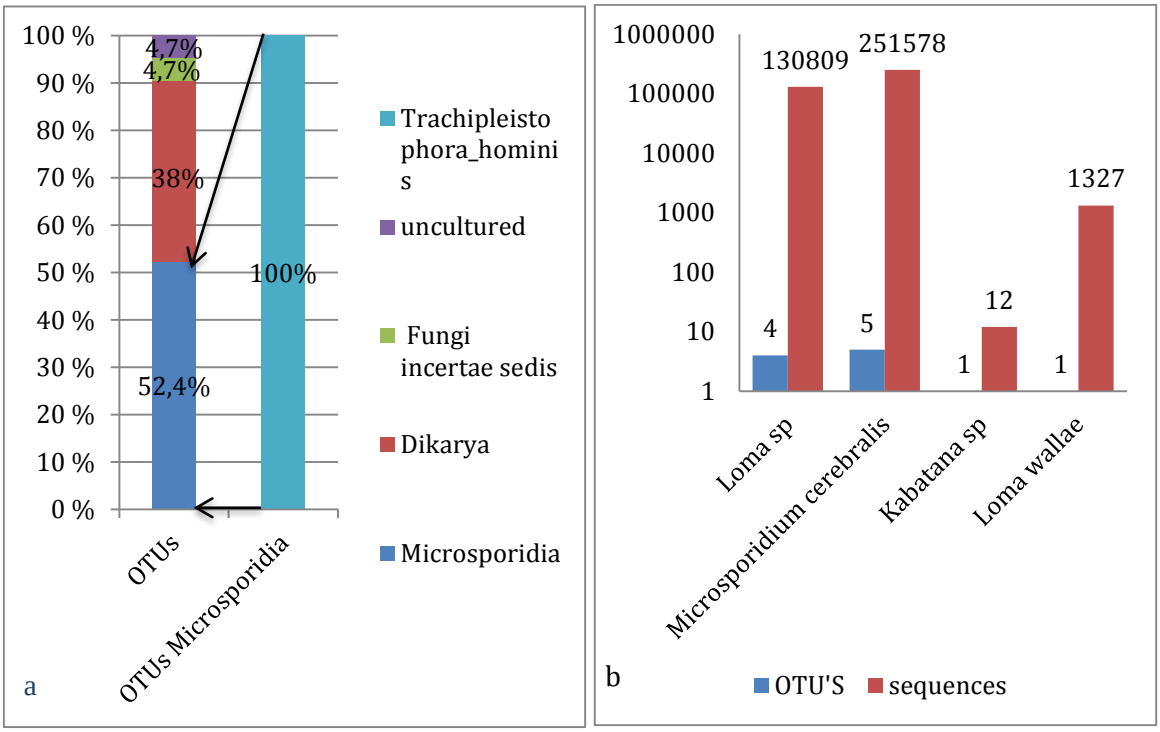


Figure 16: A) Diversity of the Fungi in urine: 52 % (11) of the total fungal OTUs (21 OTUs) are Microsporidia, and 100 % (11 OTUs) of the microsporidian OTUs are assigned to the species *Trachipleistophora hominis*. Four species was assigned to 11 microsporidian OTUs, *Loma sp, Microsporidium cerebralis, Kabatana sp* and *Loma wallae*.

Phylogeny of Microsporidia OTUs.

Based upon similarities and differences of genetic characteristics a Bayesian tree was made to show the evolutionary relationship of the OTUs from the dataset and other microsporidian reference sequences. The phylogenetic tree of Microsporidia OTUs was based on a reference tree found in the article (Vossbrinck & Debrunner-Vossbrinck, 2005a) with a supplement of sequences from Egil Karlsbakk that were not yet available on NCBI (Appendix S7). Two accession listed in the paper actually referenced to sequences of *Xenopus leavis* and *Pleistophora ovariae* and therefore had to be removed from the tree. The OTUs that comprised *Loma* and *Microsporidium* were all monophyletic grouped with a posterior probability of 84% and clustered basal to *Glugea* with a posterior probability of 83%. The OTUs assigned, as *Microsporidium cerebralis* did not groupe with their respective reference sequence based on similarity percentage from NCBI database, even though all OTUs had a similarity score over 90%. They rather made up an own monophyletic clade basal to the reference sequence *Microsporidium cerebralis* with a posterior probability score at 74 %. All of the OTUs comprising the species *Loma* clustered together with a posterior probability score

at 96 %, and are basal to the reference sequences *Loma_sp* and *Loma morhua_cod-EK* with a posterior probability score at 94%. 23_*Loma_wallae* does not cluster with its reference sequence *Loma wallae*, it clusters with the rest of the *Loma_sp* OTUs. 105_Kabatana_sp_JI-2008_99 was grouped with the reference sequence *Kabatana_takedai* with a posterior probability score at 91 %

In general the branches that include the OTUs from the dataset are very short, with a very variable posterior probability score varying from 22–93 %, meaning that several of the branches have a low support in Baysian tree. This is also the case for maximum likelihood score in the RAxML tree (Appendix S7).

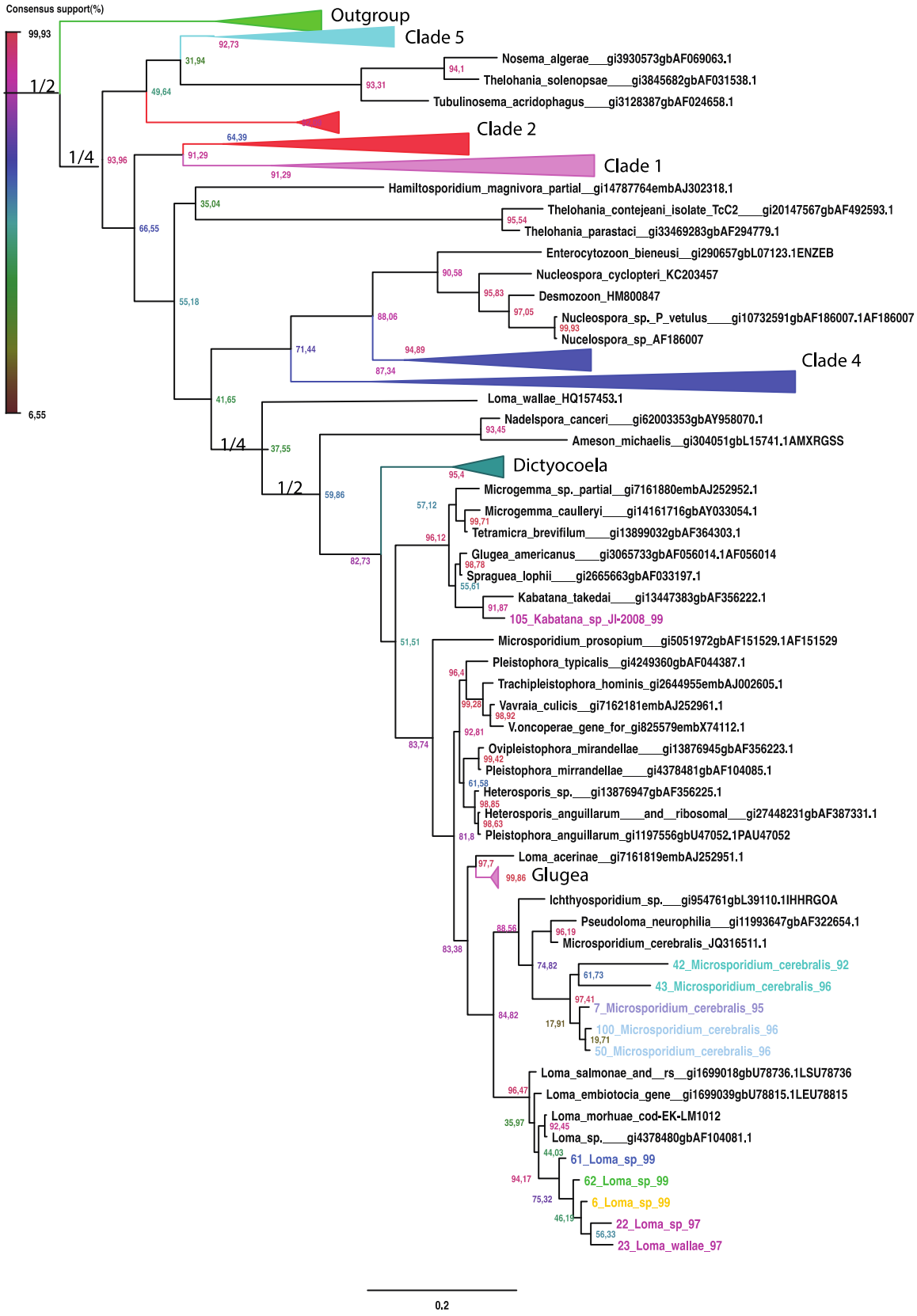


Figure17: Phylogeny of Microsporidia OTUs. This phylogeny is based on a reference tree from (Vossbrinck & Debrunner-Vossbrinck, 2005b) and supplementary sequences from Egil Karlsbakk. Each OTU has in this phylogeny been named firstly with an OTU number and then with the name and similarity percentage NCBI assigned to that OTU.

For example: 61_Loma _sp_99, 61 = OTU number, Loma sp = name assigned by blasting the sequence to NCBI and 99 = 99% similarity to Loma sp sequence that is present in the NCBI database. The names in black are reference sequences and accession numbers. Collapsed branches contain all sequences belonging to the respective taxon name. Numbers on reduced branches (e.g 1/4) is the length that is left after shortening the branches. The OTUs that comprised Loma and Microsporidium were all monophyletic grouped with a posterior probability of 84 % and clustered basal to Glugea with a posterior probability of 83 %. To easily compare trees from MrBayes and RAxMl are the OTUs from the dataset coloured. OTUs with the same colour are grouped together in the tree, and have the same colouring in both trees.

Discussion

Sampling location and visual inspection of *G. morhua*

In comparison to other fish species, *Gadus morhua* is a well-studied fish regarding parasites. An immense diversity of parasites has been detected, but the majority of these investigations have focused on studies of macro-parasites. In regards to the parasitic diversity in *G. morhua* urine, there are studies revealing that the urinary bladder of cods do contain a diversity of certain parasitic micro-eukaryotes (Heuch, 2001; Samuelsen *et al.*, 2006; Køie *et al.*, 2007; Heuch *et al.*, 2011; Glover *et al.*, 2013). Identification and taxonomic classifications of these parasites have mainly been done with the help of microscopy techniques and morphological descriptions (Schmahl *et al.*, 1990; Heuch *et al.*, 2011; Karlsbakk & Køie, 2012). These techniques are fundamental and important to confirm and identify infections in fish and other vertebrates, but in order to reveal the full extent of micro-eukaryotes diversity, molecular methods and deeper sequencing efforts as offered by current high-throughput sequencing methods are required (Pernice *et al.*, 2013).

This study represents the first time Illumina Miseq sequencing has been applied to uncover the diversity of micro-eukaryotes in the urinary bladder of *G. morhua* in the Oslofjord. By investigating the diversity of micro-eukaryotes in the cod's urine one can gain a broader knowledge of infections the fish are prone to. At the same time, we can improve our understanding of where the different parasites may reside in their hosts. Even though the sequencing depth achieved by the Illumina Miseq sequencing gives a much more detailed picture of the diversity of micro-eukaryotes in the urine of *G. morhua*, compared to previous morphological description and microscopy, it can only give an answer for what parasites are present in the particular cods that were caught and at the time and their given place of residence.

Cod in the Oslofjord has a long tradition within risk assessment in context of chemicals in the fjord. Food consumption has been a way of surveying the cod in addition with other environmental factors in the Oslofjord (Otterbech, 1953; Gjøsæter & Danielssen, 2004; Olsen *et al.*, 2004; Johansen *et al.*, 2009; Ski, 2012). Investigations of parasites on *G. morhua* the surveys are rather scarce despite the fish's high economic value. The cods in the inner Oslofjord have in recent studies been found to have a poor physical condition, regarding size, liver weight and age. Liver somatic index (LSI), determined as the ratio of liver weight to

body weight, and somatic body weight to length can be used to assess the energy reserve available to fish. LSI generally reflects the hepatic lipid content and may be affected by lack of nutrition. A low LSI may therefore indicate little food available (Nanton *et al.*, 2001; Morgan, 2004). A study conducted by Camilla Fagerli in 2006, showed that Gråøyrenna near this study's sample site Bergholmen had populations of cods with low LSI, compared to other sampling sites in the Oslofjord. Cods with poor conditions are more prone to parasite infections, as they don't have the energy to fight the infections off, making it easier for the parasites to take advantage of the fish host. As the parasites are of considerable energy cost to the fish host, it may weaken the fish substantially (Andreou *et al.*, 2012).

