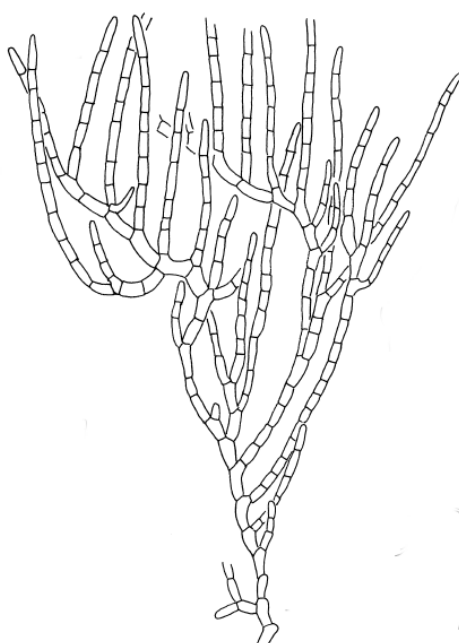# SPECIATION IN THE CLADOPHORALES SPECIES FLOCK OF ANCIENT LAKE BAIKAL

by

Sergio Diaz Martinez

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy in Marine Biology

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA
UNIVERSITY OF WELLINGTON

School of Biological Sciences
Victoria University of Wellington
Te Whare Wānanga o te Ūpoko o te Ika a Māui
2019

I

Front page image: *Cladophora compacta* from Boedeker et al. (2018)

Abstract

Understanding speciation is one of the great challenges in evolutionary biology as many of the processes involved in speciation, as well as the forces leading to morphological and genetic differentiation, are not fully understood. Three main modes of speciation have been described: allopatric, parapatric and sympatric. Sympatric speciation is the most enigmatic mode because in the absence of physical barriers, disruptive selection, assortative mating and hybridization play central roles in reproductive isolation. Although it is accepted that sympatric speciation is possible, only a few examples of this process exist to date. Another common method of speciation in plants and algae is via polyploidization. Recently, a promising system to study speciation in sympatry was discovered: the endemic Cladophorales species flock in ancient Lake Baikal, Russia. The flock consists of sixteen taxa grouped in four genera: *Chaetocladiella, Chaetomorpha, Cladophora* and *Gemmiphora*. In spite of their morphological diversity, recent molecular analyses have shown that this is a monophyletic group with low genetic variation and nested within the morphologically simple genus *Rhizoclonium*. Due to their high number of species, endemism and sympatric distribution, many interesting questions have arisen such as what processes are involved in speciation, and whether this group might be a novel example of sympatric speciation. In this study, we analysed the population genetics of the endemic Baikalian Cladophorales to infer the processes shaping the evolution of the group. First, a set of microsatellites was designed using high-throughput sequencing data. Second, species delimitation methods based on genetic clustering were performed. Third, the population genetics of three widely distributed species was analysed looking for evidence of panmixia, a common criteria to support sympatric speciation. A total of 11 microsatellites that mostly cross-amplify between most species were obtained. The genotyping revealed that most loci had more than two alleles per individual indicating polyploidy. As such, the analyses required a different approach which consisted in coding the genotypes as 'allelic phenotypes', allowing the use of individuals of different ploidy levels in the same data set. The species delimitation of 15 operative morphotaxa and 727 individuals supported reproductive isolation of five morphotaxa and two hypotheses of conspecificity. However, some morphotaxa showed unclear assignments revealing the need of further research to clarify their reproductive limits. Finally, the population genetics of

Chaetomorpha moniliformis, *Cladophora compacta* and *Cl. kursanovii* revealed patterns of genetic variation and structure that suggest different reproductive strategies and dispersal abilities. This demonstrates that contrasting biological characteristics may arise in closely related lineages: *Chaetomorpha moniliformis* with dominant asexual reproduction and long dispersal abilities; *Cladophora compacta* with high genetic diversity, no population structure and likely to reproduce sexually; *Cl. kursanovii* with a structure congruent with geographic distribution and more restricted dispersal. The results suggest that polyploidy, rather than speciation with gene flow, is the force driving the reproductive isolation and evolution of this flock. Although many questions remain to be studied, this research provides the first insights into the diversification of this Cladophorales species flock and contributes to the understanding of speciation in freshwater algae.

My science to the mankind

My faith to the nature

My loyalty to life
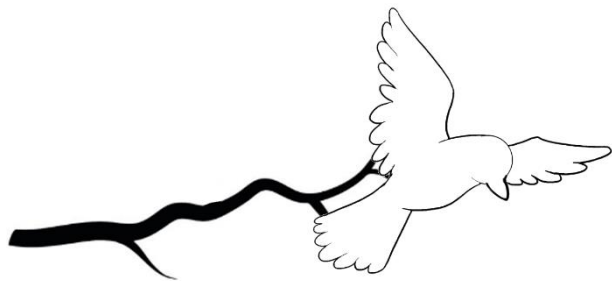
Dedicated to Edith Calixto Pérez

Dedicado a Edith Calixto Pérez

Acknowledgments

I would like to thank to Dr. Joe Zuccarello for accepting me as a student, his guidance, patience and motivation helped me to improve my skills and my knowledge in science and phycology. To Christian Boedeker, his support and comments helped to see things from a different angle. I could not have had better supervisors and I am proud of had been their student.

To the examiners of this thesis, Dr. Peter Ritchie, Dr. Stacy Krueger-Hadfield and Dr. Chrissen Gemmill, their suggestions and comments helped me to improve this manuscript.

To Edith, who walked by me during this journey. There are no words to express how grateful I feel for her support and love. Without her this would be only a dream, thank you. To my family in Mexico, thanks for all the support in spite of the distance. To my mother Juana, my father Alejandro, my brother Alejandro, my sister Diana and her husband Francisco, my uncle Javier and my grandfather Jorge.

Finally, to all the people and friends that I meet in New Zealand, for their friendship and all the time we spent together. (Andrea, Ruben "Boris", Tania, Johanna, Balam, Kim, Olivia, Maren, Sven), phycology colleagues (Wendy Nelson, Roberta D'Archino, Judy Sutherland) and Patricia Stein, for her help with the tons of questions I had during this time.

Table of contents

List of Figures

List of Tables

Appendices

Chapter One

GENERAL INTRODUCTION

1.1 Introduction

This research focused on the genetic diversity and speciation of the species flock of Cladophorales in ancient Lake Baikal, Russia (Boedeker et al. 2018). The aim of the project was to analyse the evolutionary processes involved in the diversification of this monophyletic group of green algae endemic to Lake Baikal - whether speciation occurred in sympatry, and whether processes such as polyploidy were involved. Reproductive isolation between morphological species was analysed using genetic data and cluster methods. Finaly, the genetic structure of three species that are widely distributed in the lake was studied to gain insights into their reproductive strategies and dispersal.

The general introduction contains four parts that help in understanding the questions addressed in this thesis. First, I provide an overview of known modes of speciation as well as of other speciation processes such as hybridization and polyploidy. Second, I provide a brief overview of the theoretical and practical problems involved in species delimitation and how population genetics can help to delineate species. Third, I describe the general features of ancient Lake Baikal and the biota inhabiting the lake. Finally, a general description of the order Cladophorales and the endemic species of Lake Baikal is provided, pointing out the most relevant characteristics of this interesting group of algae.

1.2 Species and speciation

1.2.1 *Modes of speciation*

Species are the basic level of organization in biology and are the fundamental unit in taxonomy, evolution, genetics, ecology and conservation. Understanding the way in which a single species diversifies into two or more independent lineages (speciation) remains one of the biggest challenges in biology, as many of the processes that lead to morphological and genetic differentiation are not fully understood (Albert and Schluter 2005, Butlin et al. 2012).

Three main modes of speciation have been described: 1) allopatric speciation, in which species originate in different geographical areas without contact (Mayr 1942); 2) parapatric speciation, in which species originate from populations with a limited contact

zone, but do not completely overlap (Smith 1955); and 3) sympatric speciation, in which species originate in the same geographic area from an ancestral species, without any physical barriers (Poulton 1904).

Among the different modes of speciation, sympatric speciation is the most controversial. In the absence of physical barriers, gene flow between populations must be interrupted by intrinsic forces. In this speciation mode, the homogenizing effect of gene flow is overcome by disruptive selection and/or other factors such as microhabitat preference, host choice, assortative mating, polyploidy and hybridization (Coyne 2007, Fitzpatrick et al. 2009). Although many of these forces are shared between animals and plants (and algae), sexual selection and behaviour are only present in animals, while speciation via hybridization and polyploidy with fertile offspring is common in plants and algae (Albert and Schluter 2005).

Part of the controversy lies in the criteria used to determine if speciation was truly sympatric rather than allopatric or parapatric. The main criteria used to determine if sympatric speciation is possible are: 1) the species have overlapping ranges; 2) speciation is complete; 3) the group is monophyletic containing sister species; and 4) allopatry is 'unlikely' (Coyne and Orr 2004). While it is now widely accepted that sympatric speciation is possible, it is debated how common this phenomenon is and how many reliable examples exist (Bolnick and Fitzpatrick 2007, Coyne 2007, Mallet et al. 2009). Difficulties arise when sympatric and allopatric speciation can explain the same observations. Secondary contact after allopatric speciation or double invasion may result in a sympatric distribution (Bolnick and Fitzpatrick 2007). However, well documented cases exist, for example: the cichlid fish in two small Nicaraguan volcanic crater lakes (Barluenga et al. 2006); the two palm species from Lord Howe Island (Savolainen et al. 2006); or the races of *Rhagoletis pomonella* that feed on different host plants (Filchak et al. 2000). Recent advances in molecular and computational techniques are helping in the search and test of new cases of speciation with gene flow (e.g. Papadopulos et al. 2011, Filatov et al. 2016).

1.2.2 *Hybrid speciation*

Generally, the speciation process leads to the development of mechanisms that prevent interbreeding between species and maintain their genetic integrity (reproductive

barriers). Prezygotic barriers prevent the fertilization of the egg and subsequent formation of zygotes, while postzygotic barriers act after fertilization producing inviable and sterile hybrids or 'hybrid breakdown' (Rieseberg and Willis 2007). However, in some cases, two species can successfully interbreed and produce viable offspring. This process is called hybridization and it is calculated that it occurs in about 10% of animal species and 11-25% of vascular plant species (Paun et al. 2006, Baack and Rieseberg 2007, Mallet 2007). Although hybridization is reported to ocurr in all gruops algae (i. e. Durand et al. 2002, Niwa et al. 2010, Moalic et al. 2011, Xu et al. 2011, Zardi et al. 2011, Montecinos et al. 2017, Ding et al. 2018, Lee et al. 2018), little is know about how frequent it is and how does this process ocurr in most of the lineages (Leliaert et al. 2014).

Hybridization may affect the evolution of plants and animals in many ways. If it occurs frequently and the offspring have good fitness, the hybrids could be assimilated by one of the parent species (with the subsequent introduction of new alleles into the parental population, i.e. introgression), or they could merge the two parental species into a single lineage (Chapman and Burke 2007). Hybridization may also lead to the formation of completely new species if the hybrids are fertile, viable and maintain some degree of reproductive isolation from the parental species (Hegarty and Hiscock 2004, Chapman and Burke 2007, Mallet 2007).

Hybrid speciation may occur without a change in chromosome number (homoploidy) (Rieseberg and Willis 2007). In these homoploid hybrids speciation is driven by ecological and genetic factors, such as suitable ecological niche, fitness promoted by positive heterosis (hybrid vigor), rapid chromosomal evolution and chromosomal rearrangements that differentiate the new species from the parental species (Paun et al. 2006). Homoploid hybrids are considered rare (Soltis and Soltis 2009) and are usually sterile, but can recover the fertility after genome doubling (polyploidization) (Rieseberg and Willis 2007, Soltis and Soltis 2009).

1.2.3 *Polyploidy as a mechanism of speciation*

Polyploidy is the multiplication of the entire genome of an organism. Depending on the origin of the extra genome, these organisms may be classified as autopolyploids or allopolyploids (Stebbins 1947). In autopolyploids, the genome of a single individual is

duplicated and this results in two copies of the same genome, for example by self-fertilization of unreduced gametes after meiotic errors. In alloployploids, the genomes derive from different species through hybridization and/or by genome doubling of hybrids (Madlung 2012, Husband et al. 2013).

Polyploidy has occurred in almost all biological groups but is more frequent in higher plants (Wood et al. 2009) and algae (Kapraun 2005, 2007). It is considered to play an important role in the evolution of these vascular plants (Soltis et al. 2010, Husband et al. 2013). Polyploid species can have some advantages over their diploid counterparts. Some of the most noticeable effects are changes in morphological traits (Comai 2005) and the increase of survival in new environmental conditions (Comai 2005, Soltis et al. 2010, te Beest et al. 2012). Perhaps the major influences of polyploidy in evolution are the modification of reproductive strategies leading to asexual reproduction (Thompson and Lumaret 1992, Comai 2005) and the origin of new species by instant speciation, where the resulting offspring may be reproductively isolated from their parents (Schluter 2001, Albert and Schluter 2005, Bolnick and Fitzpatrick 2007).

Instant speciation is considered to have contributed to about 2-4% of higher plant speciation events (Schluter 2001, Soltis et al. 2010). Although instant speciation occurs in sympatry and is considered as a mode of sympatric speciation by some authors (Coyne 2007), the processes involved differ substantially from what is considered strict sympatric speciation, due to the fact that this is not the result of selection from a single panmictic population, or in other words, it is not speciation with gene flow (Bolnick and Fitzpatrick 2007).

1.2.4 *Polyploidy in algae*

Polyploidy is considered to occur in several taxa among the algae (Kapraun 2005, 2007, Kapraun and Freshwater 2012, Husband et al. 2013). However, the occurrence of this process and ploidy level in particular species is often unknown until further analyses are performed such as chromosome counting (e.g. Del Grosso 1984, Verma 1985, 1986, Hinson and Kapraun 1991) or DNA content measurement (e.g. Kapraun 2007, Mazalova et al. 2011, Phillips et al. 2011, Kapraun and Freshwater 2012, Varela-Álvarez et al. 2018). In spite of the knowlegedge on the occurrence of polyploidy in

algae, little is known about its timing, its frequency and its influence on the evolution of each lineage (Kapraun and Freshwater 2012). The study of polyploidy in algae also face some other caveats including species with haplo-diplontic life cycles and genetic chimeras (Varela-Álvarez et al. 2018), which are organisms with cell lines of different ploidy level. Previous analyses on Cladophorales (Verma 1986, Hinson and Kapraun 1991, Kapraun 2005, 2007) have shown that polyploidy is practically universal within order consisting basically of constallations of six chromosomes with subsequent multiples of 6 (12, 18, 24, 30, 36, etc.). Therefore, the occurrence of polyploidy in species of the Cladopohorales species flock could be expected.

1.3 Species concepts and species delimitation

1.3.1 *Species and the inference of their boundaries*

Species delimitation is fundamental in studies of taxonomy, ecology and evolution. However, recognizing the boundaries between species may be a difficult task. Many different concepts have been proposed to define a species (the so called 'species problem', see Mayden 1997), each one highlighting particular attributes of the species and none applicable to all species. For example, according to Morphological Species Concept, each species is the smallest group that are consistently and persistently distinct in morphology and can be distinguish by ordinary means (Cronquist 1978). The Biological Species Concept (Mayr 2000) rests on evidence of reproductive isolation between populations (species). Others, such as the Phylogenetic Species Concept, rely on unique diagnostic characters (synapomorphies) (Cracraft 1983, Wheeler and Platnick 2000) or evidence of monophyly (Mishler and Theriot 2000) to delimit species. The Genotypic Cluster Definition, on the other hand, considers a species as a monotypic or polytypic cluster of biological entities identified by morphology or genetics (Mallet 1995). Although there is still no consensus among biologists about what a species is and the debate persists (see Wheeler and Meier 2000, De Queiroz 2007), the current conceptual framework sees species as independently evolving lineages and the other concepts are considered operational criteria to assess lineage divergence (Dayrat 2005, De Queiroz 2007).

Some of the most widely used methods to delimit species are DNA-based, such as barcode gaps (Hebert et al. 2003, Puillandre et al. 2012), coalescent-based methods (Fujisawa and Barraclough 2013) and tree-based methods (Zhang et al. 2013), which have helped not only on the inference of the species boundaries, but also in the discovery of cryptic species and to test the conspecificity taxa with huge morphological plasticity (Leliaert et al. 2014).However, these DNA-based methods have many limitations. Hybridization, incomplete lineage sorting and even low genetic signal on the selected markers may lead to the inference of incorrect species boundaries (Carstens et al. 2013, Edwards and Knowles 2014). In such case, delimitation methods based on population genetics approaches can provide more accurate species delimitation (Shaffer and Thomson 2007, Hausdorf and Hennig 2010, Carstens et al. 2013).

1.3.2 *Population genetics for species delimitation and genetic structure of species*

Population genetics is the study of the genetic structure resulting from microevolutionary processes such as selection, genetic drift, gene flow, mutation and recombination (Valero et al. 2001). Its study can help to understand some of the processes shaping the diversity, evolution and speciation of populations. This approach considers genetic diversity within species and is particularly useful in identifying recent speciation events where other methodologies lack enough power to discover species boundaries (Shaffer and Thomson 2007, Hausdorf and Hennig 2010). The most widely used methods include different types of genetic markers in combination with clustering algorithms using (Carstens et al. 2013, Edwards and Knowles 2014). These methods are model-free multivariate analysis such as Gaussian clustering or discriminant analysis of principal components (Hausdorf and Hennig 2010, Jombart et al. 2010), and others using implicit models like STRUCTURE or Structurama (Pritchard et al. 2000, Huelsenbeck et al. 2011).

Compared to DNA-sequence based methods, population genetics requires an extensive sampling and many markers to explore the genetic structure of the populations, but can reveal evidence of reproductive isolation by clustering genetically similar genotypes that have evolved independently with little or no ongoing gene flow (Hausdorf and Hennig 2010). Theoretically, these methods agree with the genotypic cluster definition of species and indirectly with the biological species concept, although

some consideration regarding their limits and possible caveats must be taken into account (Carstens et al. 2013). This approach has been used in several groups of organisms including plants (e.g. Carstens and Satler 2013, Pachschwöll et al. 2015) and many animal taxa (e.g. Jörger et al. 2012, Satler et al. 2013, Lamanna et al. 2016, Malaney et al. 2017, Posso-Terranova and Andrés 2018). In addition, population genetic approaches are also useful in the study of sympatric speciation by providing estimates of genetic differentiation and gene flow within a species and populations, and to test if the genetic structure of each daughter sympatric species follow a panmictic population pattern (Barluenga et al. 2006, Savolainen et al. 2006, Papadopulos et al. 2011).

Many different molecular markers for population genetics have been used to date, such as RAPDs (random amplified polymorphic DNA), AFLPs (amplified fragment length polymorphisms), RFLPs (restriction fragment length polymorphisms), ISSR (intersimple sequence repeats), SNPs (single nucleotide polymorphisms) and microsatellites (or simple sequence repeats, SSRs). These last two markers are the most widely used in recent years due to novel molecular technologies that allows their easier development in non-model organisms (Bräutigam and Gowik 2010, Davey et al. 2011, Cronn et al. 2012).

Although SNPs are becoming more popular among researchers (Guichoux et al. 2011, Defaveri et al. 2013), microsatellites, repetitive DNA short sequences about 2-5 base pairs scattered in the genome of many organisms, are still attractive over SNPs for population genetics studies due to their codominance, neutrality, high levels of polymorphism and easy preparation (Selkoe and Toonen 2006). In addition, recent advances in High-Throughput-Sequencing (HTS) and high computational power allow obtaining a great number of potential SSRs for a low price and short development time possible (Schoebel et al. 2013).

1.3.3 *Population genetics on algae*

Stuides on population genetics have been done in several groups of algae for the study of invasions (O'Doherty and Sherwood 2007, Andreakis et al. 2009), dispersal and migration (Van der Strate et al. 2003, Hu et al. 2013), habitat fragmentation (Billot et al. 2003, Coleman and Kelaher 2009), reproductive strategies (Engel et al. 2004, Couceiro et al. 2015, Krueger-Hadfield et al. 2016) and conservation (Couceiro et al. 2011b).

Athough the genetic variation and structure of algae population is influenced by mutation, selection, migration and genetic drift, some other factors inherent to the biology of the algae, such as polyploidy (Varela-Álvarez et al. 2018), reproduction modes (sexual or asexual) (Krueger-Hadfield et al. 2016, Dudgeon et al. 2017) and life cycle (Van der Strate et al. 2002b, Engel et al. 2004, Couceiro et al. 2015, Krueger-Hadfield et al. 2016), must be considered to avoid potential issues (Valero et al. 2001). In particular, these two last factors are important due to the different mechanism of algae reproduction and differentiation in algae with haplo-diplontic life cycle. For example, species with a bisaphic life cycle would require a specific number of individuals to capture the genetic diversity among gametophytes and sporophytes (Krueger-Hadfield and Hoban 2016), different reproduction strategies within each phase can result in genetic structure (Krueger-Hadfield et al. 2016), or the transition to asexuality, a feature common in algae, can lead to differentiation of individuals within populations (Ardehed et al. 2015, Dudgeon et al. 2017). In spite of current knowledge on population genetics in algae and its growing interest as result of the new technologies, methods and markers, studies on haplo-diplontic algae are fewer compared to diplontic algae (e.g. fucoids, kelps) (Coleman and Wernberg 2018, Liu et al. 2018, Rinne et al. 2018, Shan et al. 2019)) and the suites on freshwater algae are scarce.

1.4 Ancient Lake Baikal as system for the study of speciation

Ancient lakes are known centres of biodiversity that contain many endemic species and provide exceptional examples of ecological diversification, speciation and cases of adaptive radiation (Salzburger and Meyer 2004, Barluenga et al. 2006, Seehausen 2006, Wilke et al. 2006, Herder et al. 2008, Keller et al. 2013). Thus, they are particularly well suited to the study of speciation and the relationship between biological processes and past geological and/or environmental events (Cristescu et al. 2010).

Lake Baikal (Fig. 1.1) is an ancient lake of tectonic origin in south-eastern Siberia, Russia. It is the deepest (≈1.5 km) and oldest lake in the world having an estimated age of approximately 25-30 My (Mats and Perepelova 2011). The water of the lake is cold with an average temperature of 4° C (Sherstyankin and Kuimova 2006) and is ultraoligotrophic (Khodzher et al. 2017). Throughout its history, the lake has

experienced geological (Logatchev 1993, Mats 1993) and environmental perturbations (Sherstyankin and Kuimova 2006). During the last 5 My, the lake passed through several climate changes, from subtropical (~5 Mya), a transitional stage (from ~2.6 to 2.4 Mya), to the current subarctic climate (Sherstyankin and Kuimova 2006).

The biota of the lake is diverse and unique (Fig. 1.2). The fauna consists of more than 2600 species (Timoshkin 1999, Timoshkin et al. 2004), most of them endemic and representing many species flocks such as turbellareans (Timoshkin et al. 2010), annelids (Kaygorodova et al. 2007), amphipods (Macdonald et al. 2005), ostracods (Schön and Martens 2012), molluscs (Hausdorf et al. 2003, Stelbrink et al. 2015) and scuplins (Kontula et al. 2003, Yokoyama and Goto 2005).
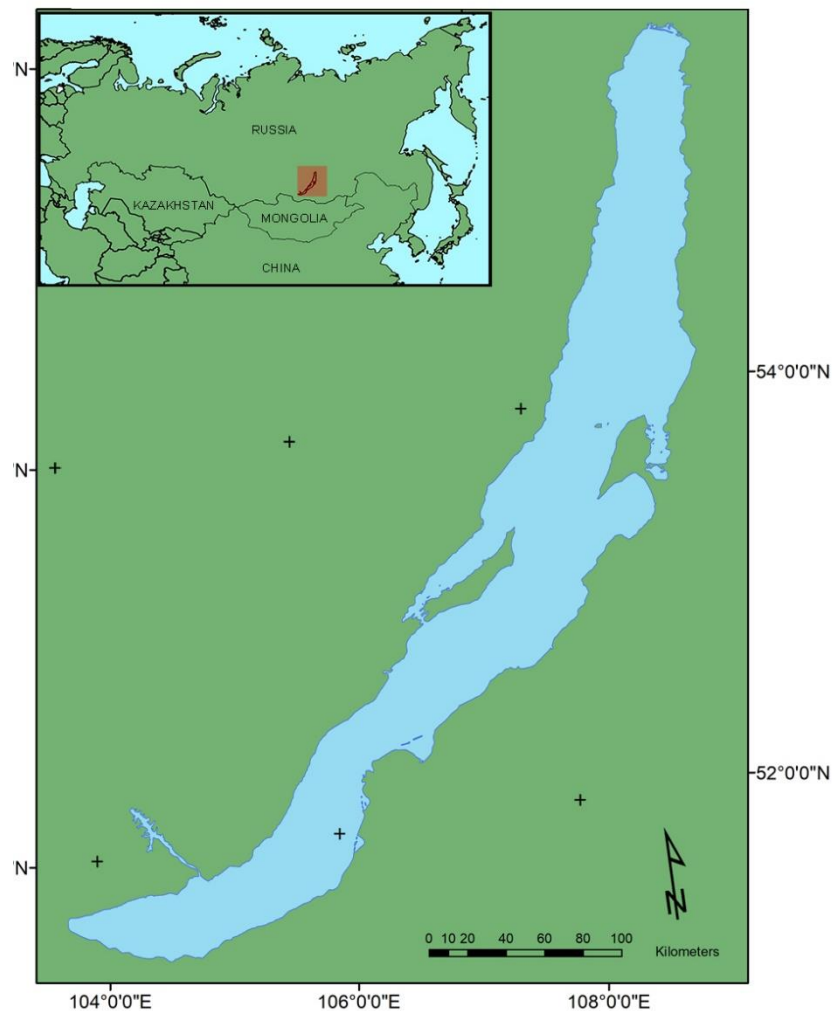
Figure 1.1. Overview of Lake Baikal, Siberia. Inset shows its location in Russia.

Although the origin of the species and the radiation timing in most groups has not been investigated, the current knowledge on some of the widely studied species flocks in Lake Baikal suggests either, lineages that radiated or invaded the basins during the cooling of the lake during the Pleistocene (~2 Mya), and lineages that may have evolved along with history of the Baikalian rift (~70 Mya)(Sherbakov 1999) producing the species assemblages that are present today. Under the first scenario, species such as sculpins (Cottoidei; Sherbakov 1999, Kontula et al. 2003, Yokoyama and Goto 2005), Lumbriculidae (Oligochaeta; Kaygorodova et al. 2007), ostracods (Crustacea; Schön and Martens 2012, Karanovic and Sitnikova 2017), limpets (Stelbrink et al. 2015) and Baicaliidae gastropods (Mollusca; Sherbakov 1999) represent young lineages. On the other hand, species of copepods (Crustacea; Mayor et al. 2010) and *Choanomphalus* pulmate gastropods (Sherbakov 1999) represent ancient lineages. In other groups such as turbellarians (Sherbakov 1999) and some amphipods (Väinölä and Kamaltynov 1999) molecular evidence indicates the occurrence of both ancient and young lineages. The factors influencing the evolution of these groups are still unclear, but it is hypothesized
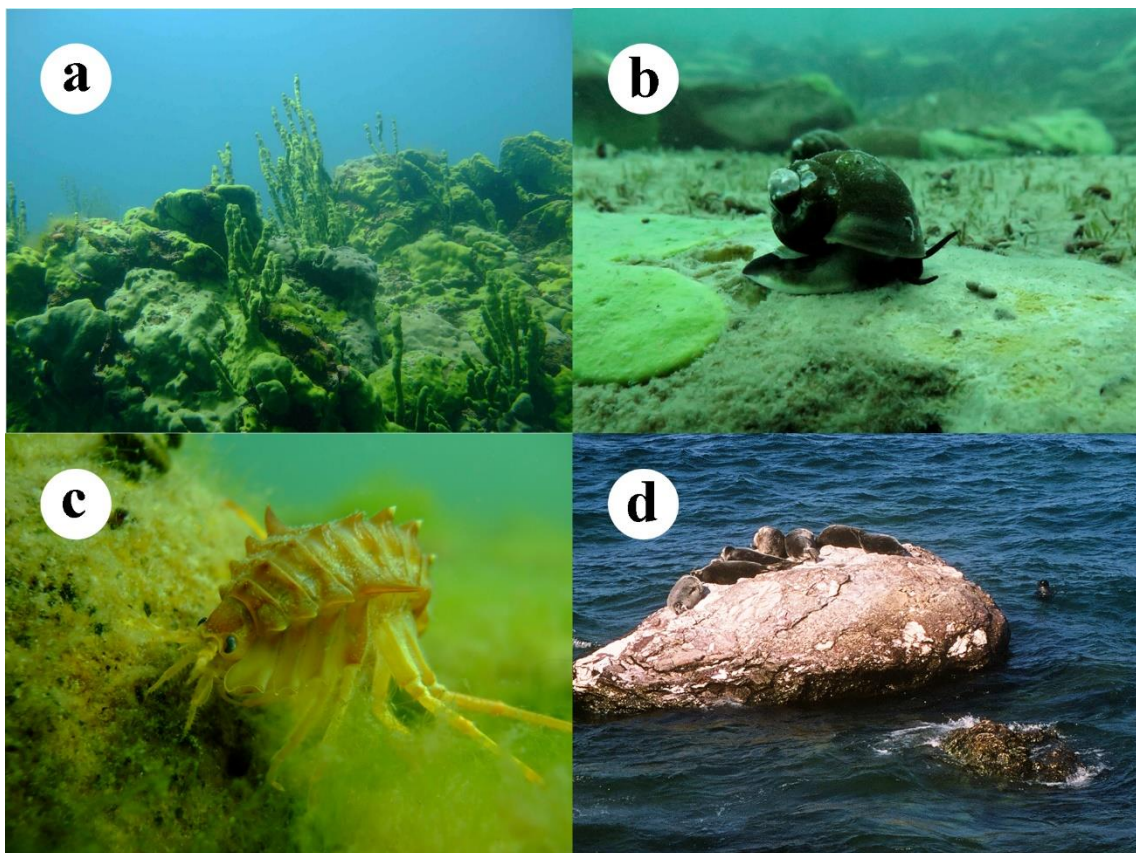
Figure 1.2. Examples of the fauna inhabiting Lake Baikal. a = Lubomirskia sponges; b = gastropods; c = isopods; d = freshwater seals. Credits: a = Holger Anlauf; b = Sven Ihnken; c = Holger Anlauf; d = Christian Boedeker.

that some groups may have radiated after the cooling of the lake (Kontula et al. 2003, Yokoyama and Goto 2005, Kaygorodova et al. 2007), the formation of the deep basin of oxygenated water (Schön and Martens 2012, Karanovic and Sitnikova 2017) or/and by multiple introduction events (Sherbakov 1999, Macdonald et al. 2005, Schön and Martens 2012).

In contrast to the widely studied groups of animals, the evolution and diversification processes of the algae inhabiting Lake Baikal remain to be studied in spite of the large number of endemic species (Izhboldina 2007, Kulikovskiy et al. 2012, Boedeker et al. 2018). In particular, the endemic green algae of the order Cladophorales (Ulvophyceae, Chlorophyta) are a group with many interesting biological questions to be solved.

1.5 Cladophorales of Lake Baikal, a case of sympatric speciation?

1.5.1 *Characteristics of the order Cladophorales*

The order Cladophorales is a cosmopolitan group of algae with more than 200 (Guiry and Guiry 2018) species that can be found in freshwater, brackish and marine waters. The thallus morphology of these algae is essentially filamentous and siphonocladous, with multinucleate cells (Hoek et al. 1995). The filaments can be unbranched (e.g. *Chaetomorpha* [Fig. 1.3a], *Okellya*, *Rhizoclonium*) or branched (e.g. *Aegagropila*, *Basicladia, Boodlea, Cladophora* [Fig. 1.3b], *Pithophora*, *Willela*, *Wittrockiella,* etc.); attached or unattached to the substratum (e.g., *Pithophora, Aegagropila* balls) and also with secondary rhizoids (e.g., some *Cladophora* spp., *Pithophora*, *Rhizoclonium* [Fig. 1.3c]). Some species may exhibit a modified body plan, developing anastomosing filaments (e.g. *Anadyomene, Microdictyon* [Fig. 1.3d]) or inflated cells (e.g. *Valonia* [Fig. 1.3e]) (Hoek et al. 1995). The typical life cycle of the species (Fig. 1.4) of this order is haplo-diplontic (alternating between sporophyte and gametophyte) and isomorphic, although some species reproduce mainly asexually by fragmentation or by producing zoospores (Škaloud et al. 2018).
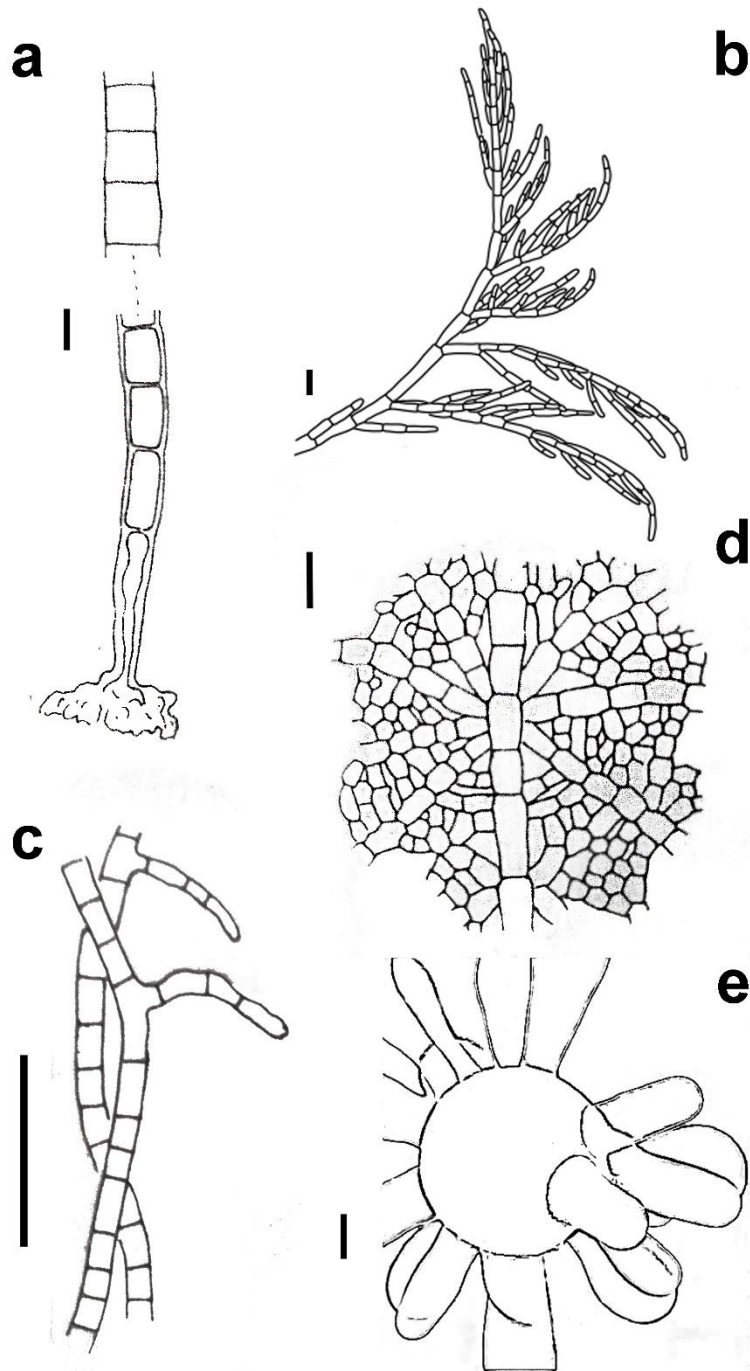
Figure 1.3. Different morphologies of some genera within the Cladophorales. a=
Chaeotomorpha sp. showing middle and basal part with rhizoids; b = *Cladophora* sp.
displaying branching filaments; c = *Rhizoclonium* sp. showing unbranched filaments and
development of rhizoids; d = *Microdictyon* sp. presenting anastomosing networks; e =
Valonia sp. showing an inflated cell. Scale bars: a, c, d = 200 µm; d = 500 µm; e = 1000
µm. a, c, d from Hoek (1984); b from Hoek (1963); e from Leliaert et al. (2007).
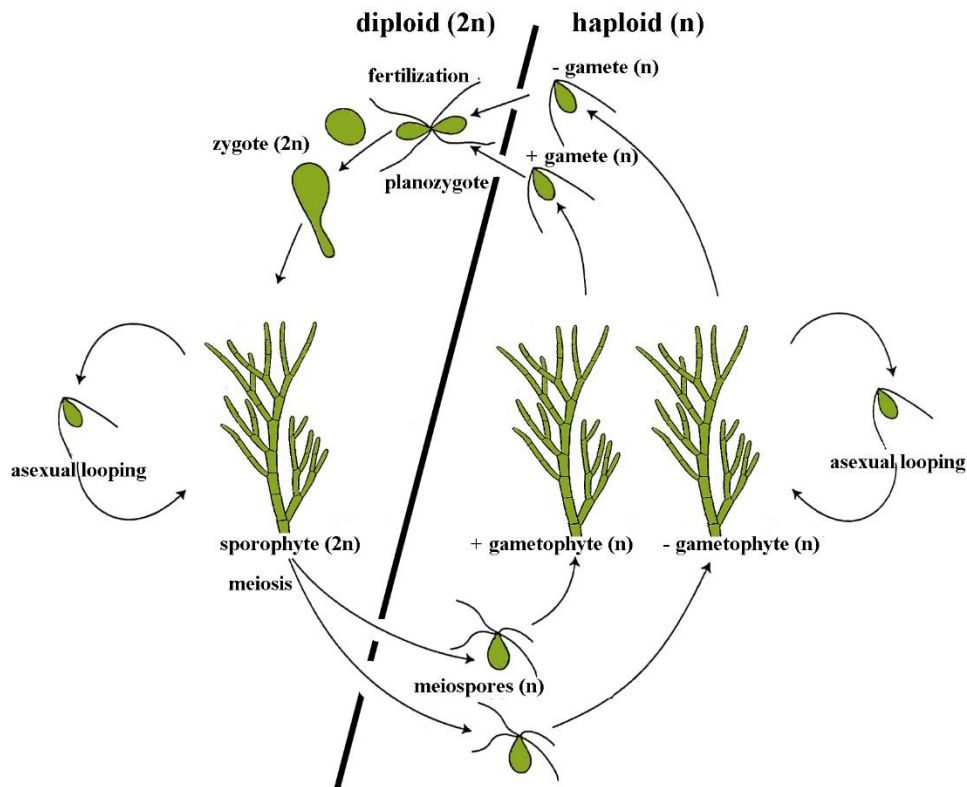
Figure 1.4. Cladophora life cycle. Isomorphic haplo-diplontic and isomorphic. Asexual reproduction may occur by fragmentation and spore production in both sporophyte and gametophyte. Modified from Škaloud et al. (2018).

The taxonomy and classification of the order was mainly based on morphological characters such as, thallus organization, mode of growth, insertion of the branches, angle of ramification, distinctness of main axes, diameter of apical cells, ultimate branches and main axes, shape and size of the cells, number of branches per node; attachment mode, and macroscopic habit and height of the plants (Hoek 1963, 1982). The low number of diagnostic characters between species and genera, parallel and convergent evolution and phenotypic plasticity has led to a confused taxonomy, but the use of molecular data (mostly ribosomal DNA sequences) has provided a new perspective on the order. The order is currently divided in five lineages (Families, Fig. 1.5) (Hanyuda et al. 2002, Leliaert et al. 2003, 2007, 2009, Boedeker and Sviridenko 2012, Boedeker et al. 2016): 1) the unbranched marine Okellyaceae which is a monotypic taxon; 2) Pseudocladophoraceae which consists of only two species inhabiting brackish and freshwaters; 3) Pithophoraceae that inhabits brackish and freshwater and includes *Aegagropila*, *Agaegropilopsis*, *Pithophora*, *Basicladia* and
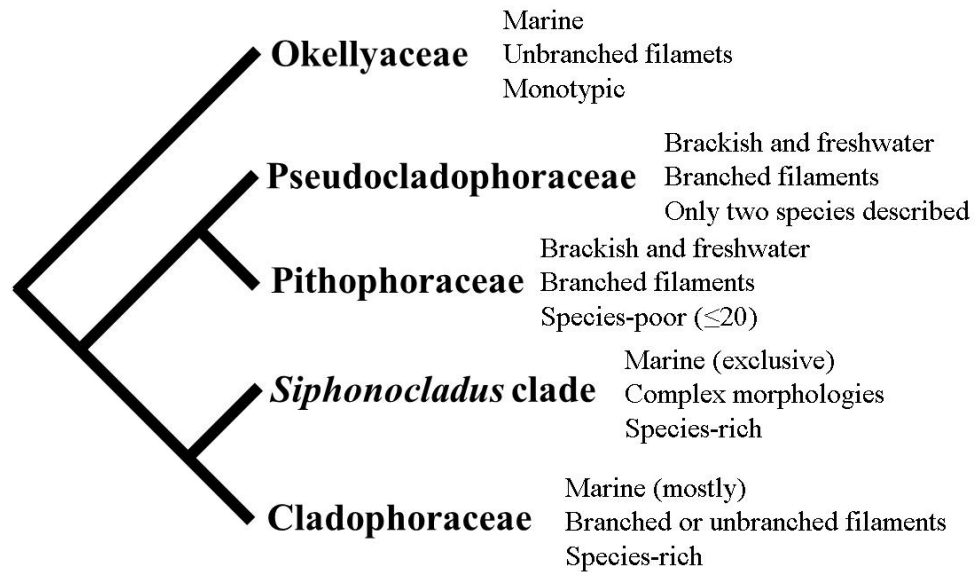
Figure 1.5. Current hypothesis of phylogenetic relationships of the main Cladophorales clades (families or undesignated clade). After Boedeker et al. (2012, 2016).

Wittrockiella; 4) the '*Siphonocladus*' clade, that consist of marine species with complex morphologies and phylogenetic relationships, which includes genera such as *Anadyomene*, *Boodlea*, *Dictyosphaeria*, *Microdictyon*, *Siphonocladus*, *Valonia*; and 5) Cladophoraceae, consisting of unbranched and branched filaments inhabiting marine, brackish and freshwater environments including many genera such as *Cladophora*, *Chaetomorpha*, *Rhizoclonium*, *Pseudorhizoclonium*, *Lurbica* and *Willeella*.

1.5.2 *The endemic Cladophorales of Lake Baikal*

Sixteen taxa of Cladophorales mostly endemic to Lake Baikal have been traditionally described based on morphology (Appendix 1.1), all of them are grouped in four genera: *Chaetocladiella* Meyer & Skabichevsky (Fig. 1.5a) which only branches at the base; the unbranched *Chaetomorpha* (= *Ch.*) Kützing (Fig. 1.5b); the abundantly branched *Cladophora* (= *Cl.*) Kützing (Fig. 1.5c,d); and the diminutive *Gemmiphora* Skabichevsky (Fig. 1.5e) (Izhboldina 2007, Boedeker et al. 2018). The morphological diversity held by the species of this group is so distinctive that it can be compared with that found in the entire order (Boedeker et al. 2018). However, recent molecular analyses have shown that a clade of endemic Cladophorales of Lake Baikal are actually monophyletic and nested within the genus *Rhizoclonium*, Cladophoraceae (Appendix 1.2) with some of the taxa likely to be conspecific. In the light of this new evidence, the
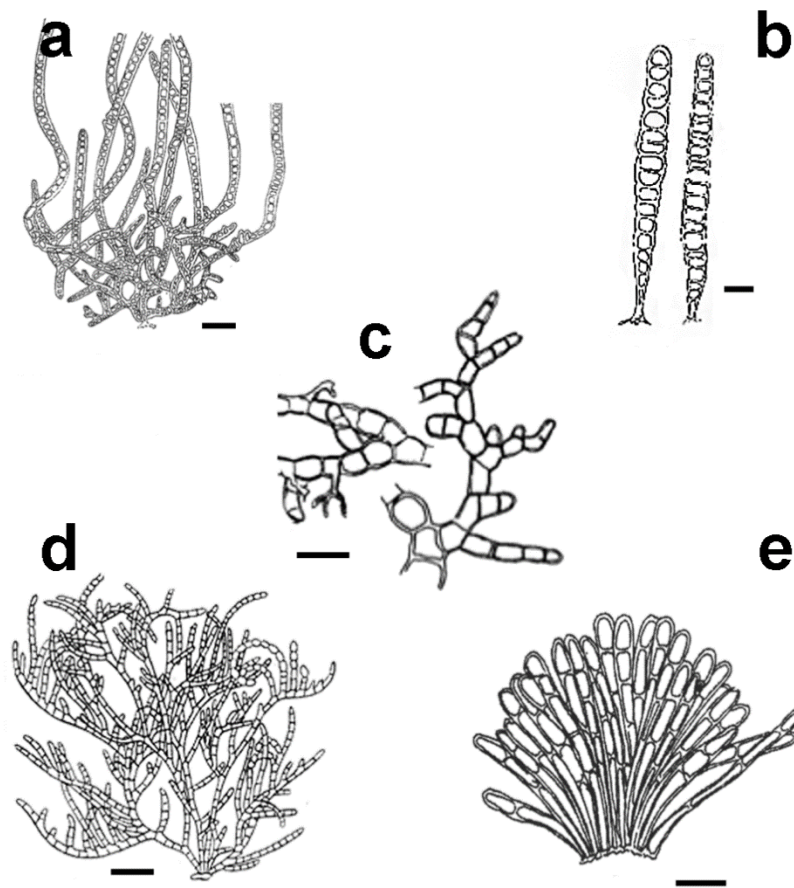
Figure 1.6. Different morphologies of the Baikalian Cladophorales genera (after Izhboldina 2007). a = *Chaetocladiella* only branched at the base; b = *Chaetomorpha* unbranched filaments; c, d = *Cladophora* branched filaments; e = *Gemmiphora* unbranched or sparsely branched filements. Scalebars: a = 800 µm; b = 500 µm; c = 300 µm; d = 350 µm; e = 200 µm.

Baikalian Cladophorales are now considered an example of a species flock in algae (Boedeker et al. 2018) due to their high number of species, their endemism and their single origin (Greenwood 1984), raising interesting questions concerning the evolutionary processes and diversification that brought them about.

Some of these questions are: How the speciation of the endemic Cladophorales in Lake Baikal occurred and whether this group might be a novel example of sympatric speciation. In this regard, the Cladophorales of Lake Baikal is likely to fulfil several criteria for sympatric speciation. The species occur together in one continuous water system and in many cases share the same microhabitat (Fig. 1.6), they are likely to be different species and they are a monophyletic group. However, adequate species delimitation must be carried out to confirm the reproductive isolation between morphospecies and the population genetics of the species must be analysed to determine

if other speciation processes such as allopatry or polyploidization can be rejected (Coyne and Orr 2004). Although the species *Cladophora floccosa* and *Chaetocladiella microscopica* are also reported for Lake Hövsgöl in Mongolia (see Boedeker et al. 2018 and references therein), their origin is likely to have occurred within Lake Baikal and no further evidence of other species inhabiting nearby lakes or rivers is available. Therefore the group as a whole is considered hereafter as the endemic Cladophorales of Lake Baikal for operative and historical reasons.
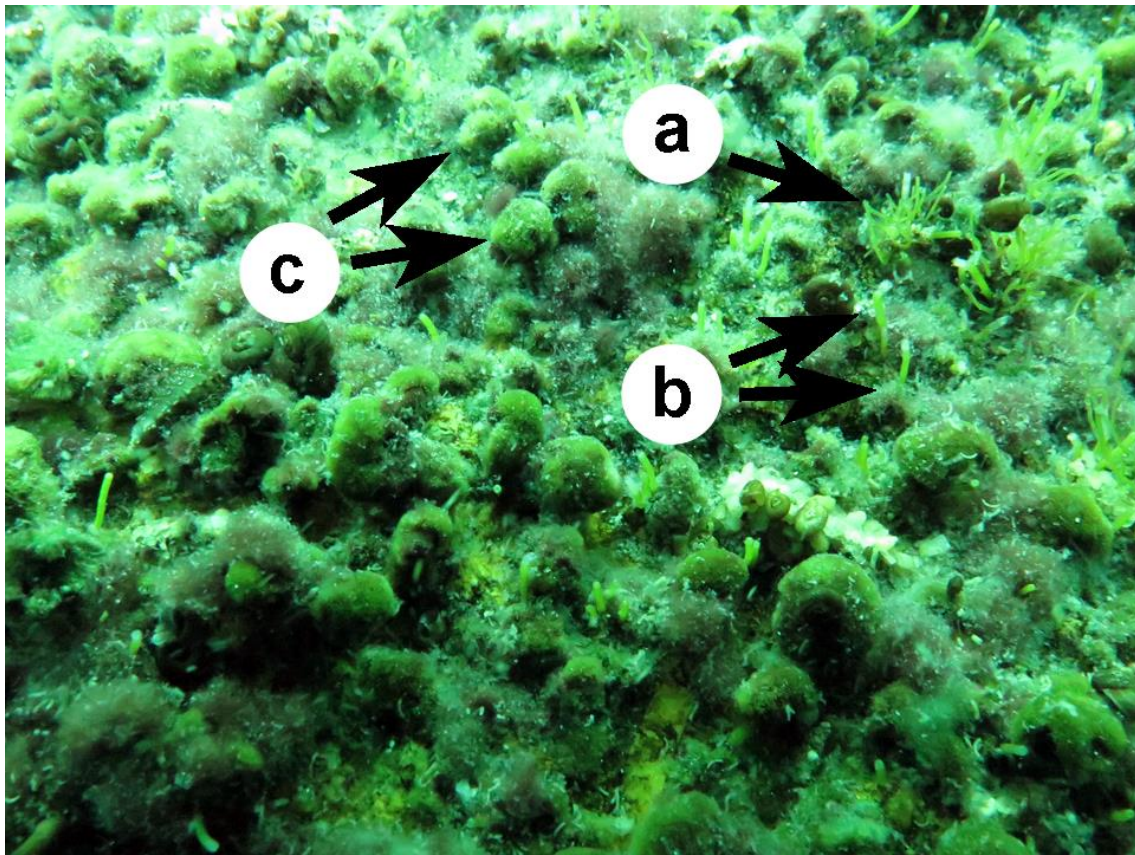
Figure 1.7. Underwater photograph showing three species inhabiting the same spot (indicated with arrows). a = *Chaetocladiella pumila*, basal branched filaments; b = *Chaetomorpha moniliformis,* unbranched filaments; c = *Cladophora compacta*, branched compact filaments forming somewhat spherical mats. Scalebar = 1 cm.

1.6 Goals of this research

The main goal of this study is to determine if the species based on morphology of the endemic Cladophorales from Lake Baikal represent a case of sympatric speciation. A multi-locus approach, using microsatellites, was implemented to recover the

population structure of the morphospecies and to detect reproductive isolation and hybridization. Polyploidy as a cause of speciation was also assessed. As there are no studies of sympatric speciation in algae, this project will provide new insights into the speciation mechanisms in algae.

The thesis is divided into three chapters, each one a step in the understanding of the group. In Chapter 2, "Microsatellite design and insights into ploidy for the Lake Baikal Cladophorales species flock", I designed a set of microsatellite markers for the study of population genetics in the endemic Baikal Cladophorales. This part of the research involved several steps such as the use of high-throughput sequencing of three species, *in silico* search and cross-validation of potential microsatellites between species, laboratory testing and amplification of the selected markers and a scoring protocol for SSRs markers from specimens of unknown life cycle phase and unknown ploidy level. Finally, it was discovered that polyploidy is occurring in the species of the endemic Baikalian Cladophorales. This chapter was submitted to *Journal of Phycology* and is currently under review (Diaz-Martinez, S., Boedeker, C. & Zuccarello, G.C. Microsatellite design and insights into ploidy for the Lake Baikal Cladophorales species flock. *J. Phycol*. Submitted). Boedeker collected the samples. Boedeker, Zuccarello and I designed the project. I conducted the methods, data collection and analyses and completed the manuscript. Boedeker and Zuccarello helped in data interpretation and improvement of the manuscript.

In Chapter 3 "Genetic data supports reproductively isolated species in the endemic Cladophorales (Chlorophyta) of Lake Baikal" I inferred the limits of most of the species of the species flock. To do this, I used microsatellites and coding protocols designed in Chapter 2 combined with cluster algorithms. A total of 727 individuals were analysed. The resulting clusters where then compared against the species based on morphology and discrete genetic units. The analyses supported reproductive isolation of five morphotaxa and two hypotheses of conspecificity. This chapter was submitted to *Journal of Phycology* and is currently under review (Diaz-Martinez, S., Boedeker, C. & Zuccarello, G.C. Genetic data supports reproductively isolated species in the endemic Cladophorales (Chlorophyta) of Lake Baikal. *J. Phycol*. Submitted). Boedeker collected the samples. Boedeker, Zuccarello and I designed the project. I conducted the methods, data collection and analyses, and completed the manuscript. Boedeker and Zuccarello helped in data interpretation and improvement of the manuscript.

In Chapter 4 "Contrasting patterns of population structure, variation and reproductive strategies of three sympatric Cladophorales species in Lake Baikal" I analysed the population structure across the lake of three putative species previously identified as unique clusters in Chapter 3. The genetic structure and diversity revealed patterns of genetic variation that suggest different reproductive strategies and dispersal patterns. This Chapter is in preparation for submission. Boedeker collected the samples. Boedeker, Zuccarello and I designed the project. I conducted the methods, data collection and analyses, and completed the manuscript. Boedeker and Zuccarello helped in data interpretation and improvement of the manuscript.

In chapter 5 "General Discussion" I provided a discussion of the results and an overview of the findings of the thesis. It is also discussed some of other biological aspects of the Baikalian Cladophorales species flock as well as other parallel research projects done in this group. Finally, new perpectives and future research pathways are described.

Chapter Two

MICROSATELLITE DESIGN AND INSIGHTS INTO PLOIDY FOR THE LAKE BAIKAL CLADOPHORALES SPECIES FLOCK

2.1 Abstract

Ancient lakes are centers of adaptive radiations and speciation. The Cladophorales endemic to ancient Lake Baikal is a morphologically diverse group nested within Rhizoclonium that may represent a case of sympatric speciation in sympatry. Recent research using ribosomal DNA markers indicates that these taxa from a monophyletic group, but was not able to resolve boundaries between all of the investigated morphospecies due to very low genetic diversity. For this reason, a population genetics approach using more variable markers was investigated. In this study, we developed a set of microsatellites (SSRs) using high-throughput sequencing (HTS) data obtained from three Cladophorales morphospecies from Lake Baikal. To increase the amplification rate of the microsatellites across taxa, we performed an *in silico* cross-validation step comparing the microsatellites retrieved from three HTS data sets and tested the most promising loci on 14 of the endemic morphospecies. We obtained 11 SSRs that cross-amplified among morphospecies, 8 SSRs in 12 taxa and 3 in only 4 taxa. Our results showed that most loci had more than two alleles, but also displayed variation between and within morphospecies. These results indicate that this group may gone through polyploidization. Polyploid systems require a different approach from standard population genetic analyses. We produced 'allelic phenotypes' (presence/absence matrices) to analyse genetic diversity. We showed that similarity indices mostly grouped morphospecies, but further work with this set of markers will aid in species delimitation and to elucidate the speciation process in this algal species flock in Lake Baikal.

2.2 Introduction

Ancient lakes are centers of great diversity which can be viewed as natural evolutionary laboratories containing examples of ecological diversification, speciation and adaptive radiations (Wilke et al. 2006, Cristescu et al. 2010). Lake Baikal in south-eastern Siberia, Russia, is more than 25 million years old and the oldest lake in the world (Mats 1993). The diversity of the lake is extraordinarily high with more than half of the animal species being endemic and considered examples of adaptive radiations (e.g. Sherbakov 1999, Yokoyama and Goto 2005, Schön and Martens 2012). However, studies on the evolution and speciation of the lakes algal flora are limited (but see Kulikovskiy et al. 2012; Boedeker et al. 2018 for examples).

One interesting group of algae are the endemic Cladophorales of Lake Baikal (Ulvophyceae, Chlorophyta). The order Cladophorales is a cosmopolitan group of algae that can be found in fresh, brackish and marine waters. The life cycle is generally haplo-diplontic and isomorphic, but asexual reproduction occurs in several species (Škaloud et al. 2018). The thalli are siphonocladous, containing multiple nuclei and chloroplasts per cell and with relatively simple morphologies consisting of uniseriate branched or unbranched filaments, but some taxa exhibit modifications of the thalli forming anastomosing networks or large spherical cells (Hoek et al. 1995, Škaloud et al. 2018). The morphological simplicity of the species within the order, phenotypic plasticity in many species and convergent evolution of characters have created confusion in the taxonomy and classification of this group. During the last decades, molecular markers have solved many of the taxonomic problems in the group as well as helping in species delimitation (i.e., Hanyuda et al. 2002, Leliaert et al. 2009, Boedeker et al. 2012, 2016, Zhu et al. 2018).

The traditional taxonomy of the endemic Baikalian Cladophorales is based on morphology (Izhboldina 2007) and reflects their wide morphological diversity with 16 distinct taxa (14 species and 2 varieties) placed in four genera: *Chaetocladiella* Meyer & Skabichevsky, which branches only near its holdfast; the unbranched *Chaetomorpha* Kützing (=*Ch.*); the abundantly branched *Cladophora* Kützing (=*Cl.*); and the diminutive *Gemmiphora* Skabichevsky (=*G.*) (Izhboldina 2007; Boedeker et al. 2018). However, recent molecular studies using nuclear ribosomal markers (SSU, LSU and ITS) have revealed that the endemic Baikalian Cladophorales are actually a monophyletic group with low genetic differentiation having a maximum of 2.8%

Table 2.1. Endemic Cladophorales morphospecies from Lake Baikal (Izhboldina 2007). The groups corresponding to the two major clades (A and B) found in Boedeker et al. (2018) using rDNA sequences. NA= not assigned.

Taxa	Group	Analyzed in this study
Chaetocladiella litoralis (Skabichevsky) Meyer & Skabichevsky	NA	X
Chaetocladiella microscopica (Meyer) Meyer	NA	
Chaetocladiella pumila (Meyer) Meyer & Skabichevsky	B	X
Chaetomorpha baicalensis Meyer	B	
Chaetomorpha curta (Skabichevsky) Skabichevsky	A	X
Chaetomorpha moniliformis Skabichevsky	A	X
Chaetomorpha solitaria Skabichevsky	NA	
Cladophora kursanovii Skabichevsky	A	X
Cladophora compacta (Meyer) Meyer ex Hollerbach	B	X
Cladophora floccosa Meyer var. *floccosa*	B	X
Cladophora floccosa Meyer var. *irregularis* Skabichevsky	B	X
Cladophora globulus (Meyer) Meyer	B	X
Cladophora meyeri Skabichevsky var. *meyeri*	B	X
Cladophora meyeri Skabichevsky var. *gracilior* (Meyer) Hollerbach	B	X
Cladophora pulvinata (Meyer) Meyer	B	X
Gemmiphora compacta Skabichevsky	B	X

pairwise differences in ITS region (Boedeker et al. 2018). These phylogenies also showed that this clade is nested within *Rhizoclonium*, a genus which shows little morphological differentiation (simple unbranched or sparsely branched filaments), but high genetic diversity (Boedeker et al. 2018). The phylogenetic signal of the ribosomal markers differentiated the Baikalian Cladophorales into two major clades (A and B; Table 2.1), but lacked support for many other clades and showed little or no genetic difference between some morphospecies.

These results suggest that this group fulfils many of the criteria of a species flock such as: monophyly, high endemism and morphological diversity (Greenwood 1984), but also questions whether the morphospecies are reproductive isolated and whether this group represents a case of sympatric speciation. To address these questions, a population genetics approach using more variable markers is needed. Among the markers available for population genetics, microsatellites have been considered a good option to assess the genetic diversity and population structure of

animals, plants and algae (Selkoe and Toonen 2006). Microsatellites, also referred to as simple sequence repeats (SSRs), consist of repeated motifs of 2 to 6 nucleotides that occur in all eukaryotic genomes. Although other genetic markers are used in population genetics (e.g. amplified fragment length polymorphisms, random amplified polymorphic DNAs, inter simple sequence repeats and single nucleotide polymorphisms), SSRs have several advantages that make them ideal for the study of population genetics, such as co-dominance (enabling differentiation of homozygotes and heterozygotes), putative neutrality, high levels of polymorphism, plus they can be applied to small and preserved tissues (Selkoe and Toonen 2006, Guichoux et al. 2011). In addition, technologies of high-throughput sequencing (HTS) and high computational power provide an easy way to obtain many potential SSRs for a low cost and short development time while exploring large portions of an organism's genome (Gardner et al. 2011, Schoebel et al. 2013).

The number of studies using SSRs in algae has increased over the last years, showing their reliability to assess many questions, such as: identification of individuals belonging to a particular population (Henrichs et al. 2013, Krueger-Hadfield et al. 2017); assessment of population structure (Krueger-Hadfield et al. 2011, Kostamo et al. 2012, Sjøtun et al. 2017); analysis of hybridization (Zardi et al. 2011, Montecinos et al. 2017); and life cycle dynamics (Guillemin et al. 2008, Couceiro et al. 2015, Krueger-Hadfield et al. 2016). However, only a few studies have used SSRs in the Cladophorales to date. These are the investigations of the population structure of *Cladophoropsis membranacea* (Hofman Bang *ex* C.Agardh) Børgesen (Van der Strate et al. 2000, 2002a, 2002b, 2003)revealing-fine scale structure and even possible cryptic species at large scale populations.

One of the main difficulties in the study of population genetics in algae is the occurrence of isomorphic haplo-diplontic life cycles (Valero et al. 2001). In some cases, it is possible to overcome this by distinguishing reproductive structures, measuring nuclear DNA content, use chemical tests, or counting the total number of alleles in each locus (Van der Strate et al. 2002b, Engel et al. 2004, Guillemin et al. 2008, Krueger-Hadfield et al. 2011, Couceiro et al. 2015). Species of the order Cladophorales mostly have an isomorphic alternation of generations (Hoek et al. 1995) and reproductive structures are similar between the sporophytic and gametophytic phases, making it impossible to distinguish morphologically between haploids and diploids. Another

25

potential challenge is the occurrence of polyploidy which has been reported to occur frequently (Hinson and Kapraun 1991, Kapraun 2007, Boedeker et al. 2012), making it even more difficult to determine the actual ploidy level of individuals.

The aim of this is study was to design a set of microsatellite markers that cross-amplifies across the Cladophorales species flock of Lake Baikal. The SSRs were developed using HTS data of three different morphospecies. We calculated statistics and similarity clusters to provide a brief overview of their diversity and genetic signal. We also explored the genetic variationof these SSRs as they will be used to address evolutionary questions of this group (e.g., species boundaries and population structure).

2.3 Materials and Methods

2.3.1 *HTS data and assembly*

Total DNA was extracted from samples of three morphospecies: *Cl. kursanovii*, *Ch. baicalensis* and *Ch. moniliformis* from Lake Baikal collected by C. Boedeker during September 2014 (collection numbers 475, 50 and 515, respectively). Many individuals from the same locality were combined to get sufficient DNA for HTS. These species were selected for the following reasons: ease in identification based on morphology; each morphospecies are likely to belong to a genetically distinct lineage; are species of both clades A and B from Boedeker et al. (2018). DNA extractions were done using a CTAB lysis protocol (Zuccarello and Lokhorst 2005). A minimum of 50 ng of DNA were sequenced (Novogene, Beijing, China) following the protocols suggested by the company. The respective genetic libraries for HTS were created by digesting the genomic DNA with the restriction enzyme *EcoR*I and sequenced in an Illumina HiSeq PE250 sequencer. The resulting reads were cleaned by removing adapters and low quality reads. Finally, the reads of the three morphospecies were assembled independently using Velvet 1.2.07 (Zerbino and Birney 2008) and Velvetopt 2.2.5 with the parameters starting with k-mers of 23 to 31 base pairs with subsequent increments of 3 base pairs per K-mer (–s 23 –e 31 –x 3).

2.3.2 *Microsatellite detection and primers development*

A first search for SSRs was done using all assembled sequences. We used MSATCOMMANDER (Faircloth 2008) to create a list of potential SSRs including their sequence, microsatellite motif type and the respective forward and reverse primers. This helped us to determine the frequency and motif types of SSRs in the genomes.

In order to increase the probability of amplifying the same loci between Baikalian Cladophorales morphospecies, we performed a cross-validation protocol *in silico* using the sequences from the three samples. All the sequences listed by MSATCOMMANDER as containing a microsatellite region were imported into Geneious R8.1 (www.geneious.com, Kearse et al. 2012) and assembled using the *de novo* option at High Sensitivity/Medium. All contigs were examined by looking for those containing the same sequences in at least two of the three datasets/morphospecies. Contigs of less than 100 bp were discarded. New primers were designed using Primer3 2.3.7 plugin (Rozen and Skaletsky 2000) favouring conserved flanking regions.

2.3.3 *Microsatellite testing*

A total of 44 samples from 14 endemic Baikalian Cladophorales morphospecies plus several ambiguously identified individuals and two *Rhizoclonium* sp. samples were tested (Appendix 2.1). The samples were collected from different localities around Lake Baikal by Boedeker (Appendices 2.2 and 2.3). DNA was extracted using a Chelex method for its fast processing time (Goff and Moon 1993). For specimens forming entangled mats each individual was examined in a stereoscope and parts of the thallus arising from the same basal portion were used. For unbranched filamentous forms only a single filament was used.

The primers were tested in two steps, first to check if the primers produced fragments of expected size, and second, to check if the loci were polymorphic. First, a standard PCR was conducted using the protocol: $94°C$ for 4 min for initial denaturation, followed by 30 cycles of 1 min at $94°C$, 1 min at $56°C$ and 1 min at $72°C$ and a final extension at $72°C$ for 10 min. The PCR mix had a volume of 20 µL and consisted of 1 µL of diluted genomic DNA (1:10), 200 µM of each dNTP, 0.4 µM of each primer, 1.5 mM of $MgCl_2$, 0.05% BSA, 1X reaction buffer and 0.7 units of BIOTAQ *Taq*

polymerase (Bioline, UK). The PCR products were electrophoresed in 1.5% agarose and post-stained with ethidium bromide to check for products.

All primers that produced amplicons were used for the next step to check fragment read quality. A three primer method for fluorescently labelling fragments (Schuelke 2000) was used for almost all the primer combinations, except for loci 5a and 6. The method consists on the addition of a third primer with the M13 (5'-TGT AAA ACG ACG GCCAGT-3') sequence ligated on its 5' end to one of four fluorescent dyes (6FAM, PET, VIC or NED) and a modified microsatellite forward primer with a M13 tail at its 5' end. A two cycle PCR program used consisted of: 94°C for 4 min for initial denaturation, followed by 25 cycles of 30 sec at 94°C, 30 sec at 56°C and 120 sec at 65°C for extension; followed by 8 cycles at 94°C for 30 sec, 30 sec at 53°C and 120 sec at 65°C for extension, followed by a final extension at 65°C for 15 min. The extension temperature was modified to reduce the occurrence of stutters (Seo et al. 2014). The PCR mix had a volume of 15µL and consisted of 1 µL of diluted genomic DNA, 200 µM of each dNTP, 0.4 µM of each M13 labelled and reverse primers, 0.1 µM of forward primer, 1.5 mM of $MgCl_2$, 0.05 % BSA, 1X reaction buffer and 0.7 units of *Taq* polymerase. For loci 5a and 6 the forward primers were directly labelled with FAM and PET respectively. A standard PCR program and mix concentrations was followed. We performed repeated PCRs on random samples using DNA polymerase with proofreading exonuclease activity (RANGER DNA Polymerase, Bioline, UK), which has an increased fidelity, to compare and confirm the chromatogram profiles of the selected loci (i.e. 5a and 6) and check if noticeable differences occurred between repeats and polymerases.

The labelled PCR products were combined with up to four different loci per vial ('poolplexed') and analysed commercially (Macrogen Inc., Seoul, Korea) on an ABI3730XL Genetic Analyzer (Life Technologies Corp.) with GeneScan 500 LIZ as a size standard. The resulting chromatograms were analysed and scored in GeneMarker 2.0.2 (SoftGenetics, LLC, State College, PA) using the default parameters for SSRs.

Each chromatogram was reviewed manually and compared with chromatograms of the same and different morphospecies, looking for repeated profiles and fragment sizes to assign and correct the alleles. In difficult profiles showing stutters the allele scoring followed a similar method to the manual routine of Pfeiffer et al. (2011) with slight modifications. This approach is considered reliable over automated scoring to

reduce noise in difficult loci due to polyploidy (Pfeiffer et al. 2011b, Cărăbuș et al. 2015, Bhandawat et al. 2016, Könyves et al. 2016). Therefore, the peaks in the chromatograms were considered as alleles depending on their putative quality. For this study, the highest main peaks were coded as alleles and peaks directly before or after a main peak with no less than 85% of the height of the main peak were considered different alleles. Finally an increase of intensity after a putative allele was coded as a different allele, as it is assumed that the intensity of stutters decreases after the "true" allele and stutter bands tend to occur before the "real" peaks. Context information, such as morphospecies and locality, was also considered, for example alleles that were found multiple times in the same morphospecies or location. The occurrence of chromatograms showing identical peaks and profiles in different samples was also used to corroborate the allele scoring.

All the alleles were scored as 'allelic phenotypes'. This method consists of coding each allele independently in a presence/absence matrix, similar to DNA fingerprints (Besnard et al. 2008, Andreakis et al. 2009, Pfeiffer et al. 2011b, García-Verdugo et al. 2013, Dufresne et al. 2014). For each locus the total number of alleles, the maximum and minimum number of alleles per individual was calculated for each locus, as well as the number of distinct multi-locus genotypes. Other genetic diversity statistics at the morphospecies level were calculated in GenAlEx 6.5 (Peakall and Smouse 2012) such as the number of alleles, the mean value of unbiased diversity (uh) and the number of private alleles (Pa). The allelic phenotypes of all the loci were combined in a single matrix. The information contained in the SSR markers was explored by creating an UPGMA tree considering all the genotypes independently and using a Jaccard similarity index in NTSYSpc 2.1 (Applied Biostatistics, Inc.).

2.4 Results

2.4.1 *HTS data*

The HTS retrieved more than 16 million reads in each of the three samples. Other genomic features are reported in Appendix 2.4. The total cleaned reads amounted to 10.75 gigabases (Gb) in *Ch. baicalensis*, 10.1 Gb in *C. moniliformis* and 9.35 Gb in *Cl. kursanovii.* The total numbers of contigs obtained after the assembly were considerably

different between species: 135,578 in *Ch. baicalensis,* 318,757 in *Ch. moniliformis and* 105,464 in *Cl. kursanovii*. The number of microsatellites found differed between species and appears to have no relation to the amount of contigs generated. In *Ch. baicalensis* we found 1,950 SSRs, in *Ch. moniliformis* 4,524 and in *Cl. kursanovii* 4,250. Three base pair repeat motifs were dominant in all species, followed by dinucleotides repeats (Table 2.2).

Table 2.2. Total number of SSRs obtained in the three species HTS datasets. Group designation from Boedeker et al. (2018) is indicated.

Motif repeat	*Chaetomorpha baicalensis* (Group B) No. of SSRs	*Chaetomorpha moniliformis* (Group A) No. of SSRs	*Cladophora kursanovii* (Group A) No. of SSRs
2	131	392	265
3	1735	3859	3367
4	49	168	406
5	21	85	191
6	14	20	21
Total	1950	4524	4250

2.4.2 *Microsatellite cross-amplification*

After filtering dubious and short assemblies, 52 SSRs were cross-validated in at least two data sets. To complement the potential markers, another 24 SSRs only present in one data set were selected, making a final SSRs list of 76 potential markers. From the total potential markers, 38 SSRs amplified in all morphospecies, 18 amplified only in morphospecies from group A and 20 SSRs did not produce amplicons.

From the total of 56 SSRs (38 that amplified in all samples and 18 only in group A) 22 SSRs produced unusable chromatograms and 23 SSRs were not polymorphic. The remaining 11 SSRs (Table 2.3) showed variation and were used with all 44 individuals exploring their variability and characteristics. Three of these loci are markers that only amplified in morphospecies of group A (Table 2.3). Unfortunately the samples of *Chaetocladiella* (=C.) *pumila* and *C. litoralis* did not produce good amplicons and were not included in the subsequent analyses. Although the primers were

successful in most of the samples, some individuals did not produce amplicons at loci 14 (4 individuals), 34 (20 individuals) and 515-35 (6 individuals). These loci were not discarded as they appear to be informative for other morphospecies, particularly those from the group A.

2.4.3 *Microsatellite diversity*

The scoring revealed consistent patterns of multiple alleles per locus per individual indicating polyploidy (Fig. 2.1). All loci had more than two alleles, except for locus 515-64 (Table 2.3). The number of alleles per individual was up to 7 in loci 14 and 34. The total number of alleles found in all 11 loci combined was 103 where the total number of alleles for each morphospecies ranged from 19 to 37 (Table 2.4). The unbiased diversity value (*uh*) had a mean value of 0.136 and a total of 30 private alleles (*Pa*) was detected (Table 2.4). All the samples had a unique genotype differing in at least one allele. It is worth to note that the chromatogram profiles of some loci were identical between samples of the same morphospecies and the fragments amplified with proof-reading polymerase did not show significant differences compared to those obtained with the standard *Taq* polymerase. Therefore, these chromatograms were reproducible and the profiles obtained from both types of polymerase were considered to be equally informative.
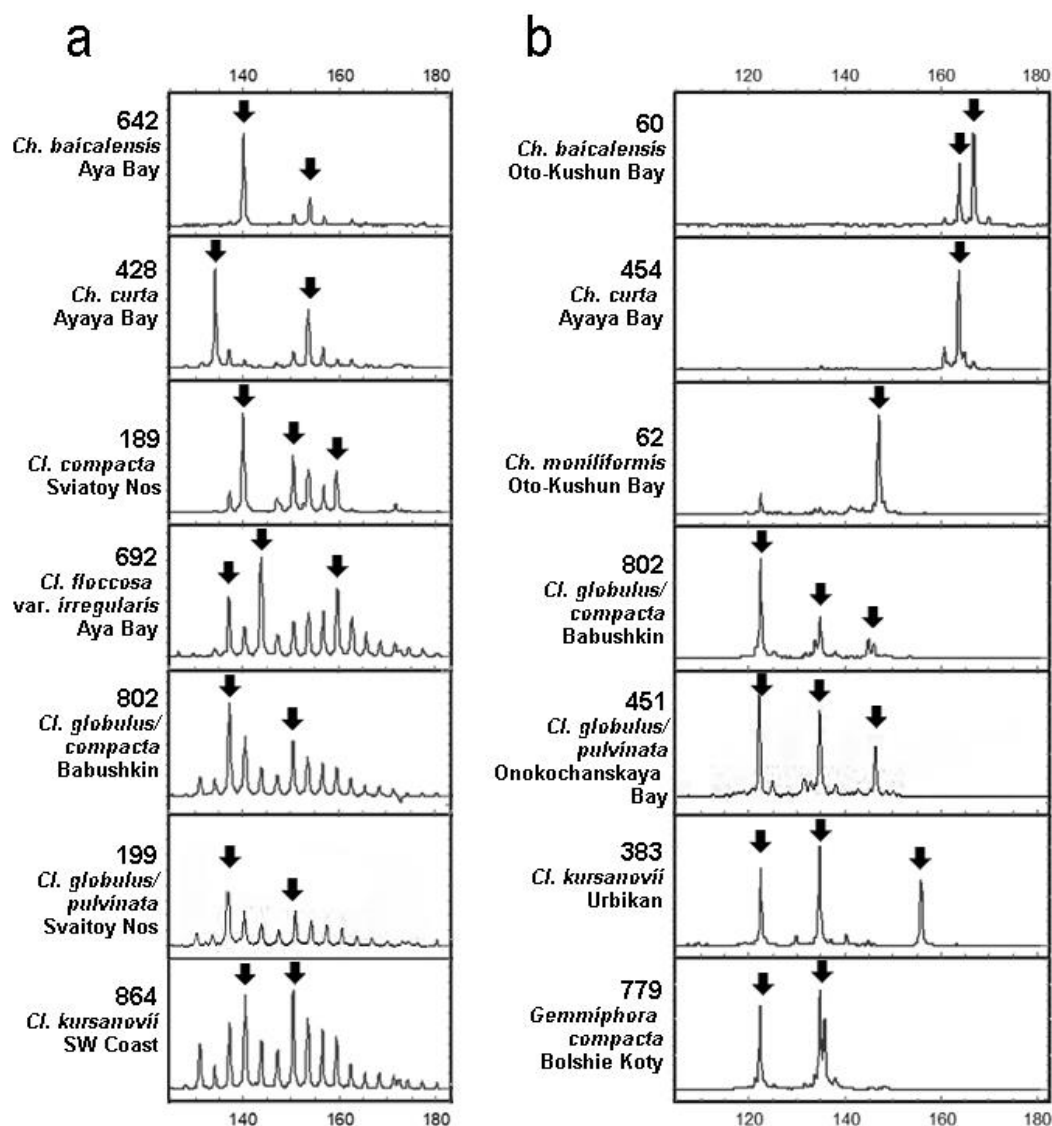
Figure 2.1. Chromatograms of several samples at locus 5a (a) and 16 (b) showing the alleles scored. Arrows indicate alleles, based on peak height and position. Numbers as sample number for different morphospecies (Appendix 2.1).

Table 2.3. Primer sequence and characteristics of the selected 11 SSRs markers. * = indicates the presence of potential null alleles. The ability to cross amplify in all samples only species in clade A (Boedeker et al. 2018) is shown.

Locus name	Primer sequence 5'-3'	Motif	Total number of alleles	Allele range	Alleles per individual	Mean number of alleles	Labeling strategy	Cross-amplification
5a	F: GTTGTCTCTGATCCGTTCTATTGCT R: CTTCGGCTACCAAGTAACCAAGAA	TAC	9	112-148	1 to 4	2.55	Direct fwd. primer	All
6	F: CGTTGATCATTATCTTGAGCACG R: ATAGCTACTTGTACAGCGGTTAT	TA	9	152-182	1 to 6	2.84	Direct fwd. primer	All
14	F: AGCTACTTTGTGTCATAACGGC R: GGTGTCTGTCCCTATCCAGG	ATA	19	155-188	0* to 7	2.24	M13 fwd. primer	All
15	F: GCACTTTATACTGCCTGTTATCG R: AATAACTCTCGTTTATAATGGCCTGG	AAT	4	151-166	1 to 4	2.50	M13 fwd. primer	Only in A
16	F: ATACCCCGCCAGAAGAGTGC R: ACGGCCACAAAACTACCTAAACT	GTA	17	117-168	1 to 5	2.81	M13 fwd. primer	All
34	F: CATTCGTATACAACAACTCCCAG R: GGTGAGTCTTCCACTCTGACTCT	TCA	13	120-168	0* to 7	1.44	M13 fwd. primer	All
46	F: AATAAACTCGTGGAGGTATGCAA R: GTAAAAAGCGCACAAGGAACTT	TA	11	130-162	1 to 4	1.89	M13 fwd. primer	All
50-15	F: GCCAGAAATGGCCAATCG R: AGGCCAGCATGAGGTACAG	TTA	12	191-257	1 to 3	2.44	M13 fwd. primer	All
515-35	F: ATCGTTGGTAGAGCGCTGG R: AGCTTGCAGAAAGGTGCG	ATT	3	242-287	1 to 3	1.33	M13 fwd. primer	Only in A
515-46	F: ATGCTTCATTGGCCCTTGG R: GGTACAACAGGGACTTGGC	TTGT	3	200-224	1 to 3	2.33	M13 fwd. primer	Only in A
515-64	F: ACCTCGCTATTGGTTTGGC R: TGACCGTATGGACCTTGGC	TAT	3	271-283	1 to 2	1.58	M13 fwd. primer	All

Table 2.4. Characteristics and genetic diversity of the alleles for each morphospecies of the endemic Baikalian Cladophorales. Total number of individuals (N); total number of alleles; mean value of unbiased diversity (uh); and the number of private alleles (Pa) are shown. * = mean value.