In this study, the body sizes of the cods were measured across the locations. The cods had some variations regarding to both parameters. The fish from Hallangspollen were on average the smallest fish in regards to length and weight, Drøbak strait had the biggest fish and Bergholmen was an intermediate between D.S and H.P. When considering the body size of the fish it is also necessary to consider the ages of the fish, as the fish naturally will become larger with age. In association with age it is not surprisingly that the fish from Hallangspollen were the smallest fish, as they are also on average the youngest (table 1). Hallangspollen is an area with little current disturbance, and is therefore an area one would expect to be habited by small cods, as predation and the energy costs to swim in in these areas are lower. Drøbak strait was the sampling area with the biggest and oldest fish on average (table 1). As Drøbak strait is an area with a lot of current it seems natural that this area would contain bigger fishes, and the age would therefore also correlate with the size. The cods caught at Bergholmen were bigger than the cods from Hallangspollen, but still on average they were 5 cm shorter and weighed 140 grams less than the cods from Drøbak strait, even though they were approximately the same age (Table 1). Bergholmen was also the sampling area where two of the cods (number 11 and 34) caught were significantly smaller in respect to their age (three and four years), (figure 7).

An explanation for the differences in body-size in the fishes from Bergholmen and Drøbak strait might be reflected in the variance of dietary composition in the sampling locations. As Bergholmen is a location with relative still water (depending on the wind directions), one could expect that the foods consumed by the cods are correlated with the substrate of the seabed. A study done by Fagerli in 2006 showed that the composition of food items in the cods' stomachs were correlated with the heterogeneous substrate in Bergholmen and was dominated by shrimps and invertebrates. Another study done by a fellow student in the same

project, Torbjørn Gylt, showed in 2014 that the food items in cods caught at Drøbak strait were more varied. Brittlestars, fish, hermit crabs, mussels and shrimps were consumed. One might speculate that in the Drøbak strait, more food in the sense of food item variety is available than in Bergholmen, making Bergholmen a place where the cods have a higher competition for food. Juvenile cods may then be out-competed by adult fish in that area. As juvenile cods usually a more stationary the first years of growth they might not move to places with higher risk for predation as for example Drøbak strait to get food, meaning they are more vulnerable for starvation. It has to be taken into consideration, that the cods were caught at different times of the day, meaning that some of the cods might just have eaten food and therefore weighed more than cods that had an empty stomach because it had not been out hunting for food yet.

The smalles fish of Bergholmen may be more infected by parasites worsening their condition. Parasitic infections caused for example by the X-cell parasite, may cause impairment for the fish as it makes lesions on the fish fins, skin and primary gill filament. Lesions at the primary gill filament may cause swelling and pale filaments (Figure 2), indicating low levels of oxygen and blood flow (Miwa *et al.*, 2004; Freeman, 2009). In a study conducted by Freeman (2009), high mortality in fish with X-cell infection of gill filaments was detected, as the fish showed a severe reduction of respiration. Reduced oxygen uptake can also be energetically costly for the fish, for example with regards to searching for food, as it will struggle with breathing while hunting for its prey. The fish may be outcompeted by healthier fish in gaining its food, hence having no spare energy in gaining weight and length.

The cods caught in this study were all checked thoroughly for parasites infections that were visible to the eye, to get a general idea of the cod's health before sending urine samples into sequencing. As seen in Figure 3 and Table 2 the cods from all locations had several visual infections. The most common infection was the black spot disease. The infection is caused by the parasite *Cryptoctyle lingua* and may cause mortality in juvenile cods if they are heavily infected. In older fish the infections are generally less severe, but can nevertheless cause massive infections that may be of distress for the fish (Karlsbakk *et al.*, 2009). Seen in Table 2 the fish were also infected by Copepods in all locations. Copepods may cause deformities in fins of fish which may drain the fish of energy, and as mentioned above making it harder to get food, making them more susceptible for other lethal infections as their condition are lower (Andreou *et al.*, 2012).

Infections from the genus *Myxobolus* was also among the infections observed in the fish sampled in this study although this was only found in the Drøbak strait, where the fish were seemingly healthier according to weight, length and age (Table 1). However, as can be seen in the appendix (Appendix S1) the *Myxobolus* infected fish had a slightly reduced weight (300 g) compared to the length of the fish (34 cm), but one cannot correlate the weight reduction with the infections as mentioned above, when it is a possibility that the fish might not have eaten before it was captured.

Studies like (Malmstrøm *et al.*, 2013) revealed that the unique immune system of *G. morhua* provides the cod with a greater tolerance for infections, and prevents diseases, parasites and bacteria from being malicious for the cod, as long as the single infections are small scaled. As the fish from each sampling location did not display a significant amount of visual parasites that could be correlated to size reduction and sickness, it was necessary to compare the visual signs with molecular data, which would reveal if the cods were carrying cryptic infections. The fish can appear to be healthy in sense of visual infections, but still bear massive internal microbal infections (Køie, 1984; Hemmingsen & MacKenzie, 2001; Nilsen, 2002; Karlsbakk *et al.*, 2009; Freeman *et al.*, 2011a).

There is very little knowledge about the parasitic fauna of the urinary bladder of *Gadus morhua*. The fish can contain internal infections that can affect the fish's health, and it is therefore important to obtain an overview of what diversity one can find in different target organs, and if the parasites are found harmful to their host. It is also important to find out what kind of harm the parasite may cause in the target organ in the light of the parasites' transmission pathways. Urinary excretion might be a transmission pathway for parasite spores, and can therefore be an important factor for spreading them. Microsporidia and Myxozoa are parasitic groups that are commonly known from earlier research to parasitize *G.morhua* and were highly prevalent and good candidates to have a special emphasize on.

Ensuring quality of the dataset to assess the parasitic fauna of the urine of cod

To get sufficient phylogenetic information for taxonomic placement, the barcoding region chosen for HTS had to be variable and short enough to be sequenced. The V4 region of the 18S rRNA gene was used, from recommendations from several studies like that say that theV4 region gives most similar results to those from full-length SSU rDNA sequences. 18s is a well known barcoding region, widely used in protists. The regions are however is too long, therefore universal general V4 eukaryot primes were used as it shows similar results in phylogenetic studies as full length (Pernice *et al.*, 2013; Decelle *et al.*, 2014).

PCRs were run on an DNA template to infer as DNA is a more stable molecule than RNA, and it would also include findings from spores and dormant cells that are known to represent a considerable proportion of the micro-eukaryotes detected in fish hosts. Spores and dormant cells can be especially abundant when they are about to spread to the environment and a new potential host to be able to detect all potential taxa, we used general primers to the V4 region(Heuch, 2001; Køie *et al.*, 2007; Holzer *et al.*, 2010; Lennon & Jones, 2011; Karlsbakk *et al.*, 2013). It is however important to mention that no primer are universal, and some species may therefore no become present (Hartikainen *et al.*, 2014a)

Furthermore, to avoid overestimation of micro-eukaryotic diversity, a conservative OTU clustering cut-off was chosen at 95 % similarity (Lepère *et al.*, 2013; Logares *et al.*, 2014a). The choice of 95 % similarity threshold was made in order to minimize inflation of diversity estimates (Kunin *et al.*, 2010; Logares *et al.*, 2014a). A high clustering threshold could inflate the species richness of the sample, by placing sequences that only differentiate with singletons, and might be the same species into different OTUs. The same way a very low similarity threshold can give a very conservative estimate of the alpha diversity. To assign a specific cut-off level is in general not an optimal strategy a no cut-off level will correctly delimit sequences into true species. Optimally different groups of micro-eukaryotes would ideally need different cut-off levels as they contain different rates of genetic variations in the 18S gene (Nebel *et al.*, 2011).The OTU concept should therefore only be considered as an estimate for species diversity, and might over or under estimate the diversity in certain linages. It is however important to notice that OTUs are not the same as a species.

To ensure that this study had sequenced deep enough a rarefaction curve was made, that showed a maximum number of OTUs have been reached in the samples. In figure 10 visual

inspection of the rarefaction curve shows that the DNA samples have reached a plateau (saturation), indicating that the sequencing depth has been sufficient to capture the species diversity present in the urine sample. When trying to uncover the micro-eukaryotic diversity of environmental samples one can expect huge diversity, as there are many groups that are not yet known and described (Logares *et al.*, 2014a). As for the urine samples it is far more difficult to know what OTUs to expect in the urine samples of *G. morhua* in the Oslofjord, as there have not been done any molecular studies on this field to my knowledge. Some studies have on the other hand exanimated the diversity of the urine found in cod in general, using techniques like microscopy, making it easier to know what species can be expected to encounter, and that is what this study make assumptions on. With the sequencing depth and length of Illumina Miseq used in this study, one can expect to uncover a rather large proportion of the diversity present in the urine samples. The challenge is to assign those recovered sequences to a meaningful taxonomy, and due to limitations of databases sequences of interest had to be put into phylogenetic content.

The choice of reference databases can have affect the outcome of taxonomy

Taxonomy assignment has been conducted against SILVA reference database. "No hits" sequences were removed after double-checking against the NCBI database to ensure that no species were overlooked. The results from SILVA taxonomy assignment showed that the Metazoa were the most dominant Kingdom in the data set in terms of number of reads (abundance) and OTUs (diversity), and on the contrary to what was expected, the number of reads/sequences originating from the host was very low, as most of the metazoan reads (and OTUs) belong to parasitic taxa Myx.ozoa

The second most abundant and diverse Kingdom was Fungi, followed by Alveolata, Stramenopiles, Streptophyta and Rhizaria. Figure 12 showed that 56% of the total OTUs were parasites. A remarkable 79% of the Metazoa OTUs were parasites, and 31% of the Fungi OTUs were parasites. Four of the OTUs from the taxa Fungi were marine Fungi, two belonging to the Phylum Basidiomycota and the two others to the phylum Ascomycota (see Appendix S6). Recent HTS studies have demonstrated that marine fungi are more abundant and taxonomically diverse than previously thought (Lepère *et al.*, 2015). Fungi are, according to recent papers by (Lepère *et al.*, 2015; Richards *et al.*, 2015), potentially parasitic associations with picophytoplankton, particularly with members of the *Prymnesiophyceae*. This emphasizes that one should not exclude the marine fungi as potential parasites in the ocean. As less than 5% of the diversity of fungal species is currently described and the

importance of marine fungi becomes more and more apparent, a further investigation of the diversity and specific roles of marine fungi is necessary to reach a greater understanding of the parasitic diversity of higher taxa (or taxonomic levels/groups) (Richards *et al.*, 2012; Lepère *et al.*, 2015).

One of the OTUs that were found was assigned to the enigmatic X-cell parasite, which is the first confirmed detection of this parasite in the urine of *G. morhua*. This group is known to infect a wide range of fish species from several linages of teleost fish, including Gadiformes (Freeman, 2009). Their phylogenetic placement is somewhat uncertain, but based on SSU rDNA, X-cells have been places as a monophyletic group within the Alveolates (Adl *et al.*, 2005), while it is represented wrongly in the SILVA as Rhizaria. Little is known about the life cycle and transmission route of X-cell, but it is hypothesized that they may need an intermediate host, to successfully infect a new fish, as experimental direct transmission from fish to a new infected fish failed (Freeman *et al.*, 2011a). I would therefore be interesting to see if X-cell has a potential release stage and uses urine to disperse the spores in to the environment of the host.

Another OTU was assigned to Alveolata, which SILVA reference database assigned as *Goussia janae*, a member of the Aicomplexa. Apicomplexans are obligate intracellular parasites that constitute a large phylum of protozoa that parasitize vertebrates, including fish amongst others (Gonzalez *et al.*, 2009).

Even-though these findings are highly interesting, unfortunately the time limited the process of further examining these groups of parasites, and therefore they will not be discussed further in this study.

The impact of the taxonomy assignment on myxozoan diversity

The Kingdom Metazoa comprised 28 OTUs with Myxozoa as the dominating Class with 22 OTUs. The remaining OTUs were the Craniata phylum comprising three OTUs, which were the OTUs from the host itself, and Myriapoda, Phoroniformea and Eleutherozoa which all were only represented by one OTU. The three last mentioned constituted a negligible part of the total sequences. They will therefore not be discussed in this thesis. Myxozoans are highly diverse, specialized endoparasites that constitute more than 2200 parasite species (Holzer *et al.*, 2010; Rocha *et al.*, 2011; Hartikainen *et al.*, 2014b). The species *Zschokella hildae* was amongst the Myxozoa that were expected to be found in this study, as it had in studies such as (Hemmingsen & MacKenzie, 2001; Køie *et al.*, 2007; Karlsbakk *et al.*, 2009; Heuch *et al.*,

2011; Kodádková *et al.*, 2014) proven to be found in the urine of *G. morhua* in Norwegian waters, even the Oslofjord. However, none of the OTUs were assigned to this species, in either NCBI or SILVA.The genus *Gadimyxa* constituted a staggering 100% of the 22 Myxozoan OTUs (Figure 13 a). The *Gadimyxa* genus is a newly described genus found in *Gadus morhua* and *Arctogadus glacialis* (Gadidae) (Køie *et al.*, 2007). Within the genus *Gadimyxa* there are three species previously found that develops in the urinary system of *G. morhua*; *Gadimyxa atlantica*, *Gadimyxa sphaerica* and *Gadimyxa arctica* (Køie *et al.*, 2007; Heuch *et al.*, 2011).