Taxa	N	No. of alleles	uh	Pa
Ch. baicalensis	3	27	0.091	3
Ch. curta	3	30	0.123	2
Ch. moniliformis	3	19	0.026	0
Cl. compacta	4	37	0.172	4
Cl. floccosa var. *floccosa*	3	24	0.097	0
Cl. floccosa var. *irregularis*	2	31	0.233	3
Cl. globulus/ Cl. compacta	2	36	0.252	1
Cl. globulus/ Cl. pulvinata	2	28	0.136	2
Cl. kursanovii	4	36	0.107	8
Cl. meyeri var. *gracilior*	2	34	0.272	2
Cl. meyeri var. *Meyeri*	3	31	0.175	4
Cl. pulvinata	3	25	0.071	0
G. compacta	2	25	0.087	1
Rhizoclonium sp.	2	28	0.068	0
Total	38	103	0.136*	30

The Jaccard (J) pairwise distances among all individuals retrieved similarity values ranging from 0.04 to 0.94 with a mean of 0.26. The similarity tree built using this index (Fig. 2.2) shows the grouping of many individuals belonging to the same morphospecies: *Ch. baicalensis* (J=0.5); the attached form of *Ch. curta* (J=0.61); *Cl. kursanovii* (J=0.55); *Cl. pulvinata* (J=0.67); *G. compacta* (J=0.64); and *Rhizoclonium* sp. (J=0.75). *Ch. moniliformis* was included in a group with high similarity (J=0.84), however this also included a sample of the unattached form of *Ch. curta* 638. The remaining individuals, with overall low levels of similarity (J<50), were not grouped in exclusive branches according to their respective morphospecies.
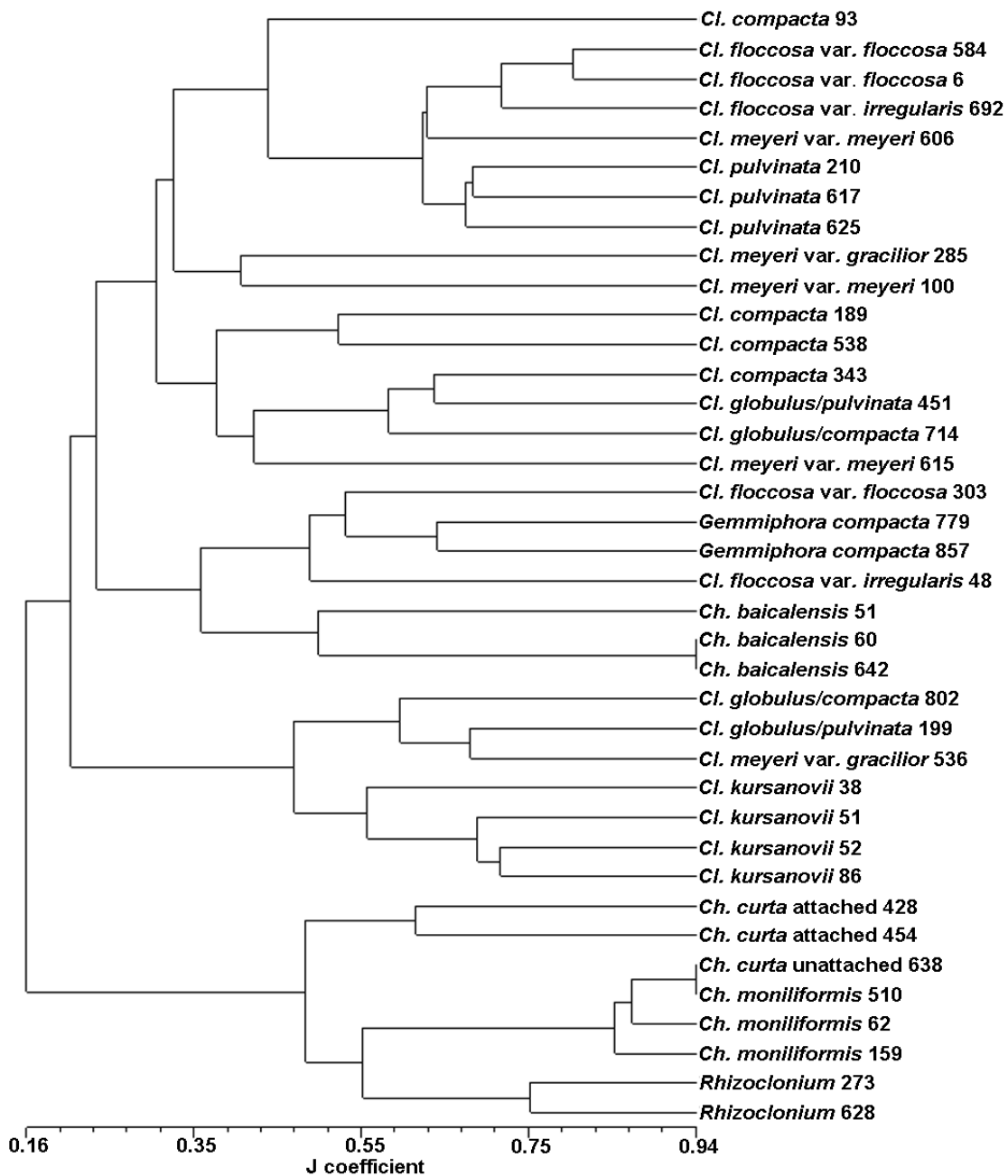
Figure 2.2. UPGMA tree based on the Jaccard similarity coefficient from allelic phenotypes of 11 loci. Full taxon full names in Table 2.1. Sample numbers in Appendix 2.1

2.5 Discussion

The HTS approach for the development of SSRs in non-model organisms allowed us to obtain a set of microsatellites that cross-amplify between morphospecies. These results showed that the markers designed displayed variation that can address evolutionary questions of these endemic algae such as their population structure and species delimitation of the morphotaxa. Microsatellites also revealed that the Baikalian Cladophorales species flock consist of a clade that could be shaped by polyploidization events and therefore needs a different approach to analyse its diversity. In addition, some preliminary insights into the partial genome of the Baikalian Cladophorales were obtained.

The total number of discovered SSRs differed between morphospecies. Although *Ch. baicalensis* and *Cl. kursanovii* had similar number of contigs, the number of SSRs in *Ch. baicalensis* was nearly 60% less (Table 2.2). The dominance of trinucleotide SSRs in Baikalian Cladophorales appears to be similar to other green algae, such as *Chlamydomonas communis* and *Volvox reinhardtii* (Zhao et al. 2014), and the brown alga *Saccharina japonica* (Li et al. 2015; Zhang et al. 2016). In contrast, dinucleotide repeats are reported as dominant in the green alga *Ulva prolifera* (Ulvophyceae, Li et al. 2016), other red algae (Couceiro et al. 2011a, Ayres-Ostrock et al. 2016) and land plants (see Schoebel et al. 2013). The reason for trinucleotide SSRs dominance in Baikalian Cladophorales is unclear.

The *in silico* cross-validation process allowed us to detect and design markers that cross-amplified rate among the Baikalian Cladophorales, either in all morphospecies or in species of the group A. However, only 11 were selected for the final tests as many were not polymorphic. The number of non-polymorphic markers (23) and the high proportion of primers that produced amplicons between Baikalian Cladophorales species may be the result of the low genetic variation between these species in spite of their morphological differences, possibly due to a recent radiation as has been suggested (Boedeker et al. 2018).

The results provided an overview of morphospecies genetic diversity and differentiation. The level of genetic variation within morphospecies was different in each one. This is observed in both the UPGMA tree (Fig. 2.2) and the unbiased diversity values (Table 2.4). However, these results should not be considered conclusive until

more data is gathered and more individuals per species and per populations are included.

The high number of alleles per locus indicates that polyploidy occurs in this group. Although this could be attributed to duplicated loci across the genome, the occurrence of more than two alleles in different loci indicates that polyploidy is most likely, as observed in other studies (Andreakis et al. 2009, Varela-Álvarez et al. 2017a). Contamination of multiple individuals in DNA extracts can also be discarded as some of the loci with more thatn two alleles occurred in extractions where crealy a single filament was used, such as in *Chaetomorpha* individuals. The occurrence of repeated profiles in different samples of the same species suggests that random amplification of fragments is also unlikely.

Polyploidy complicates the use of standard statistics for population genetic analysis (e.g., F_{st}, expected heterozygosity). In cases of low ploidy levels, such as in tetraploids, it is possible to determine the allelic dosage using the peak ratios of chromatograms (Esselink et al. 2004) or by computational algorithms (e.g. Van Puyvelde et al. 2010; Clark & Jasieniuk 2011). However, these methods require previous knowledge of the ploidy level of the samples (i.e. chromosome counts, or total DNA content per nucleus) and knowledge of whether polyploidy is due to alloploidy (product of hybridization between different species) or autoploidy (genome doubling) in order to sort allele dosage correctly (Dufresne et al. 2014).

Coding the SSRs as dominant data in 'allelic phenotypes' is a simple way for dealing with polyploids as no complex assumptions are made about the data (Cidade et al. 2013) and has the advantage of allowing the inclusion of organisms with different ploidy level in the same data set. Although this reduces the potential informative power of SSRs, it is a convenient solution when the ploidy number is unknown and when the sorting of alleles cannot be resolved (Obbard et al. 2006; Pfeiffer et al. 2011; García-Verdugo et al. 2013). In addition, some other properties of the population such as the total number of alleles and number of private alleles within populations are not affected (López-Vinyallonga et al. 2015). This approach proved to be informative for Baikalian Cladophorales as our analyses showed diversity within the microsatellite markers and the number of private alleles confirms that these markers have enough power to discriminate between most morphospecies regardless of the unknown ploidy level. Likewise, the J-similarity tree, which considers all genotypes independently, retrieved

clusters that in many cases were consistent with morphospecies (e.g., *Cl. curta*, *Cl. kursanovii*, *Cl. pulvinata*, *Ch. baicalensis* and *Ch. moniliformis*). These results are similar to the phylogenetic trees based on ribosomal markers (Boedeker et al. 2018). However, broader sampling and a more in-depth analysis are required to perform proper species delimitation and test for reproductive isolation of species.

Polyploidy could be playing an important role in the evolution of the Baikalian Cladophorales species flock. Polyploid speciation occurs when the progeny are reproductively isolated from their parents ('instant speciation') (Schluter, 2001; Albert and Schluter, 2005). Furthermore, polyploidy often results in phenotypic changes and life cycle modifications (Comai 2005, Soltis et al. 2010). Therefore, this process could explain some of the wide morphological diversity found in this group.

In conclusion, the HTS data combined with the *in silico* cross-validation procedure allowed us to explore the genome of these taxa and design a set of cross-amplifying SSRs for the study of the population genetics and reproductive isolation of the endemic Baikalian Cladophorales, a group recently discovered to represent a species flock. The markers designed here proved to be promising for the study of reproductive isolation and genetic differentiation as information was sufficient to cluster several of the morphospecies. Finally, our data revealed evidence for polyploidy, a process that could have shaped the evolution of this group.

2.6 Acknowledgments

Chapter Three

GENETIC DATA SUPPORTS REPRODUCTIVELY ISOLATED SPECIES IN THE ENDEMIC CLADOPHORALES (CHLOROPYTA) OF LAKE BAIKAL

Diaz-Martinez, S., Boedeker, C. & Zuccarello, G.C. Genetic data supports reproductively isolated species in the endemic Cladophorales (Chlorophyta) of Lake Baikal. *J. Phycol.* Submitted.

3.1 Abstract

Ancient lakes are centers of biological diversification that hold many examples of adaptive radiation and species flocks. The recently discovered Cladophorales species flock in Lake Baikal is a group of algae that exhibit low genetic divergence in ribosomal markers (LSU, SSU and ITS), but wide morphological differentiation. Microsatellite markers show evidence of polyploidy in this group. In this study, we use two clustering methods (STRUCTURE and Gaussian clustering) to delineate species within 15 distinct morphotaxa of the cladophoralean Baikal clade. The cluster analyses produced comparable results, although subtle differences in the assignment of individuals were observed. Our results indicate that many morphologically distinguishable species are discrete genetic clusters supporting reproductive isolation. This is the case for *Chaetomorpha* (*Ch.*) *baicalensis*, the attached form of *Ch. curta*, *Ch. moniliformis*, *Cladophora compacta* and *Cladophora kursanovii*. The unattached form of *Ch. curta* and a species of *Rhizoclonium* are recovered as growth forms of *Ch. moniliformis* and the attached form of *Ch. curta*, respectively. The remaining morphotaxa were not clearly delimited. While we have evidence for polyploidy within this species flock, it was not possible to determine the ploidy level of each individual with accuracy as no correlation in the number of alleles was observed between loci. A more detailed study including other sources of data, such as nuclear DNA content or chromosome counts, is required to demonstrate the ploidy changes and their role in speciation in these species.

3.2 Introduction

Ancient lakes, lakes older than 1 million years (my), are well known centers of biological diversification that contain some of the most remarkable examples of endemism, speciation, species flocks and adaptive radiations (Wilke et al. 2006, Cristescu et al. 2010). Lake Baikal, located in south east Siberia, Russia, is the oldest lake in the world (about 30 my)(Mats 1993). Like other ancient lakes, it is considered an evolutionary laboratory due to its unique biological diversity. The fauna contains approximately 2600 species, more than half are endemic and many taxonomic groups are considered the products of adaptive radiations (Sherbakov 1999, Salzburger and Meyer 2004, Barluenga et al. 2006, Seehausen 2006, Wilke et al. 2006, Herder et al. 2008, Fitzpatrick et al. 2009, Keller et al. 2013). However, the algal flora of the lake has not been extensively studied (Boedeker et al. 2018). The Baikalian Cladophrales (Ulvophyceae, Chlorophyta) species flock is a group of freshwater green algae that includes taxa endemic to Lake Baikal. The taxonomy of the species is based on morphology recognizing 16 distinct taxa (4 genera, 14 species and 2 varieties) with a morphological range as broad as the entire Cladophorales (Izhboldina 2007; Boedeker et al. 2018). However, recent studies using nuclear ribosomal markers (SSU, LSU and ITS) have revealed that the mostly endemic Baikalian Cladophorales are a monophyletic group ('Baikal clade') with low levels of genetic differentiation, nested within the genus *Rhizoclonium* (Boedeker et al. 2018). This was a striking result based on the group's morphological diversity. Phylogenetic analysis with ribosomal markers resolved the Baikal clade into two major groups (A and B, Appendix 3.1; Boedeker et al. 2018), but were not able to support most of the morphospecies.

The low genetic differentiation based on rDNA markers, the wide morphological differentiation between the taxa and the co-occurrence of many taxa in the same localities across the lake suggest that this group could be a species a flock that has undergone sympatric speciation (Boedeker et al. 2018). In addition, the recent development of SSRs indicates that polyploidy is also likely in these taxa (Chapter 2). Therefore, the Baikalian Cladophorales clade is an interesting group for the study of speciation in green algae. However, a better understanding of the species limits is required to understand the underlying evolutionary processes (i.e. possible sympatric speciation).

Species delimitation is crucial for many aspects of biology. Although there is still debate about the definition of species and how they should be delimited, there is a general agreement that species are independently evolving lineages (De Queiroz 2007, Wiens 2007). This allows for a set of operational criteria to support and delimit species, such as reproductive isolation, phylogenetic patterns and genetic clustering (De Queiroz 2007). The use of these criteria, combined with new methods, have improved our understanding in recognizing species and how to delimit them (Hausdorf and Hennig 2010, Jombart et al. 2010, Puillandre et al. 2012, Carstens et al. 2013, Fujisawa and Barraclough 2013, Zhang et al. 2013).

Many of the current phylogenetic species delimitation methods are based on DNA sequences that are used in Barcode Gap determination (Puillandre et al. 2012), coalescent-based groupings (Fujisawa and Barraclough 2013), or tree-based methods (Zhang et al. 2013). These approaches provide a good overview of species diversification and are widely used in taxonomic assessments and diversity exploration (e.g., Muangmai et al. 2014, Díaz-Martínez et al. 2016). Although these methods are useful they are not free of caveats. Problems can arise when the patterns recovered do not reflect the speciation process, for example due to hybridization or incomplete lineage sorting, or when the phylogenetic signal of the selected genes is too low due to low mutation rates or recent divergence (Carstens et al. 2013, Edwards and Knowles 2014).

Delimitation methods based on population genetics, on the other hand, can provide more accurate species delimitation in cases where sequence DNA methods have failed (Shaffer and Thomson 2007, Hausdorf and Hennig 2010, Carstens et al. 2013). Compared to DNA-sequence based methods, this approach requires more extensive sampling and more markers to capture the genetic diversity of the populations, but can reveal evidence of reproductive isolation by clustering genetically similar genotypes that have evolved independently with little or no ongoing gene flow (Hausdorf and Hennig 2010). The most commonly used methods in population genetics include clustering algorithms using model-free multivariate analysis such as Gaussian clustering or discriminant analysis of principal components (Hausdorf and Hennig 2010, Jombart et al. 2010), or explicit models like STRUCTURE or Structurama (Pritchard et al. 2000, Huelsenbeck et al. 2011) using different types of genetic markers such as SNPs and SSRs.

SSRs are repetitive short DNA sequences (about 2-6 base pairs) scattered in the genome of many organisms. Although other markers such as single nucleotide polymorphism (SNPs) have gained popularity due to the recent advances in sequencing technology (Nielsen et al. 2011, Defaveri et al. 2013), easy genotyping and low homoplasy levels (Shaffer and Thomson 2007), some of the advantages of using SSRs over other markers are their codominance, neutrality and high levels of polymorphism (Selkoe and Toonen 2006). In addition, advances in next-generation sequencing makes the obtaining of a large number of potential SSRs rapid and cost effective (Schoebel et al. 2013). Being based on polymerase chain reaction techniques, SSRs are especially useful when biological material is poorly preserved or in small quantities.

In this study, we explored the species boundaries of the Baikalian Cladophorales species flock using genetic clustering analyses for species delimitation and putatively neutral SSR markers. Finally, we determined whether the genetic clusters are consistent with the currently recognized morphospecies *sensu* Izhboldina (2007) and the results obtained by Boedeker et al. (2018).

3.3 Materials and Methods

3.3.1 *Taxon sampling*

A total of 727 samples of 15 taxa were analyzed (Appendix 3.2). The initial identification was done using morphological characters (Izhboldina 2007, Boedeker et al. 2018; see Appendix 3.3). The samples included: two distinct morphotypes of *Ch. curta* (attached and unattached form) and two dubious identifications of *Cl. globulus/Cl. compacta* and *Cl. globulus/Cl. pulvinata*. Samples of *Chaetocladiella pumila* and *Chaetocladiella litoralis* were not included as the samples failed amplification (Chapter 2). In this work we considered each taxon and morphotypes independently as a preliminary morphotaxa, making a total of 15 distinct units (Table 3.1). The samples were collected in 20 localities around Lake Baikal during September 2014 by C. Boedeker using scuba diving (Fig. 3.1; Appendix 3.4). A target number of 10 individuals per morphotaxon per locality were collected and dried in silica gel for DNA extractions.

3.3.2 *DNA extractions and SSRs amplification*

Total genomic DNA was extracted using the Chelex protocol (Goff and Moon 1993) for its short time required and to avoid losing material during the extraction. The primers and amplification conditions followed Chapter 2. Eleven SSR markers were used in samples of morphospecies belonging to clade A, while only nine were used in samples in clade B similar to previous results (Chapter 2).

The PCR products were mixed in 96-well plates, combining up to four different dyes per well and sent to Macrogen Inc. (Seoul, Korea) for fragment analysis on an ABI3730XL Genetic Analyzer (Life Technologies Corp.) with GeneScan 500 LIZ as a size standard. Chromatograms were analysed using GeneMarker 2.0.2 (SoftGenetics, LLC, State College, PA).

A previous study showed that individuals of the Baikalian Cladophorales species flock are likely to be polyploids as they showed more than two alleles (signal peaks) per locus (Chapter 2). Common statistics designed for microsatellites such as F_{st}, expected heterozygosity or linkage disequilibrium are based on diploid individuals and cannot be used in polyploids if the allele dosage is uncertain or if the ploidy level is high (Obbard et al. 2006, Pfeiffer et al. 2011b, García-Verdugo et al. 2013). To overcome this problem, the alleles were coded into a binary presence (1)-absence (0) matrix as 'allele

Table 3.1. Endemic morphospecies of the Cladophorales from Lake Baikal analyzed in this study. The groups correspond to the two major clades (A and B) found in Boedeker et al. (2018) using rDNA sequences.

Taxa	Group
Chaetomorpha baicalensis Meyer	B
Chaetomorpha curta (Skabichevsky) Skabichevsky, attached form	A
Chaetomorpha curta, unattached form	A
Chaetomorpha moniliformis Skabichevsky	A
Cladophora compacta (Meyer) Meyer *ex* Hollerbach	B
Cladophora floccosa Meyer var. *floccosa*	B
Cladophora floccosa Meyer var. *irregularis* Skabichevsky	B
Cladophora globulus (Meyer) Meyer*/ Cl. compacta*	B
Cladophora globulus/ Cl. pulvinata	B
Cladohpora kursanovii Skabichevsky	A
Cladophora meyeri Skabichevsky var. *gracilior* (Meyer) Hollerbach	B
Cladophora meyeri Skabichevsky var. *meyeri*	B
Cladophora pulvinata (Meyer) Meyer	B
Gemmiphora compacta Skabichevsky	B
Rhizoclonium sp. Kützing, 1843	A

phenotypes' an approach successfully used in population genetics of many taxa (Andreakis et al. 2009; Sampson & Byrne 2012; Cidade et al. 2013; López-Vinyallonga et al. 2015; Samah et al. 2016). In spite of the reduction of informative power, it allows the analysis of polyploids and the combination of individuals with different ploidy levels (Bruvo et al. 2004, Andreakis et al. 2009), plus other statistics, such as total number of alleles per locus and number of private alleles are not affected (López-Vinyallonga et al. 2015). The loci that failed during amplification on particular samples were coded as missing data. For loci 15, 515-35 and 515-46 in samples of group B the alleles were coded as true absences as we considered them informative for cluster analysis (Data matrix upon request).
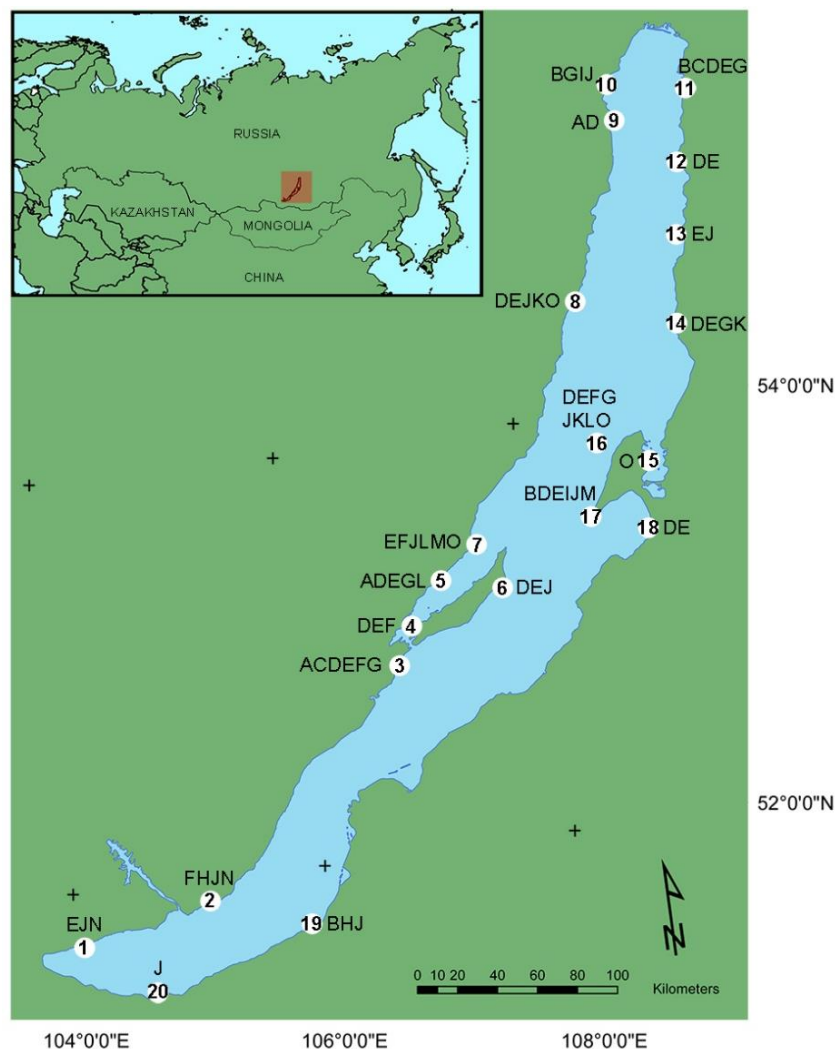
Figure 3.1. Map of Lake Baikal indicating the collection localities and morphotaxa found at each location. Taxa: A, *Ch. baicalensis*; B, *Ch. curta* attached form; C, *Ch. curta* unattached form; D, *Ch. moniliformis*; E, *Cl. compacta*; F, *Cl. floccosa* var. *floccosa*; G, *Cl. floccosa* var. *irregularis*; H, *Cl. globulus/ Cl. Compacta*; I, *Cl. globulus/ Cl. pulvinata*; J, *Cl. kursanovii*; K, *Cl. meyeri* var. *gracilior*; L, *Cl. meyeri* var. *meyeri*; M, *Cl. pulvinata*; N *G. compacta*; O, *Rhizoclonium* sp. Localities and approximate coordinates in Appendix 3.4.

45

Summary statistics for each SSR were calculated including the total number of alleles and alleles per individual. The polymorphic information content for dominant markers (PIC), a good measure of genetic diversity (Roldan-Ruiz et al. 2000), was calculated in GDdom (Abuzayed et al. 2017). Statistics for all morphotypes were calculated in GenAlEx 6.5 (Peakall and Smouse 2012) including the number of alleles (*a*), number of alleles with a frequency < 5% (*aF*), mean value of unbiased diversity (*uh*) and number of private alleles (*Pa*). The number of unique genotypes was calculated for the whole dataset and for each morphotaxon independently. The overall power of the loci to discriminate between individuals was calculated with an accumulation curve using *poprr* package (Kamvar et al. 2015) implemented in R (R Core Team 2013) with the function *genotype_curve*.

3.3.2 *Clustering analyses*

Species delimitation was performed using two different clustering methods; a model-based Bayesian clustering analysis as implemented in STRUCTURE (Pritchard et al. 2000) and the Gaussian Clustering (GC) as implemented in the R package *prabclus* (Hausdorf and Hennig 2010).

STRUCTURE uses a model-based algorithm which assigns individuals to clusters with similar genetic patterns on the assumption of Hardy-Weinberg equilibrium (Pritchard et al. 2000). The analyses were set up supporting dominant markers by adding a row of recessive alleles. Missing data was coded as '-9'. Other options used were 'correlated allele frequencies' and 'no admixture', as it is regarded that these last parameters can detect subtle population structure (Falush et al. 2007, Porras-Hurtado et al. 2013). All individuals of the full data set were considered a single population and the morphotaxon assignment was not considered as a prior to avoid bias in the analysis (Porras-Hurtado et al. 2013). An exploratory analysis was run to calculate the number of clusters (K) that best fit the data. This consisted of 10 independent runs, K= 2 – 18 and 50,000 Markov chain Monte Carlo (MCMC) iterations with a burn-in of 50,000. The best K was determined calculating the Ln(K) and Delta K values (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl and von Holdt 2012). A more extensive analysis was done to improve the cluster assignment performance using the selected K, 10 independent runs, each of 250,000 MCMC iterations and 250,000 iterations of burn-in.

A consensus Q matrix was obtained in CLUMPAK (Kopelman et al. 2015) using the greedy algorithm and plotted using STRUCTURE PLOT (Ramasamy et al. 2014).

Gaussian Clustering algorithm constructs clusters according to the normal distribution that can best explain the differences between the individuals (Hausdorf and Hennig 2010, Carstens et al. 2013, Sauer et al. 2013, Mrinalini et al. 2015). The data matrix was transformed into pairwise distance matrices using the Jaccard index. As *prabclust* cannot deal with missing data they were transformed into absences ('0'). The number of clusters was inferred with the function *prabclus* using the 'kruskal' nonmetric multidimensional scaling (NMDS), without 'noise component', with the number of clusters ranging from 2 to 18 and 99 permutations. The number of dimensions was calculated using the *stressvals* function with dimensions ranging from 2 to 10, selecting a value below 10% as suggested by the authors of the program (*prabclus* documentation).

Finally, the resulting assignments were compared, looking for clusters consistent with morphospecies and for agreement between both clustering methods. The proportion of individuals assigned to only one cluster was calculated and heat map-like bidimensional plots were created using the R package *adegraphics* to visually compare the number of individuals assigned to each cluster.

3.4 Results

3.4.1 *SSRs statistics*

The analyses of the markers revealed that individuals had more than two alleles for most of the loci (Table 3.2), except for locus 515-35 which had up to two. Although the number of alleles per individual in each locus was up to 8 (i.e. loci 14 and 34), the average number ranged from 1.4 to 2.7. PIC values ranged from 0.09 to 0.41 and had an average of 0.22. The number of missing data in all the alleles is relatively low (<1%), except for loci 14 (6.6%) and 34 (36.0%). In spite of the amount of missing data, we decided to include these loci in the analysis, as missing data mainly occurs in *Cl. floccosa*, *Cl. meyeri* and *Cl. globulus*, and the loci are informative for the remaining morphotaxa.

Table 3.2. Characteristics and number of alleles by locus. Groups according to Boedeker et al. (2018).

Locus name	Motif	Total number of alleles	Alleles per individual	Average number of alleles	Polymorphic information content	Cross-amplification
5a	TAC	12	1 to 5	2.5	0.24	All
6	TA	23	1 to 6	2.1	0.26	All
14	ATA	27	1 to 8	2.6	0.14	All
15	AAT	4	1 to 4	2.4	0.33	Group A only
16	GTA	21	1 to 7	2.7	0.16	All
34	TCA	31	1 to 8	2.2	0.09	All
46	TA	16	1 to 5	1.8	0.09	All
50-15	TTA	40	1 to 7	2.5	0.09	All
515-35	ATT	2	1 to 2	1.4	0.37	Group A only
515-46	TTGT	3	1 to 3	2.3	0.41	Group A only
515-64	TAT	4	1 to 3	1.7	0.25	All
Range or means	-	16.6	1 to 5	2.2	0.22	-

The total number of alleles found among all 11 loci was 170 and the number of alleles per morphospecies ranged from 31 to 101 (Table 3.3). The overall unbiased diversity value (uh) was 0.068 ranging from 0.029 to 0.118. The proportion of alleles with a frequency <5% (aF) ranged from 21.1% to 52.5%. A total of 21 private alleles (Pa) were detected, ranging from 0 to 4. A total of 600 different genotypes were detected. The majority of the morphospecies showed a large proportion of unique genotypes, except for *Ch. moniliformis* (34% unique) and the unattached form of *Ch. curta* (57%) (Table 3.3). The cumulative curve plot of loci versus number of genotypes shows that the markers have sufficient power to discriminate each individual as the curve plateaus (Appendix 3.5).

3.4.2 *Species delimitation by cluster analysis*

The results of the exploratory analysis in STRUCTURE revealed that K = 12 was the number of clusters that best fit the data (mean Ln P(K) = -12517.6) based on the ΔK plot (Appendix 3.6). The number of individuals by cluster ranged from 16 to 133. Most morphotaxa were recovered as independent clusters containing a large proportion of the pre-identified individuals (Fig. 3.2 and 3.3a; Appendix 3.7). 100% assignment of morphotaxa to clusters were found in *Ch. baicalenesis* (cluster = 3_{strc}); the attached form of *Ch. curta* (6_{strc}); *Ch. moniliformis* (5_{strc}); *Cl. kursanovii* (2_{strc}); and a high value of

76% for *Cl. compacta* (1_{strc}). A good proportion of individuals of both varieties *Cl. meyeri* var. *meyeri* and *Cl. meyeri* var. *gracilior* (66.67% each) were assigned to the same cluster (4_{strc}). For both varieties of *Cl. floccosa*, the assignment to independent clusters is not clear. Instead, they appear in different clusters mixed with samples of *Cl. compacta* and *Gemmiphora compacta* (7_{strc}) and *Cladophora pulvinata* (9_{strc}), respectively. *Rhizoclonium* sp. was assigned to cluster 6_{strc} which includes the attached form of *Ch. curta* and some individuals of the unattached form. The samples of *Cl. globulus/Cl. compacta* and *Cl. globulus/Cl. pulvinata* were mostly assigned to cluster 12_{strc} with the inclusion of a few individuals of other morphotaxa (Fig. 3.3a). The cluster 8_{strc} contained only a few individuals of various morphotaxa. The cluster assignment for each individual is shown in Appendix 3.8.

The Gaussian clustering was used with NMDS = 4, which had an optimum imprecision of 8.18%. A total of 14 clusters were recovered. Although there appears to be no clear distinction of each cluster in the 2D plot, this is only a false impression as two other dimensions are not displayed (Appendix 3.9). The clusters contained from 14 to 112 individuals. The overall assignment was similar to STRUCTURE but slight differences were observed (Fig. 3.3b). For example, *Ch. moniliformis* and *Cl.*

Table 3.3. Characteristics and genetic diversity of the SSR in each morphospecies. Total number of individuals (N), number of alleles (*a*), alleles with a frequency <5% (*aF*), mean value of unbiased diversity (*uh*), number of unique genotypes and the number of private alleles (*Pa*) are shown.

Morphotaxa	N	*a*	*aF*	*uh*	Unique genotypes	*Pa*
Ch. baicalensis	30	39	6	0.029	25	1
Ch. curta attached	41	50	12	0.049	39	0
Ch. curta unattached	21	34	2	0.048	12	0
Ch. moniliformis	122	38	16	0.012	41	0
Cl. compacta	127	101	53	0.084	126	4
Cl. floccosa var. *floccosa*	65	86	34	0.098	62	3
Cl. floccosa var. *irregularis*	65	73	19	0.118	65	2
Cl. globulus/ Cl. Compacta	31	72	18	0.105	31	2
Cl. globulus/ Cl. pulvinata	21	63	19	0.076	21	1
Cl. kursanovii	100	55	16	0.042	76	3
Cl. meyeri var. *gracilior*	30	72	11	0.115	30	3
Cl. meyeri var. *meyeri*	30	72	8	0.096	30	2
Cl. pulvinata	20	60	0	0.072	20	0
G. compacta	20	42	0	0.040	20	0
Rhizoclonium sp.	4	31	0	0.034	4	0
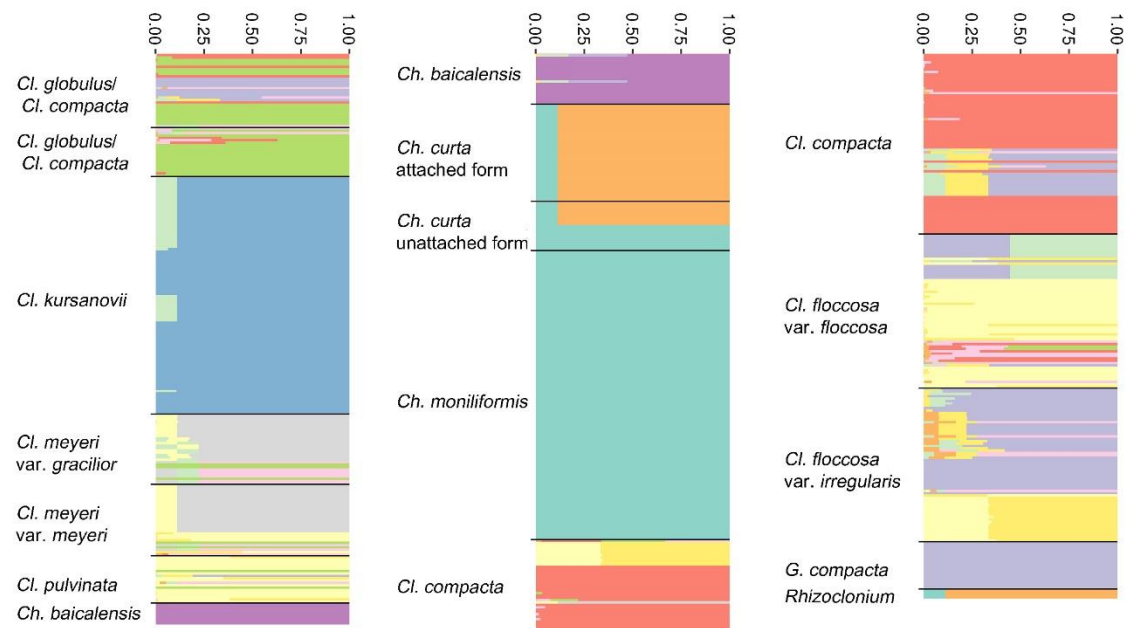
Figure 3.2. Assignment of individuals by STRUCTURE (K=12). Bar plot indicating the probability of assignment of each individual to a given cluster (by color).

kursanovii, previously recognized in STRUCTURE as single clusters, were further divided (7_{GC}, GC and 9_{GC}, and 2_{GC} and 3_{GC}, respectively) and the assignment of a few individuals also differed (Appendices 3.8 and 3.10).