The assignment of these OTUs was first performed with both SILVA reference database but did not provide a very good diversity, as only two species were assigned to the OTUs, *Gadimyxa atlantica* and *Gadimyxa_sp_EK-2009* (Figure 13 b). Therefore the sequences were blasted manually against NCBI. The results from NCBI demonstrated that most of the OTUs gave significantly better taxonomically hits against other species, according to Bit scores, E-value, and identity values than the previous blast against SILVA database. To use a highly reduced reference database like SILVA that only comprises high quality sequences, will increase the risk not get significant taxonomical hits. When using reference databases such as NCBI, one might however find plenty of wrongly annotated sequences, thereby a validation of the taxonomic assignment of sequences has to be made.

In order to validate the taxonomic assignment of the sequences and putting them in to phylogenetic context, a phylogenetic tree was calculated. A phylogenetic tree from the research article (Kodádková *et al.*, 2014) with a supplement of sequences from Egil Karlsbakk that were not yet available on NCBI see (Appendix S7) was used as a reference, to give an impression of where in the phylogeny the *Gadimyxa* sequences would appear relative to the described sequences and to each other.

In the resulting phylogenetic tree, all 22 sequences from the urine OTUs comprise a monophyletic group with a posterior probability of 99.39. The Myxozoa OTU reference sequences group are divided into two groups, with a probability score at 85.99%, where one group of OTU-sequences are clustered around the reference sequence Parvicapsula-indet-Sek6, and the others cluster around the remaining reference sequences from the genus *Gadimyxa*. The fish parasitic Parvicapsula group cluster basal to both of the Myxozoa groups from the urine OTUS with a posterior probability of 96.82. According to (Køie *et al.*, 2007; Karlsbakk & Køie, 2009; Kodádková *et al.*, 2014; Okamura, 2015) *Gadimyxa* are within the

family parvicapsualidae. A support of this clade grouping is found in a study conducted by (Køie *et al.*, 2007), where they position the clade *Gadimyxa spp* . together with the core group *Parvicapsula spp* . As the topology was recovered, we feel like one cantrust the tree up to a certain degree. Each OTU in the phylogeny was numbered with species name the similarity percentage assigned from NCBI. The OTUs from the dataset cluster somewhat with their respective reference sequences based on similarity from NCBI. OTU 14, 52, 3 and 110 from the dataset have all a similarity score between 93-95% to the reference sequence Gadimyxa_sp_ EK- 2007 according to NCBI blasting. OTU 13 had a similarity score of 94% to the reference sequence *Gadimyxa_atlantica*. In the phylogenetic tree however, one can see that they cluster around a new reference sequence Parvicapsulidae-indet-Sek6 supplemented from Egil Karlsbakk. As this reference is not yet available on NCBI or SILVA it explains the low smiliarity scores obtained with blast and why they do not cluster to their closest hit. Never the less the tree topology confirms a relation between *Gadimyxa* and those sequences. This emphasizes once more the importance of broad taxonomic sampling in reference databases used for taxonomy assignment. In this clade, one can also see that OTU 14 and 52 form their own clade with a posterior probability at 90.8 %. With these supports new species, was found within the Parvicapsulidae. OTU 3 and 110 are grouped closely to the reference sequence Parvicapsulidae, they both seems very alike and are probably a variant of the reference sequence.

In the second Myxozoa grouping of urine OTUs below, the general outcome is that the OTUs are sub-groups according to the reference sequence. This means that they are probably variants of the reference species, and not new species. OTU 1 in the tree gives a wrong impression however. When looking at the posterior probability score (35), this position has very low support and it could be located anywhere in the tree. The same is true for OTUs 31 and 32 which group together with a probability score at 98.1. It could look like they are two new species that resemble each other, but as the branching support to the other *Gadimyxa* species has only a posterior probability score of 63.9, it could as well have been placed within the other *Gadimyxa* species. Nevertheless, these two OTUs resemble each other more than the surrounding *Gadimyxa* species. Most of the OTUs taxonomically assigned to *Gadimyxa arctica* and *G. atlantica*, are placed near their closest blast hit, but there are a few that do not, OTU 33, which reached 100% similarity score with *Gadimyxa atlantica* (EU163421) in NCBI, clusters more closely together with another OTU 45 from the dataset, or OTU 136 G. arctica that clusters more closely to the reference of sequence *Gadimyxa atlantica* than to the

arctica reference sequence. These OTUs are very much alike, and the reference alignments are also according to NCBI blast a 100% similar. As the posterior probabilities are low within all these groups containing the OTUs, their branching pattern within the *Gadimyxa atlantica / arctica* clade is somewhat uncertain, but they are all variants of the reference sequences, and resemble each other closely as can be seen by the very short branches.

An explanation of why reference sequences G. *arctica* and G. *atlantica* are set as two individual species even though they are 100% alike according to NCBI blasting, is that they morphologically do not resemble each other. This is supported by a research done by Køie et al (2007) who claims that in many genera of the Myxosporeans, the morphology and molecular data often disagree in classification of species. Although one can based on 18S molecular data alone, consider them to be cryptic species.

In the tree there are also two branches that are significantly longer than the others (OTU 123 and 129). The alignments of these OTUs were thoroughly inspected in Geneious and blasted one more time against NCBI to check for errors in the annotation. In OTU 123, it seemed like the first 50 nucleotides where not sequenced and therefore mis-aligned. This resulted in misleading tree branch length and probably a misleading placement within the tree. OTU 129 on the other hand did not have any error, so it is rather a result of a lot of genetic variation, which might indicate that it is another species of *Gadimyxa* without representatives in this tree. As the posterior probability is so low (nine), it might be another genus or even differ on a higher taxonomic level. This is however probably not a finding of a new species of *Gadimyxa*.

Within the myxozoan marine lineage, there are five clades: the marine *Myxidium* clade, the *Ceratomyxa* clade, the *Enteromyxum* clade, the *Kudoa* clade and the marine urinary clade, divided into the *Parvicapsula* and the *Zschokkella* subclade (Kodádková *et al.*, 2014). The reference tree follows a tissue tropism criteria rather than myxospore morphology and in this respect, the species found in this study do also phylogenetically relate closely to myxosporea of the genera *Parvicapsula* and *Gadimyxa* (Holzer *et al.*, 2004; Kodádková *et al.*, 2014). There are not much information found regarding the reference sequence Gadimyxa EK-2007 and 2009. They are however, according to molecular findings in this study and according to NCBI, variants of the genus *Gadimyxa*, and are found to infect the urine bladder and kidneys of gadoid fish like *G. morhua* and whiting (*Merlangius merlangus*).

The new genus *Gadimyxa* that contains the species *Gadimyxa atlantica* and *Gadimyxa arctica* have previously been observed to have two morphological spore forms (wide spores and (sub) spherical spores), and both are found in the urinary bladder as well as in the posterior and anterior kidneys (Køie *et al.*, 2007). The new species of *Gadimyxa* are most similar to the Myxozoa species *Ortholinea sp*, in form of myxospores but clearly differs from the genus in the 18S sequence information (Køie *et al.*, 2007). *Gadimyxa atlantica* and *Gadimyxa arctica* differ from one another morphologically in size and shape, and they also have different hosts according to Køie et al (2007). *Gadimyxa atlantica* infects the polychaetes *Spirorbis spp ., S. inornatus, S.tridentatus* and *S. corallinae* with the actinospore (tetractinomyxon). In the same publication, (Køie *et al.*, 2007) confirmed that the actionspore in *Spirorbis spp* . were conspecific to the myxospore found in *G. morhua*. For *Gadimyxa arctica* the invertebrate host is unknown. They both however are only recorded in vertebrate hosts of the family Gadidae. But contrary to the findings in this study, (Køie *et al.*, 2007), claims that *Gadimyxa arctica* is only found in the gadoid *A.glacialis*. However, they do also say that the differences between *Gadimyxa arctica* and *Gadimyxa atlantica* are only 16 substitutions and a 4 bp insert. This makes both species highly similar on a molecular level, thereby explaining the inconsistent positioning of the OTUs that were annotated as one of those two species. Due to the molecular similarities, information on spore morphology in the samples taken here could have been helpful, but unfortunately unrealistic duto the very limited sample size.

Cods that were heavily infected by these *Gadimyxa* species showed symptoms in the kidney tubules, that were dilated and tubular epithelium that were flattened. The experiment that were conducted by (Køie *et al.*, 2007) showed that the infection of *Gadimyxa atlantica* were limited, and that once the infection was gone, the cod did not get re-infected immediately. Even though *Gadimyxa* infections have to date not been associated with pathogenicity, the closely related genus *Parvicapsula* includes two species that are serious pathogens for both wild and farmed fish like salmonids (Heuch *et al.*, 2011). Myxosporeans in the genus *Parvicapsula* have been found to destruct glomeruli, tubules and pseudobranches that causes death in salmonid fish (Sterud *et al.*, 2003). In this study, one can see that two of the OTUs might be new species of *Parvicapsula*. It is noteworthy that they also display similarity to *Gadimyxa*, which make them interesting species to study further , as they potentially can have the same invertebrate and vertebrate hosts as *Gadimyxa*. In addition they might be able to cause as serious infections as *Parvicapsula*.

The occurrence of wide spore and the (sub) spherical spores of *Gadimyxa* in the urine bladder indicate that the urine might be used as excretion route in their transmission pathway. If both forms are infective to their invertebrate host, it is likely that thin walled wide spores have a limited life span in the water, whereas the (sub) spherical spores with thick walls might act as resistant spores, and may have a long life span in the sea. In general there are some myxospores that may live for more than 20 years in the outer environments, and have shown to resist freezing as well as the passage through the alimentary tract of animals. Furthermore, they seem to be highly adapted to environmental changes (Kent *et al.*, 2001). The research from (Køie *et al.*, 2007)also showed that the prevalence of the myxospores of *Gadimyxa atlantica* increased proportionally in *G. morhua* and the polychaete final host when they were exposed to each other. So one can predict that this could become a serious threat for fish that are living in close proximity to each other for example wild fish, fish farms and fish that are stationary in general, like the coastal cod (Heuch *et al.*, 2011).

Assignment of Microsporidian diversity

The Kingdom Fungi were the second largest group with 21 OTUs, whereof 11 of the OTUs were assigned to the phylum Microsporidia. Dikarya was the second most abundant with 8 OTUs, and the rest of the OTUs were divided between Fungi incertae sedis (1 OTU) and "uncultured" (1 OTU). The three last mentioned were not of further interest to this study, so they are not addressed for the rest of the study. Microsporidia is a well-known genus that parasitizes fish hosts, and at least 18 genera have been recorded in fish. There have also been studies that found Microsporidia in urine of *G. morhua*, but the species was unidentified (Heuch *et al.*, 2011). The 11 OTUs found in this study were taxonomically assigned to *Trachipleistophora hominis*. As *Trachipleistophora hominis* is usually found to infect human patients with AIDS (Cheney *et al.*, 2001), it was therefore again suspected that the SILVA database did not contain sufficiently broad information (reference sequences) for microsporidian species. *G. morhua* is, as mentioned, known from previous studies to be infected by microsporidian parasites. The sequences where therefore blasted manually against NCBI. Four OTUs were assigned to the *Loma sp.*, five OTUs to *Microsporidium cerebralis* , one OTU to *Kabatana sp*. and the last OTU was assigned to *Loma wallae* (Figure 16 a). According to Powell & Gamperl (2015), infection caused by the Microsporidium *Loma* can lead to a reduced body condition (liver somatic index), and a mild anaemia and leukaemia in fish like *G. morhua* and salmonids such as rainbow trout *Oncorhynchus mykiss*. *Kabatana sp*.

are known to infect bottom dwelling fish as two-spotted goby *Gobiusculus flavescens* and *Lophius piscatorius* and fresh water salmonids. Subcutaneous creamy-white patches in the body musculature are associated with *Kabatana sp.* infection (Barber *et al.*, 2009).

Validation of the taxonomic assignment of sequences was done by inserting them into a phylogenetic reference tree. A phylogenetic tree assembled from the research article (Vossbrinck & Debrunner-Vossbrinck, 2005b) with a supplement of sequences from Egil Karlsbakk was used as a reference, to obtain an impression of where in the phylogeny the Microsporidia OTUs would appear relative to the to the reference sequences and the other sample sequences.

With the exception of OTU 105 that grouped monophyletic with *Kabatana takedai* with a posterior probability score at 91.8%, the rest of the OTUs comprised one monophyletic group with *Glugea* basal to them with a posterior probability at 83 %. Within the monophyletic group, there was a division into two groups, One group consisted of OTUs that were assigned to *Microsporidia cerebralis* and the other group was composed of the OTUs assigned to *Loma sp*. and *Loma wallae*. The OTUs assigned to *Microsporidia cerebralis* form two sister clades to their reference sequence with a posterior probability score at 74.8%. The two clades split with a probability score at 97.4%, which shows strong evidence that these are two new species that not only differ from their reference sequences but also from each other. They do resemble one another more than their reference a. OTU 42 and 43 have a posterior probability score at 61.7 % meaning that they are quite divergent. For OTU 7, 100 and 50, there is a slightly different picture. Their probability scores for this branching pattern are 17.9 and 19.7%, which is very low support, so the order within this clade is unresolved. One can explain this finding by looking at the clustering and similarity scores that were assigned to the dataset. The clustering was relatively conservative (95 %/ g =5), which says that the species had to be under 95 % similar to form different OTUs.

The second group which consists of the *Loma* OTUs and their reference alignment, are considered to be one group, where the OTUs are just variants of the reference alignment. All OTUs (61, 62, 5, 22 and 23) are sister to the reference sequences *Loma_sp* and *Loma_ morhuae* cod-EK-LM1012. The probability score is quite low for all the OTUs that cluster with these reference sequences, meaning that they in actuality switch places with each other. This would not make much difference, as they are very much alike. OTU 23 *_Loma_ wallae_97* is in this tree not closely related to *Loma wallae* reference sequencesat all, *Loma*

wallae is basal to a completely different clade, described as clade 4 in the Vossbrinck et al paper.

All OTUs found in the sampling, were of the phylum *Microsporidia*, and are all known to infect internal organs and gill filaments of their vertebrate host. Research such as (Heuch, 2001; Heuch *et al.*, 2011) have found Microsporidia infecting the urine bladder of *G. morhua*, whilst inspecting the fish with microscopes, however these have not yet been identified in the sense of molecular and morphological studies. *Microsporidia cerebralis* itself that many OTUs in the sample sequenced were assigned to here, has never been described in the urine of *G. morhua*. *Microsporidia cerebralis* is only known to infect the brains of Atlantic salmon. Essential information on development and generic status are still missing, to this day, so it is difficult to say more about this species (Brocklebank *et al.*, 1995). Their findings showed that in addition to differences in hosts and geographic location, the spores from the brains of Atlantic salmon are distinctly larger than those of other microsporidian species infecting nervous tissues of fishes, and were found in massive numbers (>100) in the neuronal cells. The salmons that were infected showed abnormal swimming behavior, in sense that they swam slowly. There were no consistent external finding on the fish, but some showed hemorrhages in the web of the pectoral, pelvic, and caudal fins (Brocklebank *et al.*, 1995). The findings in my study suggest that the OTUs assigned to *Microsporidium cerebralis* were not the same species, but rather one or two new species that do resemble the reference sequences more that other reference sequence. These findings are quite interesting, as there has not been done much research about which host these microsporidian species infect. This study suggest that it is possible that new species resembling *Microsporidium cerebralis* infects hosts as *G. morhua*, and are found present in the urine bladder.

Among commercially valuable fish species, Microsporidia of the genus *Loma* are particularly problematic, with many identified host-parasite pairs. Affected species include the Atlantic haddock (host to *L. branchialis*) the pollock (*L. wallae)* salmonids of the genus *Oncorhynchus* (*L. salmonae*), and the Atlantic cod (*L. morhua)* (Baldwin *et al.*, 2000; Powell & Gamperl, 2015).

Microsporidians belonging to the genus *Loma* are intracellular, spore-forming parasites. The life-cycle of *Loma* is relatively unknown but it is thought that it begins with host ingesting the infective spores (Rodriguez-Tovar *et al.*, 2003). The spores infect epithelia cells and

intraepithelial cells (leucocytes) and are transported to the gills, where they form cyst-like nodules filled with xenomas. Xenomas occlude the blood vessels and cause inflammatory response and death. (Rodriguez-Tovar *et al.*, 2003; Brown *et al.*, 2010). To my knowledge there are no publications available today about findings of *Loma sp*. specifically in the urine bladder of *G. morhua*. Five of the OTUs in the phylogenetic tree were variants of genus *Loma*. They were assigned to and grouped with the reference alignments *Loma morhua* and *Loma sp*. *Loma morhua* is a well-known microsporidian in aqua culture of cod (hence *morhua*), where they cause substantial infections that causes mortality and reduced growth in both grown and juvenile cods (Frenette *et al.*, 2011). The host is exposed to the infections either by ingestion of spores present in the water column, when cohabiting with other exposed fish, or feeding on smaller infected fish or reservoir host as invertebrates (Frenette *et al.*, 2011). The parasite infects the host cell by extruding genetic material into their cytoplasma, and then proliferate, leading to cell death and thereby releasing infective spores into the environment (Powell & Gamperl, 2015). The spores are, similar to other species of the genus *Loma*, enormous and they form xenomas in the gills and epithelium (Hemmingsen & MacKenzie, 2001; Powell & Gamperl, 2015). They are however also found to infect organs as the heart, liver and spleen of the host *G. morhua* (Hemmingsen & MacKenzie, 2001). In this present study it also gives indications that these species (or variants of *Loma* species) are also present in the urine of *G. morhua*. As the spores are found in the blood path (heart and blood vessels), it is possible that through this vascular system, via the kidneys the spores end up in the urine of cods, as a way of excreting spores which is a possible transmission path.

The last microsporidian OTU that was found in the samples was assigned to *Kabatana sp,* and grouped monophyletically with *Kabatana takedai.* The good posterior probability support shows that the OTU was not the same species, but do however resemble the one reference sequence. As there were no other *Kabatana* reference sequences in the tree, it is impossible without further research to know if it is a new species or a variant of another *Kabatana* species. *Kabatana* are prone to infect several species of bottom dwelling fish, as mentioned above. They cause creamy-white patches in the body musculature and enlarged hearts have been observed in heavily infected hosts (Barber *et al.*, 2009). *Kabatana takedai* is known as an important and endemic pathogen of wild and cultured salmonid fish, where it infects the heart and trunk muscle of the fish (Fujiyama *et al.*, 2002). Studies have proven that *Kabatana takedai* form cysts when the environment (water) are below 13°C, which indicates that the Norwegian waters likely are suitable for this form of infection. It is presumed that *Kabatana*

takedai uses the rotifer *Euchlanis dilatata* as an intermediate organism in natural transmission (Fujiyama *et al.*, 2002). *Kabatana sp*. is a well-known parasite in Japan, but as seen in this study, it is also present in the Oslofjord, in the urine of *G. morhua*. A possible explanation can be that the parasite has been introduced to Norwegian water with the *Euchlanis dilatata* or as spores that can be present in the ballast water of container ships (Pagenkopp Lohan *et al.*, 2015). According to Artsdatabanken, there are finding of *Euchlanis dilatata* in Norwegian waters (Artsdatabanken, 2015). However, further studies like mine should be conducted as a risk evaluation of new parasite species, to clarify the transmission stage and infective period of the parasites.

The parasitic findings in this study are of ecological importance for the cods in Oslofjord, as well as other fish species. There are many parasitic species that are not described in forms of morphology, genetics and life cycles. This study presents un-described diversity of parasites that might be of concern for *G. morhua*, when parasites are present in high numbers.

In particular juvenile fish are susceptible to health problems and have higher chances of mortality than adults, when infected by the parasites mentioned above (Molnar, 2005). The juvenile fish are usually stationary in their first years, and because they also are more vulnerable to infections they might are more exposed to pathogens present in the environment, they also live close to each other, which might facilitate the risk of infections. The juvenile fish are also, if heavily infected with parasites, in higher risk of being eaten by predators, as they can become weaker of the infections or by the lack of food consumption caused by the infection.

Given the fact that we know that some of the genera *Gadimyxa, Loma sp* and *Microsporidium cerebralis* are found in several invertebrate hosts (rotifers in the case of *Kabatana*), and that they have a somewhat unknown life cycle, there might several unknown reservoirs of parasites out there which can be athreat to fish. The parasitic fauna in the urine bladder of *G. morhua* in the Oslofjord is a field that has not been investigated to my knowledge. Proven several times in this study, it contains parasitic species diversity that might be of concern, as the spores of these parasitic species can be excreted from the urine as a transmission pathway, and therefore be a way of spreading micro-eukaryotic parasites that can cause harm to fish.

As new techniques as used in this study help to find species that also have an unknown life cycle, there are indeed a lot more work that has to be done it will be necessary to predict plausible pathogen infections that might cause harm for important fish as *G. morhua*, and potential other vertebrate host, as well as wild populations.

Methodological aspects

Sampling

Every single step from fieldwork, to sequencing and data analysis, can create bias that would influence the composition of the final taxonomic composition of the research. In the present study, attempts have been made to minimize these biases as far as it is possible. There are though some biased steps that could have influenced the final outcome of the study, and some problems that happened during the study.

Samplings of the fish in the different locations was done by fishing rods and hooks with bait, under controlled conditions. This resulted in capturing of the wanted juvenile cods as seen in table one.The warm weather was a factor that had an impact on the sampled fish, as the water in the containers became warm very fast. When fishing with rods, a long time can pass between capturing of the fishes, and in connection with the fact that the laboratory facility was at least 45 minutes away with the boat, it resulted in a high mortality rate particular in the smallest cods, which might have been even more interesting for this study. The bigger fish on the other hand seemed to tolerate being in the containers under transportation and sampling better. A faster and stronger boat would have helped on this problem, as some days the transportation could take up to an hour, when the current in the Drøbak strait was strong. At the research facility, the cods were stored in tanks outside the laboratory, making it hard to survey the cods at all times. Several of the fishes that were stressed during transportation died before dissection, and could not be used in this study as the cods excreted their urine before death. The tanks also lacked a lid, so without surveillance, many of the fishes were taken from the tanks by predators such as seagulls and minks.

Biases during DNA-extraction and PCR

During extraction of DNA the main bias occurring is because various groups of organisms have different cell lysis efficiency, therefore the nucleic acids from some organisms will be harder to extract than others.

The main source of bias though is usually the amplification step (PCR) (Logares *et al.*, 2013). Detecting the wrong nucleotide at the single position (base call errors), can result artificial richness of diversity. Error can be ignored by removing low abundant OTUs (singletons). Mutation errors are also a well-known error that occurs in amplification, where nucleotides are incorporated in the wrong way. By using proofreading polymerase enzymes like Phusion polymerase used in this study, one can reduce this error (Zinger *et al.*, 2012). Bias correlated with polymerase enzymes, is the uneven amplification of the AT (adenine and thymine) and CG (cytosine and guanine) rich sequences. Some groups may occur more abundant because they are more easily amplified, therefore increasing the chances of them being sequenced later on (Polz & Cavanaugh, 1998). A problem that appears when studying amplicons of SSU rRNA, is the PCR-generated chimeras. They can comprise a big portion of unique sequences even though they don't constitute much of the total sequence number, and thereby inflate species richness observed in the sample. However, certain tools exist in order to identify and exclude such sequences like chimera checking within USEARCH. In the present study 3 % of the total reads were annotated as chimeras, and removed by the checking mentioned above.

Choice of primers

The choice of primers is also an artefact than can determine what kinds of organisms are being amplified. It is essential to find the right genetic marker for ones samples to get the organisms of interest for the study. The V4 region of the 18S ribosomal small subunit (SSU) was used in this study. V4 region is known to be the most appropriate region for capturing the diversity of micro-eukaryotes compared to other regions in the 18S (Dunthorn *et al.*, 2014; Hadziavdic *et al.*, 2014). Due to its highly variable region, the "barcode" marker SSU allows the use of universal primers in the amplification process. Although the primer used are called "universal primer", one can not expect that all species are found (Hartikainen *et al.*, 2014b). In order to get a better picture of the general diversity several primers should have been used. The same regarding the diversity of Myxozoa and Microsporidia, if specific primers were used, one could have expected a higher diversity than the results showed.