3.5 Discussion

Our results indicate that many morphologically distinguishable species are discrete genetic clusters. While traditional molecular markers based on DNA sequences are not able to resolve species in the Baikalian Cladophorales clade, population genetics using SSRs provided evidence to delimit these species. Species delimitation is fundamental in several biological disciplines but remains a difficult task due to the complexity of the speciation process and to uncertainty on which species concepts and criteria are the most appropriate to distinguish between species (Duminil and Di Michele 2009, Leache and Fujita 2010, Carstens et al. 2013). Under the 'species as lineages framework' (De Queiroz 2007) the use of different sources of data is crucial to delimit species correctly (Carstens et al. 2013).
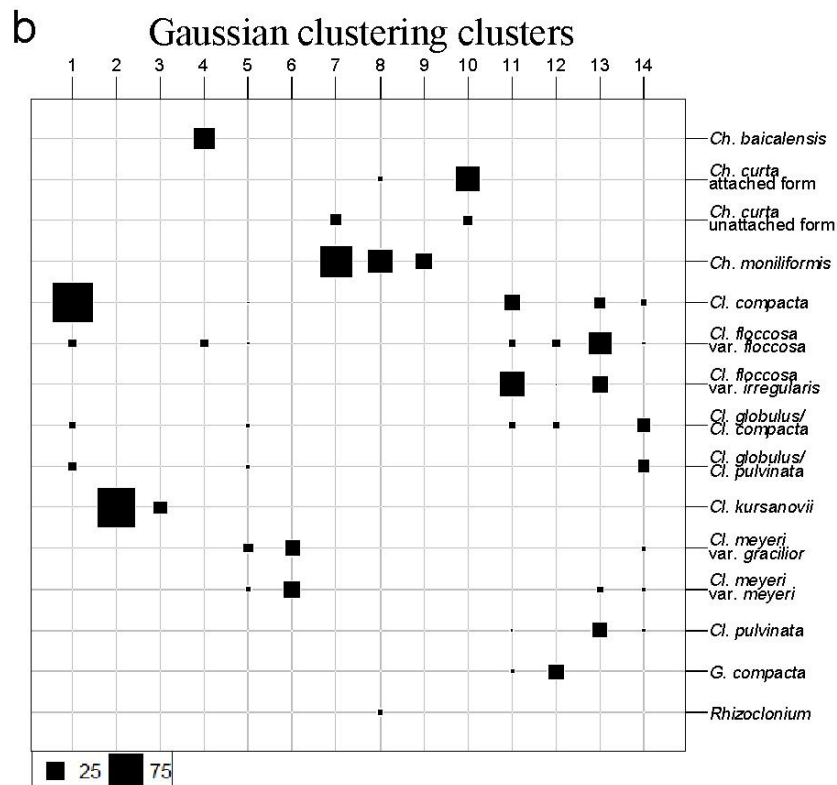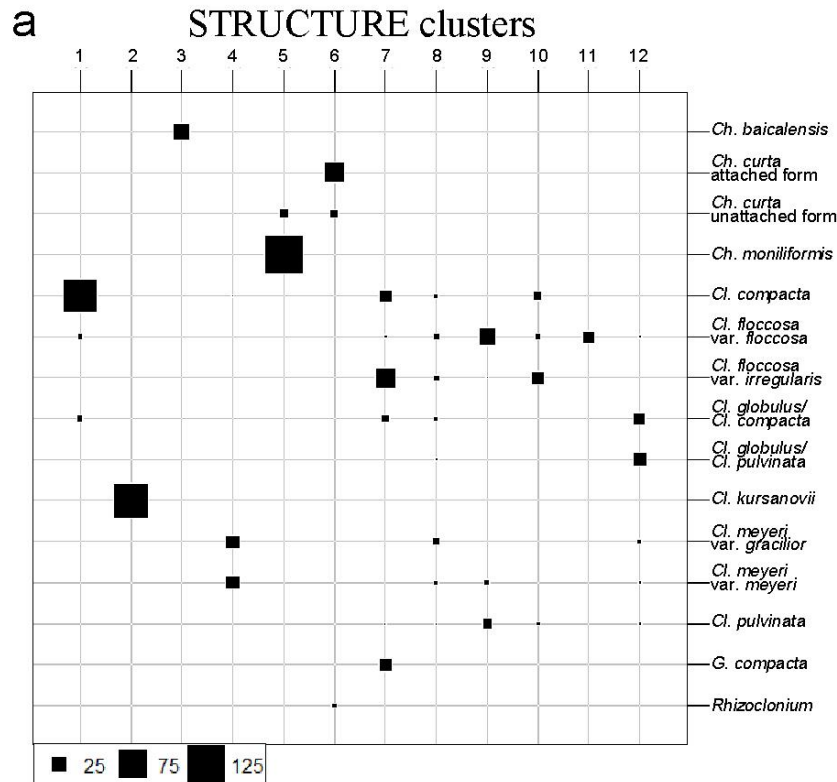
Figure 3.3. Bidimensional plot showing the number of individuals of each morphospecies (rows) assigned to a given cluster (columns). a) STRUCTURE, 12 clusters recovered; b) Gaussian clustering, 14 clusters recovered. For total number of individuals assigned to each cluster see Appendix 3.7 and Appendix 3.9.

Most of the loci had more than two alleles (Table 3.2) supporting the occurrence of polyploidy in this group and complicating population genetic analyses (Chapter 2). It was not possible to determine the ploidy level of each individual (including the haploid stages) as no correlation was observed in the number of alleles between loci. Two loci had up to 8 alleles suggesting up to octoploidy (loci 14 and 34, Table 3.2). An analysis of chromosome numbers (Hinson and Kapraun 1991) or total DNA content (Kapraun 2007) may elucidate ploidy level changes. However, this was not possible with our samples and will need further work on new collections.

The PIC values (Table 3.2) showed that many of the SSRs used have a relatively high diversity (the maximum value expected for dominant markers is 0.5)(Roldan-Ruiz et al. 2000) and are therefore adequate for the analyses. Noteworthy, loci 14 and 34 were the SSRs that showed the highest number of alleles and also had the largest proportion of missing data. However, this appears to have had no significant effects on the genetic clustering methods (see below). We speculate these results could be related to mutations in flanking regions.

The proportion of unique genotypes is high in most morphotaxa. (Table 3.3). However, this is not the case for *Ch. moniliformis* where only 34% of samples had unique multi-locus genotypes and had a low unbiased diversity value ($uh = 0.012$), event hough this species had one of the largest sample sizes. This suggests that clonal reproduction may be prevalent in this species. This finding is in contrast to *Cl. compacta* and *Cl. kursanovii* ($uh = 0.084$ and $uh = 0.042$ repectively), morphospecies also with large distribution ranges which exhibit a much higher proportion of unique genotypes (99% and 76% respectively). It has been assumed that most of the species in the lake reproduce asexually (Izhboldina 2007), but the different diversity values may indicate other reproductive strategies. Further analyses will be performed to explore the genetic diversity and reproductive strategies of these species in the lake.

STRUCTURE has been used widely in studies where polyploidy is present and the alleles coded as allelic phenotypes. This may be attributed to its extensive use among population geneticists and the relatively easy and explicit input of binary data matrices obtained from dominant markers (Meudt et al. 2009). However, it is reported that STRUCTURE can fail in obtaining the 'correct' K (clusters) if deviations from Hardy-Weinberg equilibrium are present (Hausdorf and Hennig 2010, Huelsenbeck et al. 2011). For this reason, the use of complementary and independent methods is

recommended (Hausdorf and Hennig 2010, Kamvar et al. 2014). Gaussian clustering uses mixtures of normal allele frequency distributions and therefore is regarded as a good alternative to model-based methods for dominant data (Hausdorf and Hennig 2010, Sauer and Hausdorf 2012, Sauer et al. 2013).

Another potential bias to the species delimitation by genetic clustering might be the decoupling of the life cycles (dominance of sporophyte or gametophyte in populations)(Krueger-Hadfield et al. 2016) and the isolation of asexual lineages from the original populations during the transition from sexuality to asexuality (Dudgeon et al. 2017). This is a possible scenario for those morphotaxa that were divided into two or more genetic clustres (e.g. *Ch. moniliformis* and *Cl. kursanovii*; Fig. 3.3b). However, this effect appears to be neglible due to the low level of genetic differentiation observed in most of the species.

Both cluster analyses (STRUCTURE and GC) produced comparable results although GC recovered two more clusters than STRUCTURE and other subtle differences in the assignment of individuals were observed (Figure 3.3 and Appendix 3.8). These genetic clusters likely correspond to reproductively isolated lineages as the assignment did not lump individuals based on locality or geographic regions, but mostly by morphospecies. The cluster delimitation supports many of the morphotaxa identified in this study, such as *Ch. baicalenesis*, *Ch. curta* (attached form), *Ch. moniliformis*, *Cl. compacta and Cl. kursanovii*. Individuals of the unattached form of *Chaetomorpha curta* and *Ch. moniliformis* were recovered in the same cluster in both methods supporting their conspecificity. Both thalli appear to be similar in shape and size with the only difference the attached or unattached habit. However, another group of samples of unattached *Ch. curta* was assigned to the same cluster as the attached form, which may indicate that the unattached morphology is influenced by the environment and can arise in different species. In spite of their morphological differences, the attached form of *Ch. curta* and *Rhizoclonium sp*. were placed in the same cluster by STRUCTURE (Fig. 3.3a; 6_{strc}) supporting the ITS phylogeny obtained by Boedeker et al. (2018). The GC placed the *Rhizoclonium* samples in a *Ch. moniliformis* cluster (Fig. 3.3b; 8_{GC}) indicating slight genetic differences with *Ch. curta*. A more detailed analysis is required to confirm if this is evidence of hybridization or an artifact of the data. For *Cl. meyeri* the analysis suggests no genetic distinction between either variety, similar to the ITS phylogeny (Boedeker et al. 2018). Our results also show that both *Cl. globulus/Cl.*

compacta and *Cl. globulus/Cl. pulvinata* may actually belong to the same species. The remaining morphotaxa, namely *Cl. foccosa* var. *irregularis*, *Cl. floccosa* var. *floccosa*, *Cl. pulvinata* and *Gemmiphora compacta*, have unclear assignments with both methods (Fig. 3.3; clusters 7-11$_{strc}$ and 4-12$_{GC}$) and a reexamination of their morphological characteristics is needed. Furthermore, our results indicate that a reassessment of the taxonomy and nomenclature of the Baikalian Cladophorales species flock is required. Unfortunately, this requires a more intensive taxonomic review not only for these species but for the entire order, as shown by previous phylogenetic analyses of the Cladophorales (Boedeker et al. 2016, 2018).

Polyploidy is likely to have occurred in this group making it interesting for the study of speciation and evolution. Polyploidy is regarded as a form of instant speciation as the offspring become reproductively incompatible with their parents (Schluter 2001, Albert and Schluter 2005). This may also explain the morphological variation in the group as polyploidy is also often related to morphological changes (Comai 2005, Herben et al. 2017). The formation of polyploids via genome doubling (autopolyploids) or hybridization (allopolyploids) could have led to the diversification of Baikalian Cladphorales, as it is reported that polyploidization favors the establishment of species in new environments (Soltis et al. 2010, te Beest et al. 2012). The role of polyploidy was not solved and studies including other sources of data (i.e. DNA content or chromosome counting) will be necessary to reveal ploidy changes and their influence in speciation and adaption in these species.

Lake Baikal has undergone several geological (Logatchev 1993, Mats 1993) and climatic changes (Sherstyankin and Kuimova 2006) and the transition of the region into a subarctic climate (~2.4 my) coincides with a major change in the diversification of many species in the lake (Sherbakov 1999, Yokoyama and Goto 2005, Schön and Martens 2012, Stelbrink et al. 2015, Karanovic and Sitnikova 2017). An accurate timing of the diversification of these morphospecies would indicate whether this diversification correlates with these environmental changes.

The Baikalian Cladphorales species flock shows a complex evolutionary history that remains to be fully understood. One of the most intriguing questions about this group is if the species have diversified in sympatry. A finer population genetics analysis may help to understand the population structure, the genetic diversity within species,

and further evidence of speciation in sympatry such as gene flow between populations and lack of isolation by distance (Coyne and Orr 2004).

The species delimitation performed on the Baikalian Cladophorales allowed us to delimit many of species as genetic clusters that support reproductive isolation of lineages. This is also supported by the cluster analyses which did not lump individuals based on locality or geographic regions, but mostly by morphotaxa. It was possible to delimit five out of 15 morphotaxa with 100% confidence (Fig. 3.3 and Appendix 3.8); two morphotaxa pairs seem to be conspecific; and two other morphotaxa are growth forms of other species. The remaining morphotaxa could not be fully resolved and need a deeper review and additional data. The endemic Baikalian Cladophorales species is a recently discovered group with many properties of a species flock (Boedeker et al. 2018) that may have evolved in sympatry or via polyploidization. Therefore, it might represent one of the first examples of sympatric speciation in algae.

3.6 Acknowledgements

CONTRASTING PATTERNS OF POPULATION STRUCTURE, GENETIC VARIATION AND REPRODUCTION STRATEGIES OF THREE SYMPATRIC SPECIES OF CLADOPHORALES ENDEMIC TO LAKE BAIKAL

4.1 Abstract

Studies on population genetics are relevant for the understanding of the factors driving the genetic variability and structure of populations, as well as ecological and evolutionary processes of a species. To date, many studies on population genetics have been done in marine algae, but the number of studies on population genetics of haplo-diplontic algae are scarce compared to diplontic species and those done in freshwater algae are few. The endemic Cladophorales of Lake Baikal are a group with many interesting characteristics such as high species richness, monophyly, low genetic differentiation and apparently recent speciation, making it ideal for the study of the genetic structure of an alga in a large freshwater environment. In this study we analyzed the population genetics of three species of the Cladophorales species flock of Lake Baikal *Chaetomorpha* (*Ch.*) *moniliformis*, *Cladophora* (*Cl.*) *compacta* and *Cl. kursanovii*, using a set of eleven microsatellites coded as allelic phenotypes. General statistics for microsatellites were calculated and the genetic variability of the populations was explored. Analyses of molecular variance were performed to account for the partitioning of the genetic variation between populations. Clustering methods were implemented to infer the genetic structure of the species. Finally, Mantel tests were performed to test isolation-by-distance. The analyses of population genetics of *Ch. moniliformis*, *Cl. compacta* and *Cl. kursanovii* revealed patterns of genetic variation and structure that suggest different reproductive strategies and dispersal abilities. In *Ch. moniliformis* the analysis suggests asexual reproduction and long-distance dispersal. *Cladophora compacta* is likely to reproduce sexually with high gene flow between populations. In *Cl. kursanovii* the analysis indicated high differentiation and structure between populations. The results obtained in this study demonstrate that contrasting biological characteristics may arise in closely related species. Finally, the results of this study will contribute to the understanding of the evolution and diversification of the Cladophorales species flock of Lake Baikal.

4.2 Introduction

Population genetics, the study of the genetic structure and variability of populations, is relevant to recognize the factors influencing the genetic variability of a species as well as many aspects of their ecology and evolution (e.g. Van der Strate et al. 2002, Billot et al. 2003, Bouza et al. 2006). Population genetics in algae, although not as developed as in animals or plants, has been helpful in understanding species invasions (O'Doherty and Sherwood 2007, Andreakis et al. 2009), dispersal and migration (Van der Strate et al. 2003, Hu et al. 2013), habitat fragmentation (Billot et al. 2003, Coleman and Kelaher 2009), reproductive strategies and life cycles (Engel et al. 2004, Couceiro et al. 2015, Krueger-Hadfield et al. 2016), conservation (Couceiro et al. 2011b) and species delimitation (Chapter 3).

Genetic variation and population structure of algae species is directly influenced by mutation, selection, migration and genetic drift (Valero et al. 2001) as well as other factors including their reproduction modes (sexual or asexual) (Krueger-Hadfield et al. 2016, Dudgeon et al. 2017) and polyploidy (Varela-Álvarez et al. 2018). In sexual species the genetic diversity is usually high and is mainly driven by recombination whereas in asexual species the genetic diversity is predominantly influenced by mutation (Faugeron et al. 2001, Guillemin et al. 2008). Migration can also play a significant role in the genetic structure of asexual species as a single genotype can represent a large proportion of the individuals in a population (Tatarenkov et al. 2005, Chan et al. 2013, Ardehed et al. 2015). Genetic differentiation between populations at fine scales is frequently linked to the poor dispersal of gametes and spores producing a pattern of isolation by distance (Van der Strate et al. 2002b, Billot et al. 2003, Engel et al. 2004, Couceiro et al. 2011, Li et al. 2016). On the other hand, weak genetic structure at large spatial scales is usually associated with asexual reproduction and/or drifting of vegetative structures (Kinlan et al. 2005, Tatarenkov et al. 2005, Chan et al. 2013) which can establish in a new site (Engel et al. 2004, Alberto et al. 2009, Chan et al. 2013). In spite of the growing interest in the study of population genetics of algae as a result of new technologies and methods, including the design of microsatellites (e.g. Krueger-Hadfield et al. 2011, Couceiro et al. 2015, Ayres-Ostrock et al. 2016) and single nucleotide polymorphism genotyping (e.g. Provan et al. 2013, Zhang et al. 2015, Yesson et al. 2018), such studies on freshwater algae are scarce.

Lake Baikal (Siberia, Russia) is a unique freshwater body due to its age, size, depth (Mats and Perepelova 2011), low concentration of nutrients (Khodzher et al. 2017) and its overall stable environmental conditions (Mikhailov et al. 2018). Inhabiting the lake, the endemic Cladophorales represent an opportunity for the study of population structure in a large lake. This endemic group of algae contains the highest number of described Cladophoracean morphospecies inhibiting a single freshwater environment (16 taxa) (Izhboldina 2007) and have a morphological diversity that can be compared to that found in the entire order, including: upright tufts and cushions, small tufts, crusts of procumbent branched filaments, upright filaments branched only at the base and stiff unbranched filaments (see Boedeker et al. 2018). However, molecular data revealed that this monophyletic group of algae has very little genetic divergence between morphotaxa (Boedeker et al. 2018). Currently, the endemic Baikalian Cladophorales is considered a species flock (Boedeker et al. 2018) due to its monophyly and species richness (Greenwood 1984) in a similar way as other taxa of this lake, such as turbellareans, annelids, amphipods, ostracods, gastropods and sculpins (Sherbakov 1999, Hausdorf et al. 2003, Kontula et al. 2003, Macdonald et al. 2005, Yokoyama and Goto 2005, Kaygorodova et al. 2007, Timoshkin et al. 2010, Schön and Martens 2012, Stelbrink et al. 2015, Karanovic and Sitnikova 2017). Subsequent studies on microsatellite development and species delimitation (Chapter 2 and 3) have shown reproductive isolation between some Baikalian Cladophoracean morphotaxa, confirmed the conspecificity of some other morphospecies, and provided evidence for polyploidy playing a significant role in the diversification of this group.

In spite of our current knowledge on the morphology and genetics on the Baikalian Cladophorales species flock, little is known about the reproduction and dispersal abilities of the species. The life cycle of Cladophorales is generally haplo-diplontic and isomorphic (Hoek et al. 1995), but some species can reproduce asexually by biflagellate or quadriflagellate zoospores as well as by fragmentation of the thalli (Škaloud et al. 2018). In Baikalian species vegetative reproduction is thought to be dominant in most taxa (Izhboldina 2007) as no sexual reproduction has been observed in all the taxa. The dispersal mechanisms of the Cladophorales are poorly understood but long distance dispersal of thalli fragments and short-distance dispersal of gametes and spores are both reported (Van der Strate et al. 2002b, 2002a, 2003, Boedeker et al. 2010b, 2010a).

Because the species of the Baikalian Cladophorales species flock appear to have diversified recently (Boedeker et al. 2018), occur in sympatry even growing side by side and as the current environmental conditions of Lake Baikal are considered relatively homogeneous across the lake, the analysis of the genetic diversity and population structure might reveal whether the reproductive strategies and gene flow within these species are similar or not. In addition, the assessment of the population genetics of the endemic Cladophorales of Lake Baikal will not only improve our knowledge of the processes involved in the diversification of other freshwater algae, but also of other Baikalian organisms as the population structure of only a few taxa are known (Yampolsky et al. 1994, Mashiko et al. 2000, Gomanenko et al. 2005, Peretolchina et al. 2007, Kravtsova et al. 2015).

In this study, we analyzed the population genetics of *Chaetomorpha* (*Ch.*) *moniliformis*, *Cladophora* (*Cl.*) *compacta* and *Cl. Kursanovii*, three species of Cladophorales speices flock of Lake Baikal. These three species are widely distributed across the lake and were previously recovered as independent genetic clusters (Chapter 3). We used a set of eleven simple sequence repeats (SSRs) coded as allelic phenotypes (Chapter 2 and 3). The species were independently analyzed to determine if they have similar population structure, infer their reproductive strategies and how they relate to their respective genetic structure.

4.3 Materials and Methods

4.3.1 *Sampling*

A total of 317 individuals (Appendix 4.1) of three morphospecies was used: *Ch. moniliformis* with 122 individuals from 12 populations (i.e., localities; Fig. 4.1); *Cl. compacta* with 95 individuals from 10 populations; and *Cl. kursanovii* with 100 individuals from 11 populations. The individuals for each species were chosen based on previous results showing that they belong to independent genetic clusters which putatively corresponding to reproductively isolated species that are congruent with their morphological species (Chapter 3).

4.3.2 *DNA extractions and SSRs amplification*

Total genomic DNA was extracted using the Chelex protocol (Goff and Moon 1993) adding 100 to 150 µl per sample. The extractions were stored at -20°C. Dilutions 1:10 in ultra pure water were used for amplification. The amplification protocol, primers, amplification and fragment analysis followed Chapater 2. Eleven SSRs markers were used in *Ch. moniliformis* and *Cl. kursanovii* and nine in *Cl. compacta* (see Chapters 2 and 3).
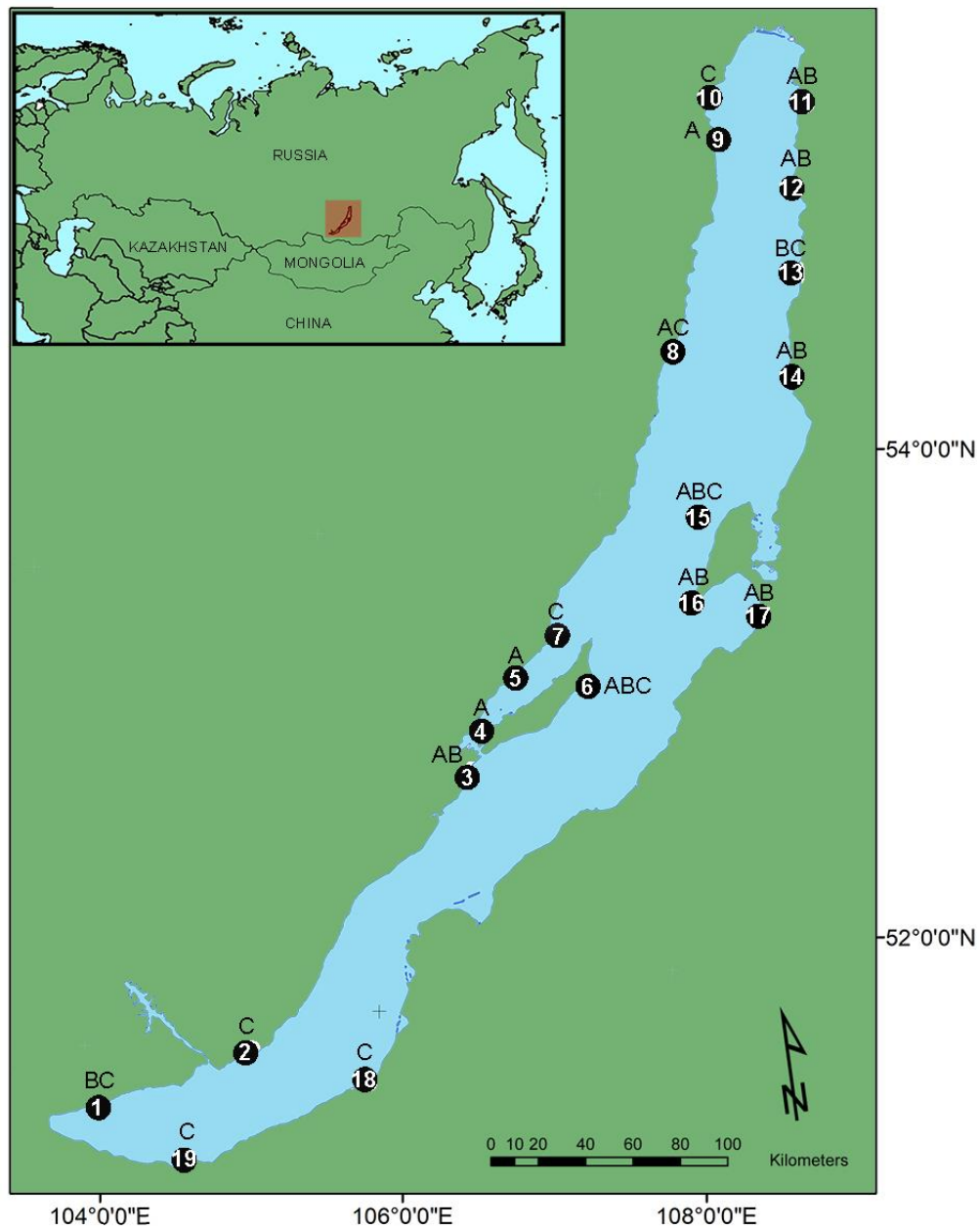
Figure 4.1. Collection sites and species distribution across Lake Baikal. A = *Ch. moniliformis*; B = *Cl. compacta*; C = *Cl. kursanovii*. Names of localities and approximate coordinates in Appendix 4.2.

As previously reported, the species of the Baikalian Cladophorales show allele patterns consistent with polyploidy (Chapter 2). Therefore, the alleles were coded as 'allelic phenotypes' into a binary presence (1)/absence (0) matrix (Data matrices upon request). This approach has proven useful for the species delimitation of Baikalian Cladophorales (Chapter 3).

4.3.3 *Genetic analyses*

The three species were analyzed independently using the same methods to compare their diversity and population structure.

General statistics for SSRs such as: number of alleles, average number of alleles, fixed alleles and number of alleles with a frequency $< 5\%$ (aF) were determined. The polymorphic information content (PIC) (Roldan-Ruiz et al. 2000) was calculated in GDdom (Abuzayed et al. 2017) for each locus. Other population statistics were calculated using GenAlEx 6.5 (Peakall and Smouse 2012). These included the number of alleles in each population (a), mean value of unbiased diversity (uh), the number of private alleles (Pa) and the number of identical genotypes. An accumulation curve to examine the power of the loci to discriminate between individuals was calculated in *poppr* (Kamvar et al. 2015) implemented in R (R Core Team 2013) using the function *genotype_curve*.

Analyses of molecular variance (AMOVA) were performed using GenAlEx (Peakall and Smouse 2012) to account for the partitioning of genetic variation within the species. The analysis estimates Φ_{PT} which is a measure of genetic differentiation analogous to F_{st} (Peakall and Smouse 2012). Analyses were performed in both ways, using all the individuals in the data set and including only distinct multi-locus genotypes in each population to estimate the effects of clonality in the genetic variation (Grünwald and Hoheisel 2006). A total of 9,999 non-parametric replicates were conducted to calculate the significance level for Φ_{PT}.

Three different clustering methods were used to examine the structure within species: 1) model-based Bayesian clustering analysis using STRUCTURE (Pritchard et al. 2000); 2) discriminant analysis of principal components (DAPC) implemented in the R package *adegenet* (Jombart 2008); and 3) Gaussian clustering (GC) implemented in

the package *prabclus* (Hausdorf and Hennig 2010). The last two methods are multivariate-based analyses but rely on different parameters for clustering and optimization. In addition, these methods are model-free making them useful for asexual or clonal populations (Halkett et al. 2005, Jombart 2008). The number of clusters estimated between algorithms were compared looking for congruencies.

STRUCTURE analyses were run as dominant markers adding a row of recessive alleles. The selected parameters were correlated allele frequencies, admixture model and '-9' for missing data. Ten parallel runs of 100,000 MCMC iterations preceded by 100,000 of burn-in iterations were performed. To infer the number of clusters (K), independent analysis were run assuming K values ranging from 1 to 13 calculating the Ln(K) and ΔK values (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl and VonHoldt 2012). A consensus Q matrix of the best K was obtained in CLUMPAK (Kopelman et al. 2015) using the greedy algorithm. The final consensus Q matrix with the individual probabilities was plotted using STRUCTURE PLOT (Ramasamy et al. 2014).

The number of clusters in DAPC were calculated using the function *find.clusters* following the package documentation. The analysis was executed using the function *dapc,* retaining the majority of the components (PCs), number of clusters estimated and considering all the linear discriminants.

For the Gaussian Clustering (GC) algorithm the data matrix was transformed into pairwise distance matrices using the *Jaccard* index. The number of clusters were inferred with the function *prabclus* using the 'kruskal' nonmetric multidimensional scaling (NMDS). Other parameters were: noise component = 0, number of clusters ranging from 2 to 18 and 99 permutations. The number of dimensions were calculated using the *stressvals* function with dimensions ranging from 2 to 5, selecting a tolerance value up to 20% (*prabclus* documentation).

Finally, a Mantel test in GenAlEx was performed to test isolation by distance (IBD). The correlation between the linearized pair-wise population genetic population distance ($\Phi_{PT}/(1- \Phi_{PT})$) and the geographic distance (kilometers) was calculated. A total of 1000 replicates was used to test significance. Only unique genotypes per population were used to avoid biases on the estimation of the correlation, as the removal of

identical genotypes approaches that of a sexual population (Grünwald and Hoheisel 2006).

4.4 Results

4.4.1 *Microsatellite statistics*

The characteristics of each locus varied across each species. In *Ch. moniliformis* the total number of alleles per locus ranged from 1 to 20. (Appendix 4.3). The number of alleles at each locus per individuals ranged from 1 to 5. Although some individuals failed to produce PCR products (coded as absences or 0 depending on the analysis). This only occurred in low frequency in loci 46, 50-15 and 515-46. Four loci (15, 46, 5015 and 515-46) had individiduals with more than two alleles. Fixed alleles within populations were observed, resulting in the absence of polymorphisms in some populations. Low frequency alleles (<5%, *aF)* were also present, mainly in locus 50-15. The PIC values were relatively low, from 0 (loci with all the aleles fixed) to 0.1082.

In *Cl. compacta* the total number of alleles per locus ranged from 3 to 24 (Appendix 4.3). The number of alleles at each locus per individual ranged from 1 to 6, although missing data was present. The frequency of failed PCR amplifications was low (<5%) occurring only in loci 6 and 34 in a few of samples and was not expected to influence the analyses. All loci had more than two alleles per sample depending the individual. Fixed alleles were low and all tested loci had some level of polymorphism (PIC ranging from 0.0337 to 0.1841.) although the number of low frequency alleles was high, particularly for loci 34 and 50-15 that had 13 both.

For *Cl. kursanovii*, the total number of alleles per locus ranged from 1 to 17 (Appendix 4.3). The number of alleles at each locus per individual ranged from 1 to 7. Six loci (5a, 6, 16, 34, 46, 50-15) had more than two alleles per individual. The number of fixed alleles was also high (Appendix 4.3), resulting in loci with medium levels of polymorphism compared to the other two speices (PIC ranging from 0 to 0.2419). Low frequency alleles were absent in most of the loci except for 50-15 (*aF=* 14) and 5a (*aF=* 1).

The number of alleles per population (*a*) ranged from 20 to 30 in *Ch. moniliformis* and the overall unbiased diversity value (*uh*) was 0.052 (Appendix 4.4). A

total of 8 private alleles among populations (*Pa*) were found. The overall proportion of unique genotypes (G-1/N-1) defined as the proportion of unique genotypes (G) minus 1 over the total number of individuals (N) minus 1 (Arnaud-Haond et al. 2007) was 0.66 indicating a large number of duplicated multi-locus genotypes. The genotype accumulation curve did not reach a plateau indicating a lack of resolution to discriminate individuals possibly due to the amount of identical genotypes and low genetic variation (Appendix 4.5). *Cl. compacta* had an *a* value ranging from 25 to 48. This species showed the highest diversity value (average *uh* = 0.099) and highest number of private alleles (*Pa* = 25) (Appendix 4.4). All individuals within this species had unique genotypes (G-1/N-1 = 1). The genotype accumulation curve reached a plateau indicating sufficient information to discriminate individuals (Appendix 4.6). In *Cl. kursanovii* the total *a* value ranged from 27 to 36, the overall *uh* was 0.066 and *Pa*=11 (Appendix 4.4). The (G-1/N-1) proportion was 0.78 indicating some populations with identical genotypes. The markers had enough resolution to discriminate individuals as the genotype accumulation curve reached a plateau (Appendix 4.7).

4.4.2 *Analysis of molecular variance*

In *Ch. monilliformis*, the AMOVA including all the individuals, or with identical genotypes removed, produced similar results (Table 4.1). Both analyses indicate that most of the genetic variation occurs within populations as the proportion of variance with or without identical genotypes was > 91% (Table 4.1). This species also has low differentiation between populations as the Φ_{PT} values were close to 0 (0.089 including all individuals and 0.05 only with unique multi-locus genotypes). Only for *Cl. compacta* the analyses included all individuals as identical multi-locus genotypes were not found. The analysis showed that most of the genetic variation occurs within populations (84%, Table 4.1). The Φ_{PT} value showed slight level of differentiation among populations with a value of 0.163. In *Cl. kursanovi*, the levels of genetic variation within and among populations, although similar, varied depending on the inclusion or exclusion of identical genotypes. The analysis including all the individuals had a value of 53% among populations, while excluding identical genotypes the major proportion of variation occurred within populations, 51%. Both analyses showed great differentiation between populations, with a Φ_{PT} value > 0.492 (Table 4.1).

66

Table 4.1. AMOVA results for the three species showing the distribution of variance within and between populations and Φ_{PT} values for the full data set and after identical genotypes removal (= *). Other values df = degrees of freedom, SS= sum of squares, MS = mean squares, Est. Var. = estimate of variance, % = percentage of variation and p value after 9999 permutations are shown.

Species	Source	df	SS	MS	Est. Var.	%	Φ_{PT}	p value
Ch. moniliformis	Among populations	11	21.535	1.958	0.096	9	0.089	< 0.001
	Within populations	110	108.227	0.984	0.984	91		
	Total	121	129.762		1.080	100		
*Ch. moniliformis	Among populations	11	19.541	1.776	0.069	5	0.050	0.027
	Within populations	69	90.311	1.309	1.309	95		
	Total	80	109.852		1.378	100		
Cl. compacta	Among populations	9	108.566	12.063	0.825	16	0.163	< 0.001
	Within populations	85	360.044	4.236	4.236	84		
	Total	94	468.611		5.061	100		
Cl. kursanovii	Among populations	10	210.112	21.011	2.118	53	0.534	< 0.001
	Within populations	89	164.498	1.848	1.848	47		
	Total	99	374.610		3.967	100		
*Cl. kursanovii	Among populations	10	169.479	16.948	2.077	49	0.492	< 0.001
	Within populations	68	145.635	2.142	2.142	51		
	Total	78	315.114		4.219	100		

4.4.3 *Cluster analyses*

STRUCTURE analysis yielded different structure patterns for each species. *Chaetomorpha moniliformis* had an optimum K = 2, where 48 individuals were assigned to cluster 1 and 74 to cluster 2. However, the differences with other K value estimates were small (Appendix 4.8a). The STRUCTURE bar plot did not show clear differentiation of clusters and individual probabilities of membership to either clusters 1 or 2 appears to be very similar (Fig. 4.2a). Other results with increasing *K* values had similar results (data not shown). *Cladophora compacta* had a K optimum = 11

(Appendix 4.8b), but the combined Q-matrices of the 10 runs in CLUMPACK resulted in incongruent membership probabilities and unclear cluster assignment (Fig. 4.2b), except for individuals from Ighimey Bay and some from Barguzinsky Bay which were assigned to a cluster which was recovered consistently, even when different K values were used. The cluster results using K = 3 following the result of DAPC (see below) showed unclear cluster assignments as well (data not shown). *Cl. kursanovii* had an optimum K = 3 (Appendix 4.8c). The number of individuals per cluster were 39, 10 and 51 respectively and the STRUCTURE bar plot showed a clear differentiation between clusters (Fig. 4.2c).