Pooling

One of the main aims was to get a general overview of the urine fauna, focusing on cods in the Oslofjord. It can be difficult to get sufficient amount of DNA in samples, as DNA tend to break down quite fast in harsh environments like urine (Juen & Traugott, 2005). The replicates were therefore pooled to have sufficient amount of DNA, and to avoid biases introduced during PCR amplification before Illumina sequencing. The triplicates from the

amplifications were equimolarly pooled in one tube. Subsequently, the pooled PCR products from multiple individuals were also pooled into a single DNA mixture, which was then prepared as a single library and sequenced on the Illumina MiSeq platform. Many studies have used this approach because pooling of samples reduces the cost of library preparation as only one library is prepared per pool instead of one library per sample. The pooling approach has also been proven successful to detect the rare variants in samples (Craig *et al.*, 2009; Rivas *et al.*, 2011; Orgiazzi *et al.*, 2012). Nevertheless, this limits the analyses of the resulting sequence data to be on a general overview level, as the information about parasite infection of individual fish is lost.

Data analysis

The bioinformatics part of the study may introduce a lot of biases, as there are several steps and many choices along the way before the end results. Each step (seen in the bioinformatics part) has several parameters one can choose that affect the outcome of clustering for example the choice of similarity threshold. It might be that 95% similarity threshold and g 5 as parameter was too strict in this study, and thereby resulting in an underestimating of the diversity in the urine sample. As mentioned Throughout the discussion part, the choice of reference database has a significant impact on the assignment of the species. If the database does not contain broad enough diversity to assign sequences to, the resulting taxonomy and phylogenetic tree will then be far from accurate. The outcome may look like there have been found a lot of new species as they don't resemble their reference alignment.

Illumina

Although the Illumina sequencing platform is known to have less sequencing errors, compared to other NGS, there are some issues that still are present (Loman *et al.*, 2012). The occurrences of sample leakage were of risk in this study, as my samples were sequenced with another master students' samples (Nelson *et al.*, 2014). This may give a wrong picture of diversity in the results. The diversity in of the samples should be inspected with other similar studies employing other methods compared carefully before drawing conclusions about ones findings.

Conclusion and further prospects

The present study is the first to investigate micro-eukaryotic parasite diversity in the urine bladder of *Gadus morhua* in the Oslofjord, using Illumina MiSeq sequencing. Surprisingly, a large concentration of parasites was found in the sample, 56 % of the detected OTUs were micro-eukaryotic parasites. Among these, over 50% of the OTUs were assigned to Myxozoa and Microsporidia. Within these groups, a vast majority were assigned to the genus *Gadimyxa*, *Loma* and *Microsporidium.* Potential new species within the genus *Micrsporidium* and *Gadimyxa* were discovered, demonstrating that there are probably more micro-eukaryotic parasites remaining to be discovered. The parasitic findings in the urine of *G. morhua*, may be a first step in the right direction of resolving several unfamiliar transmission pathways of parasitic species infecting fish, and thereby also get a better understanding of the life cycle of these parasites, in the sense of how the spores spread from one host to another.

By applying the Illumina MiSeq technology most of the species richness in the urine seemed to be successfully recovered.

A master thesis is very time limited it was unfortunately not possible to conduct all the ideas and on inspirations derived from the obtained data. However, here are some additional investigations that could be done, if such research should be repeated or continued.

As spores of parasites could be active and dormant, it would be useful to sequence both rDNA and reverse transcribed rRNA, as a combination of those two would distinguish between metabolic active and inactive cells. A morphological screening of the parasites' spores would also be important to add, as the morphology and molecular data often disagree in classification of species. To see if the fish actually contains parasites in the urine in general, or if parasites only are present in sick fish, it would be recommended to sequence each urine sample from the sampled fish separately. One could thereby also see if in a specific area fish sampled were more or less sick. By sequencing the urine samples separately, it would be easier to see if there was a correlation between visual parasitic infections and parasites in urine, and thereby see if parasites in the urine could be a generally good way for checking the health of the fish. A liver somatic index should also be measured to conduct the general health of the cods, as low liver somatic index usually is a symptom of poor health in fish.

To further confirm if urine samples are a good way to check the fishes for parasites inside the fish, it would be necessary to do a study were the whole fish both inside and outside is checked for parasites, and compare the finding to the urine diversity.

Even-though universal primers were used as they in general are able to detect most species, the general primers may have missed some species, as no primer set is truly universal (Hartikainen *et al.*, 2014a). Therefore the use of specific primers for the genus and species found in urine samples would be recommended.

Finally, DNA from both head and body kidneys of all sampled fish were extracted to support a possible transmission path way of parasites into the urinary bladder, but as the sequencing of these samples are both expensive and time consuming it was not possible to conduct it within the scope of this master thesis. It would be an excellent way to compare parasites from the head and body kidney to the urine samples though to see if the same parasites are present in these body part as well, with the rDNA and rRNA comparison as mentioned in the section above, and thereby having a stronger support that urine is a transmission pathway for parasites.

References:

Adl SM, Leander BS, Simpson AGB, Archibald JM, Anderson OR, Bass D, Bowser SS, Brugerolle G, Farmer MA, Karpov S, *et al.* 2007. Diversity, nomenclature, and taxonomy of protists. *Systematic biology* **56**: 684–689.

Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, *et al.* 2005. The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists. *The Journal of Eukaryotic Microbiology* **52**: 399–451.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. **1997**. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389–3402.

Andreou D, Arkush K, Guegan JF, Gozlan RE. **2012**. Introduced Pathogens and Native Freshwater Biodiversity: A Case Study of *Sphaerothecum destruens*. *PloS one* **7**: 1–6.

Artsdatabanken. **2015**. Artsdatabanken- Nasjonal kunnskapkilde for biologisk mangfolk.

Azam F, Fenchel T, Fiels J, Gray J, Meyer-Rei L, Thingstad F. **1983**. The Ecological Role of Water-column Microbes in the Sea. *Marine Ecology Progess Serie* **10**: 257–263.

Baldwin TJ, Vincent ER, Silflow RM, Stanek D. **2000**. *Myxobolus cerebralis* infection in rainbow trout (*Oncorhynchus mykiss*) and brown trout (*Salmo trutta*) exposed under natural stream conditions. *Journal of Veterinary Diagnostic Investigation* **12**: 312–321.

Barber I, Davies A, Ironside J, Forsgren E, Amundsen T. **2009**. First record of a *Kabatana sp*. microsporidium infecting fish in the Atlantic Ocean. *Diseases of Aquatic Organisms* **83**: 145–152.

Bellay S, Oliveira EF, Almeida-Neto M, Mello MAR, Takemoto RM, Luque JL, Drevon CA, Hadziavdic K, Lekang K, Lanzen A, *et al.* 2008. Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. *Journal of Invertebrate Pathology* **278**: 186–188.

Beveridge M, Charles A, Dieckmann U, Fock HO, Froese R, Keller M, Löwenberg U, Merino G, Möllmann C, Ommer RE, *et al.* 2013. The Future of the Fishes. *World Ocean Review* **2**: 145.

Birkbeck T, Feist S, Verner, Jeffreys D. **2011**. Francisella infections in fish and shellfish. *Journal of Fish Diseases* **34**: 173–187.

Bittner L, Gobet A, Audic S, Romac S, Egge ES, Santini S, Ogata H, Probert I, Edvardsen B, de

Vargas C. **2013**. Diversity patterns of uncultured Haptophytes unravelled by pyrosequencing in Naples Bay. *Molecular Ecology* **22**: 87–101.

Boenigk J, Ereshefsky M, Hoef-Emden K, Mallet J, Bass D. **2012**. Concepts in protistology: Species definitions and boundaries. *European Journal of Protistology* **48**: 96–102.

Bråte J, Logares R, Berney C, Ree DK, Klaveness D, Jakobsen KS, Shalchian-Tabrizi K. **2010a**. Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *The ISME journal* **4**: 1144–1153.

Bråte J, Logares R, Berney C, Ree DK, Klaveness D, Jakobsen KS, Shalchian-Tabrizi K. **2010b**. Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *The ISME journal* **4**: 1144–53.

Bricknell IR, Bron JE, Bowden TJ. **2006**. Diseases of gadoid fish in cultivation: A review. *ICES Journal of Marine Science* **63**: 253–266.

Brocklebank JR, Speare DJ, Kent ML. **1995**. Microsporidian encephalitis of farmed Atlantic salmon (*Salmo salar*) in British Columbia. *The Canadian Veterinary Journal* **36**: 631–633.

Brooker RJ, Widmaier EP, Graham LE., Stiling PD. **2008**. *Biology*. New York: McGraw-Hill.

Brooks F, Rindi F, Suto Y, Ohtani S, Green M. **2015**. The Trentepohliales (Ulvophyceae, Chlorophyta): An Unusual Algal Order and its Novel Plant Pathogen— *Cephaleuro*s. *Plant Disease* **99**: 740–753.

Brothers EB, Mathews CP, Lasker R. **1976**. Daily growth increments in otoliths from larval and adult fishes. *Fishery Bulletin* **74**: 1–8.

Brown AM V, Kent ML, Adamson ML. **2010**. Description of five new *Loma* (Microsporidia) species in pacific fishes with redesignation of the type species *Loma morhua* Morrison & Sprague, 1981, based on morphological and molecular species-boundaries tests. *J Eukaryot Microbiol* **57**: 529–553.

Du Buit MH. **1995**. Food and feeding of cod (Gadus morhua L.) in the Celtic Sea. *Fisheries Research* **22**: 227–241.

Burt M, MacKinnon B. **1997**. Parasites and marine aquaculture. *Aquaculture Association of Canada Special Publication Number 2*: 65–68.

Bylen E. **1994**. *ultrastructure and taxonomy of the microsporidia (protozoa: Microsporidia) with emphasis on the microsporidia infecting midges (Diptera: Chironomidae).* Lund: Department of

Zoology Division of Systematics Lund Univerity.

Canning E, Okamura B. **2003**. Biodiversity and Evolution of the Myxozoa. *Advances in Parasitology* **56**: 43–131.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich K, Gordon JI, *et al.* 2011. QIIME allows analysis of high troughtput community sequencing data Intensity normalization improves color calling in Solid sequencing. *Nature Publishing Group* **7**: 335–336.

Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A. **2012**. Marine Protistan Diversity. *Annual Review of Marine Science* **4**: 467–493.

Cavalier-Smith T, Lewis R, Chao EE, Oates B, Bass D. **2009**. Helkesimastix marina n. sp. (Cercozoa: Sainouroidea superfam. n.) a Gliding Zooflagellate of Novel Ultrastructure and Unusual Ciliary Behaviour. *Protist* **160**: 452–479.

Cheney SA, Lafranchi-Tristem NJ, Canning EU. **2001**. Serological differentiation of microsporidia with special reference to *Trachipleistophora hominis*. *Parasite-journal* **8**: 91–97.

Churko JM, Mantalas GL, Snyder MP, Wu JC. **2013**. Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. *Circulation Research* **112**: 1613–1623.

Craig JE, Hewitt AW, McMellon AE, Henders AK, Ma L, Wallace L, Sharma S, Burdon KP, Visscher PM, Montgomery GW, *et al.* 2009. Rapid inexpensive genome-wide association using pooled whole blood. *Genome research* **19**: 2075–2080.

Decelle J, Romac S, Sasaki E, Not F, Mahé F. **2014**. Intracellular Diversity of the V4 and V9 Regions of the 18S rRNA in Marine Protists (Radiolarians) Assessed by High-Throughput Sequencing. *PLoS ONE* **9**: e104297.

Didier ES, Becnel JJ, Kent ML, Sanders JL, Weiss LM. **2014**. Microsporidia. Systematics and Evolution.115–133.

Dunthorn M, Otto J, Berger SA, Stamatakis A, Mahe F, Romac S, de Vargas C, Audic S, Biomarks C, Stock A, *et al.* 2014. Placing Environmental Next-Generation Sequencing Amplicons from Microbial Eukaryotes into a Phylogenetic Context. *Molecular Biology and Evolution* **31**: 993–1009.

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

Egge E, Bittner L, Andersen T, Audic S, de Vargas C, Edvardsen B. **2013**. 454 Pyrosequencing To Describe Microbial Eukaryotic Community Composition, Diversity and Relative Abundance: a Test for Marine Haptophytes. *PloS one* **8**: 1–15.

Foox J, Siddall ME. **2015**. The Road To Cnidaria: History of Phylogeny of the Myxozoa. *Journal of Parasitology* **101**: 269–274.

Freeman M. **2009**. X-cell parasites in the European dab *Limanda limanda* are related to other X-cell organisms: a discussion on the potential identity of this new group of parasites. *Parasitology* **136**: 967–980.

Freeman M, Eydal M, Yoshimizu M, Watanabe K, Shinn A, Miura K, Ogawa K. **2011a**. Molecular identification and transmission studies of X-cell parasites from Atlantic cod *Gadus morhua* (Gadiformes: Gadidae) and the northern black flounder *Pseudopleuronectes obscurus* (Pleuronectiformes: Pleuronectidae). *Parasites & vectors* **4**: 1–3.

Freeman MA, Kristmundsson Á. **2015**. Histozoic myxosporeans infecting the stomach wall of elopiform fishes represent a novel lineage, the Gastromyxidae. *Parasites & Vectors* **8**: 1–13.