Cluster analyses with DAPC showed similar patterns to STRUCTURE for all three species. In *Ch. moniliformis* the inference of clusters using the BIC values curve was unclear as no characteristic 'elbow' showing a lowest value was obtained (Appendix 4.9a). For this reason, we ran the analysis choosing arbitrary values of 2 to 4 clusters, considering the findings in STRUCTURE. The parameters used to compute the DAPC were 40 PCAs to avoid over-fitting (Jombart 2008) and all the discriminant eigenvalues (DE). The analysis using 3 clusters (Appendix 4.10a) produced identical results as in GC (Fig. 4.2a) but the individuals showed an admixed pattern between populations. In *Cl. compacta* the inference of clusters returned a value of 3 according to the lowest BIC value (Appendix 4.9b). A total of 15 PCAs and all the DE were used for the DAPC (Appendix 4.10b). The three clusters recovered individuals admixed between populations (Fig. 4.2b). For *Cl. kursanovii* the total number of clusters retrieved was unclear (Appendix 4.9c). For this reason, we used the same number of clusters as STRUCTURE, 15 PCAs and all the DE (Appendix 4.10c). The assignment of individuals to clusters was similar to STRUCTURE with only slight differences in the assignment of some individuals (Fig. 4.2c).
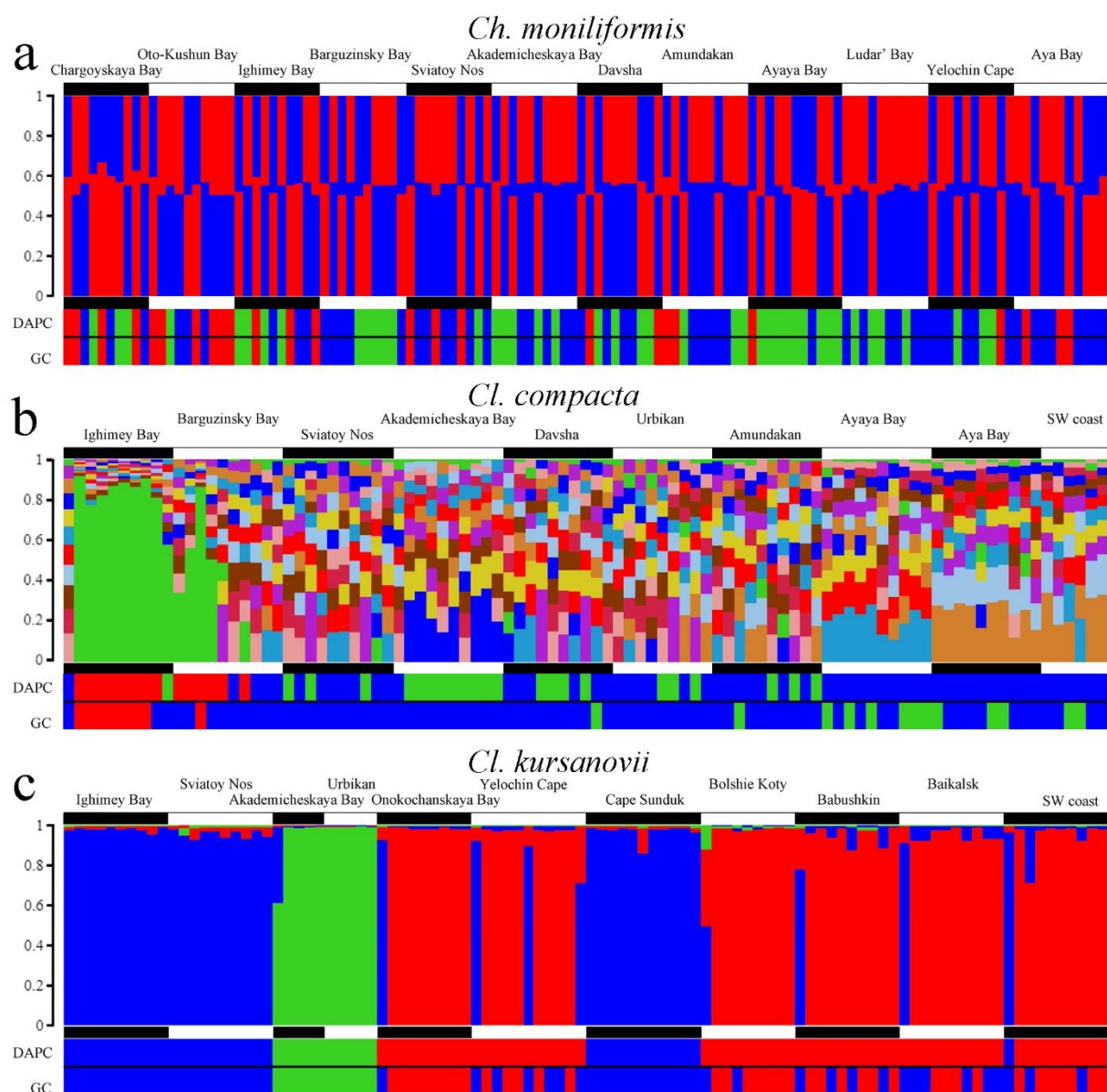
Figure 4.2. Results of the cluster analysis in STRUCTURE and comparison with DAPC and GC; a) *Ch. moniliformis*; b) *Cl. compacta*; and c) *Cl. kursanovii*. Collection sites marked by black and white bars. Number of clusters for each analysis are detailed in the text. Colors in DAPC and GC (red, blue and green) indicate one of the three different clusters (see Figure 4.3).
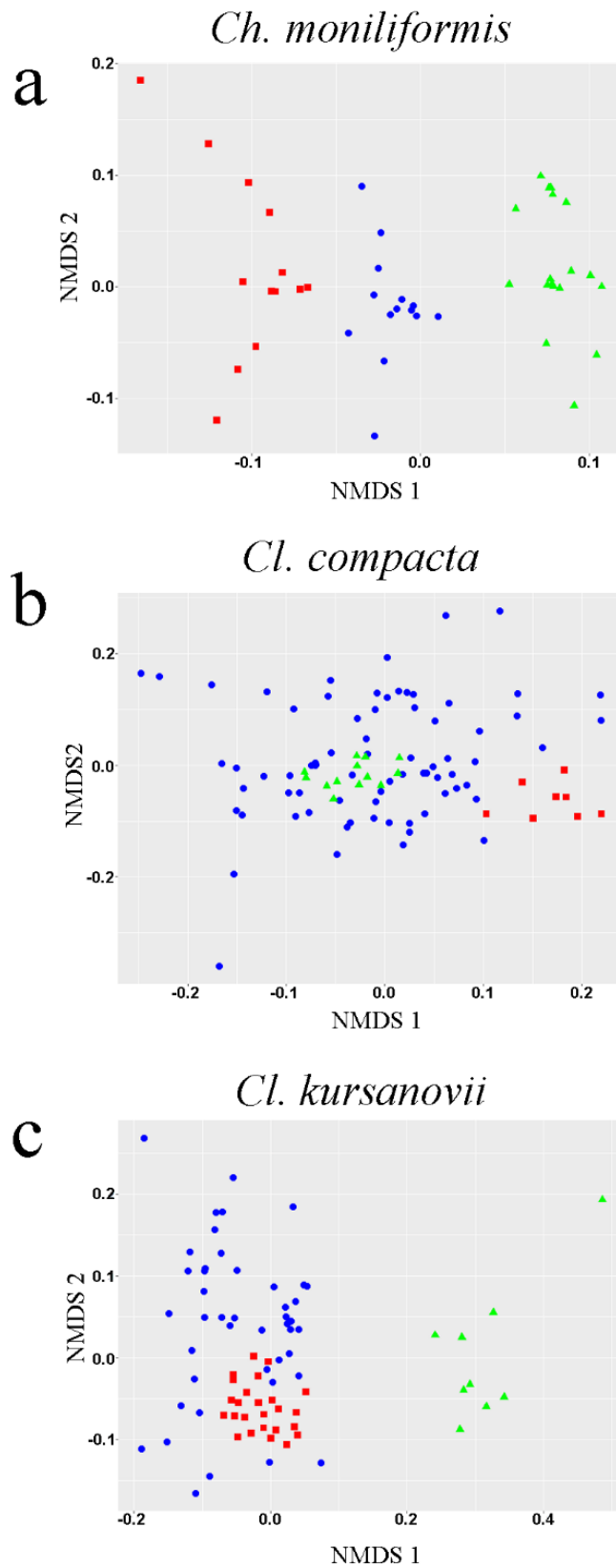
Figure 4.3. Bidimensional scatterplot of the results of the Gaussian Clustering (NMDS 1 and 2) showing three clusters in each speces. a) *Ch. moniliformis*; b) *Cl. compacta*; and c) *Cl. kursanovii.* Blue circles = cluster 1, green triangles = cluster 2, and red squares = cluster 3. The clusters are also indicated in Figure 4.2.

The Gaussian clustering produced different results depending on the number of NMDS used. Selecting NMDS with values of stress below 10% as recommended by Hausdorf and Hennig (2010) resulted in different numbers of clusters for each run. For this reason, the analyses were run tolerating up to 20% stress and using the same number of clusters as in DAPC. The cluster assignment in *Ch. moniliformis* (4 NDMS and 3 clusters, Fig. 4.3a) produced the same results as in DAPC (Fig. 4.2a); in *Cl. compacta* (5 NDMS and 3 clusters; Fig. 4.3b) the assignment produced a different pattern to STRUCTURE and DAPC, with more individuals placed in the same cluster (Fig. 4.2b); and finally, in *Cl. kursanovii* the analysis retrieved a defined number of clusters (3 NDMS and 3 clusters; Fig. 4.3c) and produced similar assignment to STRUCTURE and DAPC (Fig. 4.2c), but assigned more individuals to cluster 1 (blue) and more mixture between this cluster and cluster 3 (red). The geographic distribution of the genetic clusters are shown in Fig. 4.4.
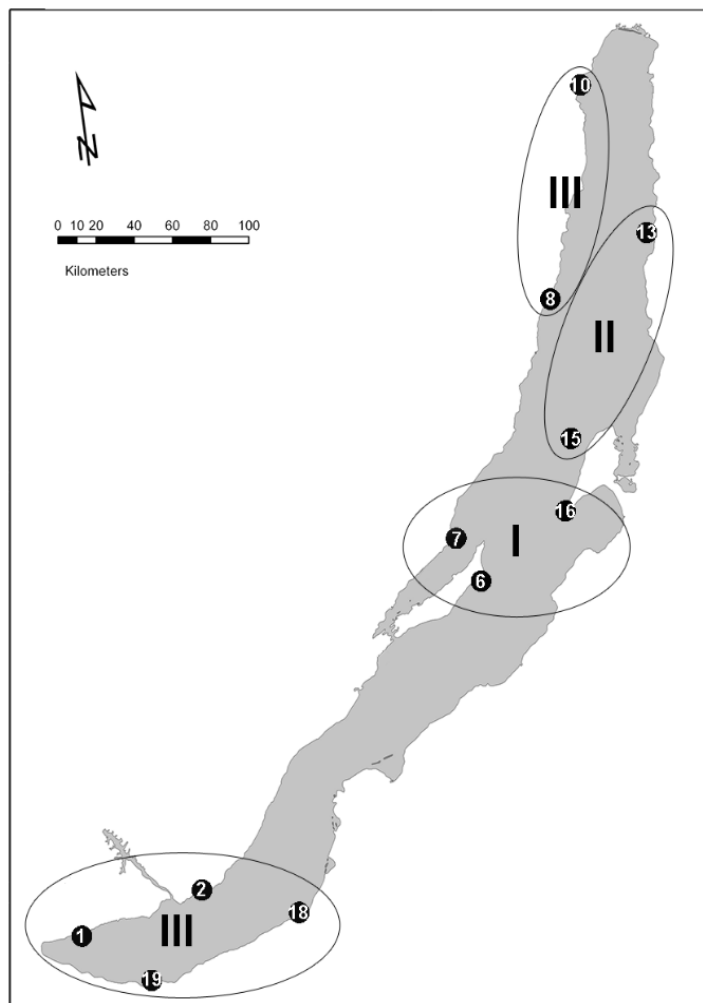
Figure 4.4. General geographic distribution of the three clusters of *Cl. kursanovii* inferred by STRUCTURE, DAPC and GC. I, II, III indicate the cluster number. For locality numbers (1-18) see Appendix 4.2.

71
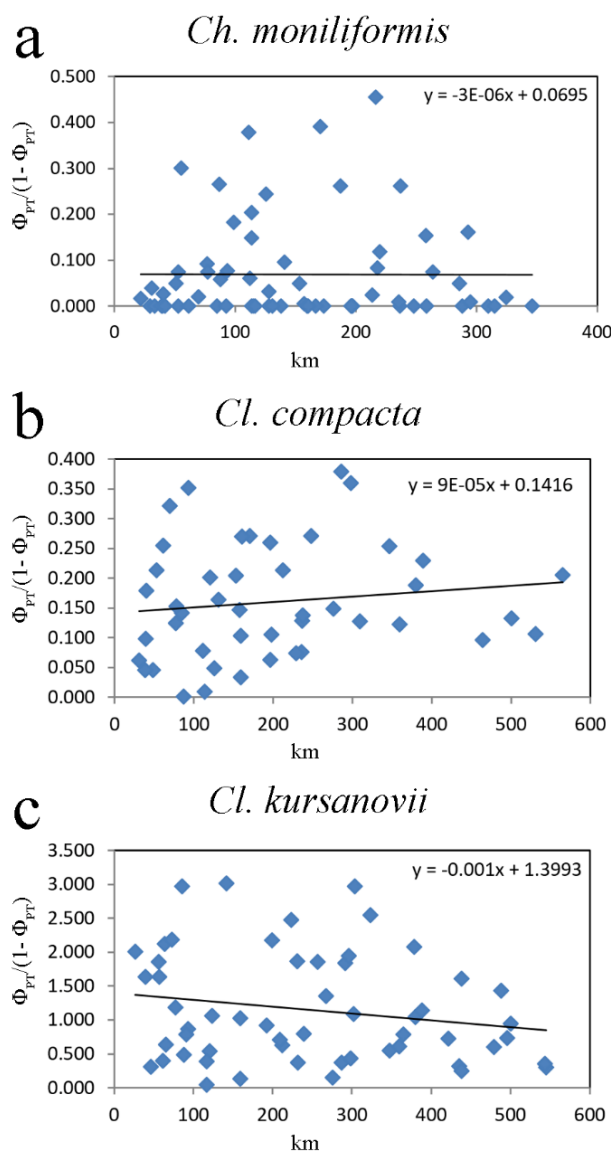
Figure 4.5. Results of the Mantel test for isolation-by-distance. Geographic distance in km versus genetic distance ($\Phi_{PT}/(1- \Phi_{PT})$) ; a) *Ch. moniliformis*; b) *Cl. compacta*; and c) *Cl. kursanovii*.

The isolation by distance test (IBD) did not show significant correlation between the linearized genetic distance ($\Phi_{PT}/(1- \Phi_{PT})$) and geographic distance in all three species (Fig. 4.5, Table 4.2). Although in *Cl. compacta* a slight positive correlation and in *Cl. kursanovii* a negative correlation was observed.

Table 4.2. Mantel results isolation-by-distance for the three species. Rxy and *p* values for each of the species are shown.

Species	Correlation coefficient (Rxy)	p
Ch. moniliformis	-0.002	0.543
Cl. compacta	0.133	0.278
Cl. kursanovii	-0.190	0.072

4.5 Discussion

The genetic diversity of the three species analyzed in this study showed patterns of genetic structure that suggest different reproductive strategies and dispersal abilities. *Ch. moniliformis* with low diversity (uh =0.052), high proportion of identical genotypes ((G-1)/(N-1)=0.66) and weak population structure (Table 4.1, Figs. 4.2 and 4.3) suggests a strong influence of asexual reproduction and the propagation of only a few genotypes over long distances across the lake. *Cl. compacta* is likely to reproduce sexually due to its high diversity (uh =0.052), unique genotypes ((G-1)/(N-1)=1) and weak genetic structure (Table 4.1, Figs. 4.2 and 4.3) which is linked to high gene flow between populations. Finally, *Cl. kursanovii* which had intermediate diversity compared to the other two species (uh =0.066) but with high differentiation between populations

, indicating local structure and reduced gene flow between some populations (Table 4.1, Figs. 4.2 and 4.3).

       Ch. moniliformis had low genetic diversity as demonstrated by its low number of alleles, number of private alleles and low PIC values per locus (Appendix 4.3), in spite of the large distribution range of the species (up to 346 km between populations). The removal of identical genotypes in each population did not greatly influence the diversity and the estimates of variance. Both AMOVA results (Table 4.1) indicated that most of the variance occurs within populations and there is little differentiation between populations. The results of the three clustering methods also indicated weak genetic structure (Figs. 4.2a and 4.3a) and no IBD (Fig. 4.5a). This might indicate high gene flow between populations, which is supported by the high number of the same genotypes across populations. This homogeneity pattern is also congruent with asexual propagation combined with long distance dispersal (Chan et al. 2013, Ardehed et al. 2015) and mutation by drift. Another potential explanation is the recent expansion of individuals of a recently formed species that has not accumulated mutations to produce structured populations, by drift or selection. However, we lack other molecular evidence to infer either speciation times or mutation rates.

       Cl. compacta had high genetic diversity as demonstrated by the high number of alleles per individual, low number of fixed alleles, high PIC values at each locus and all genotypes unique (Appendix 4.3). The AMOVA results indicated that this species had the majority of genetic variance within populations but with low levels of genetic

differentiation between populations (Table 4.1). The cluster analyses did not reveal clear differentiation between populations and were inconsistent between each other (Fig. 4.2b). STRUCTURE, in spite of showing a clear Δ increment at K=11 (Appendix 4.8), produced a consensus Q-matrix with individuals having low probabilities to multiple clusters and providing only a good assignment to individuals of Ighimey Bay. Examination of other results using K = 3 as number of clusters following the GC and DAPC indicates weak structure as well (Appendix 4.11). On the other hand, the non-model methods recovered three groups with weak structure (Figs. 4.2b and 4.3b), but no clear pattern related to populations or geography was observed, again except for Ighimey Bay. This weak genetic structure, high diversity (Appendix 4.3) and no IBD (Fig. 4.5b) suggest extensive gene flow between populations in spite of their large geographic distance (distances between populations up to 600 km), a pattern observed in many marine algae in which populations appear to have gene flow at scales of >100 km (Olsen et al. 2010, Neiva et al. 2012, J.-J. Li et al. 2016). Sampling more individuals per population and the analysis of intermediate populations might shed some light in the mid-/low-scale genetic structure and diversity of this species.

For *Ch. kursanovii* the number of total alleles, number of fixed alleles, PIC values per locus (Appendix 4.3) and number of identical genotypes (Appendix 4.4) suggests moderate genetic diversity in this species. The AMOVA including and excluding identical genotypes (Table 4.1) showed similar results, both indicating that roughly half of the genetic variation occurs within populations and half between populations. This pattern was observed in species that have some degree of differentiation between populations (Zhao et al. 2010, Cidade et al. 2013). The structure analyses (Figs. 4.2c) were congruent with the AMOVA (Table 4.1) creating three clusters with individual assignments in general consistent between clustering methods. In addition, the STRUCTURE bar plot showed a clear differentiation between clusters due to good individual probabilities. In general, it can be observed that the clusters follow a simple but distinguishable geographical pattern with a few individuals mixing in from other clusters (Figs. 4.2c and 4.4). The first cluster corresponds to populations located in the central basin; the second cluster was of populations on the east of the northern basin; and the third correspond to populations located in the southern basin and the western part of the northern basin. Some individuals from cluster three were assigned to cluster one depending on the method, possibly as a result of shared ancestral

alleles or low gene flow between individuals of these clusters. Although clusters one and two show a defined geographic pattern, the third cluster had a disjunct distribution. It is unclear what could cause this pattern and the degree of connectivity between these distant regions.

Some insights regarding the reproductive strategies of these three species can also be inferred from our data. The life cycle of most Cladophorales species remains unknown (Hoek et al. 1995), but the occurrence of clonal individuals or individuals having the same genotype is not unexpected due to their known ability for asexual reproduction (Škaloud et al. 2018). The proportion of unique genotypes indicated that *Ch. moniliformis* had a large number of identical genotypes (G-1/N-1 = 0.66). This result, in addition to the low genetic diversity, the lack of genetic structure and shared genotypes between populations, suggest that asexuality is likely to play a major role in the reproduction and dispersal of this species. In contrast, in *Cl. compacta* the results suggest large populations with sexual reproduction. This is supported by the lack of identical genotypes (G-1/N-1=1) and the high proportion of variation within populations (Wang et al. 2008, Andreakis et al. 2009). Another possible scenario is the increase of heterzygozity by mutations as observed in asexual algae populations (Guillemin et al. 2008, Krueger-Hadfield et al. 2016), but it is unlikely due to the degree of genetic variation. For *Cl. kursanovii*, sexual reproduction within populations is also suggested due to its diversity, but asexual reproduction is also occuring (G-1/N-1=0.78). Low gene flow between some populations appears to occur as well (see above). The transition from sexual to asexual reproduction my have an impact in the recovered clusters of *Ch. moniliformis* and particularly of *Cl. kursanovii*. It has been observed that decouplings of life cycles and asexual reproduction in algae could produce clustering within a population as each asexual lineage will evolve isolated from the other individuals within the population (Dudgeon et al. 2017).

As little is known about the life cycle of freshwater Cladophorales, these findings are interesting considering that Baikalian Cladophorales species flock, which are all closely related and monophyletic (Boedeker et al. 2018), are thought to reproduce mainly asexually (Izhboldina 2007). What has induced different reproductive mechanism remains to be discovered, but two main factors that might influence the reproduction in these algae are: 1) environmental conditions, such as low light during winter and low nutrients levels, known for Lake Baikal (Sherstyankin and Kuimova

2006, Khodzher et al. 2017) might favor asexual over sexual reproduction (Agrawal 2012); and 2) genetic factors such as polyploidy and hybridization (Comai 2005, Soltis and Soltis 2009, Soltis et al. 2010). For instance, it is known that polyploidy followed by loss of sex with subsequent clonal expansion can produce rapidly expanding populations with low genetic diversity (Halkett et al. 2005), a pattern observed in *Ch. moniliformis*. What remains to be discovered is if polyploidy, also suggested in the other species from previous results (Chapter 2 and 3), is impairing sexual reproduction in the other two species.

The weak genetic structure in *Ch. moniliformis* and *Cl. compacta*, the disjunction of the cluster 3 in *Cl. kursanovii* and lack of IBD raised questions about the dispersal potential and connectivity of these populations. The distances between localities are large, especially for a lake, making migration of gametes and spores less likely as it is observed that they are short lived and mainly establish close to the original plant (Krueger-Hadfield et al. 2013, J.-J. Li et al. 2016, Thibaut et al. 2016). The long dispersal of clones of *Cladophoropsis membranacea* was observed up to 10 km apart (Van der Strate et al. 2002b) and was attributed to dispersal of thallus fragments (Van der Strate et al. 2002a, Van der Strate et al. 2003). This is also likely to occur in Baikalian Cladophorales. For instance, it was possible to find identical genotypes even between distant populations in *Ch. moniliformis* (up to 340 km) and *Cl. kursanovii* (up to 346 km). Further studies of these genotypes might reveal insights on their origin, dominance in populations and the correlation between their dispersal and spatial structure (Ardehed et al. 2015).

Multiple other factors might be influencing the dispersal and connectivity of these species, such as surface currents, depth of the lake and the intrinsic characteristics of each species. For instance, *Ch. moniliformis* can survive detached from the substratum, as demonstrated by the conspecificity of the morphospecies *Ch. moniliformis* with the unattached morphospecies *Ch. curta* (Boedeker et al. 2018, Chapter 3). A similar process might occur in *Cl. compacta* which produce branched tufts that could fragment and be transported by currents.

The genetic diversity of the three species were considerably different between each other and no common structure patterns were found. This is an intriguing result considering that these species have similar distributions, the environmental conditions are more or less homogenous across the lake and it is suggested that all this group of

Cladophoralean Baikal species have evolved recently (Boedeker et al. 2018). In this regard, the few other studies on population genetics of aquatic fauna such as gammarids (Yampolsky et al. 1994, Mashiko et al. 2000), amphipods (Gomanenko et al. 2005); molluscs (Peretolchina et al. 2007) and diptera (Kravtsova et al. 2015) from Lake Baikal suggest that there is not a common and unique driver for the genetic differentiation of species.

In conclusion, the three species analyzed here revealed different genetic structures and degrees of genetic variability in spite of their close phylogenetic relationships and apparently recent diversification. The results indicate different degrees of gene flow between populations in each species and the presence of different reproductive strategies that range from highly vegetative to highly sexual. The dispersal abilities of these species by asexual means appear to play an important role on their genetic structure at large scales. Finally, this study not only contributes to our knowledge of the species of the endemic Cladopohorales of Lake Baikal, but to an understanding of the diversity and evolution of the order Cladophorales as few population genetics studies in the order have been done to date.

4.6 Acknowledgments

Chapter Five

GENERAL DISCUSSION

5.1 Main findings of this research

This research provides some of the first insights into the diversification and evolution of the endemic Cladophorales of Lake Baikal and contributes to the understanding of speciation in freshwater algae. This study also contributes to the knowledge of an overlooked part of the diversity in ancient lakes, as most of the studies on evolution and population genetics are focused on animals.

The first part of this research (Chapter 2) revealed that polyploidy is occurring in this endemic group of algae and a methodology for genotyping of polyploids was implemented. In Chapter 3, the species delimitations conducted in this clade of Baikalian Cladophorales revealed some genetic clusters consistent with morphology, supporting reproductive isolation of five taxa. Two hypotheses of conspecificity between morphotaxa, previously inferred with ribosomal markers (Boedeker et al. 2018) are also supported. Some individuals formed clusters with different morphotaxa revealing the need of further research to clarify if they belong to a different species or if hybridization is occurring. Finally, Chapter 4 revealed aspects of the genetic structure, genetic diversity, dispersal and reproduction modes in three selected species of these sympatric algae: 1) a species with genetic variation patterns consistent with dominant asexual reproduction, no population structure, low genetic diversification and long dispersal capabilities (*Chaetomorpha* [=*Ch.*] *moniliformis*); 2) a species with high genetic diversity, no population structure, which is likely to reproduce sexually (*Cladophora* [=*Cl.*] *compacta*); and 3) a species with a pattern consistent with moderate diversity, but also with genetic structure congruent with geographic distribution and more restricted dispersal (*Cladophora kursanovii*). Therefore, this study demonstrates that this group of algae is following different life modes and dispersal abilities in spite of their close genetic relationships.

In this research, new technologies of HTS allowed the design of a set of microsatellites to study the population structure and diversity of the Baikalian Cladophorales (Chapter 2). The good success of cross-amplification is attributed to the overall low genetic differentiation between morphotaxa, but also to the *in silico* cross-validation performed. This step allowed for PCR tests with a reduced number of potential SSRs instead of performing trial and error over a list of hundreds of SSRs retrieved from the NGS data. For this reason, integrating this approach is a great

improvement when analyzing multiple species or even when working with a single species with datasets from different populations.

Discovering more than two alleles in most loci was an interesting result as it indicates polyploidy in these taxa. Although this could be the outcome of amplification of duplicated loci across the genome, the occurrence of multiple alleles in different loci indicates that polyploidy is most likely, as observed in other studies (Varela-Álvarez et al. 2017a, 2017b). The occurrence of polyploidy in this group required adjustments on the population genetics methods and analyses. The most important was the method of allele scoring as 'allelic phenotypes'. This method to score alleles is difficult if stutter peaks and other PCR artifacts are produced. However, these problems were tackled adapting a previous published method (see Chapter 2, Pfeiffer et al. 2011) and manually analyzing the trace patterns produced in all the chromatograms for every single individual and locus.

The use of allelic phenotypes allowed me to overcome the difficulties of analyzing individuals with unknown ploidy level (n, 2n, to xn). However, this procedure has the cost of reducing the genetic signal and prevents the use of some of the most common statistics such as F_{st} or expected heterozygosity, as it is not possible to calculate the allele dosage (Obbard et al. 2006, Pfeiffer et al. 2011b, García-Verdugo et al. 2013). In spite of the drawbacks of using SSRs as allelic phenotypes, the genetic signal of these markers was powerful enough to delimitate most of the species of the Baikalian Cladophorales and to study the population structure of *Ch. moniliformis*, *Cl. compacta* and *Cl. kursanovii*.

The genetic species delimitation implemented in the Baikalian Cladophorales distinguished morphospecies as genetic clusters supporting the reproductive isolation of *Ch. baicalensis*, the attached form of *Ch. curta*, *Ch. moniliformis*, *Cl. compacta* and *Cl. kursanovii*. The unattached form of *Ch. curta* is a growth form of *Ch. moniliformis* and a species of *Rhizoclonium* is growth form of the attached form of *Ch. curta*, as suggested in a previous study (Boedeker et al. 2018). These species are also easy to identify using morphological characters. The dubious individuals of *Cl. globulus/compacta* and *Cl. globulus/pulvinata* were recovered as conspecific, as were the two varieties of *Cl. meyeri* (var. *meyeri* and var. *gracilior*). The remaining morphotaxa (*Cl. floccosa* var. *floccosa, Cl. floccosa* var. *irregularis, Cl. pulvinata* and *Gemmiphora compacta*) had inconsistencies, with samples placed in different clusters.

Some reasons that can explain these results are: 1) misidentification of samples; 2) not enough genetic differentiation due to recent speciation, 3) shared genetic signal due to hybridization, or 4) cryptic speciation.

The analysis of the population genetics of *Ch. moniliformis*, *Cl. compacta* and *Cl. kursanovii* was important to understand some of the evolutionary processes occurring within species. These species are assumed to have evolved relatively recently from a common ancestor, based on low diversity of DNA sequence markers and phylogenetic reconstruction (Boedeker et al. 2018) and occupy similar niches (i.e. sympatric at small scales, centimeters apart). Therefore, finding such distinct genetic structure and reproduction strategies in these species was unexpected. This research also revealed that some of these species may disperse over large distances (Chapter 4), a property that is apparent in species such as *Ch. moniliformis*, which can also grow unattached (= *Ch. curta* unattached, Chapter 3) which might facilitate migration. On the other hand, *Cl. kursanovii* showed clear evidence of genetic structure between populations that can be the result of thalli strongly attached to the substratum by rhizoids reducing dispersal. However, further studies are needed to analyze if the growth forms are related to the dispersal abilities of the species.

5.2 Tackling population genetics in Cladophorales

5.2.1 *Consequences of the life cycle of Cladophorales on population genetics*

The difficulty of population genetics in algae arises to a large extent because of the diversity in life cycles of many species. Life cycles of algae are diverse but in general can be divided into 4 major types: 1) haplo-diplontic, containing two vegetative phases, sporophyte and gametophyte; 2) haplontic, where the vegetative phase is haploid; 3) diplontic, where the vegetative phase is diploid; or 4) triphasic, which only occurs in red algae and includes the formation of diploid carpospophytes, a second diploid phase (Hoek et al. 1995). However, many other differences to the 'typical' life cycles may occur depending on the taxon (see Hoek et al. 1995, Graham et al. 2009). In many cases, i.e. haplontic, diplontic and heteromorphic haplo-diplontic species, allows researchers to differentiate between gametophytes (n) and sporophytes (2n) by eye, making it possible to analyze individuals with the same ploidy level and assess the

contribution of each phase to the genetic structure of a species (Van der Strate et al. 2002b, Engel et al. 2004, Couceiro et al. 2015, Krueger-Hadfield et al. 2016). The life cycle of species of the Cladophorales is haplo-diplontic and isomorphic (Hoek et al. 1995, Škaloud et al. 2018), with no morphological differentiation between sporophytes and gametophytes, thus unfortunately making it impossible to assign *a priori* the ploidy level of individuals, including algae of the Baikalian Cladophorales.

Studies in *Cladophoropsis membranacea* using microsatellites (Van der Strate et al. 2002b, 2002a) were done using the total number of alleles per locus to infer the phase of each individual. However, in our study the occurrence of polyploidy and fixed alleles in some loci within species (see chapter 4) prevented the assignment of ploidy level to individuals. For this reason, in spite of the reduction of information, coding the genotypes as allelic phenotypes was considered the best approach as many other studies have shown (e.g. Besnard et al. 2008, Andreakis et al. 2009, Pfeiffer et al. 2011a, García-Verdugo et al. 2013, Samah et al. 2016).

5.2.2 *Polyploidy in Cladophorales*

This research contributes to the knowledge of polyploidy in algae. The number of studies dealing with population genetics of polyploids are few due to the complexity of dealing with multiple alleles per locus (Sampson and Byrne 2012). The occurrence of polyploidy in algae is assumed to occur in many taxa (Kapraun 2005, 2007, Husband et al. 2013). However, most of the polyploids and their frequency remains unexplored as studies on chromosome counting and comparisons of nuclear DNA content are few.

Research on chromosome counting in the Cladophorales has demonstrated that polyploidy is practically universal in the order from a basic constellation of six chromosomes with subsequent multiples of 6 (12, 18, 24, 30, 36, etc.) found (Verma 1986, Hinson and Kapraun 1991, Kapraun 2005, 2007). Thus, the occurrence of polyploidy in the endemic Cladophorales of Lake Baikal is highly probable and this inferred polyploidy may have led to speciation. However, some intriguing questions about polyploidy in Baikalian Cladophorales remain to be studied and answered: 1) whether polyploidy occurred by genome doubling (autopolyploid) or by hybridization (allopolyploid), 2) the timing and frequency of this event; 3) the evolutionary consequences (ecological, morphological) linked to this process.

It was impossible to confirm whether these species are autopolyploids or allopolyploids. However, high levels of fixed heterozygosity, in other words, two (or more) fixed alleles in a locus may indicate hybrid polyploidy as both alleles are not following a disomic inheritance model (Sampson and Byrne 2012, Dar and Rehman 2017). In this study, the taxa analyzed (Chapter 4) showed high levels of fixed alleles in many loci suggesting hybrid origin. However, the low allele diversity, high number of identical genotypes and the unknown life cycle phase (sporophyte or gametophyte) hinders firm conclusions.

The timing of the polyploidization events is also important to understand the process involved in the diversification and speciation of these algae. Polyploid events may occur at different times in the history of a lineage and can be classified as ancient (also referred as paleo-polyploids), which occurred millions of years ago (Bennett 2004, Van De Peer et al. 2009, 2017), mesopolyploid, which occurred about 20 mya (Mayland-Quellhorst et al. 2016), and neo-polyploids, which are relatively 'recent' polyploidization (Soltis et al. 2010). In all cases, it has been possible to detect evidence of these processes due to incremental changes in chromosome number compared to related taxa (Soltis et al. 2010), or by using molecular data such as complete nuclear genomes sequencing and expressed sequence tags (Cui et al. 2006, Soltis and Soltis 2009). Polyploids reduce their genomes over time deleting redundant DNA regions and accumulating mutation in duplicated regions (Bennett 2004, Adams and Wendel 2005, Soltis et al. 2010) becoming functional diploids by a process called diploidization (Van De Peer et al. 2017). Therefore, finding putative polyploids with multiple copies of the same locus might indicate that polyploidization have occurred relatively recently as not enough time has passed for diploidization or for mutations to accumulate in a given locus and primer binding site. This recent diversification is also supported by the low sequence divergence found in ribosomal genes and spacers (Boedeker et al. 2018). So, the Baikal Cladophorales species may represent a case of neopolyploidization. In plants, it is estimated that 25-30% of polyploid species are young polyploids (Van De Peer et al. 2017). For instance, some potential cases of recent polyploids in algae are *Asparagopsis taxiformis* (Andreakis et al. 2009), *Porphyra* (Varela-Álvarez et al. 2017b) and *Hesperophycus/Pelvetiopsis* (Neiva et al. 2017) where the effects of genome reduction and mutation are still incipient and many alleles per locus are observed.

Polyploidization can produce substantial changes in the evolutionary history of a lineage as polyploids are known to have advantages over diploids that allow the species to persist during environmental changes and allowing a species to survive events of colonization in unfavorable conditions (Comai 2005, Soltis et al. 2010, te Beest et al. 2012). At the genetic level, polyploids may increase the level of diversity because of the genome doubling and/or hybridization and show signs of heterosis (Paun et al. 2006). Polyploidy can also produce changes in morphology (Comai 2005), favor asexual reproduction (Thompson and Lumaret 1992) and produce instant speciation due to gametes incompatibilities (Schluter 2001, Albert and Schluter 2005, Bolnick and Fitzpatrick 2007). Polyploidy is suggested as one of the main drivers in the evolution of this species flock as the following features can be found in the Baikalian Cladophorales: the species exhibit a wide range of different morphological body plans, the genetic evidence suggests recent and rapid speciation (Boedeker et al. 2018) and there is evidence of asexual reproduction (Chapter 4).

5.2.3 *Is clonal reproduction occurring in Baikalian Cladophorales?*

The high number of identical genotypes in the Baikalian Cladophorales, particularly in *Ch. moniliformis* and *Cl. kursanovii* (Chapter 4) suggest that asexual reproduction is playing an important role in the genetic structure of the populations. It is known that some Cladophorales species can reproduce asexually by forming zoospores or through fragmentation (Škaloud et al. 2018). In this regard, it was reported that Baikalian species mainly reproduced asexually (Izhboldina 2007) and therefore the presence of individuals having the same genotype may not be unexpected. However, whether these genotypes are produced by asexual reproduction (fragmentation or asexual spores), or produced sexually is still not completely solved due to the low allele diversity in many species, unknown ploidy levels and unknown allele sorting patterns, thus preventing analyses of sex ratio or producing statistical tests of clonality.

5.3 Is there evidence of sympatric speciation?

This research provides evidence that is congruent with the sympatric speciation model: a group of closely related species that shows evidence of reproductive isolation and

shares overlapping distributions (Coyne and Orr 2004). However, the fourth criterion "testing if other modes of speciation are unlikely" remains unsupported. First, *Ch. moniliformis* and *Cl. compacta* show evidence of panmictic populations, a characteristic usually linked to sympatric speciation (e.g. Seehausen and Van Alphen 1999, Barluenga et al. 2006; reviewed in Turelli et al. 2001, Fitzpatrick et al. 2009, Mallet et al. 2009) as isolation-by-distance and subsequent secondary contact is rejected. In spite of the fact that the phylogenetic relationships of the Baikalian Cladophorales are not fully resolved, *Cl. kursanovii*, which is closely related to *Ch. moniliformis* (Boedeker et al. 2018), has a genetic structure congruent with a geographic pattern instead of a panmictic population. Additionally, the fact that *Ch. moniliformis* has a populations with high asexual reproduction and fixed alleles leads to the possibility that other forces, such as polyploidy and hybridization, are acting in the reproductive isolation of these taxa rather than speciation with gene flow (Mallet 2007, Soltis and Soltis 2009). Whether speciation via polyploidy could be considered as sympatric speciation or not remains a debate (see Jiggins 2006, Coyne 2007, Fitzpatrick et al. 2008, Mallet et al. 2009), nevertheless, finding evidence of polyploidy driving the evolution of Baikalian species and their genetic (and morphological) diversity is no less important and intriguing.

5.4 Future research for the study of speciation in the endemic Baikal Cladophorales

5.4.1 *New perspectives*

This research has opened many questions concerning speciation and evolution of the Baikalian Cladophorales: How common is polyploidy in this group? Is there hybridization and how frequent it is? When did polyploidization occur? Is the morphology of each species related to polyploidization? Is the reproductive strategy linked to polyploidy? What are the dispersal methods of the species- by gametes, spores and/or fragmentation? Is there evidence of adaptation or niche preference between the species at smaller scales? To solve these questions, basic knowledge of the biological aspects of these species (e.g. ploidy level, life cycle, direct evidence of sexual reproduction, evidence of adaptation and fitness to a specific niche) must be addressed in order to provide stronger foundations for further studies. In this regard, the development of a successful culture protocol would help to understanding their life

cycle and reproductive mechanisms, as well as designing experiments focused on the effects of polyploidy in the development of the species. To the best of my knowledge, culturing attempts for this group of endemic algae have been unsuccessful.

5.4.2 *Estimation of polyploidy*

Nuclear DNA quantification may be an easy way to determine ploidy level in organisms with small chromosomes or without an established protocol for producing condensed chromosomes. Knowing the ploidy level of the organisms may reveal if one particular level of polyploidy can be linked to a given genetic cluster, to a specific morphology, or determine the life cycle phase of individuals (which can be used to infer if sexual reproduction occurs). Estimates of nuclear DNA content of some of the species studied here was attempted under a fluorescent microscope using DAPI staining (Fig. 5.1) and measuring the fluorescence intensity of the nuclei in ImageJ (Collins 2007). However, no significant results were obtained. Some of the factors that could have affected the measurements were the different fixation conditions of the samples (ethanol versus formaldehyde) that lead to variation in staining intensity and background fluorescence, variation in the DNA content of each sample due to its cell cycle stage (G1 versus G2), unknown life cycle phase (gametophyte or sporophyte). The large size of the cells in species such as *Ch. moniliformis* and *Cl. kursanovii* also produces inconsistent results as the fluorescence of nuclei at different focal planes interfered with each other and the
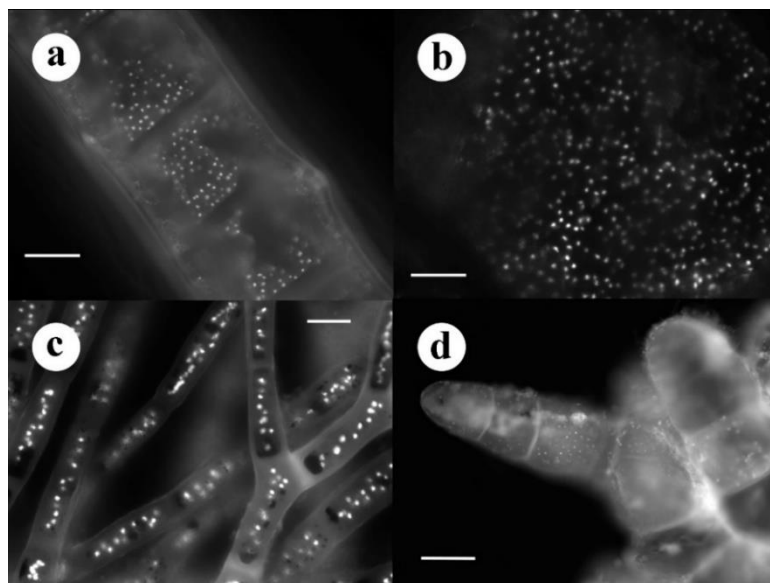
Figure 5.1. Fluorescent micrographs of Baikal Cladophorales stained with DAPI. Multiple nuclei seen in a) *Chaetomorpha curta*; b) *Ch. moniliformis*; c) *Cladophora compacta*; d) *Cl. kursanovii*. Scale bars: a, b = 100μm; c = 40μm; d = 200μm.

fluorescence of the cytoplasm produced high background. Thus, fresh sampling is required to measure several samples either fresh or fixed under the same conditions for each species and to implement other methods like flow cytometry that allows measurement of many nuclei (Oppliger et al. 2007, Guillemin et al. 2008, Varela-Álvarez et al. 2018).

5.4.3 *Is the morphological diversity related to polyploidy?*

The Baikalian Cladophorales display variation in body plans that is as great as is found in the entire order (Boedeker et al. 2016, 2018, Škaloud et al. 2018). Due to the morphological range of the species, members of this clade were originally classified into four different genera (Izhboldina 2007), two of them unique and endemic (*Chaetocladiella* and *Gemmiphora*). The current knowledge now indicates that the two non-endemic genera (*Chaetomorpha* and *Cladophora*) are in fact due to parallel/convergent characteristics with similar species around the world. Why there are different morphologies remain unknown. At first glance there is no signs of differential environmental pressure at large scales as the conditions of the lake appears to be highly homogeneous and different morphotaxa coexist in sympatry (see Chapter 3 and 4). Therefore, these morphologies might be related to unidentified genetic factors, life cycle stage, or environmental differences at small scales. It is remarkable that the '*Chaetomorpha*' and '*Rhizoclonium*' body plans were recovered as polyphyletic in the order (Boedeker et al. 2016) indicating that different body plans have emerged at different times in the evolution of the Cladophorales. Therefore, determining if the origin of such morphological differentiation is related to polyploidization or hybridization events might help us to understand the evolution in the order more generally.