Freeman M, Yokoyama H, Osada A, Yoshida T, Yamanobe A, Ogawa K. **2011b**. *Spraguea* (Microsporida: Spraguidae) infections in the nervous system of the Japanese anglerfish, *Lophius litulon* (Jordan), with comments on transmission routes and host pathology. *Journal of fish diseases* **34**: 445–452.

Frenette A, O'Neil M, Bryne H, Benfey N, Trippel E, Benfey T, Duffy M, Neil S, O'Keefe R. **2011**. Research to mitigate the negative impacts of *Loma morhua* on cod farming. *Aquaculture Association of Canada Special Publication 109-1* **1**: 1–32.

Fujiyama I, Urawa S, Yokoyama H, Ogawa K. **2002**. Investigation of the transmission stage of the microsporidian *Kabatana takedai* in salmonids. *Bulletin of the National Salmon* **5**: 1–6.

Gjøsæter J, Danielssen SD. **2004**. *Fiskeressurser i Ytre Oslofjord*.

Glover KA, Sørvik Eide AG, Karlsbakk E, Zhang Z, Skaala Ø. **2013**. Molecular genetic analysis of stomach contents reveals wild Atlantic cod feeding on piscine reovirus (PRV) infected Atlantic salmon originating from a commercial fish farm. *PloS one* **8**: 1–8.

Gonzalez V, Combe A, David V, Malmquist NA, Delorme V, Leroy C, Blazquez S, Ménard R, Tardieux I. **2009**. Host Cell Entry by Apicomplexa Parasites Requires Actin Polymerization in the Host Cell. *Cell Host & Microbe* **5**: 259–272.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C,

Decelle J, *et al.* 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research* **41**: 597–604.

Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C. **2014**. Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. *PLoS ONE* **9**: 186–188.

Haeckel E. **1866**. Generelle Morhologie der organismen, Zweiter band.

Hartikainen H, Ashford OS, Berney C, Okamura B, Feist SW, Baker-Austin C, Stentiford GD, Bass D. **2014a**. Lineage-specific molecular probing reveals novel diversity and ecological partitioning of haplosporidians. *The ISME journal* **8**: 177–186.

Hartikainen H, Gruhl A, Okamura B. **2014b**. Diversification and repeated morphological transitions in endoparasitic cnidarians (Myxozoa: Malacosporea). *Molecular Phylogenetics and Evolution* **76**: 261–269.

Håstein T. **2011**. Fiskesykdommer. *Store Norske Leksikon*.

Hedrick RP, Baxa D V, De Kinkelin P, Okamura B. **2004**. Malacosporean-like spores in urine of rainbow trout react with antibody and DNA probes to *Tetracapsuloides bryosalmonae*. *Parasitology research* **92**: 81–8.

Hemmingsen W, MacKenzie K. **2001**. The parasite fauna of the atlantic cod, *Gadus morhua* L. *Advances in Marine Biology* **40**: 1–80.

Hendriks L, Goris A, Neefs J, Y VDP, Hennebert G, De Wacter R. **1998**. The Nucleotide Sequence of the Small Ribosomal Subunit RNA of the Yeast *Candida albicans* and the Evolutionary Position of the Fungi among the Eukaryotes. *Systematic and applied Microbiology* **12**: 223–229.

Hervio M, Khattra J, Devlin H, Kent M, Sakanari J, Yokoyama H. **1997**. Taxonomy of *Kudoa* species (Myxosporea), using a small-subuint ribosomal DNA sequence. *Canadian Journal of Zoology* **75**: 2112–2119.

Heuch PA. **2001**. Parasitter på vill og oppdrettet torsk - en sammenligning. *Veterinærinstituttets rapportserie* **12**: 1–21.

Heuch PA, Jansen P, Hansen H, Sterud E, MacKenzie K, Haugen P, Hemmingsen W. **2011**. Parasite faunas of farmed cod and adjacent wild cod populations in Norway: a comparison. *Aquaculture Environment Interactions* **2**: 1–13.

Holland JW, Okamura B, Hartikainen H, Secombes CJ. **2011**. A novel minicollagen gene links cnidarians and myxozoans. *Proceedings of the Royal Society Biological Sciences* **278**: 546–553.

Holzer AS, Sommerville C, Wootten R. **2004**. Molecular relationships and phylogeny in a community of myxosporeans and actinosporeans based on their 18S rDNA sequences. *International Journal for Parasitology* **34**: 1099–1111.

Holzer AS, Wootten R, Sommerville C. **2010**. *Zschokkella hildae* Auerbach, 1910: phylogenetic position, morphology, and location in cultured Atlantic cod. *Parasitology international* **59**: 133–140.

Hopla CE, Durden LA, Keirans JE. **1994**. Ectoparasites and classification. *Revue scientifique et technique (International Office of Epizootics)* **13**: 985–1017.

Hylis M, Weiser J, Oborník M, Vávra J. **2005**. DNA isolation from museum and type collection slides of microsporidia. *Journal of invertebrate pathology* **88**: 257–60.

Janeway CA, Travers P, Walport M, Mark S. **2001**. *Immunobiology, 5th edition : The major histocompability complex and its functions*. New York: Garland Science.

Jimenez-Guri E, Phillippe H, Okamura B, Holland, W, H P. **2007**. *Buddenbrockia* Is a Cnidarian Worm. **317**: 116–118.

Johansen T, Berg E, Dahle G. **2009**. Kysttorsk–et tema med mange variasjoner. *Fisken og havet, særnummer 2-2009* **Kyst og ha**: 1–4.

Jonsson B. **2006**. *Fisker*. Latvia: Cappellen.

Jørstad KE, Fjalestad KT, Àgùstsson T, Marteinsdottir G. **2003**. *Atlantic cod-Gadus morhua*.

Juen A, Traugott M. **2005**. Detecting predation and scavenging by DNA gut-content analysis: a case study using a soil insect predator-prey system. *Oecologia*: 344–352.

Kallert DM. **2006**. *Life cycle studies and transmission mechanisms of myxozoan parasites*.

Kallert DM, El-Matbouli M, Haas W. **2005**. Polar filament discharge of *Myxobolus cerebralis* actinospores is triggered by combined non-specific mechanical and chemical cues. *Parasitology* **131**: 609–616.

Karlsbakk E, Einen Bardsgjaere A-C, Bartosova P. **2013**. Sphaeromyxa artedielli sp. n. (myxozoa: sphaeromyxidae), a parasite of sculpins (cottidae) in northern Norway. *Folia Parasitologica* **60**: 425–432.

Karlsbakk E, Isaksen TE, Hamre LA. **2009**. Hva vet vi om parasitter og oppdrett av torsk? *kyst og*

havbruk: 1–4.

Karlsbakk E, Køie M. **2009**. *Bipteria formosa* (Kovaleva et Gaevskaya, 1979) comb. n. (Myxozoa: Myxosporea) in whiting *Merlangius merlangus* (Teleostei: Gadidae) from Denmark. *Folia parasitologica* **56**: 86–90.

Karlsbakk E, Køie M. **2012**. The marine myxosporean *Sigmomyxa sphaerica* (Thélohan, 1895) gen. n., comb. n. (syn. *Myxidium sphaericum*) from garfish (*Belone belone* (L.)) uses the polychaete *Nereis pelagica* L. as invertebrate host. *Parasitology research* **110**: 211–218.

Karlsen HE. **2012**. *Anatomi og funksjon bruskfisk og beinfisk*. Oslo.

Katoh K, Misawa K, Kuma K, Miyata T. **2002**. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research* **30**: 3059–3066.

Kaur H, Singh R. **2010**. A new myxosporean species *Mysxobolus sclerii* p. nov. and one known species *M. stomum* Ali et al. 2003 from two Indian major carp fishes. *Journal of Parasitic Diseases* **34**: 33–39.

Keeling PJ. **1998**. A kingdom's progress: Archezoa and the origin of eukaryotes. *BioEssays* **20**: 87–95.

Kent MI, Andree KB, Bartholomew JI, El-Matbouli M, Desser SS, Devlin RH, Feist SW, Hedrick RP, Hoffmann RW, Khattra J, *et al.* 2001. Recent Advances in Our Knowledge of the Myxozoa. *The Journal of Eukaryotic Microbiology* **48**: 395–413.

Kodádková A, Dyková I, Tyml T, Ditrich O, Fiala I. **2014**. Myxozoa in high Arctic: Survey on the central part of Svalbard archipelago. *International Journal for Parasitology: Parasites and Wildlife* **3**: 41–56.

Køie M. **1984**. Digenetic trematodes from *Gadus morhu*a L. (Osteichthyes, Gadidae) from Danish and adjacent waters, with special reference to their life-histories. *Ophelia* **23**: 195–222.

Køie M, Karlsbakk E, Nylund A. **2007**. (Myxozoa, Parvicapsulidae) parasitic in marine fish (Gadidae) and the two-host life cycle of *Gadimyxa atlantica* s.p. **93**: 1459–1467.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. **2013**. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology* **79**: 5112–5120.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. **2010**. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Enviromental microbiology*

12: 118–123.

Lennon J, Jones S. **2011**. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature reviews. Microbiology*: 119–130.

Lepère C, Domaizon I, Taïb N, Mangot J-F, Bronner G, Boucher D, Debroas D. **2013**. Geographic distance and ecosystem size determine the distribution of smallest protists in lacustrine ecosystems. *FEMS Microbiology Ecology* **85**: 85–94.

Lepère C, Ostrowski M, Hartmann M, Zubkov M V., Scanlan DJ. **2015**. In situ associations between marine photosynthetic picoeukaryotes and potential parasites - a role for fungi? *Environmental Microbiology Reports*: 1–7.

Liu H, Probert I, Uitz J, Claustre H, Aris-Brosou S, Frada M, Not F, de Vargas C. **2009**. Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *PNAS*: 12803–12808.

Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie J-M, Decelle J, Dolan JR, Dunthorn M, *et al.* 2014a. Patterns of rare and abundant marine microbial eukaryotes. *Current biology : CB* **24**: 813–821.

Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie J-M, Decelle J, Dolan JR, Dunthorn M, *et al.* 2014b. Patterns of rare and abundant marine microbial eukaryotes. *Current biology : CB* **24**: 813–21.

Logares R, Sunagawa S, Salazar G, Cornejo-Castillo FM, Ferrera I, Sarmento H, Hingamp P, Ogata H, de Vargas C, Lima-Mendez G, *et al.* 2013. Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environmental Microbiology* **16**: 2659–2671.

Lom J, Nilsen F. **2003**. Fish microsporidia: Fine structural diversity and phylogeny. *International Journal for Parasitology* **33**: 107–127.

Loman NJ, Misra R V, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. **2012**. Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology* **30**: 434–439.

Malmstrøm M, Jentoft S, Gregers TF, Jakobsen KS. **2013**. Unraveling the Evolution of the Atlantic Cod's (*Gadus morhua* L.) Alternative Immune Strategy. *PLoS ONE* **8**: 1–10.

van der Meeren T. **2015**. Torskens livshistorie. *havforskningsinstituttet*: 1–4.

de Meeus T, Renaud F. **2002**. Parasites within the new phyologeny of eukaryotes. *Trends Parasitol*: 247–251.

Miller K, Kaukinen K, Schulze A. **2002**. Expansion and contraction of major histocompability complex genes: a teleostean example. *Immunogenetics*: 941–963.

Miwa S, Nakayasu C, Kamaishi T, Yoshiura Y. **2004**. X-cells in fish pseudotumors are parasitic protozoans. *Diseases of aquatic organisms* **58**: 165–170.

Molnar K. **2005**. Protistan parasites and Myxozoa. Marine parasitology.26–46.

Morgan MJ. **2004**. The relationship between fish condition and the probability of being mature in American plaice (*Hippoglossoides platessoides*). *ICES Journal of Marine Science* **61**: 64–70.

Morrison CM, Sprague V. **1981**. *Electron microscopical study of a new genus and new species of microsporidia in the gills of Atlantic cod Gadus morhua L*. Maryland: journal of fish diseases.

Nanton D, Lall S, McNiven M. **2001**. Effects of dietary lipid level on liver and muscle lipid deposition in juvenile haddock, Melanogrammus aeglefinus L. *Aquaculture Research* **32**: 225–234.

Nebel M, Pfabel C, Stock A, Dunthorn M, Stoeck T. **2011**. Delimiting operational taxonomic units for assessing ciliate enviromental diversity using small-subunit rRNA gene sequences. *Enviromental microbiology reports* **3**: 154–158.

Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J. **2014**. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. *PloS one* **9**: e94249.

Nilsen F. **2000**. Small subunit ribosomal DNA phylogeny of microsporidia with particular reference to genera that infect fish. *The Journal of parasitology* **86**: 128–33.

Nilsen F. **2002**. *Parasitter hos torsk - hvilke problemer kan man møte i et fremtidig torskeoppdrett ?*

Ohtsuka S, Toshinobu S, Horiguchi T, Not F. **2015**. *Marine Protist Diversity and Dynamics*. Tokyo: Springer Japan.

Okamura B. **2015**. *Myxozoan Evolution, Ecology and Development*.

Olsen EM, Knutsen H, Gjøsæter J, Jorde PE, Knutsen JA, Stenseth NC. **2004**. Life-history variation among local populations of Atlantic cod from the Norwegian Skagerrak coast. *Journal of Fish Biology* **64**: 1725–1730.

Orgiazzi A, Lumini E, Nilsson RH, Girlanda M, Vizzini A, Bonfante P, Bianciotto V. **2012**. Unravelling Soil Fungal Communities from Different Mediterranean Land-Use Backgrounds. *PLoS*

ONE **7**: e34847.

Otterbech F. **1953**. The Cod Population of the Oslofjord. *Conseil International pour l'Exploration de la Mer* **29**: 1–10.

Pagenkopp Lohan KM, Fleischer RC, Carney KJ, Holzer KK, Ruiz GM. **2015**. Amplicon-Based Pyrosequencing Reveals High Diversity of Protistan Parasites in Ships' Ballast Water: Implications for Biogeography and Infectious Diseases. *Microbial Ecology*.

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, *et al.* 2012. CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. *PLoS Biology* **10**: e1001419.

Pernice MC, Logares R, Guillou L, Massana R. **2013**. General Patterns of Diversity in Major Marine Microeukaryote Lineages. *PLoS ONE* **8**: e57170.

Persson A, Stet R, Pilstrom L. **1999**. Characterization of MCH class I and Beta(2). microglobulin sequences in Atlantic cod reveals an unusually high number of expressed class I genes. *Immunogenetics*: 49–59.

Pethon P. **1998**. *Saltvannsfisker*. *Aschehoughs store fiskebok: Norges fisker i farger*. Oslo: Aschehough.

Polz MF, Cavanaugh CM. **1998**. Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology* **64**: 3724–3730.

Powell MD, Gamperl AK. **2015**. Effects of *Loma morhua* (Microsporidia) infection on the cardiorespiratory performance of Atlantic cod *Gadus morhua* (L). *Journal of Fish Diseases*: n/a–n/a.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. **2013**. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**: 590–596.

Richards TA, Jones MDM, Leonard G, Bass D. **2012**. Marine Fungi: Their Ecology and Molecular Diversity. *Annual Review of Marine Science* **4**: 495–522.

Richards TA, Leonard G, Mahe F, del Campo J, Romac S, Jones MDM, Maguire F, Dunthorn M, De Vargas C, Massana R, *et al.* 2015. Molecular diversity and distribution of marine fungi across 130 European environmental samples. *Proceedings of the Royal Society B: Biological Sciences* **282**: 1–10.

Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, Boucher G, Ripke S,

Ellinghaus D, Burtt N, *et al.* 2011. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet* **43**: 1066–1073.

Rocha S, Casal G, Matos P, Matos E, Dkhil M. **2011**. Description of Triangulamyxa psittaca sp. nov. (Myxozoa: Myxosporea), a New Parasite in the Urinary Bladder of Colomesus psittacus (Teleostei) from the Amazon River, with Emphasis on the Ultrastructure of Plasmodial Stages. *Protozoologica*: 327–338.

Rodriguez-Tovar LE, Wadowska DW, Wright GM, Groman DB, Speare DJ, Whelan DS. **2003**. Ultrastructural evidence of autoinfection in the gills of Atlantic cod *Gadus morhua* infected with *Loma sp* (phylum Microsporidia). *Diseases of Aquatic Organisms* **57**: 227–230.

Samuelsen OB, Nerland AH, Jørgensen T, Schrøder MB, Svåsand T, Bergh Ø. **2006**. Viral and bacterial diseases of Atlantic cod *Gadus morhua*, their prophylaxis and treatment: a review. *Diseases of Aquatic Organisms* **71**: 239–254.

Schlegel M, Hülsmann N. **2007**. Protists - A textbook example for a paraphyletic taxon. *Organisms Diversity and Evolution* **7**: 166–172.

Schmahl G, Toukhy A El, Ghaffar FA. **1990**. Transmission electron microscopic studies on the effects of toltrazuril on Glugea anomMa, Moniez, 1887 (Microsporidia) infecting the three-spined stickleback Gasterosteus aculeatus. *Parasitology*.

Schmieder R, Edwards R. **2011**. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**: 863–864.

Ski S. **2012**. Demografi, Populasjonsstørrelse og Vandringer hos Atlantisk torsk i indre Oslofjord - en Merke-Gjenfangststudie.

Skovgaard A. **2014**. Dirty tricks in the plankton: Diversity and role of marine parasitic protists. *Acta Protozoologica* **53**: 51–62.

Slamovits C. **2004**. Transfer of *Nosema locustae* (Microsporidia) to *Antonospora locustae* n. comb. Based on Molecular and Ultrastructural Data1. *Journal of Eukaryotic …* **51**: 207–213.

Sleigh MA. **1992**. *Protozoa and other protists*. New York: Cambridge University Press.

Smothers J, von Dohlen C, Smith L, Spall R. **1994**. Molecular evidence that the myxozoa protist are metazoa. *Science*: 19–21.

Sokolova YY, Paskerova GG, Rotari YM, Nassonova ES, Smirnov a V. **2014**. Description of Metchnikovella spiralis sp. n. (Microsporidia: Metchnikovellidae), with notes on the ultrastructure of

metchnikovellids. *Parasitology* **141**: 1108–22.

Southern TR, Jolly CE, Russell Hayman J. **2006**. Augmentation of microsporidia adherence and host cell infection by divalent cations. *FEMS microbiology letters* **260**: 143–149.

Star B, Haverkamp THA, Jentoft S, Jakobsen KS. **2013**. Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *Bmc Microbiology* **13**: 248.

Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, Gregers TF, Rounge TB, Paulsen J, Solbakken MH, Sharma A, *et al.* 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**: 207–210.

Stechmann A, Cavalier-Smith T. **2002**. Rooting the eukaryote tree by using a derived gene fusion. *Science*: 89–91.

Stentiford GD, Feist SW, Stone DM, Peeler EJ, Bass D. **2014**. Policy, phylogeny, and the parasite. *Trends in Parasitology* **30**: 274–281.

Sterud E, Simolin P, Kvellestad A. **2003**. Infection by *Parvicapsula sp.* (Myxozoa) is associated with mortality in sea-caged Atlantic salmon *Salmo salar* in northern Norway. *Diseases of aquatic organisms* **54**: 259–63.

Teshima K, Tomonaga S, Yamamoto K. **1986**. Structure of the Urinary Bladder in the Pacific Cod.

Timi JT, Mackenzie K. **2015**. Parasites in fisheries and mariculture. *Parasitology* **142**: 1–4.

Vávra J, Lukeš J. **2013**. Microsporidia and 'the art of living together'. *Advances in parasitology* **82**: 253–319.

Vivares G, Gouy M, Thomarat F, Metenier G. **2002**. Functional and evolutionay analysis of a eukaryotic parasitic genome. *Curr Opin Microbiology*: 499–505.

Vøllestad A. **2009**. Torsk. *Store Norske Leksikon*.

Vossbrinck CR, Debrunner-Vossbrinck B a. **2005a**. Molecular phylogeny of the Microsporidia: ecological, ultrastructural and taxonomic considerations. *Folia parasitologica* **52**: 131–42; discussion 130.

Vossbrinck CR, Debrunner-Vossbrinck BA. **2005b**. Molecular phylogeny of the Microsporidia: ecological, ultrastructural and taxonomic considerations. *Folia parasitologica* **52**: 131–42.

Wilmot C. **2005**. Animal diversity.

Windsor, D A. **1998**. Controversies in parasitology. Most of the species on Eart are parasites. *International Journal for Parasitology*: 1939–1941.

Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ. **2015**. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**: 1–10.

WWF-Norway. **2008**. *Management and Technical Measures in the Norwegian Cod and Groundfish Fisheries*.

Zhang J, Kobert K, Flouri T, Stamatakis A. **2014**. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**: 614–620.

Zinger L, Gobet A, Pommier T. **2012**. Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology* **21**: 1878–1896.

sampling	date	area	blood	pellet	Labelnum	fish nr	sampl nr	age	length cm	weight kg	sex	DNA for seqencing	W spots eye	black spot	copepodes	tumors	other
Drøbak	5.7.13	B.H	x		1	4	3	2	31	0,3	female	x		yes	yes		
	5.7.13	H.P	x		2	5	2	1	20	0,08	infant	x			1		
	5.7.13	H.P	x		3	6	2	1	18	0,018	female	0					
	5.7.13	H.P		x	37	7	2	2	40	0,53	male	x			1		
	5.7.13	B.H		x	38	8	3	2	30	0,28	female	x					
	5.7.13	B.H	x		4	9	3	1	24	0,14	female	x					
Drøbak	6.7.13	B.H		x	39	10	4	1	22	0,06	female	x		yes	3		
	6.7.13	B.H	x		5	11	4	3	29	0,22	male	x		yes	9		
	6.7.13	D.S		x	40	12	5	2	39	0,51	female	x	yes	yes	7		
Drøbak	7.7.13	B.H	x		6	13	6	1	28	0,23	female	x		yes	4		
	7.7.13	B.H		x	41	14	6	2	34	0,33	female	x		yes			
	7.7.13	B.H	x		7	15	7	2	26	0,06	female	x		yes	4		
	7.7.13	B.H		x	42	16	7	2	31	0,31	male	x		yes			
	7.7.13	B.H		x	43	17	7	2	23	0,11	female	x				yes	picture
	7.7.13	B.H	x		8	18	7	1	22	0,09	male	x			1		
Drøbak	8.7.13	H.P	x		9	19	8	1	20	0,06	female	x		yes			
	8.7.13	H.P	x		10	20	8	1	17	0,05	male	x		yes			
	9.7.13	D.S		x	44	21	10	2	40	0,67	female	x		yes			pale,lots of dots
Drøbak	9.7.13	D.S		x	45	22	10	2	38	0,57	female	x		yes	2		
	9.7.13	D.S	x		11	23	10	1	24	0,15	male	x					
	9.7.13	D.S	x		12	24	10	2	40	0,64	female	x			4		
	9.7.13	D.S	x		13	25	10	1	26	0,17	female	x		yes			
	9.7.13	D.S		x	46	26	10	2	28	0,24	male	x					
	9.7.13	B.H			x	27	9	2	44	0,82	male	x		yes		yes	picture
	29.7.13	B.H	x		14	28	11	2	28	0,24	male	x			1		
	29.7.13	B.H	x		15	29	11	2	34	0,3	female	x		yes	1		white spots on the stomach
	30.7.13	D.S	x		16	30	12	1	20	0,07	female	x		yes			
	30.7.13	D.S		x	47	31	12	3	40	0,69	male	x	yes	yes	11		
	30.7.13	D.S	x		17	32	12	2	30	0,31	male	x		yes	8		
	30.7.13	D.S		x	48	33	12	2	38	0,47	male	x					
	31.7.13	B.H	x		18	34	13	4	34	0,36	female	x		yes	1		
	31.7.13	B.H	x		19	35	13	2	31	0,32	female	x		yes			
	31.7.13	B.H		x	49	36	13	2	33	0,32	female	x		yes			
	31.7.13	D.S	x		20	37	14	2	32	0,27	male	x					
	31.7.13	D.S	x		21	38	14	2	32	0,31	male	x					tail looks eaten
	31.7.13	D.S	x		22	39	14	3	40	0,67	female	x	yes	yes		mabye	white spots in cartilage in head+ eyes
	1.8.13	K.k	x		23	40	15	2	30	0,24	female	x			1		
	1.8.13	K.k	x		24	41	15	2	28	0,2	female	x			2		
	1.8.13	B.H		x	50	42	16	2	29	0,28	female	x					
	1.8.13	B.H	x		25	43	16	1	28	0,15	male	x					
	19.8.13	B.H	x		26	44	17	1	18	0,11	male	x					
	19.8.13	B.H	x		27	45	17	1	28	0,15	male	x		yes	1		
	20.8.13	D.S	x		28	46	18	2	32	0,35	male	x		yes	2		Flesh anchor, smelled
	20.8.13	D.S	x		29	47	18	2	34	0,35	female	x			2		
	20.8.13	D.S	x		30	48	18	1	29	0,14	male	x					lots of parasites in stomach
	20.8.13	H.P	x		31	49	19	2	35	0,3	female	x		yes			
	20.8.13	H.P	x		32	50	19	1	11	0,09	male	x		yes			
	21.8.13	D.S	x		33	51	20	2	32	0,3	female	x		yes			pale,white spots on tail, fungus?
	21.8.13	D.S	x		34	52	20	2	33	0,38	male	x					little urine
	21.8.13	D.S	x		35	53	20	2	40	0,63	female	x		yes	2		
	21.8.13	D.S	x		36	54	20	2	34	0,3	female	x					

Table S1: Table showing sampling areas, weight, length , age, sex, visual parasites and how the sampled was treated before DNA-extraction.