5.4.4 *Design of haplotype markers*

The study of the evolution and speciation of the Baikalian Cladophorales would benefit from the use of haplotype markers (plastid and/or mitochondrial) allowing us to contrast our findings and provide new findings (e.g., molecular clock dating, sequence-based species delimitation and genetic structure). During this research, the design and amplification of sequences of mitochondrial genes (e.g. cytochrome oxidase 1 = COI

and several others) were attempted using the HTS sequences used for microsatellite development and known DNA regions from *Boodlea* (Del Cortona et al. 2017) as reference. Unfortunately, most of the attempts did not yield PCR products and those that amplified produced unusable sequences.

A parallel study on the design of mitochondrial markers using the same HTS database used in this study (Villsen 2016) was able to obtain fragments of three regions: COI, *RRSA* (ribosomal small subunit) and *RRLC* (ribosomal large subunit). The COI sequences analyzed were highly conserved and *RRLC* produced chromatograms with multiple peaks and intra-individual variation (Villsen 2016). *RRSA* was recovered with high variability, the phylogenetic signal of the marker recovered the Baikalian Cladophorales species as polyphyletic, also producing branches including samples of other phylogenetically distant species. These results corroborate the difficulty of designing mitochondrial markers for this group. To date, there is no universal markers for any chloroplast or mitochondrial genomes in the Cladophorales. A potential explanation lies in the architecture of these organelle genomes which might be hindering the design of markers and their ability to be amplified. Recent studies have revealed that the chloroplast (and possibly mitochondria) genomes within Cladophorales have a peculiar arrangement including many single-stranded DNA molecules forming hairpin structures (Del Cortona et al. 2017).

5.5 Final considerations

This research has revealed important aspects on the evolution of the endemic Cladophorales species of Lake Baikal. First, polyploidy, a major evolutionary force (Soltis et al. 2010, Van De Peer et al. 2017), was found in this group probably influencing speciation and diversification. Second, most of the species were recovered in independent genetic clusters indicating their reproductive isolation. Finally, the assessment of the population structure of three species suggests different reproductive strategies and dispersal abilities within the group. The approach of this study combined with other traditional methodologies such as SNPs, ISSRs or AFLPs, new approaches including low-copy nuclear makers and Fluidigm microfluidic PCRs (Mayland-Quellhorst et al. 2016), wider sampling, and a better understanding of the environmental

conditions of the lake at different scales, could improve our understanding on the evolution and genetic structure of the Cladophorales species flock.

The current trend in population genetics and phylogenetics is the use of the state-of-the-art technologies to study the evolution and diversification of species. However, some non-model taxa may not follow the expected models and assumptions (e.g. disomic inheritance and allele sorting, no polyploidy, known life cycles, haploid markers) leading to a change in the analyses and interpretation of the results. It is also clear that the need exists to develop analytical tools to overcome the problems of working on species with complex biology. For example, only a few methods exist to date to analyze the population genetics of polyploids (Dufresne et al. 2014). In this regard, this study demonstrates that innovative modes of analysis can be applied to overcome the challenges associated with organisms with unusual biology and particularly for the Cladophorales species flock, provide insights of the speciation and diversification of these interesting algae inhabiting the oldest lake in the world.

REFERENCES

Abuzayed, M., El-Dabba, N., Frary, A. & Doganlar, S. 2017. GDdom: An online tool for calculation of dominant marker gene diversity. *Biochem. Genet.* 55:155–7.

Adams, K.L. & Wendel, J.F. 2005. Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8:135–41.

Agrawal, S.C. 2012. Factors controlling induction of reproduction in algae-review: The text. *Folia Microbiol. (Praha).* 57:387–407.

Albert, A.Y.K. & Schluter, D. 2005. Selection and the origin of species. *Curr. Biol.* 15:R283–8.

Alberto, F., Whitmer, A., Coelho, N.C., Zippay, M., Varela-Alvarez, E., Raimondi, P.T., Reed, D.C. et al. 2009. Microsatellite markers for the giant kelp *Macrocystis pyrifera*. *Conserv. Genet.* 10:1915–7.

Andreakis, N., Kooistra, W.H.C.F. & Procaccini, G. 2009. High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Mol. Ecol.* 18:212–26.

Ardehed, A., Johansson, D., Schagerström, E., Kautsky, L., Johannesson, K. & Pereyra, R.T. 2015. Complex spatial clonal structure in the macroalgae *Fucus radicans* with both sexual and asexual recruitment. *Ecol. Evol.* 5:4233–45.

Ayres-Ostrock, L.M., Mauger, S., Plastino, E.M., Oliveira, M.C., Valero, M. & Destombe, C. 2016. Development and characterization of microsatellite markers in two agarophyte species, *Gracilaria birdiae* and *Gracilaria caudata* (Gracilariaceae, Rhodophyta), using next-generation sequencing. *J. Appl. Phycol.* 28:653–62.

Baack, E.J. & Rieseberg, L.H. 2007. A genomic view of introgression and hybrid speciation. *Curr. Opin. Genet. Dev.* 17:513–8.

Barluenga, M., Stölting, K.N., Salzburger, W., Muschick, M. & Meyer, A. 2006. Sympatric speciation in Nicaraguan Crater Lake cichlid fish. *Nature*. 439:719–23.

Bennett, M.D. 2004. Perspectives on polyploidy in plants - ancient and neo. *Biol. J. Linn. Soc.* 82:411–23.

Besnard, G., Garcia-Verdugo, C., Rubio De Casas, R., Treier, U.A., Galland, N. & Vargas, P. 2008. Polyploidy in the olive complex (*Olea europaea*): Evidence from flow cytometry and nuclear microsatellite analyses. *Ann. Bot.* 101:25–30.

Bhandawat, A., Singh, G., Raina, A.S., Kaur, J. & Sharma, R.K. 2016. Development of genic SSR marker resource from RNA-Seq data in *Dendrocalamus latiflorus*. *J. Plant Biochem. Biotechnol.* 25:179–90.

Billot, C., Engel, C.R.C., Rousvoal, S., Kloareg, B. & Valero, M. 2003. Current patterns, habitat discontinuities and population genetic structure: the case of the kelp *Laminaria digitata* in the English Channel. *Mar. Ecol. Prog. Ser.* 253:111–21.

Boedeker, C., Eggert, A., Immers, A. & Wakana, I. 2010a. Biogeography of *Aegagropila linnaei* (Cladophorophyceae, Chlorophyta): A widespread freshwater alga with low effective dispersal potential shows a glacial imprint in its distribution. *J. Biogeogr.* 37:1491–503.

Boedeker, C., Leliaert, F., Timoshkin, O.A., Vishnyakov, V., Diaz Martinez, S. & Zuccarello, G.C. 2018. The endemic Cladophorales (Ulvophyceae) of ancient Lake Baikal represent a monophyletic group of very closely related but morphologically diverse species. *J. Phycol.* 54:616-29.

Boedeker, C., Leliaert, F. & Zuccarello, G.C. 2016. Molecular phylogeny of the Cladophoraceae (Cladophorales, Ulvophyceae), with the resurrection of *Acrocladus* Nägeli and *Willeella* Børgesen, and the description of *Lurbica* gen. nov. and *Pseudorhizoclonium* gen. nov. *J. Phycol.* 52:905–28.

Boedeker, C., O'Kelly, C.J., Star, W., Leliaert, F., Kelly, C.J., Star, W. & Leliaert, F. 2012. Molecular phylogeny and taxonomy of the *Aegagropila* clade (Cladophorales, Ulvophyceae), including the description of *Aegagropilopsis* gen. nov. and *Pseudocladophora* gen. nov. *J. Phycol.* 48:808–25.

Boedeker, C., Ramírez, M.E. & Nelson, W.A. 2010b. *Cladophoropsis brachyartra* from southern South America is a synonym of *Wittrockiella lyallii* (Cladophorophyceae, Chlorophyta), previously regarded as endemic to New Zealand. *Phycologia.* 49:525–36.

Boedeker, C. & Sviridenko, B.F. 2012. *Cladophora koktschetavensis* from Kazakhstan is a synonym of *Aegagropila linnaei* (Cladophorales, Chlorophyta) and fills the gap in the disjunct distribution of a widespread genotype. *Aquat. Bot.* 101:64–8.

Bolnick, D.I. & Fitzpatrick, B.M. 2007. Speciation : Sympatric Models and Empirical Evidence. *Annu. Rev. Ecol. Evol. Syst.* 38:459–87.

Bouza, N., Caujapé-Castells, J., González-Pérez, M.Á. & Sosa, P.A. 2006. Genetic structure of natural populations in the red algae *Gelidium canariense* (Gelidiales, Rhodophyta) investigated by random amplified polymorphic DNA (RAPD) markers. *J. Phycol.* 42:304–11.

Bräutigam, A. & Gowik, U. 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biol.* 12:831–41.

Bruvo, R., Michiels, N.K., D'Souza, T.G. & Schulenburg, H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Mol. Ecol.* 13:2101–6.

Butlin, R., Debelle, A., Kerth, C., Snook, R.R., Beukeboom, L.W., Castillo Cajas, R.F., Diao, W. et al. 2012. What do we need to know about speciation? *Trends Ecol. Evol.* 27:27–39.

Cărăbuş, M.C., Leinemann, L., Curtu, A.L. & Şofletea, N. 2015. Preliminary results on the genetic diversity of *Carpinus belutus* in carpathian populations. *Bulletin of the Transilvania University of Braşov.* 8:1–6.

Carstens, B.C., Pelletier, T.A., Reid, N.M. & Satler, J.D. 2013. How to fail at species delimitation. *Mol. Ecol.* 22:4369–83.

Chan, S.W., Cheang, C.C., Chirapart, A., Gerung, G., Tharith, C. & Ang, P. 2013. Homogeneous Population of the Brown Alga *Sargassum polycystum* in Southeast Asia: Possible Role of Recent Expansion and Asexual Propagation. *PLoS One.* 8:1–9.

Chapman, M.A. & Burke, J.M. 2007. Genetic divergence and hybrid speciation. *Evolution (N. Y).* 61:1773–80.

Cidade, F.W., Vigna, B.B., de Souza, F.H., Valls, J.F.M., Dall'Agnol, M., Zucchi, M.I., de Souza-Chies, T.T. et al. 2013. Genetic variation in polyploid forage grass: Assessing the molecular genetic variability in the *Paspalum* genus. *BMC Genet.* 14:50.

Clark, L. V. & Jasieniuk, M. 2011. polysat: An R package for polyploid microsatellite analysis. *Mol. Ecol. Resour.* 11:562–6.

Coleman, M.A. & Kelaher, B.P. 2009. Connectivity among fragmented populations of a habitat-forming alga, *Phyllospora comosa* (Phaeophyceae, Fucales) on an urbanised coast. *Mar. Ecol. Prog. Ser.* 381:63–70.

Collins, T.J. 2007. ImageJ for microscopy. *Biotechniques.* 43:25–30.

Comai, L. 2005. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 6:836–46.

Couceiro, L., Le Gac, M., Hunsperger, H.M., Mauger, S., Destombe, C., Cock, J.M., Ahmed, S. et al. 2015. Evolution and maintenance of haploid-diploid life cycles in natural populations: The case of the marine brown alga *Ectocarpus. Evolution (N. Y).* 69:1808–22.

Couceiro, L., Maneiro, I., Mauger, S., Valero, M., Ruiz, J.M. & Barreiro, R. 2011a. Microsatellite development in Rhodophyta using high-throughput sequence data. *J. Phycol.* 47:1258–65.

Couceiro, L., Maneiro, I., Ruiz, J.M. & Barreiro, R. 2011b. Multiscale genetic structure of an endangered seaweed *Ahnfeltiopsis pusilla* (Rhodophyta): Implications for its conservation. *J. Phycol.* 47:259–68.

Coyne, J.A. 2007. Sympatric speciation. *Curr. Biol.* 17:R787–8.

Coyne, J.A. & Orr, H.A. 2004. Speciation. Sinauer Associates, Sunderland, MA, USA, 480 pp.

Cracraft, J. 1983. Species Concepts and Speciation Analysis. *In* Johnston, R.F. [Ed.] *Current Ornithology*. Springer US, Boston, MA, USA, pp. 159–87.

Cristescu, M.E., Adamowicz, S.J., Vaillant, J.J. & Haffner, D.G. 2010. Ancient lakes revisited: From the ecology to the genetics of speciation. *Mol. Ecol.* 19:4837–51.

Cronn, R., Knaus, B.J., Liston, A., Maughan, P.J., Parks, M., Syring, J. V. & Udall, J. 2012. Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.* 99:291–311.

Cronquist, A. 1978. Once again, what is a species? *In* Knutson, L. V. [Ed.] *BioSystematics in Agriculture*. Allenheld Osmin, Montclair, New Jersey, USA, pp. 3–20.

Cui, L., Wall, P.K., Leebens-Mack, J.H., Lindsay, B.G., Soltis, D.E., Doyle, J.J., Soltis, P.S. et al. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Res.* 16:738–49.

Davey, J.W., Hohenlohe, P. A, Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12:499–510.

Dayrat, B. 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.* 85:407–15.

De Queiroz, K. 2007. Species concepts and species delimitation. *Syst. Biol.* 56:879–86.

Defaveri, J., Viitaniemi, H., Leder, E. & Merilä, J. 2013. Characterizing genic and nongenic molecular markers: Comparison of microsatellites and SNPs. *Mol. Ecol. Resour.* 13:377–92.

Del Cortona, A., Leliaert, F., Bogaert, K.A., Turmel, M., Boedeker, C., Janouskovec, J., Lopez-Bautista, J.M. et al. 2017. The plastid genome in Cladophorales green algae is encoded by hairpin chromosomes. *Curr. Biol.* 27:3771–82

Díaz-Martínez, S., Zuccarello, G.C., Chávez, G.A.S., Pedroche, F.F. & Avila-Ortiz, A.G. 2016. Species of *Padina* (Dictyotales, Phaeophyceae) in tropical Mexican waters based on molecular-assisted taxonomy. *Phycologia*. 55:673–87.

Dufresne, F., Stift, M., Vergilino, R. & Mable, B.K. 2014. Recent progress and challenges in population genetics of polyploid organisms: An overview of current state-of-the-art molecular and statistical tools. *Mol. Ecol.* 23:40–69.

Duminil, J. & Di Michele, M. 2009. Plant species delimitation: A comparison of morphological and molecular markers. *Plant Biosyst. - An Int. J. Deal. with all*

Asp. Plant Biol. 143:528–42.

Earl, D.A. & VonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359–61.

Edwards, D.L. & Knowles, L.L. 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? *Proc. R. Soc. B Biol. Sci.* 281:20132765.

Engel, C.R., Destombe, C. & Valero, M. 2004. Mating system and gene flow in the red seaweed *Gracilaria gracilis*: Effect of haploid-diploid life history and intertidal rocky shore landscape on fine-scale genetic structure. *Heredity (Edinb).* 92:289–98.

Esselink, G.D., Nybom, H. & Vosman, B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theor. Appl. Genet.* 109:402–8.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2611–20.

Faircloth, B. 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8:92–4.

Falush, D., Stephens, M. & Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol. Ecol. Notes*. 7:574–8.

Faugeron, S., Valero, M., Destombe, C., Martinez, E.A., Correa, J.A. & Martínez, E.A. 2001. Hierarchical spatial structure and discriminant analysis of genetic diversity in the red alga *Mazzaella laminarioides* (Gigartinales, Rhodophyta). *J. Phycol.* 37:705–16.

Filatov, D.A., Osborne, O.G. & Papadopoulos, A.S. 2016. Demographic history of speciation in a *Senecio* altitudinal hybrid zone on Mt. Etna. *Mol. Ecol.* 25:2467-81.

Filchak, K.E., Roethele, J.B. & Feder, J.L. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*. 407:739–42.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. 2008. What, if anything, is sympatric speciation? *J. Evol. Biol.* 21:1452–9.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. 2009. Pattern, process and geographic modes of speciation. *J. Evol. Biol.* 22:2342–7.

Fujisawa, T. & Barraclough, T.G. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation

on simulated data sets. *Syst. Biol.* 62:707–24.

García-Verdugo, C., Calleja, J.A., Vargas, P., Silva, L., Moreira, O. & Pulido, F. 2013. Polyploidy and microsatellite variation in the relict tree *Prunus lusitanica* L.: How effective are refugia in preserving genotypic diversity of clonal taxa? *Mol. Ecol.* 22:1546–57.

Gardner, M.G., Fitch, A.J., Bertozzi, T. & Lowe, A.J. 2011. Rise of the machines - recommendations for ecologists when using next generation sequencing for microsatellite development. *Mol. Ecol. Resour.* 11:1093–101.

Goff, L. & Moon, D. 1993. PCR amplification of nuclear and plastid genes from algal herbarium specimens and algal spores. *J. Phycol.* 29:381–4.

Gomanenko, G. V., Kamaltynov, R.M., Kuzmenkova, Z. V., Bêrênos, K. & Sherbakov, D.Y. 2005. Population structure of the Baikalian amphipod *Gmelinoides fasciatus* (Stebbing). *Russ. J. Genet.* 41:907–12.

Graham, L.E., Graham, J.M. & Wilcox, L.W. 2009. Algae. Benjamin Cummings. 616 pp.

Greenwood, P.H. 1984. What is a species flock? *In* Echelle, A. A. & Kornfield, I. [Eds.] *Evolution of Fish Species Flocks*. University of Maine, Orono Press, pp. 13–9.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C. et al. 2011. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* 11:591–611.

Guillemin, M.L., Faugeron, S., Destombe, C., Viard, F., Correa, J.A. & Valero, M. 2008. Genetic variation in wild and cultivated populations of the haploid- diploid red alga *Gracilaria chilensis*: How farming practices favor asexual reproduction and heterozygosity. *Evolution (N. Y).* 62:1500–19.

Halkett, F., Simon, J.C. & Balloux, F. 2005. Tackling the population genetics of clonal and partially clonal organisms. *Trends Ecol. Evol.* 20:194–201.

Hanyuda, T., Wakana, I., Arai, S., Miyaji, K., Watano, Y. & Ueda, K. 2002. Phylogenetic relationships within Cladophorales (Ulvophyceae, Chlorophyta) inferred from 18s rRNA gene sequences, with special reference to *Aegagropila linnaei*. *J. Phycol.* 38:564–71.

Hausdorf, B. & Hennig, C. 2010. Species delimitation using dominant and codominant multilocus markers. *Syst. Biol.* 59:491–503.

Hausdorf, B., Röpstorf, P. & Riedel, F. 2003. Relationships and origin of endemic Lake Baikal gastropods (Caenogastropoda: Rissooidea) based on mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 26:435–43.

Hegarty, M.J. & Hiscock, S.J. 2004. Hybrid speciation in plants: new insights from molecular studies. *New Phytol.* 165:411–23.

Henrichs, D.W., Renshaw, M.A., Gold, J.R. & Campbell, L. 2013. Genetic diversity among clonal isolates of *Karenia brevis* as measured with microsatellite markers. *Harmful Algae*. 21–22:30–5.

Herben, T., Suda, J. & Klimešová, J. 2017. Polyploid species rely on vegetative reproduction more than diploids: A re-examination of the old hypothesis. *Ann. Bot.* 120:341–9.

Herder, F., Pfaender, J. & Schliewen, U.K. 2008. Adaptive sympatric speciation of polychromatic "roundfin" sailfin silverside fish in Lake Matano (Sulawesi). *Evolution (N. Y).* 62:2178–95.

Hinson, T.K. & Kapraun, D.F. 1991. Karyology and nuclear DNA quantification of four species of *Chaetomorpha* (Cladophorales, Chlorophyta) from the western Atlantic. *Helgol. Mar. Res.* 45:273–85.

Hoek, C. van den 1963. Revision of the European Species of Cladophora. E.J. Brill, Leiden, NL. 248 pp.

Hoek, C. van den 1982. A taxonomic revision of the American species of *Cladophora* (Chlorophyceae) in the North Atlantic Ocean and their geographic distribution. *Verh. Kon. Ned. Akad. Wetensch., Afd. Natuurk., Tweede Sect.* 78:1–236.

Hoek, C. van den, Mann, D.G. & Jahns, H.M. 1995. Algae: An introduction to phycology. Cambridge University Press, Cambridge, UK. 623 pp.

Hu, Z.M., Zhang, J., Lopez-Bautista, J. & Duan, D.L. 2013. Asymmetric genetic exchange in the brown seaweed *Sargassum fusiforme* (Phaeophyceae) driven by oceanic currents. *Mar. Biol.* 160:1407–14.

Huelsenbeck, J.P., Andolfatto, P. & Huelsenbeck, E.T. 2011. Structurama: Bayesian inference of population structure. *Evol. Bioinforma.* 2011:55–9.

Husband, B.C., Baldwin, S.J. & Suda, J. 2013. The incidende of plyploidy in natural plant populations: major patterns and evolutionary processes. *In* Greilhuber, J., Dolezel, J. & Wendel, J. F. [Eds.] *Plant genome diversity Volume 2*. Springer Vienna, Vienna. 255-76 pp.

Izhboldina, L. 2007. Guide and key to benthonic and periphyton algae of Lake Baikal (meio- and macrophytes) with short notes of their ecology. Nauka-Center, Novosibirsk, 248 pp.

Jiggins, C.D. 2006. Sympatric Speciation: Why the Controversy? *Curr. Biol.* 16:R333–4.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic

markers. *Bioinformatics*. 24:1403–5.

Jombart, T., Devillard, S., Balloux, F., Falush, D., Stephens, M., Pritchard, J., Pritchard, J. et al. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94.

Kamvar, Z.N., Brooks, J.C. & Grünwald, N.J. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* 6:1–10.

Kamvar, Z.N., Tabima, J.F. & Grünwald, N.J. 2014. Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ.* 2:e281.

Kapraun, D.F. 2005. Nuclear DNA content estimates in multicellular green, red and brown algae: Phylogenetic considerations. *Ann. Bot.* 95:7–44.

Kapraun, D.F. 2007. Nuclear DNA content estimates in green algal lineages: *Chlorophyta* and *Streptophyta*. *Ann. Bot.* 99:677–701.

Karanovic, I. & Sitnikova, T.Y. 2017. Phylogenetic position and age of Lake Baikal candonids (Crustacea, Ostracoda) inferred from multigene sequence analyzes and molecular dating. *Ecol. Evol.* 7:7091–103.

Kaygorodova, I.A., Sherbakov, D.Y. & Martin, P. 2007. Molecular phylogeny of Baikalian Lumbriculidae (Oligochaeta): evidence for recent explosive speciation. *Comp. Cytogenet.* 1:71–84.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S. et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647–9.

Keller, I., Wagner, C.E., Greuter, L., Mwaiko, S., Selz, O.M., Sivasundar, A., Wittwer, S. et al. 2013. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.* 22:2848–63.

Khodzher, T. V., Domysheva, V.M., Sorokovikova, L.M., Sakirko, M. V. & Tomberg, I. V. 2017. Current chemical composition of Lake Baikal water. *Inl. Waters*. 7:250–8.

Kinlan, B.P., Gaines, S.D. & Lester, S.E. 2005. Propagule dispersal and the scales of marine community process. *Divers. Distrib.* 11:139–48.

Kontula, T., Kirilchik, S. V. & Väinölä, R. 2003. Endemic diversification of the monophyletic cottoid fish species flock in Lake Baikal explored with mtDNA sequencing. *Mol. Phylogenet. Evol.* 27:143–55.

Könyves, K., David, J.C. & Culham, A. 2016. Microsatellite markers for hoop-petticoat daffodils (*Narcissus* sect. *Bulbocodii* ; Amaryllidaceae). *Appl. Plant Sci.* 4:1500127.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. & Mayrose, I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15:1179–91.

Kostamo, K., Korpelainen, H. & Olsson, S. 2012. Comparative study on the population genetics of the red algae *Furcellaria lumbricalis* occupying different salinity conditions. *Mar. Biol.* 159:561–71.

Kravtsova, L.S., Bukin, Y.S., Peretolchina, T.E. & Shcherbakov, D.Y. 2015. Genetic differentiation of populations of Baikal endemic *Sergentia baicalensis* Tshern. (Diptera, Chironomidae). *Russ. J. Genet.* 51:707–10.

Krueger-Hadfield, S.A., Collén, J., Daguin-Thiébaut, C. & Valero, M. 2011. Genetic population structure and mating system in *Chondrus crispus* (Rhodophyta). *J. Phycol.* 47:440–50.

Krueger-Hadfield, S.A., Kollars, N.M., Byers, J.E., Greig, T.W., Hammann, M., Murray, D.C., Murren, C.J. et al. 2016. Invasion of novel habitats uncouples haplo-diplontic life cycles. *Mol. Ecol.* 25:3801–16.

Krueger-Hadfield, S.A., Kollars, N.M., Strand, A.E., Byers, J.E., Shainker, S.J., Terada, R., Greig, T.W. et al. 2017. Genetic identification of source and likely vector of a widespread marine invader. *Ecol. Evol.* 7:4432–47.

Krueger-Hadfield, S.A., Roze, D., Mauger, S. & Valero, M. 2013. Intergametophytic selfing and microgeographic genetic structure shape populations of the intertidal red seaweed *Chondrus crispus*. *Mol. Ecol.* 22:3242–60.

Kulikovskiy, M.S., Lange-Bertalot, H., Metzeltin, D. & Witkowski, A. 2012. Lake Baikal: Hotspot of Endemic Diatoms. *Iconogr. Diatomol.* 23.

Leache, A.D. & Fujita, M.K. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc. R. Soc. B Biol. Sci.* 277:3071–7.

Leliaert, F., De Clerck, O., Verbruggen, H., Boedeker, C. & Coppejans, E. 2007. Molecular phylogeny of the Siphonocladales (Chlorophyta: Cladophorophyceae). *Mol. Phylogenet. Evol.* 44:1237–56.

Leliaert, F., Rousseau, F., De Reviers, B. & Coppejans, E. 2003. Phylogeny of the Cladophorophyceae (Chlorophyta) inferred from partial LSU rRNA gene sequences: Is the recognition of a separate order Siphonocladales justified? *Eur. J. Phycol.* 38:233–46.

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J.M., Zuccarello, G.C. & De Clerck, O. 2014. DNA-based species delimitation in algae.

Eur. J. Phycol. 49:179–96.

Leliaert, F., Verbruggen, H., Wysor, B. & Clerck, O.D. 2009. DNA taxonomy in morphologically plastic taxa: Algorithmic species delimitation in the *Boodlea* complex (Chlorophyta: Cladophorales). *Mol. Phylogenet. Evol.* 53:122–33.

Li, J.J., Hu, Z.M., Liu, R.Y., Zhang, J., Liu, S.L. & Duan, D.L. 2016. Phylogeographic surveys and apomictic genetic connectivity in the North Atlantic red seaweed *Mastocarpus stellatus*. *Mol. Phylogenet. Evol.* 94:463–72.

Li, Q., Zhang, J., Yao, J., Wang, X. & Duan, D. 2015. Development of *Saccharina japonica* genomic SSR markers using next-generation sequencing. *J. Appl. Phycol.* 28:1387-90.

Li, Y., Huang, H. jia, Li, H., Liu, J. & Yang, W. 2016. Genetic diversity of *Ulva prolifera* population in Qingdao coastal water during the green algal blooms revealed by microsatellite. *Mar. Pollut. Bull.* 111:237–46.

Logatchev, N.A. 1993. History and geodynamics of the Lake Baikal rift in the context of the eastern Siberia rift system: A review. *BCREDP*. 17:353–70.

López-Vinyallonga, S., Soriano, I., Susanna, A., Montserra, J.M., Roquet, C. & Garcia-Jacas, N. 2015. The polyploid series of the *Achillea millefolium* aggregate in the Iberian Peninsula investigated using microsatellites. *PLoS One*. 10:1–19.

Macdonald, K.S.I., Yampolsky, L. & Duffy, J.E. 2005. Molecular and morphological evolution of the amphipod radiation of Lake Baikal. *Mol. Phylogenet. Evol.* 35:323–43.

Madlung, A. 2012. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity (Edinb)*. 110:99.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.* 10:294–9.

Mallet, J. 2007. Hybrid speciation. *Nature*. 446:279–83.

Mallet, J., Meyer, A., Nosil, P. & Feder, J.L. 2009. Space, sympatry and speciation. *J. Evol. Biol.* 22:2332–41.

Mashiko, K., Kamaltynov, R., Morino, H. & Sherbakov, D.Y. 2000. Genetic differentiation among gammarid (*Eulimnogammarus cyaneus*) populations in Lake Baikal, East Siberia. *Fundam. Appl. Limnol.* 148:249–61.

Mats, V. 1993. The structure and development of the Baikal rift depression. *Earth-Science Rev.* 34:81–118.

Mats, V.D. & Perepelova, T.I. 2011. A new perspective on evolution of the Baikal Rift. *Geosci. Front.* 2:349–65.

Mayden, R.L. 1997. A hierarchy of species concepts: the denounement in the saga of

the species problem. *Species units Divers.* 381–423.

Mayr, E. 1942. Systematics and the Origin of Species from the Viewpoint of a Zoologist. Columbia University Press, New York. 334 pp.

Mayr, E. 2000. The Biological Species Concept. *In* Wheeler, Q. & Meier, R. [Eds.] *Species Concepts and Phylogenetic Theory*. Columbia University Press, pp. 17–29.

Meudt, H.M., Lockhart, P.J. & Bryant, D. 2009. Species delimitation and phylogeny of a New Zealand plant species radiation. *BMC Evol. Biol.* 9:1–17.

Mikhailov, I.S., Zakharova, Y.R., Bukin, Y.S., Galachyants, Y.P., Petrova, D.P., Sakirko, M. V & Likhoshway, Y. V 2018. Co-occurrence networks among bacteria and microbial eukaryotes of Lake Baikal during a spring phytoplankton bloom. *Microb. Ecol.* 77:96-109.

Mishler, B.D. & Theriot, E.C. 2000. The Phylogenetic Species Concept (sensu Mishler and Theriot): *In* Wheeler, Q. & Meier, R. [Eds.] *Species Concepts and Phylogenetic Theory*. Columbia University Press, pp. 44–54.

Montecinos, A.E., Guillemin, M.L., Couceiro, L., Peters, A.F., Stoeckel, S. & Valero, M. 2017. Hybridization between two cryptic filamentous brown seaweeds along the shore: analysing pre- and postzygotic barriers in populations of individuals with varying ploidy levels. *Mol. Ecol.* 26:3497–512.

Mrinalini, Thorpe, R.S., Creer, S., Lallias, D., Dawnay, L., Stuart, B.L. & Malhotra, A. 2015. Convergence of multiple markers and analysis methods defines the genetic distinctiveness of cryptic pitvipers. *Mol. Phylogenet. Evol.* 92:266–79.

Muangmai, N., West, J.A. & Zuccarello, G.C. 2014. Evolution of four Southern Hemisphere *Bostrychia* (Rhodomelaceae, Rhodophyta) species: phylogeny, species delimitation and divergence times. *Phycologia*. 53:593–601.

Neiva, J., Pearson, G.A., Valero, M. & Serrão, E.A. 2012. Drifting fronds and drifting alleles: Range dynamics, local dispersal and habitat isolation shape the population structure of the estuarine seaweed *Fucus ceranoides*. *J. Biogeogr.* 39:1167–78.

Neiva, J., Serrão, E.A., Anderson, L., Raimondi, P.T., Martins, N., Gouveia, L., Paulino, C. et al. 2017. Cryptic diversity, geographical endemism and allopolyploidy in NE Pacific seaweeds. *BMC Evol. Biol.* 17:30.

Nielsen, R., Paul, J.S., Albrechtsen, A. & Song, Y.S. 2011. Genotype and SNP calling from next-generation sequencing data. *Nat. Rev. Genet.* 12:443–51.

O'Doherty, D. & Sherwood, A. 2007. Genetic population structure of the Hawaiian alien invasive seaweed *Acanthophora spicifera* (Rhodophyta) as revealed by DNA sequencing and ISSR analyses. *Pacific Sci.* 61:223–33.

Obbard, D.J., Harris, S.A. & Pannell, J.R. 2006. Simple allelic-phenotype diversity and

differentiation statistics for allopolyploids. *Heredity (Edinb).* 97:296–303.

Olsen, J.L., Zechman, F.W., Hoarau, G., Coyer, J.A., Stam, W.T., Valero, M. & Åberg, P. 2010. The phylogeographic architecture of the fucoid seaweed *Ascophyllum nodosum* : an intertidal 'marine tree' and survivor of more than one glacial-interglacial cycle. *J. Biogeogr.* 37:842–56.

Oppliger, L.V., Correa, J.A. & Peters, A.F. 2007. Parthenogenesis in the brown alga *Lessonia nigrescens* (Laminariales, Phaeophyceae) from central Chile. *J. Phycol.* 43:1295–301.

Papadopulos, A.S.T., Baker, W.J., Crayn, D., Butlin, R.K., Kynast, R.G., Hutton, I. & Savolainen, V. 2011. Speciation with gene flow on Lord Howe Island. *Proc. Natl. Acad. Sci.* 108:13188–93.

Paun, O., Stuessy, T.F. & Hörandl, E. 2006. The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytol.* 171:223–36.

Peakall, R. & Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics.* 28:2537–9.

Peretolchina, T.E., Bukin, I.S., Sitnikova, T.I. & Shcherbakov, D.I. 2007. Genetic differentiation of the endemic Baikalian mollusk *Baicalia carinata* (Mollusca: Caenogastropoda). *Genetika.* 43:1667–75.

Pfeiffer, T., Klahr, A., Heinrich, A. & Schnittler, M. 2011a. Does sex make a difference? Genetic diversity and spatial genetic structure in two co-occurring species of *Gagea* (Liliaceae) with contrasting reproductive strategies. *Plant Syst. Evol.* 292:189–201.

Pfeiffer, T., Roschanski, A.M., Pannell, J.R., Korbecka, G. & Schnittler, M. 2011b. Characterization of microsatellite loci and reliable genotyping in a polyploid plant, *Mercurialis perennis* (Euphorbiaceae). *J. Hered.* 102:479–88.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á. & Lareu, M. V. 2013. An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Front. Genet.* 4:1–13.

Poulton, E.B. 1904. What is a species? *Proc. Entomol. Soc. London.* 1903:lxxvii–cxvi.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics.* 155:945–59.

Provan, J., Glendinning, K., Kelly, R. & Maggs, C.A. 2013. Levels and patterns of population genetic diversity in the red seaweed *Chondrus crispus* (Florideophyceae): a direct comparison of single nucleotide polymorphisms. *Biol. J. Linn. Soc.* 108:251–62.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21:1864–77.

Ramasamy, R.K., Ramasamy, S., Bindroo, B.B. & Naik, V.G. 2014. STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus*. 3:1–3.

Rehman, R. 2017. Polyploidy: Recent Trends and Future Perspectives. Springer India, New Delhi.104 pp.

Rieseberg, L.H. & Willis, J.H. 2007. Plant Speciation. *Science.* 317:910–4.

Roldan-Ruiz, I., Dendauw, J., Van Bockstaele, E., Depicker, A. & De Loose, M. 2000. AFLP markers reveal high polymorphic rates in ryegrass (*Lolium* spp.). *Mol. Breed.* 6:125–34.

Rozen, S. & Skaletsky, H.J. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* Krawetz, S. & Misener, S. [Eds.] *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365–86.

Salzburger, W. & Meyer, A. 2004. The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften*. 91:277–90.

Samah, S., De Teodoro Pardo, C.V., Serrato Cruz, M.A. & Valadez-Moctezuma, E. 2016. Genetic diversity, genotype discrimination, and population structure of Mexican *Opuntia* sp., determined by SSR markers. *Plant Mol. Biol. Report.* 34:146–59.

Sampson, J.F. & Byrne, M. 2012. Genetic diversity and multiple origins of polyploid *Atriplex nummularia* Lindl. (Chenopodiaceae). *Biol. J. Linn. Soc.* 105:218–30.

Sauer, J. & Hausdorf, B. 2012. A comparison of DNA-based methods for delimiting species in a Cretan land snail radiation reveals shortcomings of exclusively molecular taxonomy. *Cladistics*. 28:300–16.

Sauer, J., Oldeland, J. & Hausdorf, B. 2013. Continuing fragmentation of a widespread species by geographical barriers as initial step in a land snail radiation on Crete. *PLoS One*. 8: e62569.

Savolainen, V., Anstett, M.-C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M. V, Powell, M.P. et al. 2006. Sympatric speciation in palms on an oceanic island. *Nature*. 441:210–3.

Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol. Evol.* 16:372–80.

Schoebel, C.N., Brodbeck, S., Buehler, D., Cornejo, C., Gajurel, J., Hartikainen, H., Keller, D. et al. 2013. Lessons learned from microsatellite development for

nonmodel organisms using 454 pyrosequencing. *J. Evol. Biol.* 26:600–11.

Schön, I. & Martens, K. 2012. Molecular analyses of ostracod flocks from Lake Baikal and Lake Tanganyika. *Hydrobiologia*. 682:91–110.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18:233–4.

Seehausen, O. 2006. African cichlid fish: a model system in adaptive radiation research. *Proc. R. Soc. B*. 273:1987–98.

Seehausen, O. & Van Alphen, J.J.M. 1999. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* 2:262–71.

Selkoe, K. a. & Toonen, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9:615–29.

Seo, S.B., Ge, J., King, J.L. & Budowle, B. 2014. Reduction of stutter ratios in short tandem repeat loci typing of low copy number DNA samples. *Forensic Sci. Int. Genet.* 8:213–8.

Shaffer, H.B. & Thomson, R.C. 2007. Delimiting species in recent radiations. *Syst. Biol.* 56:896–906.

Sherbakov, D.Y. 1999. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. *Trends Ecol. Evol.* 14:92–5.

Sherstyankin, P.P. & Kuimova, L.N. 2006. Hydrophysical processes in Lake Baikal in its transition from subtropical to modern climates. *Hydrobiologia*. 568:253–7.

Sjøtun, K., Heesch, S., Lluch, J.R., Martín, R.M., Garreta, G., Brysting, A.K. & Coyer, J.A. 2017. Unravelling the complexity of salt marsh '*Fucus cottonii*' forms (Phaeophyceae, Fucales). *Eur. J. Phycol.* 52:360–70.

Škaloud, P., Rindi, F., Boedeker, C. & Leliaert, F. 2018. Freshwater Flora of Central Europe, Vol 13: Chlorophyta: Ulvophyceae (Süßwasserflora von Mitteleuropa, Bd. 13: Chlorophyta: Ulvophyceae). Springer Berlin Heidelberg, Berlin, Heidelberg. 288 pp.

Smith, H.M. 1955. The perspective of species. *Turtox News*. 33:74–7.

Soltis, D.E., Buggs, R., Doyle, J. & Soltis, P.S. 2010. What we still don't know about polyploidy. *Taxon*. 59:1387–403.

Soltis, P.S. & Soltis, D.E. 2009. The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* 60:561–88.

Stebbins, G.L. 1947. Types of polyploids: Their classification and significance. *Advances in Genetics.* 1:403–29.

Stelbrink, B., Shirokaya, A.A., Clewing, C., Sitnikova, T.Y., Prozorova, L.A. & Albrecht, C. 2015. Conquest of the deep, old and cold: an exceptional limpet radiation in Lake Baikal. *Biol. Lett.* 11:20150321.