Table 3: table showing DNA extraction and quality values after extraction measured by Nanodrop. 260/280 scores are highly variable, indicating a problem with inhibitory substances. Values to the far right are quantity of DNA of DNA used in the PCR reaction for each fish.

			Nanodrop of DNA extraction		
tube lables	fish number	microliter Urine	ng/uL	260/280	260/230
1	4	10	-3,6	1,38	0,69
2	5	5	-0,4	-0,74	4,63
3	6	10	5,1	1,39	1,01
4	9	10	9,7	1,24	1,18
5	11	10	7,7	1,99	2,49
6	13	10	13,6	1,26	0,83
7	15	5	47,4	1,38	0,79
8	18	8	0,1	0,39	-0,03
9	19	10	0,8	0,77	-0,18
10	20	15	2,1	0,88	-3,57
11	23	10	2,6	1,52	1,5
12	24	10	4	2	1,65
13	25	10	2,5	1,07	3,66
14	28	10	0,1	-0,27	-0,05
15	29	10	0	-0,15	0
16	30	5	1,2	1,36	-0,64
17	32	10	-0,1	0,12	0,02
18	34	10	-5,4	1,28	0,44
19	35	10	1,4	1,43	-0,62
20	37	10	0,8	1,13	-0,27
21	38	10	3,8	1,67	2,78
22	39	10	1,7	1,7	-3,04
23	40	10	0,6	-1,35	-0,23
24	41	10	1	2,23	-1,01
25	43	10	0,9	2,49	-2,6
26	44	10	1,3	-5,44	0,46
27	45	10	2,1	1,21	1,27
28	46	5	1,7	3,16	0,75
29	47	10	3	2,1	
30	48	10	2,4	1,98	1,44
31	49	10	0,7	0,6	-0,36
32	50	10	3,1	1,2	0,28
33	51	10	1,5	0,96	1,27
34	52	6	2,8	1,36	1,39
35	53	10	1,7	1,26	4,89
36	54	10	1,2	1,63	-3,24
37	7	600	2	1,14	-0,28
38	8	400	5,3	1,34	1,12
39	10	200	3,1	1,41	1,08
40	12	600	17,1	1,56	0,66
41	14	300	17,7	1,44	0,17
42	16	600	1,2	1,81	-0,37
43	17	400	5,7	1,84	0,71
44	21	600	10,5	1,36	0,8
45	22	250	11,4	1,53	1,02
46	26	400	5	1,73	6,35
47	31	500	1,3	1,05	-0,12
48	33	400	3,7	1,88	2,79
49	36	300	3,2	1,35	1,12
50	42	250	3,2	2,93	0,48

Table 4: Pipetting scheme developed to create an equimolar sample for Illumina sequencing.

Total final volume was 120 µL.

Samples	Qf	Concentration (ng/µl)	Vol. needed for 40 ng		optimal quantity for Miseq	2000	ng
1	0,506	20,24	2,01661692		Nb of sample	49	
2	0,239	9,56	4,26949022		Quantity per samples	40,81632653	ng
4	0,529	21,16	1,92893793				
5	0,485	19,4	2,10393436				
6	0,549	21,96	1,85866696				
7	0,535	21,4	1,90730498				
8	0,149	5,96	6,84837693				
9	0,648	25,92	1,57470396				
10	0,445	17,8	2,29305205				
11	0,54	21,6	1,88964475				
12	0,367	14,68	2,78040371				
13	0,338	13,52	3,01895906				
14	0,436	17,44	2,3403857				
15	0,425	17	2,40096038				
16	0,541	21,64	1,88615187				
17	0,289	11,56	3,53082409				
18	0,655	26,2	1,55787506				
19	0,438	17,52	2,329699				
20	0,166	6,64	6,14703713				
21	0,754	30,16	1,35332648				
22	0,27	10,8	3,77928949				
23	0,551	22,04	1,85192044				
24	0,385	15,4	2,65041081				
25	0,819	32,76	1,24591961				
26	0,664	26,56	1,53675928				
27	0,88	35,2	1,15955473				
28	0,358	14,32	2,85030213				
29	0,171	6,84	5,9672992				
30	0,552	22,08	1,84856551				
31	0,44	17,6	2,31910946				
32	0,646	25,84	1,5795792				
33	0,257	10,28	3,97045978				
34	0,669	26,76	1,52527379				
35	0,291	11,64	3,50655726				
36	0,162	6,48	6,29881582				
37	0,519	20,76	1,96610436				
38	0,955	38,2	1,06849022				
39	0,776	31,04	1,31495897				
40	0,559	22,36	1,82541711				
41	0,632	25,28	1,61456988				
42	0,703	28,12	1,45150521				
43	0,729	29,16	1,39973685				
44	0,456	18,24	2,2377372				
45	0,647	25,88	1,57713781				
46	0,396	15,84	2,57678829				
47	0,469	18,76	2,17571037				
48	0,486	19,44	2,09960527				
49	0,564	22,56	1,80923433				
50	0,686	27,44	1,48747546				
			120,730639				

Table 5: Phylogenetic tree of Myxozoa done with RAxML: This phylogeny is based on a reference tree from (Kodádková *et al.*, 2014) and supplementary sequences from Egil Karlsbakk. Each OTU has in this phylogeny been named firstly with an OTU number and then with the hit name and similarity percentage NCBI assigned to that OTU.OTUs that are coloured in the tree, have the same colouring in both MR bayes and RAxML trees. All the clades to the left of.

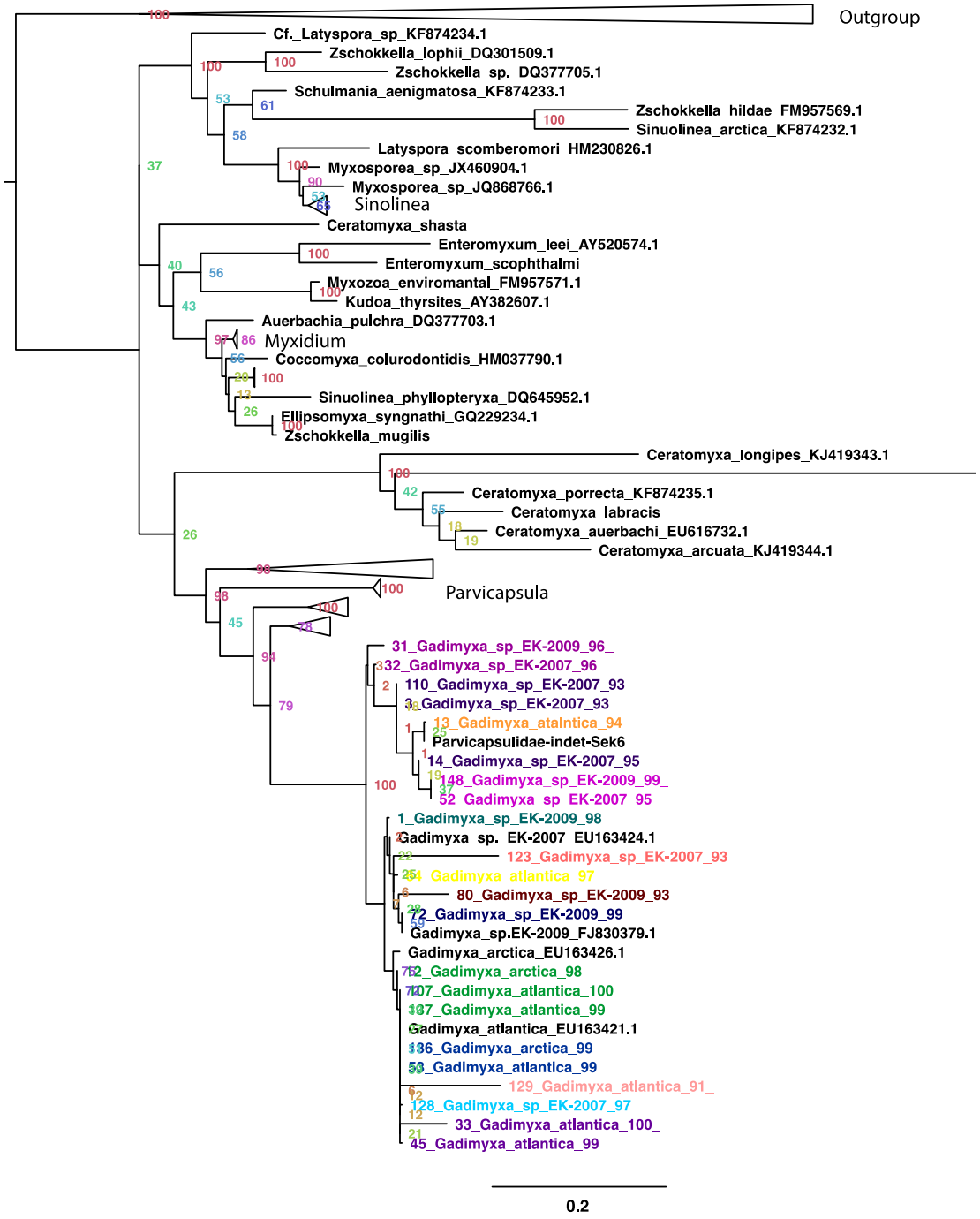


Table 6 Phylogenetic tree of Microsporidia done with RAxML. This phylogeny is based on a reference tree from (Vossbrinck & Debrunner-Vossbrinck, 2005b) and supplementary sequences from Egil Karlsbakk. Each OTU has in this phylogeny been named firstly with an OTU number and then with the hit name and similarity percentage NCBI assigned to the OTUs.

Table 7: Taxonomy assignment of marine fungi. The table shows that five OTUs are assigned to marine fungi, with 100 % similarity. Both SILVA and NCBI taxonomy are included in the table, ass well as Accession number.

OTU#	tax	Silva Taxonomy	note	NCBI Taxonomy	QC	e-value	%ID	Acc.no
59	Fungi	Penicillium marneffei ATCC 18224	marine yeast (environmental, and in sea cucumber intestine)	Debaryomyces hansenii strain Nc6HA-1 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	KR336844.1
83	Fungi	Hyphodontia rimosissima	skin fungi	Uncultured fungus clone nco39f04c1 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	KC670759.1
118	Fungi	Coniosporium perforans	rock inhabiting fungi	Chaetothyriales sp. TRN1 isolate TRN 1 small subunit ribosomal RNA gene, partial sequence	100 %	0.0	99 %	FJ358319.1
30	Fungi	Aspergillus fumigatus	marine deep sea fungi	Aspergillus restrictus isolate NIOCC F47 18S ribosomal RNA gene, partial sequence	100 %	0.0	100 %	EU723495.1
34	Fungi	uncultured Agaricales	marine fungi	Gloiocephala aquatica voucher SFSU:MR 10772 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	NG_013175.1
96	Fungi	Graphium adansoniae	marine fungi	Doratomyces stemonitis strain MF505 18S ribosomal RNA gene, partial sequence	100 %	0.0	100 %	KM096312.1
48	Fungi	Sistotrema raduloides	marine fungi	Sistotrema brinkmannii strain CB37 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	KM222227.1
70	Fungi	Lentinus tigrinus	air fungi?	Hyphodermella rosae strain MA-Fungi 88071 18S ribosomal RNA (SSU) gene, partial sequence	100 %	0.0	99 %	JN940191.1
46	Fungi	Rhizopus stolonifer	?	Rhizopus stolonifer strain FSU9872 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	KJ408540.1
18	Fungi	Uncultured	marine fungi (or human skin..)	Uncultured Malassezia isolate LT85_J8 18S ribosomal RNA gene, partial sequence	100 %	0.0	99 %	KC487833.1

Table 8: Reference sequences. These are the reference sequences the OTUs of Myxozoa and Microsporidia clustered to in their respective phylogenetic trees. The reference sequence Parvicapsulidae-indet- Sek6 is not yet published, and are therefore missing Accession number.

Reference sequences	Accession number
Gadimyxa sp EK-2007	EU163424
Gadimyxa sp EK-2009	FJ830379
Gadimyxa arctica	EU164426
Gadimyxa atlantica	EU163421
Parvicapsulidae-indet-Sek6	unpublished
Microsporidium cerebralis	JQ316511
Ichtyosporidium sp	BL39110
Pseudoloma neurophila	AF322654
Loma morhuae cod-EK	LM1012
Loma sp	AF104081
Kabatana takedai	AF356222