Tatarenkov, A., Bergström, L., Jönsson, R.B., Serrão, E.A., Kautsky, L. & Johannesson, K. 2005. Intriguing asexual life in marginal populations of the brown seaweed *Fucus vesiculosus*. *Mol. Ecol.* 14:647–51.

Te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubešová, M. & Pyšek, P. 2012. The more the better? The role of polyploidy in facilitating plant invasions. *Ann. Bot.* 109:19–45.

Thibaut, T., Bottin, L., Aurelle, D., Boudouresque, C.-F., Blanfuné, A., Verlaque, M., Pairaud, I. et al. 2016. Connectivity of Populations of the Seaweed *Cystoseira amentacea* within the Bay of Marseille (Mediterranean Sea): Genetic Structure and Hydrodynamic Connections. *Cryptogam. Algol.* 37:233–55.

Thompson, J.D. & Lumaret, R. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends Ecol. Evol.* 7:302–7.

Timoshkin, O.., Sitnikova, T.Y., Rusinek, O.T. & Al., E. 2004. Index of animal species inhabiting Lake Baikal and its catchment area. Vol. I: Lake Baikal, Book 2. Nauka-Center, Novosibirsk.

Timoshkin, O.A. 1999. Biology of Lake Baikal: "White spots" and Progress in Research. *Berliner Geowissenschaftliche Geowissenschaftliche*. 30:333–48.

Timoshkin, O.A., Lukhnev, A.G. & Zaytseva, E.P. 2010. First data on the endemic fauna of *Turbellaria Proseriata* (Platyhelminthes, Otomesostomidae) from Lake Baikal. *Biol. Bull.* 37:861–75.

Turelli, M., Barton, N.H. & Coyne, J. a. 2001. Theory and speciation. *Trends Ecol. Evol.* 16:330–43.

Valero, M., Engel, C., Billot, C., Kloareg, B. & Destombe, C. 2001. Concepts and issues of population genetics in seaweeds. *Cah. Biol. Mar.* 42:53–62.

Van De Peer, Y., Maere, S. & Meyer, A. 2009. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* 10:725–32.

Van De Peer, Y., Mizrachi, E. & Marchal, K. 2017. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18:411–24.

Van der Strate, H.J., Boele-Bos, S.A., Olsen, J.L., Van de Zande, L. & Stam, W.T. 2002a. Phylogeographic studies in the tropical seaweed *Cladophoropsis membranacea* (Chlorophyta, Ulvophyceae) reveal a cryptic species complex. *J. Phycol.* 38:572–82.

Van der Strate, H.J., Van De Zande, L., Stam, W.T. & Olsen, J.L. 2002b. The

contribution of haploids, diploids and clones to fine-scale population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta). *Mol. Ecol.* 11:329–45.

Van der Strate, H.J., Van de Zande, L. V., Stam, W.T., Haroun, R.J. & Olsen, J.L. 2003. Within-island differentiation and between-island homogeneity: Non-equilibrium population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta) in the Canary Islands. *Eur. J. Phycol.* 38:15–23.

Van Puyvelde, K., Van Geert, A. & Triest, L. 2010. Atetra, a new software program to analyse tetraploid microsatellite data: Comparison with tetra and tetrasat. *Mol. Ecol. Resour.* 10:331–4.

Varela-Álvarez, E., Balau, A.C., Paulino, C., Berecibar, E., Pearson, G.A. & Serrão, E.A. 2017a. Isolation and characterization of microsatellite markers for the red alga *Porphyra umbilicalis*. *Plant Genet. Resour.* 1–4.

Varela-Álvarez, E., Paulino, C. & Serrão, E.A. 2017b. Development and characterization of twelve microsatellite markers for *Porphyra linearis* Greville. *Genetica*. 145:127–30.

Varela-Álvarez, E., Loureiro, J., Paulino, C. & Serrão, E.A. 2018. Polyploid lineages in the genus *Porphyra*. *Scientific Reports* 8:8696.

Verma, B.N. 1986. New counts for *Rhizoclonium riparium* complex. *Cytologia (Tokyo).* 51:507–12.

Villsen, K. 2016. Novel mitochondrial markers for the Cladophorales (Ulvophyceae) provides insights into mitochondrial gene migration and duplication. Victoria University of Wellington, 35 pp.

Wang, X., Zhao, F., Hu, Z., Critchley, A.T., Morrell, S.L. & Duan, D. 2008. Inter-simple sequence repeat (ISSR) analysis of genetic variation of *Chondrus crispus* populations from North Atlantic. *Aquat. Bot.* 88:154–9.

Wheeler, Q. & Meier, R. 2000. Species Concepts and Phylogenetic Theory: A Debate. Columbia University Press. 256 pp.

Wheeler, Q. & Platnick, N. 2000. The Phylogenetic Species Concept (sensu Wheeler and Platnick). *In* Wheeler, Q. & Meier, R. [Eds.] *Species Concepts and Phylogenetic Theory : A Debate*. Columbia University Press, pp. 55–69.

Wiens, J.J. 2007. Species delimitation: new approaches for discovering diversity. *Syst. Biol.* 56:875–8.

Wilke, T., Väinölä, R. & Riedel, F. 2006. Patterns and processes of speciation in ancient lakes. Springer, Netherlands. 235 pp.

Yampolsky, L.Y., Kamaltynov, R.M., Ebert, D., Filatov, D.A. & Chernykh, V.I. 1994. Variation of allozyme loci in endemic gammarids of Lake Baikal. *Biol. J. Linn.*

Soc. 53:309–23.

Yesson, C., Jackson, A., Russell, S., Williamson, C.J. & Brodie, J. 2018. SNPs reveal geographical population structure of *Corallina officinalis* (Corallinaceae, Rhodophyta). *Eur. J. Phycol.* 53:180–8.

Yokoyama, R. & Goto, A. 2005. Evolutionary history of freshwater sculpins, genus *Cottus* (Teleostei; Cottidae) and related taxa, as inferred from mitochondrial DNA phylogeny. *Mol. Phylogenet. Evol.* 36:654–68.

Zardi, G.I., Nicastro, K.R., Canovas, F., Ferreira Costa, J., Serrão, E.A. & Pearson, G.A. 2011. Adaptive traits are maintained on steep selective gradients despite gene flow and hybridization in the intertidal zone. *PLoS One*. 6:e19402.

Zerbino, D.R. & Birney, E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–9.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*. 29:2869–76.

Zhang, L., Peng, J., Li, X., Cui, C., Sun, J. & Yang, G. 2016. Characterization of genome-wide microsatellites of *Saccharina japonica* based on a preliminary assembly of Illumina sequencing reads. *J. Ocean Univ. China*. 15:523–32.

Zhang, N., Zhang, L., Tao, Y., Guo, L., Sun, J., Li, X., Zhao, N. et al. 2015. Construction of a high density SNP linkage map of kelp (*Saccharina japonica*) by sequencing Taq I site associated DNA and mapping of a sex determining locus. *BMC Genomics*. 16:1–11.

Zhao, J., Jiang, P., Li, N., Wang, J.F., Liu, Z.Y. & Qin, S. 2010. Analysis of genetic variation within and among *Ulva pertusa* (Ulvaceae, Chlorophyta) populations using ISSR markers. *Chinese Sci. Bull.* 55:705–11.

Zhao, Z., Guo, C., Sutharzan, S., Li, P., Echt, C.S., Zhang, J. & Liang, C. 2014. Genome-Wide Analysis of Tandem Repeats in Plants and Green Algae. *Genes|Genomes|Genetics*. 4:67–78.

Zhu, H., Zhao, Z.-J., Liu, X., Song, H.-Y., Liu, G.-X. & Hu, Z.-Y. 2018. Molecular phylogeny and morphological diversity of inland *Cladophora* (Cladophorales, Ulvophyceae) from China. *Phycologia*. 57:191–208.

Zuccarello, G.C. & Lokhorst, G.M. 2005. Molecular phylogeny of the genus *Tribonema* (Xanthophyceae) using *rbc*L gene sequence data: monophyly of morphologically simple algal species. *Phycologia*. 44:384–92.

APPENDICES

Appendix 1.1. Species of the endemic Baikalian Cladophorales including synonyms (see Boedeker et al. 2018 and references therein)

Taxa	Synonyms
Chaetocladiella litoralis (Skabichevsky) Meyer & Skabichevsky	*Chaetomorpha litoralis* Skabichevsky, *Cladochaete litoralis* (Skabichevsky) Meyer & Skabichevsky
Chaetocladiella microscopica (Meyer) Meyer	*Chaetomorpha microscopica* Meyer, *Chaetomorpha solitaria* Skabichevsky, *Cladochaete microscopica* (Meyer) Meyer & Skabichevsky
Chaetocladiella pumila (Meyer) Meyer & Skabichevsky	*Chaetomorpha pumila* Meyer, *Cladochaete pumila* (Meyer) Meyer & Skabichevsky
Chaetomorpha baicalensis Meyer	
Chaetomorpha curta Skabichevsky	*Chaetomorpha baicalensis* var. *curta* Skabichevsky
Chaetomorpha moniliformis Skabichevsky	
Cladophora compacta (Meyer) Meyer ex Hollerbach	*Aegagropila compacta* Meyer
Cladophora floccosa Meyer	
Cladophora floccosa Meyer var. *irregularis* Skabichevsky	
Cladophora globulus (Meyer) Meyer	*Aegagropila globulus* Meyer
Cladophora kursanovii Skabichevsky	*Aegagropila pygmaea* Meyer
Cladophora kusnetzowii Meyer = *Aegagropila linnaei* Kützing	*Aegagropila linnaei*
Cladophora meyeri Skabichevsky	*Cladophora humilis* Meyer
Cladophora meyeri Skabichevsky var. *gracilior* (Meyer) Hollerbach	*Cladophora humilis* var. *gracilior* Meyer
Cladophora pulvinata (Meyer) Meyer	*Aegagropila pulvinata* Meyer
Gemmiphora compacta Skabichevsky	

Appendix 1.2. Bayesian Inference phylogram of the Baikalian Cladophorales inferred from rDNA small subunit (SSU) and partial large subunit (LSU) sequences including several samples of other Cladophoraceae and Pseudocladophoraceae as outgroup (collapsed for easy visualization). Phylogeny shows monophyly of the Baikal clade, nested within *Rhizoclonium*. Numbers in branches indicate posterior probabilities from Bayesian Inference (below) and maximum-likelihood bootstrap values (above). Only bootstrap values > 50% and posterior probabilities > 0.9 shown (from Boedeker et al. 2018).

Appendix 2.1. Samples used in the SSR trials in Chapter 2. * = individuals which were not clearly identified. Locality approximate coordinates in Appendix 2.2.

Taxon	Collection number	Locality	Collection date
Chaetocladiella litoralis	111	Oto-Kushun Bay, Maloe More	13/09/2014
Chaetocladiella litoralis	549	Yelochin Cape	24/09/2014
Chaetocladiella litoralis	554	Yelochin Cape	24/09/2014
Chaetocladiella pumila	573	Cape Sunduk	25/09/2014
Chaetocladiella pumila	583	Cape Sunduk	25/09/2014
Chaetocladiella pumila	664	Aya Bay	25/09/2014
Chaetomorpha baicalensis	51	Oto-Kushun Bay, Maloe More	13/09/2014
Chaetomorpha baicalensis	60	Oto-Kushun Bay, Maloe More	13/09/2014
Chaetomorpha baicalensis	642	Aya Bay	25/09/2014
Chaetomorpha curta attached form	428	Ayaya Bay	19/09/2014
Chaetomorpha curta unattached form	454	Onokochanskaya Bay	22/09/2014
Chaetomorpha curta unattached form	638	Aya Bay	25/09/2014
Chaetomorpha moniliformis	62	Oto-Kushun Bay, Maloe More	13/09/2014
Chaetomorpha moniliformis	159	Barguzinsky Bay	15/09/2014
Chaetomorpha moniliformis	510	Yelochin Cape	24/09/2014
Cladophora compacta	93	Oto-Kushun Bay, Maloe More	13/09/2014
Cladophora compacta	189	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	16/09/2014
Cladophora compacta	343	Urbikan (ca. 50 nautical miles N. of Davsha)	18/09/2014
Cladophora compacta	538	Yelochin Cape	24/09/2014
Cladophora floccosa var. *floccosa*	303	Akademicheskaya Bay, Ushkani Islands	16/09/2014
Cladophora floccosa var. *floccosa*	584	Cape Sunduk	25/09/2014
Cladophora floccosa var. *floccosa*	675	Aya Bay	25/09/2014
Cladophora floccosa var. *irregularis*	48	Oto-Kushun Bay, Maloe More	13/09/2014
Cladophora floccosa var. *irregularis*	692	Aya Bay	25/09/2014
*Cladophora globulus/C. compacta**	714	Bolshie Koty	27/09/2014

*Cladophora globulus/C. compacta**	802	near Babushkin	29/09/2014
*Cladophora globulus/C. pulvinata**	199	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	16/09/2014
*Cladophora globulus/C. pulvinata**	451	Onokochanskaya Bay	22/09/2014
Cladophora kursanovii	383	Urbikan (ca. 50 nautical miles N of Davsha)	18/09/2014
Cladophora kursanovii	516	Yelochin Cape	24/09/2014
Cladophora kursanovii	526	Yelochin Cape	24/09/2014
Cladophora kursanovii	864	SW coast	30/09/2014
Cladophora meyeri var. *gracilior*	285	Akademicheskaya Bay, Ushkani Islands	16/09/2014
Cladophora meyeri var. *gracilior*	536	Yelochin Cape	24/09/2014
Cladophora meyeri var. *meyeri*	100	Oto-Kushun Bay, Maloe More	13/09/2014
Cladophora meyeri var. *meyeri*	606	Cape Sunduk	25/09/2014
Cladophora meyeri var. *meyeri*	615	Cape Sunduk	25/09/2014
Cladophora pulvinata	210	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	16/09/2014
Cladophora pulvinata	617	Cape Sunduk	25/09/2014
Cladophora pulvinata	625	Cape Sunduk	25/09/2014
Gemmiphora compacta	779	Bolshie Koty	28/09/2014
Gemmiphora compacta	857	SW coast	30/09/2014
Rhizoclonium sp.	273	Akademicheskaya Bay, Ushkani Islands	16/09/2014
Rhizoclonium sp.	628	Cape Sunduk	25/09/2014

Appendix 2.2. Approximate locality coordinates

Locality	Longitude	Latitude
Akademicheskaya Bay, Ushkani Islands	108° 39' 42" E	53° 50' 29" N
Aya Bay	106° 48' 17" E	52° 55' 01" N
Ayaya Bay	109° 52' 31" E	55° 27' 25" N
Barguzinsky Bay	108° 57' 31" E	53° 23' 55" N
Bolshie Koty	105° 04' 59" E	51° 54' 01" N
Cape Sunduk	107° 33' 18" E	53° 26' 30" N
near Babushkin	105° 50' 50" E	51° 43' 28" N
Onokochanskaya Bay	109° 13' 18" E	55° 32' 17" N
Oto-Kushun Bay, Maloe More	107° 13' 58" E	53° 17' 51" N
Sviatoy Nos (southern cape of Chivyrkui Peninsula)	108° 30' 57" E	53° 29' 36" N
SW coast	104° 03' 14" E	51° 44' 06" N
Urbikan (ca. 50nm N of Davsha)	109° 35' 28" E	54° 46' 38" N
Yelochin Cape	108° 39' 57" E	54° 32' 12" N
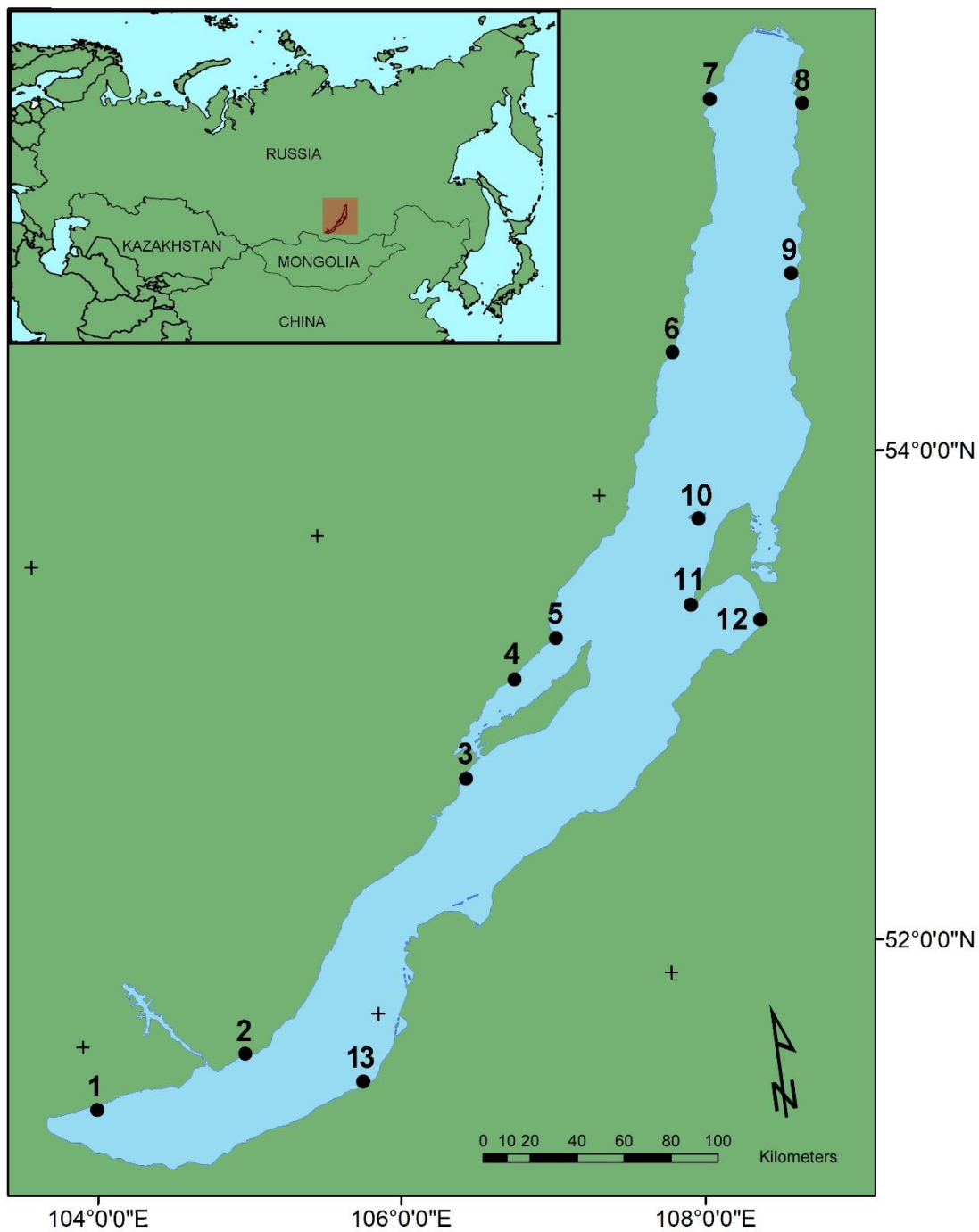
Appendix 2.3. Sampling locations in Lake Baikal. 1 = SW coast; 2 = Bolshie koty; 3 = Aya Bay; 4 = Oto-Kushun Bay, Maloe More; 5 = Cape Sunduk; 6 = Yelochin Cape; 7 = Onokochanskaya Bay; 8 = Ayaya Bay; 9 Urbikan; 10 = Akademicheskaya Bay, Ushkani Island; 11 = Sviatoy Nos; 12 = Barguzinsky Bay; 13 = Babushkin. Approximate coordinates in Appendix 2.2.

Appendix 2.4. Properties of the sequencing data. Gb = gigabases.

Sample	Raw bases (Gb)	Clean bases (Gb)	GC content in clean reads (%)	Clean reads pairs	Unique reads	Total contigs assembled	Average contig length	Total bases in contigs	GC content in contigs (%)
Chaetomorpha baikalensis	12.85	10.75	51.21	21326541	16052785	135578	376	51082622	51.54
Chaetomorpha moniliformis	12.08	10.1	43.12	20192401	17157064	318757	273	87045334	41.13
Cladophora kursanovii	12.78	9.35	40.3	18693438	16679149	105464	389	41086739	43.52

Appendix 3.1. Maximum likelihood phylogenetic tree inferred from internal transcribed spacer (ITS) sequences of the Baikalian Cladophorales species flock. Maximum-likelihood bootstrap support (1000 replicates) is indicated (>85%). The scale bar indicates substitutions per site.
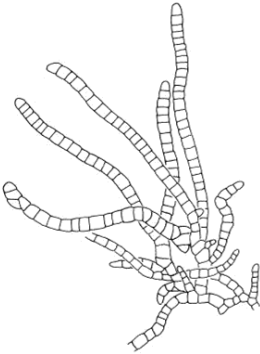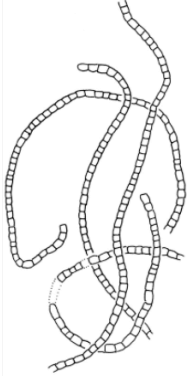
Appendix 3.2. Samples used for species delimitation
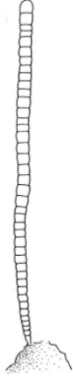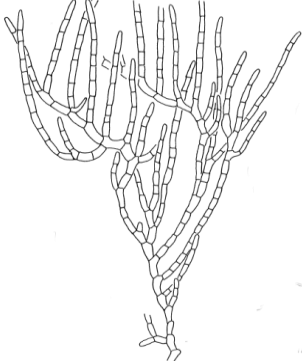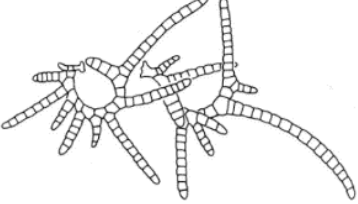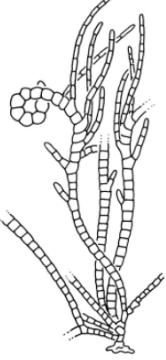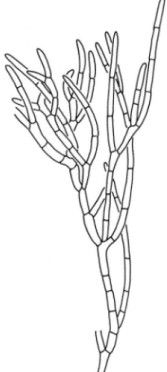
Taxon	Localities	No. of samples	Collection numbers	Collection date
Cladophora compacta	Akademicheskaya Bay, Ushkani Islands	10	242-251	16/09/2014
Cladophora compacta	Amundakan	10	363-372	19/09/2014
Cladophora compacta	Aya Bay	10	681-690	25/09/2014
Cladophora compacta	Ayaya Bay	10	396-405	19/09/2014
Cladophora compacta	Barguzinsky Bay	10	169-178	15/09/2014
Cladophora compacta	Cape Sunduk	10	562-571	25/09/2014
Cladophora compacta	Chargoyskaya Bay, Olkhon Island, Maloe More	1	4	12/09/2014
Cladophora compacta	Ighimey Bay, Eastern Olkhon Island	10	116-125	14/09/2014
Cladophora compacta	Davsha village	10	322-331	18/09/2014
Cladophora compacta	Oto-Kushun Bay, Maloe More	10	84-93	13/09/2014
Cladophora compacta	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	189-198	16/09/2014
Cladophora compacta	SW coast	6	847-852	30/09/2014
Cladophora compacta	Urbikan	10	343-352	18/09/2014
Cladophora compacta	Yelochin Cape	10	538-547	24/09/2014
Cladophora floccosa var. floccosa	Akademicheskaya Bay, Ushkani Islands	9	301, 303-310	16/09/2014
Cladophora floccosa var. floccosa	Aya Bay	6	675-680	25/09/2014
Cladophora floccosa var. floccosa	Bolshie Koty	30	736-755, 757-766	27/09/2014
Cladophora floccosa var. floccosa	Cape Sunduk	10	584-593	25/09/2014
Cladophora floccosa var. floccosa	Chargoyskaya Bay, Olkhon Island, Maloe More	10	28-37	12/09/2014
Cladophora floccosa var. irregularis	Akademicheskaya Bay, Ushkani Islands	10	274-283	16/09/2014
Cladophora floccosa var. irregularis	Aya Bay	20	692-701, 703-712	25/09/2014
Cladophora floccosa var. irregularis	Ayaya Bay	7	388-394	19/09/2014
Cladophora floccosa var. irregularis	Davsha	10	353-362	18/09/2014
Cladophora floccosa var. irregularis	Onokochanskaya Bay	8	476-483	22/09/2014
Cladophora floccosa var. irregularis	Oto-Kushun Bay, Maloe More	10	39-48	13/09/2014
Cladophora globulus/compacta	Bolshie Koty	20	714-723, 725, 727-735	27/09/2014
Cladophora globulus/compacta	near Babushkin	11	192-802	29/09/2014
Cladophora globulus/pulvinata	Onokochanskaya Bay	11	442-452	22/09/2014
Cladophora globulus/pulvinata	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	199-208	16/09/2014
Cladophora kursanovii	Akademicheskaya Bay, Ushkani Islands	5	296-300	16/09/2014
Cladophora kursanovii	Baikalsk	10	825-834	30/09/2014
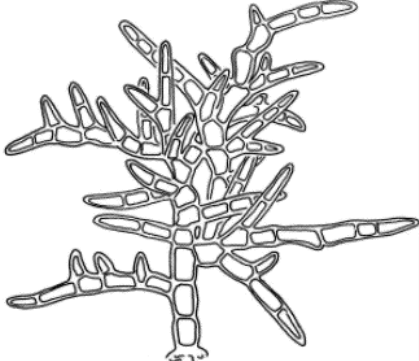
Cladophora kursanovii	Bolshie Koty	9	769, 771-778	27/09/2014
Cladophora kursanovii	Cape Sunduk	11	595-605	25/09/2014
Cladophora kursanovii	Ighimey Bay, Eastern Olkhon Island	10	146-155	14/09/2014
Cladophora kursanovii	near Babushkin	10	803-812	29/09/2014
Cladophora kursanovii	Onokochanskaya Bay	9	465, 466, 468-474	22/09/2014
Cladophora kursanovii	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	221-230	16/09/2014
Cladophora kursanovii	SW coast	10	864-873	30/09/2014
Cladophora kursanovii	Urbikan	5	383-387	18/09/2014
Cladophora kursanovii	Yelochin Cape	11	516-526	24/09/2014
Cladophora meyeri var. *gracilior*	Akademicheskaya Bay, Ushkani Islands	10	285-294	16/09/2014
Cladophora meyeri var. *gracilior*	Davsha	10	333-342	18/09/2014
Cladophora meyeri var. *gracilior*	Yelochin Cape	10	527-536	24/09/2014
Cladophora meyeri var. *meyeri*	Akademicheskaya Bay, Ushkani Islands	10	263-272	16/09/2014
Cladophora meyeri var. *meyeri*	Cape Sunduk	10	606-615	25/09/2014
Cladophora meyeri var. *meyeri*	Oto-Kushun Bay, Maloe More	10	95-104	13/09/2014
Cladophora pulvinata	Cape Sunduk	10	617-626	25/09/2014
Cladophora pulvinata	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	210-219	16/09/2014
Chaetomorpha baikalensis	Aya Bay	10	642-651	25/09/2014
Chaetomorpha baikalensis	Ludar' Bay	10	484-493	22/09/2014
Chaetomorpha baikalensis	Oto-Kushun Bay, Maloe More	10	51-60	13/09/2014
Chaetomorpha curta attached form	Ayaya Bay	11	419-429	19/09/2014
Chaetomorpha curta attached form	Babushkin	10	814-823	29/09/2014
Chaetomorpha curta attached form	Onokochanskaya Bay	10	454-463	22/09/2014
Chaetomorpha curta attached form	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	231-240	16/09/2014
Chaetomorpha curta unattached form	Aya Bay	11	629-639	25/09/2014
Chaetomorpha curta unattached form	Ayaya Bay	10	430-439	19/09/2014
Chaetomorpha moniliformis	Akademicheskaya Bay, Ushkani Islands	10	243-262	16/09/2014
Chaetomorpha moniliformis	Amundakan	10	373-382	19/09/2014
Chaetomorpha moniliformis	Aya Bay	11	653-663	25/09/2014
Chaetomorpha moniliformis	Ayaya Bay	11	407-417	19/09/2014
Chaetomorpha moniliformis	Barguzinsky Bay	10	159-168	15/09/2014
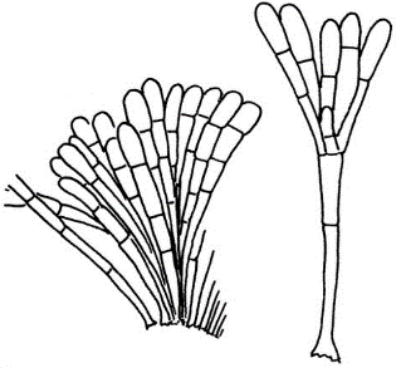
Chaetomorpha moniliformis	Chargoyskaya Bay, Olkhon Island, Maloe More	10	16-25	12/09/2014
Chaetomorpha moniliformis	Ighimey Bay, Eastern Olkhon Island	10	126-135	14/09/2014
Chaetomorpha moniliformis	Ludar' Bay	10	495-504	22/09/2014
Chaetomorpha moniliformis	Davsha	10	312-321	18/09/2014
Chaetomorpha moniliformis	Oto-Kushun Bay, Maloe More	10	62-71	13/09/2014
Chaetomorpha moniliformis	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	179-188	16/09/2014
Chaetomorpha moniliformis	Yelochin Cape	10	505-514	24/09/2014
Gemmiphora compacta	Bolshie Koty	10	779-788	28/09/2014
Gemmiphora compacta	SW coast	10	853-862	30/09/2014
Rhizoclonium sp.	Akademicheskaya Bay, Ushkani Islands	1	273	16/09/2014
Rhizoclonium sp.	Cape Sunduk	1	628	25/09/2014
Rhizoclonium sp.	Chivyrkui Bay	1	Q10	19/06/2011
Rhizoclonium sp.	Ushkani Islands	1	Q18	20/06/2011
Total		727		

Appendix 3.3. Morphological characteristics of the Baikalian Cladophorales species flock. Descriptions and drawings adapted from Boedeker et al. (2018).

Species	Brief description	Drawing
Chaetocladiella litoralis (Skabichevsky) Meyer & Skabichevsky	Branched filaments near the base but unbranched long filaments in the upper parts. Attached by hapteroid holdfast and secondary rhizoids growing from the basal cells.	
Chaetocladiella pumila (Meyer) Meyer & Skabichevsky	Branched filaments near the base, unbranched long filaments in the upper parts, or entirely unbranched. Solitary filaments of 1-2 mm high and up to up to 40 µm cell diameter Attached on small pebbles.	
Chaetomorpha baicalensis Meyer	Unbranched filaments forming entangled wooly masses. Unattached. Can be found rolled up forming balls balls.	
Chaetomorpha curta Skabichevsky	Rigid unbranched filaments. Attached and unattached forms. Unattached filaments of 500 µm thick with uniform diameter and light green. Attached filaments with a small basal cell. Less than 300 µm thick, darker compared to the unattached form. Found in deeper.	

Chaetomorpha moniliformis Skabichevsky	Unbranched filaments. Attached on rocks in solitary or forming small tufts. Filaments of 2-3 cm high with cells widest in middle of thalli, up to 500 µm cell diameter.	
Cladophora compacta (Meyer) Meyer ex Hollerbach	Densely branched filaments forming attached tufts less than 6 mm high. Often dense and hemispherical, falcate, apical branch insertion, basal parts less branched, long basal cells brush-shaped. Without secondary rhizoids.	
Cladophora floccosa Meyer	Densely branched filaments forming tufts up to 1.5cm in diameter attached to rocks by simple primary holdfast. Branches falcate, typically every cell producing laterals of similar length. Delayed cell wall formation and cells with thin walls.	
Cladophora floccosa Meyer var. *irregularis* Skabichevsky	Branched filaments forming tufts up to 1.5cm in diameter attached to rocks by simple primary holdfast. Can be found forming dense branches in basal parts, but typically forming very long unbranched filaments that occasionally produce short laterals. Delayed cell wall formation, cells with thin walls, pointy apical cells.	
Cladophora globulus (Meyer) Meyer	Densely branched forming small tufts with thin filaments. Basal branching growing in all directions. Falcate above, apical insertion of branches.	

Cladophora kursanovii Skabichevsky	Branched filaments forming creeping turfs less than 1mm high, insertion. Strongly attached to the substrate. Short cells with relatively large cell diameter up to 100 µm. Thick cell walls. Forming secondary rhizoids, some with hapteres Unique branch insertion.	
Cladophora meyeri Skabichevsky	Branched filaments forming minuscule spherical tufts up to 1mm. Attached to rocks and pebbles with simple primary rhizoids. Branches inserted irregularly.	
Cladophora meyeri Skabichevsky var. *gracilior* (Meyer) Hollerbach	Branched filaments forming small tufts. Attached to rocks and pebbles. Differentiates Differs from the typical *C. meyeri* by having thinner apical cells, longer cells and higher tufts.	
Cladophora pulvinata (Meyer) Meyer	Densely branched filaments forming small tufts up to 6mm high. Attached to the substratum. Branching is falcate with apical branch insertion. Secondary rhizoids present.	

Gemmiphora compacta Skabichevsky	Branched or unbranched filaments forming a minuscule thalli growing up to 1mm high and less than 10 cells high. Packed thalli attached inside cracks of rocks by tiny primary holdfast and rarely additional secondary rhizoids from basal cells. Displaying rounded apical cells.	

Appendix 3.4. Approximate locality coordinates.

Map number	Localities	Longitude	Latitude
1	SW coast	104° 03' 14.14" E	51° 44' 06.6" N
2	Bolshie Koty	105° 04' 59.59" E	51° 54' 01.1" N
3	Aya Bay	106° 48' 17.17" E	52° 55' 01.1" N
4	Chargoyskaya Bay	106° 56' 46.46" E	53° 05' 41.41" N
5	Oto-Kushun Bay	107° 13' 58.58" E	53° 17' 51.51" N
6	Ighimey Bay	107° 42' 51.51" E	53° 13' 08.8" N
7	Cape Sunduk	107° 33' 18.18" E	53° 26' 30.30" N
8	Yelochin Cape	108° 39' 57.57" E	54° 32' 12.12" N
9	Ludar Bay	109° 13' 37.37" E	55° 21' 57.57" N
10	Onokochanskaya Bay	109° 13' 18.18" E	55° 32' 17.17" N
11	Ayaya Bay	109° 52' 31.31" E	55° 27' 25.25" N
12	Amudakan	109° 41' 55.55" E	55° 07' 13.13" N
13	Urbikan	109° 35' 28.28" E	54° 46' 38.38" N
14	Davsha	109° 27' 28.28" E	54° 21' 10.10" N
15	Chivyrkui Bay	109° 01' 50.50" E	53° 40' 39.39" N
16	Akademicheskaya	108° 39' 42.42" E	53° 50' 29.29" N
17	Sviatoy Nos	108° 30' 57.57" E	53° 29' 36.36" N
18	Barguzinsky	108° 57' 31.31" E	53° 23' 55.55" N
19	Babushkin	105° 50' 50.50" E	51° 43' 28.28" N
20	Baikalsk	104° 35' 19.19" E	51° 29' 11.11" N
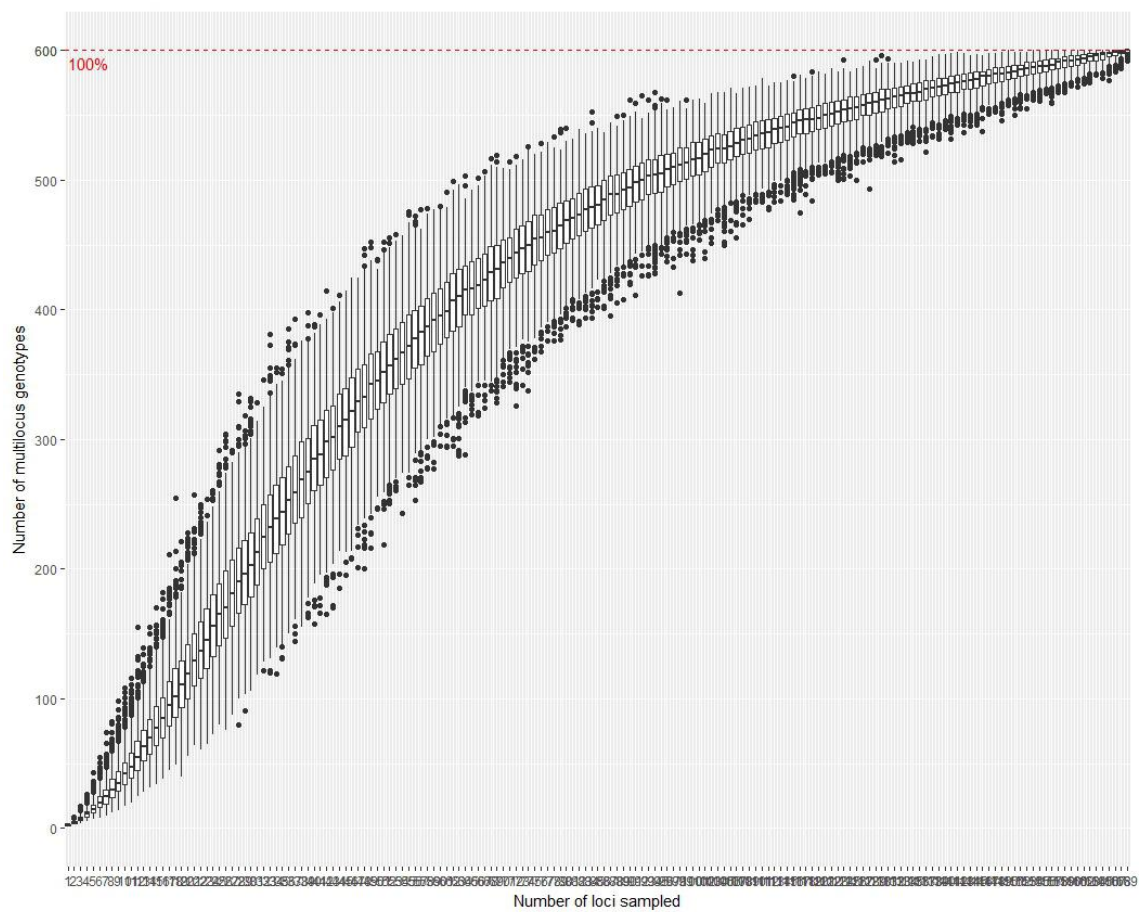
Appendix 3.5. Genotype accumulation curve (Loci versus genotypes) obtained in *poppr* for all the individuals and using all the loci combined.

DeltaK = mean(|L''(K)|) / sd(L(K))
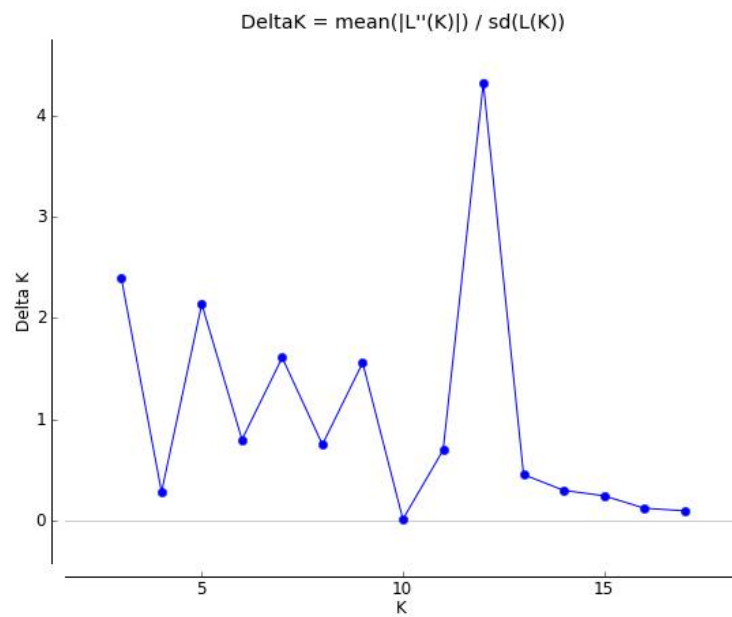
Appendix 3.6. Delta K estimation in STRUCTURE, indicating that K = 12 has the highest value and the number of clusters that fits the data set.

Appendix 3.7. STRUCTURE assignment of individuals to a given cluster. Total number of clusters = 12. Total # = total number of individuals per morphotaxa analyzed. Major cluster = number of the cluster with most assigned individuals of a single morphotaxon. # assigned to major cluster = individuals assigned to major cluster. # assigned to another clusters = number of individuals assigned to another clusters from other morphotaxa. Graphic representation in Fig. 3.3a.

Morphotaxa	Total #	Major cluster	# assigned to major cluster	# assigned to another clusters
Ch. baicalensis	30	3	30	-
Ch. curta, attached	41	6	41	-
Ch. curta, unattached	21	6	11	10
Ch. moniliformis	122	5	122	-
Cl. compacta	127	1	96	31
Cl. floccosa var. *floccosa*	64	9	31	33
Cl. floccosa var. *irregularis*	65	7	40	25
Cl. globulus/ Cl. compacta	31	12	15	16
Cl. globulus/ Cl. pulvinata	21	12	19	2
Cl. kursanovii	100	2	100	-
Cl. meyeri var. *gracilior*	30	4	20	10
Cl. meyeri var. *meyeri*	30	4	20	10
Cl. pulvinata	20	9	13	7
G. compacta	20	7	20	-
Rhizoclonium sp.	4	6	4	-

Appendix 3.8. Cluster assignment for each sample obtained with STRUCTURE (STR) and Gaussian Clustering (GC).

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
Chaetomorpha baicalensis	51	3	4		240	6	10
	52	3	4		419	6	10
	53	3	4		420	6	10
	54	3	4		421	6	10
	55	3	4		422	6	10
	56	3	4		423	6	10
	57	3	4		424	6	10
	58	3	4		425	6	10
	59	3	4		426	6	10
	60	3	4		427	6	10
	484	3	4		428	6	8
	485	3	4		429	6	10
	486	3	4		454	6	8
	487	3	4		455	6	10
	488	3	4		456	6	10
	489	3	4		457	6	10
	490	3	4		458	6	10
	491	3	4		459	6	10
	492	3	4		460	6	10
	493	3	4		461	6	10
	642	3	4		462	6	10
	643	3	4		463	6	10
	644	3	4		814	6	10
	645	3	4		815	6	10
	646	3	4		816	6	10
	647	3	4		817	6	10
	648	3	4		818	6	10
	649	3	4		820	6	10
	650	3	4		821	6	10
	651	3	4		822	6	10
Chaetomorpha curta attached form	231	6	10		823	6	10
	232	6	10		824	6	8
	233	6	10	*Chaetomorpha curta* unattached form	430	6	10
	234	6	10		431	6	10
	235	6	10		432	6	10
	236	6	10		433	6	10
	237	6	10		434	6	10
	238	6	10		435	6	10
	239	6	10		436	6	10

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	437	6	10		135	5	9
	438	6	10		159	5	8
	439	6	10		160	5	7
	629	5	7		161	5	7
	630	5	7		162	5	8
	631	5	7		163	5	7
	632	5	7		164	5	7
	633	5	7		165	5	7
	634	5	7		166	5	7
	635	5	7		167	5	7
	636	5	7		168	5	8
	637	5	7		179	5	9
	638	5	7		180	5	7
	639	5	7		181	5	7
Chaetomorpha moniliformis	16	5	9		182	5	9
	17	5	9		183	5	7
	18	5	7		184	5	7
	19	5	8		185	5	9
	20	5	8		186	5	7
	21	5	8		187	5	8
	22	5	8		188	5	7
	23	5	7		253	5	8
	24	5	8		254	5	7
	25	5	7		255	5	7
	62	5	9		256	5	7
	63	5	9		257	5	7
	64	5	7		258	5	8
	65	5	8		259	5	7
	66	5	7		260	5	7
	67	5	9		261	5	7
	68	5	7		262	5	7
	69	5	9		312	5	8
	70	5	9		313	5	9
	71	5	9		314	5	8
	126	5	7		315	5	7
	127	5	7		316	5	7
	128	5	8		317	5	7
	129	5	7		318	5	7
	130	5	8		319	5	8
	131	5	7		320	5	8
	132	5	9		321	5	9
	133	5	8		373	5	9
	134	5	7		374	5	9

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	375	5	8		657	5	7
	376	5	7		658	5	9
	377	5	7		659	5	8
	378	5	7		660	5	7
	379	5	8		661	5	7
	380	5	7		662	5	7
	381	5	7		663	5	8
	382	5	8	*Cladophora compacta*	4	8	14
	407	5	9		84	10	13
	408	5	8		85	10	13
	409	5	8		86	10	13
	410	5	7		87	10	13
	411	5	8		88	10	13
	412	5	8		89	10	13
	413	5	8		90	10	13
	414	5	8		91	10	13
	415	5	7		92	10	13
	416	5	8		93	10	13
	417	5	8		116	1	1
	495	5	8		117	1	1
	496	5	7		118	1	1
	497	5	7		119	1	1
	498	5	8		120	1	1
	499	5	8		121	1	1
	500	5	7		122	1	1
	501	5	7		123	1	1
	502	5	7		124	1	1
	503	5	8		125	1	1
	504	5	7		169	1	1
	505	5	8		170	1	1
	506	5	7		171	1	1
	507	5	7		172	1	1
	508	5	7		173	1	1
	509	5	7		174	4	5
	510	5	8		175	1	1
	511	5	7		176	1	1
	512	5	7		177	1	1
	513	5	9		178	1	1
	514	5	7		189	1	1
	653	5	7		190	1	1
	654	5	9		191	1	1
	655	5	8		192	1	1
	656	5	7		193	1	1

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	194	1	1		371	1	1
	195	1	1		372	1	1
	196	1	1		396	1	1
	197	1	1		397	1	1
	198	1	1		398	1	1
	242	1	1		399	1	1
	243	1	1		400	1	1
	244	1	1		401	1	1
	245	1	1		402	1	1
	246	1	1		403	1	1
	247	1	1		404	1	1
	248	1	1		405	1	1
	249	1	1		538	7	11
	250	1	14		539	8	11
	251	1	14		540	7	11
	322	1	1		541	7	11
	323	1	1		542	7	11
	324	1	1		543	1	1
	325	1	1		544	7	11
	326	1	14		545	7	11
	327	1	1		546	7	11
	328	1	1		547	1	1
	329	1	1		562	7	11
	330	1	1		563	7	11
	331	1	1		564	7	11
	343	1	1		565	7	11
	344	1	1		566	7	11
	345	1	1		567	7	11
	346	1	1		568	7	11
	347	1	1		569	7	11
	348	1	1		570	7	11
	349	8	11		571	7	11
	350	1	1		681	1	1
	351	1	1		682	1	1
	352	1	1		683	1	1
	363	1	1		684	1	1
	364	1	1		685	1	1
	365	1	1		686	1	1
	366	1	1		687	1	1
	367	1	1		688	1	1
	368	1	1		689	1	1
	369	1	1		690	1	1
	370	1	1		847	1	1

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	848	1	1		739	10	13
	849	1	1		740	9	13
	850	1	1		741	9	13
	851	1	1		742	9	13
	852	1	1		743	10	13
Cladophora floccosa var. *flocossa*	28	11	4		744	9	13
	29	11	4		745	9	13
	30	11	4		746	8	5
	31	11	4		747	1	1
	32	11	4		748	12	14
	33	11	12		749	12	1
	34	11	4		750	1	1
	35	11	4		751	8	1
	36	11	11		752	8	1
	37	11	11		753	1	1
	301	10	13		754	1	1
	303	7	11		755	8	11
	304	10	13		757	7	11
	305	11	12		758	9	13
	306	11	12		759	9	13
	307	11	12		760	9	13
	308	11	12		761	9	13
	309	11	12		762	9	13
	310	11	12		763	9	13
	584	9	13		764	8	14
	585	9	13		765	9	13
	586	9	13		766	10	13
	587	9	13	*Cladophora floccosa* var. *irregularis*	39	7	12
	588	9	13		40	7	11
	589	9	13		41	7	11
	590	9	13		42	7	11
	591	9	13		43	7	11
	592	9	13		44	7	11
	593	9	13		45	7	11
	675	9	13		46	7	11
	676	9	13		47	7	11
	677	9	13		48	7	11
	678	9	13		274	7	11
	679	9	13		275	7	11
	680	9	13		276	7	11
	736	9	13		277	7	11
	737	9	13		278	8	11
	738	9	13		279	7	11

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	280	7	11		707	10	13
	281	7	11		708	10	13
	282	7	11		709	10	13
	283	7	11		710	10	13
	353	8	11		711	10	13
	354	7	11		712	10	13
	355	7	11	*Cladophora globulus/compacta*	714	1	1
	356	7	11		715	1	14
	357	7	11		716	12	1
	358	7	11		717	12	14
	359	7	11		718	12	1
	360	8	11		719	1	1
	361	8	11		720	12	14
	362	7	11		721	12	14
	388	7	11		722	12	14
	389	7	11		723	1	14
	390	7	11		725	7	12
	391	7	11		727	7	11
	392	7	11		728	7	12
	393	7	11		729	7	12
	394	7	11		730	8	5
	476	7	11		731	7	11
	477	7	11		732	7	11
	478	7	11		733	7	12
	479	7	11		734	8	11
	480	7	11		735	7	11
	481	7	11		792	1	14
	482	8	13		793	12	14
	483	7	11		794	12	14
	692	9	13		795	12	14
	693	10	13		796	12	14
	694	10	13		797	12	14
	695	10	13		798	12	14
	696	10	13		799	12	1
	697	10	13		800	12	14
	698	10	13		801	12	14
	699	10	13		802	8	5
	700	10	13	*Cladophora globulus/pulvinata*	199	8	5
	701	10	13		200	12	14
	703	10	13		201	8	5
	704	10	13		202	12	14
	705	10	13		203	12	14
	706	10	13		204	12	1

134

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	205	12	14		386	2	3
	206	12	1		387	2	3
	207	12	1		465	2	3
	208	12	14		466	2	2
	442	12	14		468	2	2
	443	12	14		469	2	2
	444	12	14		470	2	2
	445	12	14		471	2	2
	446	12	14		472	2	2
	447	12	1		473	2	2
	448	12	14		474	2	2
	449	12	1		516	2	2
	450	12	14		517	2	2
	451	12	1		518	2	2
	452	12	1		519	2	2
Cladophora kursanovii	146	2	2		520	2	2
	147	2	2		521	2	2
	148	2	2		522	2	2
	149	2	2		523	2	2
	150	2	2		524	2	2
	151	2	2		525	2	2
	152	2	2		526	2	2
	153	2	2		595	2	2
	154	2	2		596	2	2
	155	2	2		597	2	2
	221	2	3		598	2	2
	222	2	2		599	2	2
	223	2	2		600	2	2
	224	2	2		601	2	2
	225	2	2		602	2	2
	226	2	2		603	2	2
	227	2	2		604	2	2
	228	2	2		605	2	3
	229	2	2		769	2	3
	230	2	2		771	2	2
	296	2	3		772	2	2
	297	2	3		773	2	2
	298	2	3		774	2	2
	299	2	3		775	2	2
	300	2	3		776	2	2
	383	2	3		777	2	2
	384	2	3		778	2	2
	385	2	3		803	2	2

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	804	2	2		337	4	6
	805	2	2		338	4	6
	806	2	2		339	4	6
	807	2	2		340	4	6
	808	2	2		341	4	6
	809	2	2		342	4	6
	810	2	2		527	8	5
	811	2	2		528	12	14
	812	2	2		529	12	14
	825	2	2		530	8	5
	826	2	2		531	8	5
	827	2	2		532	8	5
	828	2	2		533	8	5
	829	2	2		534	12	14
	830	2	2		535	8	5
	831	2	2		536	8	5
	832	2	2		95	4	6
	833	2	2		96	4	6
	834	2	2		97	4	6
	864	2	2		98	4	6
	865	2	2		99	4	6
	866	2	2		100	4	6
	867	2	2		101	4	6
	868	2	2		102	4	6
	869	2	2		103	4	6
	870	2	2		104	4	6
	871	2	2		263	4	6
	872	2	2		264	4	6
	873	2	2		265	4	6
Cladophora meyeri var. *gracilior*	285	4	6		266	4	6
	286	4	6		267	4	6
	287	4	6		268	4	6
	288	4	6		269	4	6
	289	4	5		270	4	6
	290	4	6		271	4	6
	291	4	6		272	4	6
	292	4	6		606	9	13
	293	4	6		607	9	13
	294	4	6		608	9	13
	333	4	6		609	9	13
	334	4	6		610	12	14
	335	4	6		611	8	5
	336	4	6		612	12	14

Morphotaxa	Sample number	STR cluster	GC cluster	Morphotaxa	Sample number	STR cluster	GC cluster
	613	8	5		780	7	12
	614	9	13		781	7	12
	615	8	5		782	7	12
Cladophora pulvinata	210	10	13		783	7	12
	211	9	13		784	7	12
	212	9	13		785	7	12
	213	9	13		786	7	12
	214	9	13		787	7	12
	215	9	13		788	7	11
	216	12	14	*Gemmiphora compacta*	853	7	12
	217	9	13		854	7	12
	218	7	11		855	7	12
	219	10	13		856	7	12
	617	9	13		857	7	12
	618	8	13		858	7	12
	619	9	13		859	7	12
	620	12	14		860	7	12
	621	9	13		861	7	12
	622	9	13		862	7	12
	623	9	13	*Rhizoclonium* sp.	Q10	6	8
	624	9	13		Q18	6	8
	625	10	13		273	6	8
	626	9	13		628	6	8
	779	7	11				

Appendix 3.9. Gaussian clustering scatterplot of the morphotaxa analyzed (by color).
Nonmetric multidimensional scaling using 4 dimensions. Only the two first
dimensions were used for the plot.

Appendix 3.10. Gaussian clustering assignment of individuals to a given cluster. Total number of clusters = 14. Total # = total number of individuals per morphotaxa analyzed. Major cluster = number of the cluster with most assigned individuals of a single morphotaxon (subcluster = number of the clusters with high morphotaxon assignment). # assigned to major cluster (subcluster) = individuals assigned to major cluster (or subcluster). # assigned to another clusters = number of individuals assigned to another clusters from other morphotaxa. Graphic representation in Fig. 3.3b.

Morphotaxa	Total #	Major cluster (subcluster)	# assigned to major cluster (subcluster)	# assigned to another clusters
Ch. baicalensis	30	4	30	-
Ch. curta, attached	41	10	38	3
Ch. curta, unattached	21	7 (10)	11(10)	-
Ch. moniliformis	122	7 (8,9)	62(60)	-
Cl. compacta	127	1	93	34
Cl. floccosa var. *floccosa*	64	13	35	29
Cl. floccosa var. *irregularis*	65	11	43	22
Cl. globulus/ Cl. compacta	31	14	31	16
Cl. globulus/ Cl. pulvinata	21	14	21	9
Cl. kursanovii	100	2 (3)	86(14)	-
Cl. meyeri var. *gracilior*	30	6	19	11
Cl. meyeri var. *meyeri*	30	6	20	10
Cl. pulvinata	20	13	17	3
G. compacta	20	12	18	2
Rhizoclonium sp.	4	8	4	-

Appendix 4.1. Samples used for the analyses for each species.

Taxon	Localities	No. of samples	Collection number	Collection date
Cladophora compacta	Akademicheskaya Bay, Ushkani Islands	10	242-251	16/09/2014
	Amundakan	10	363-372	19/09/2014
	Aya Bay	10	681-690	25/09/2014
	Ayaya Bay	10	396-405	19/09/2014
	Barguzinsky Bay	10	169-178	15/09/2014
	Ighimey Bay, Eastern Olkhon Island	10	116-125	14/09/2014
	Davsha village	10	322-331	18/09/2014
	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	189-198	16/09/2014
	SW coast	6	847-852	30/09/2014
	Urbikan	9	343-348, 350-352	18/09/2014
Subtotal		95		
Cladophora kursanovii	Akademicheskaya Bay, Ushkani Islands	5	296-300	16/09/2014
	Baikalsk	10	825-834	30/09/2014
	Bolshie Koty	9	769, 771-778	27/09/2014
	Cape Sunduk	11	595-605	25/09/2014
	Ighimey Bay, Eastern Olkhon Island	10	146-155	14/09/2014
	near Babushkin	10	803-812	29/09/2014
	Onokochanskaya Bay	9	465, 466, 468-474	22/09/2014
	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	221-230	16/09/2014
	SW coast	10	864-873	30/09/2014
	Urbikan	5	383-387	18/09/2014
	Yelochin Cape	11	516-526	24/09/2014
Subtotal		100		
Chaetomorpha moniliformis	Akademicheskaya Bay, Ushkani Islands	10	243-262	16/09/2014
	Amundakan	10	373-382	19/09/2014
	Aya Bay	11	653-663	25/09/2014
	Ayaya Bay	11	407-417	19/09/2014
	Barguzinsky Bay	10	159-168	15/09/2014
	Chargoyskaya Bay, Olkhon Island, Maloe More	10	16-25	12/09/2014
	Ighimey Bay, Eastern Olkhon Island	10	126-135	14/09/2014
	Ludar' Bay	10	495-504	22/09/2014
	Davsha	10	312-321	18/09/2014
	Oto-Kushun Bay, Maloe More	10	62-71	13/09/2014
	Sviatoy Nos (southern cape of Chivyrkui Peninsula)	10	179-188	16/09/2014
	Yelochin Cape	10	505-514	24/09/2014
Subtotal		122		
Total		317		

Appendix 4.2. Approximate locality coordinates of the studied species.

Map number	Localities	Longitude	Latitude
1	SW coast	104° 03' 14.14" E	51° 44' 06.6" N
2	Bolshie Koty	105° 04' 59.59" E	51° 54' 01.1" N
3	Aya Bay	106° 48' 17.17" E	52° 55' 01.1" N
4	Chargoyskaya Bay	106° 56' 46.46" E	53° 05' 41.41" N
5	Oto-Kushun Bay	107° 13' 58.58" E	53° 17' 51.51" N
6	Ighimey Bay	107° 42' 51.51" E	53° 13' 08.8" N
7	Cape Sunduk	107° 33' 18.18" E	53° 26' 30.30" N
8	Yelochin Cape	108° 39' 57.57" E	54° 32' 12.12" N
9	Ludar Bay	109° 13' 37.37" E	55° 21' 57.57" N
10	Onokochanskaya Bay	109° 13' 18.18" E	55° 32' 17.17" N
11	Ayaya Bay	109° 52' 31.31" E	55° 27' 25.25" N
12	Amudakan	109° 41' 55.55" E	55° 07' 13.13" N
13	Urbikan	109° 35' 28.28" E	54° 46' 38.38" N
14	Davsha	109° 27' 28.28" E	54° 21' 10.10" N
15	Akademicheskaya	108° 39' 42.42" E	53° 50' 29.29" N
16	Sviatoy Nos	108° 30' 57.57" E	53° 29' 36.36" N
17	Barguzinsky	108° 57' 31.31" E	53° 23' 55.55" N
18	Babushkin	105° 50' 50.50" E	51° 43' 28.28" N
19	Baikalsk	104° 35' 19.19" E	51° 29' 11.11" N

Appendix 4.3. Characteristics and number of alleles by locus. Total number of alleles per lucus, alleles per indicidual per locus, average number of alleles per locus, number of fixed alleles per locus, alleles with with a frequency <5% (aF) and polymorphic information content (PIC) are shown. * = Loci where individuals had no PCR product.

Locus name	Total number of alleles	Alleles per individual	Average number of alleles	Number of fixed alleles	aF	PIC
Ch. moniliformis						
5a	1	1	1	1	0	0.00
6	1	1	1	1	0	0.00
14	1	1	1	1	0	0.00
15	3	3	3	3	0	0.00
16	2	1 to 2	1.01	1	1	0.0161
34	1	0* to 1	0.94	0	0	0.1082
46	3	2 to 3	2	2	1	0.0054
50-15	20	1 to 5	2.1	0	12	0.0924
515-35	1	1	1	1	0	0.00
515-46	3	3	3	3	0	0.00
515-64	2	2	2	2	0	0.00
Cl. compacta						
5a	7	2 to 4	3.28	0	2	0.111
6	4	0* to 3	1.35	0	0	0.1841
14	13	1 to 6	3.55	0	6	0.1734
15	-	-		-	-	-
16	6	2 to 5	3.17	1	2	0.0794
34	20	0* to 6	1.85	0	13	0.0943
46	7	1 to 5	1.65	1	3	0.1506
50-15	24	1 to 6	2.79	0	13	0.1238
515-35	-	-		-	-	-
515-46	-	-		-	-	-
515-64	3	2 to 3	2.01	2	1	0.0337
Cl. kursanovii						
5a	7	2 to 4	3.72	1	1	0.2063
6	3	3	3	3	0	0.00
14	1	1	1	1	0	0.00
15	1	1	1	1	0	0.00
16	7	3 to 5	3.9	2	0	0.1964
34	11	3 to 7	4.76	0	0	0.2419
46	3	2 to 3	2.3	1	0	0.1833
50-15	17	1 to 5	1.71	1	14	0.0707
515-35	2	2	2	2	0	0.00
515-46	1	1	1	1	0	0.00
515-64	2	2	2	2	0	0.00

Appendix 4.4. Characteristics and genetic diversity of the alleles in each population for each species. Total number of individuals (N), number of alleles (*a*), mean value of unbiased diversity (*uh*), corrected ubiased diversity (removing identical genotypes), number of private alleles (*Pa*), number of identical genotypes and proportion of unique genotypes (G-1/N-1). * = The total data set is used for calculations.

| Population	N	*a*	*uh*	Corrected *uh*	*Pa*	Identical genotypes	G-1/N-1
Ch. moniliformis							
Chargoyskaya Bay	10	30	0.112	0.121	5	1	0.89
Oto-Kushun Bay	10	23	0.046	0.060	0	3	0.67
Ighimey Bay	10	22	0.058	0.064	0	2	0.78
Barguzinsky Bay	10	22	0.050	0.065	1	3	0.67
Sviatoy Nos	10	22	0.039	0.068	1	5	0.44
Akademicheskaya Bay	10	20	0.036	0.053	0	5	0.44
Davsha	10	22	0.049	0.058	0	3	0.67
Amundakan	10	21	0.043	0.063	0	4	0.56
Ayaya Bay	11	22	0.052	0.060	1	3	0.70
Ludar' Bay	10	22	0.044	0.056	0	4	0.56
Yelochin Cape	10	22	0.048	0.063	0	3	0.67
Aya Bay	11	22	0.047	0.072	0	5	0.50
Total*	122	38	0.052	0.067	8	41	0.66
Cl. compacta							
Ighimey Bay	10	36	0.081	-	2	0	1.00
Barguzinsky Bay	10	45	0.137	-	4	0	1.00
Sviatoy Nos	10	41	0.126	-	4	0	1.00
Akademicheskaya Bay	10	48	0.121	-	2	0	1.00
Davsha	10	43	0.115	-	5	0	1.00
Urbikan	9	37	0.118	-	2	0	1.00
Amundakan	10	40	0.115	-	3	0	1.00
Ayaya Bay	10	28	0.053	-	1	0	1.00
Aya Bay	10	33	0.068	-	2	0	1.00
SW coast	6	25	0.056	-	0	0	1.00
Total*	95	84	0.099	-	25	0	1.00
Cl. kursanovii							
Ighimey Bay	10	31	0.070	0.070	0	1	0.89
Sviatoy Nos	10	33	0.061	0.061	2	0	1.00
Akademicheskaya Bay	5	27	0.087	0.087	0	0	1.00
Urbikan	5	29	0.040	0.048	1	1	0.75
Onokochanskaya Bay	9	30	0.042	0.065	0	4	0.50
Yelochin Cape	11	36	0.101	0.108	4	1	0.90
Cape Sunduk	11	36	0.062	0.101	1	5	0.50
Bolshie Koty	9	32	0.062	0.066	1	1	0.88
Babushkin	10	35	0.085	0.094	2	2	0.78
Baikalsk	10	31	0.055	0.061	0	3	0.67
SW coast	10	31	0.063	0.080	0	3	0.67
Total*	100	55	0.066	0.076	11	21	0.79
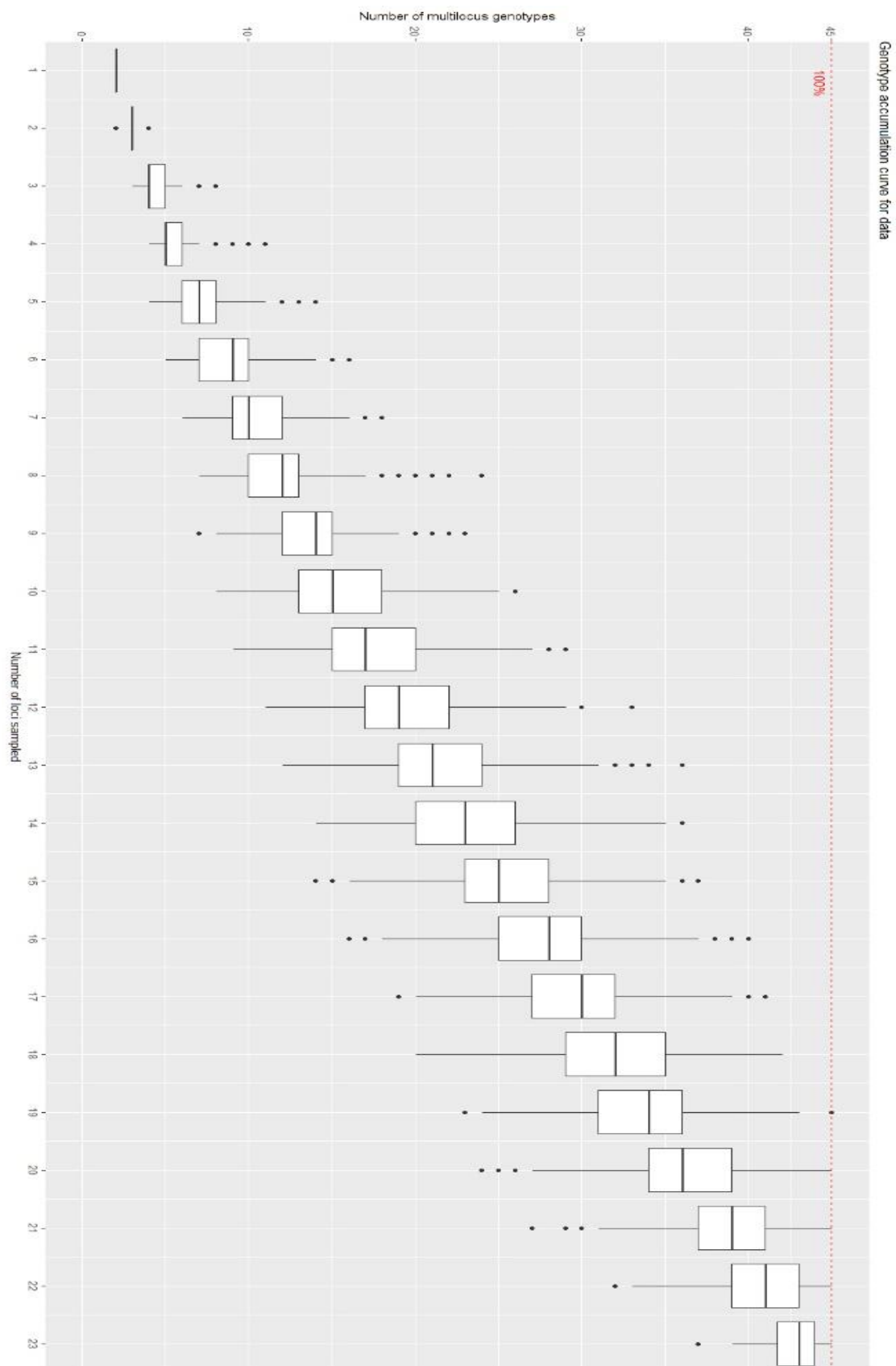
Appendix 4.5. Genotype accumulation curve obtained in *poprr* for *Chaetomorpha moniliformis*. Loci versus genotypes.
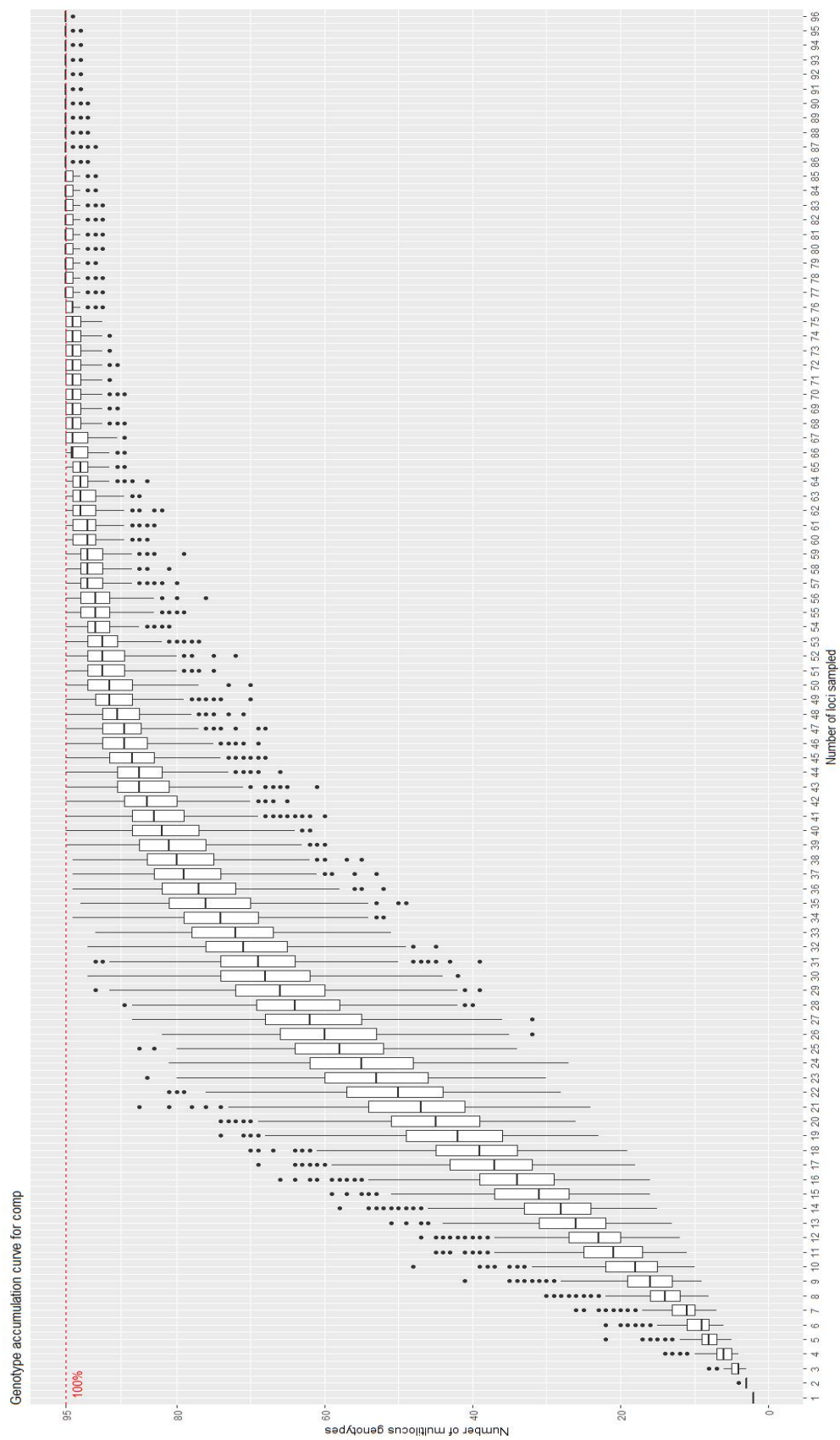
Appendix 4.6. Genotype accumulation curve obtained in *poppr* for *Cladophora compacta*. Loci versus genotypes.
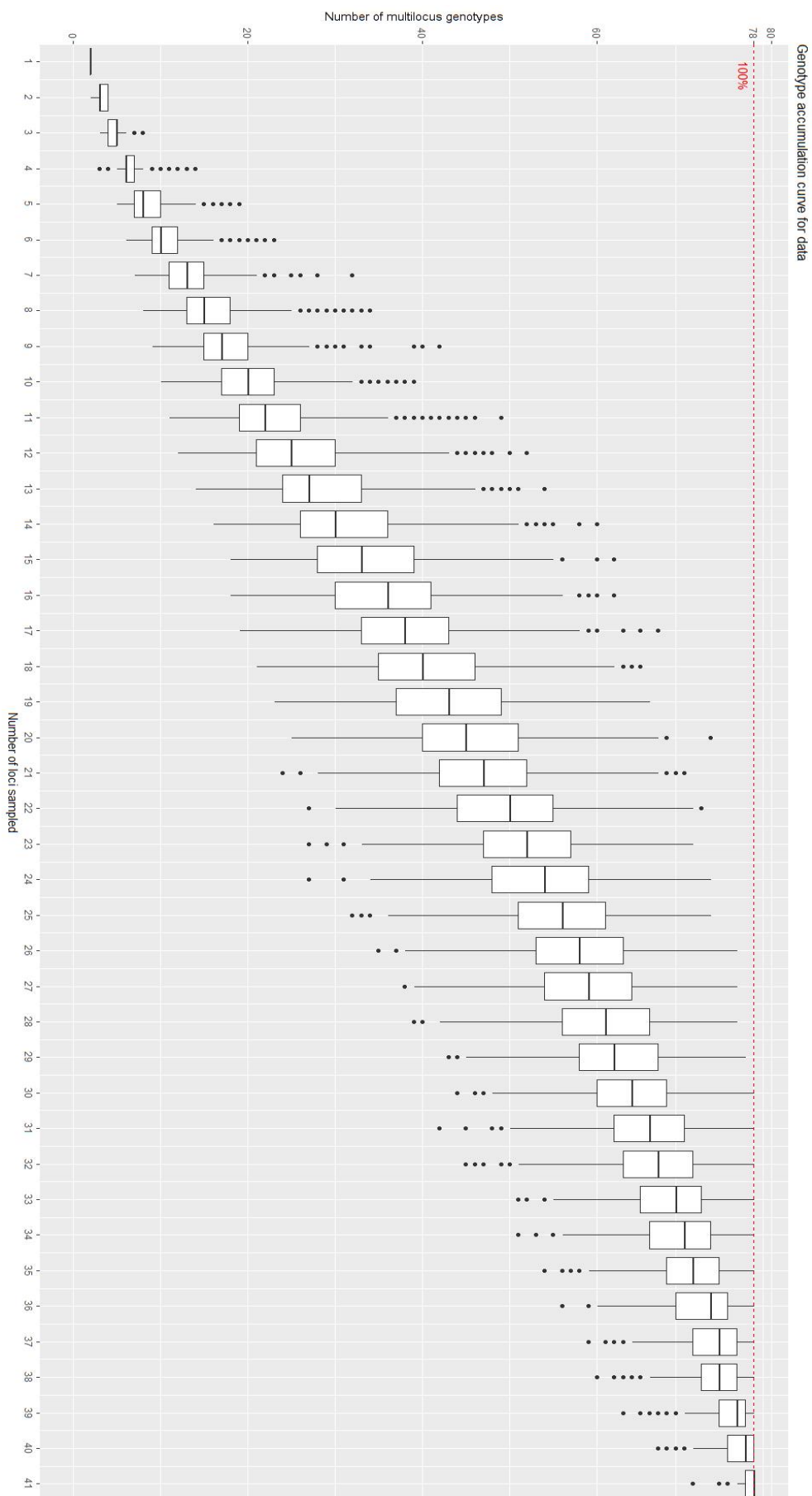
Appendix 4.7. Genotype accumulation curve obtained in *poprr* for *Chaetomorpha kursanovii*. Loci versus genotypes.
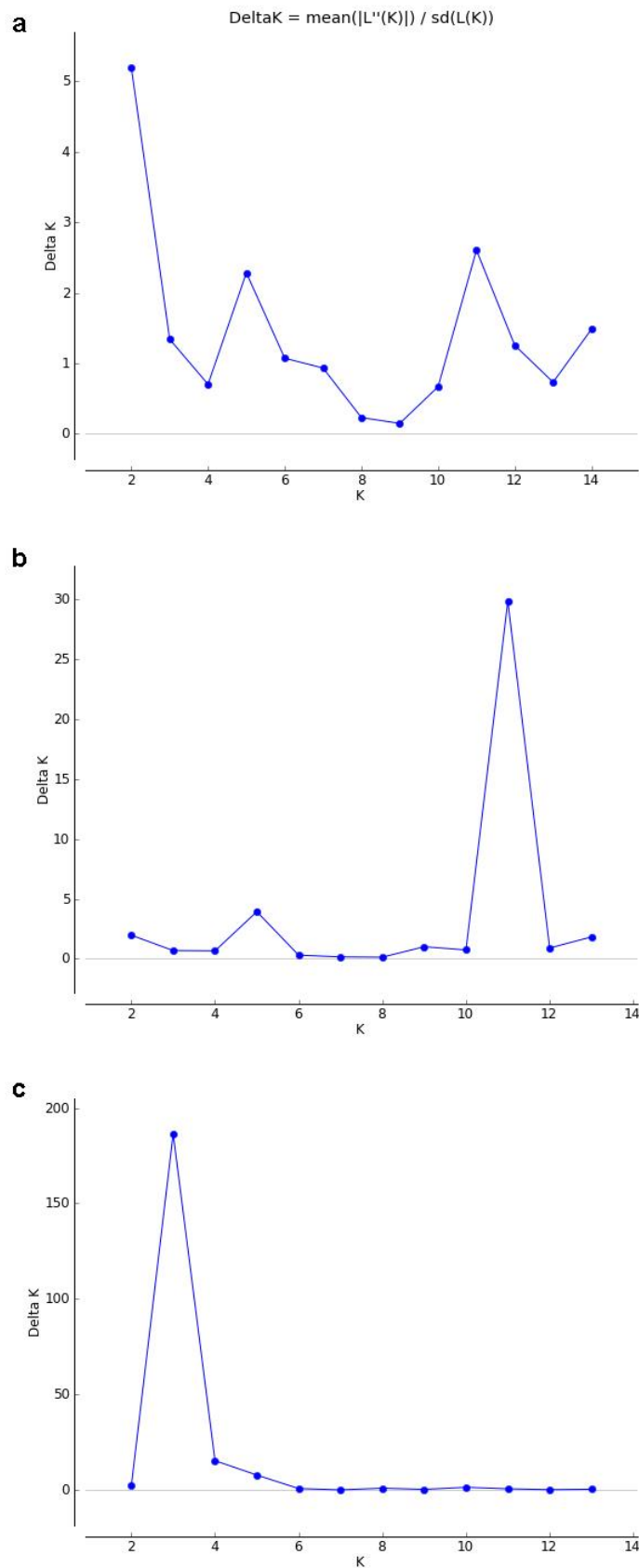
Appendix 4.8. Delta K estimation in STRUCTURE, indicating the highest value for K clusters for each species analyzed. a = *Chaetomorpha moniliformis*, b = *Cladophora compacta*, c = *Cladophora kursanovii*
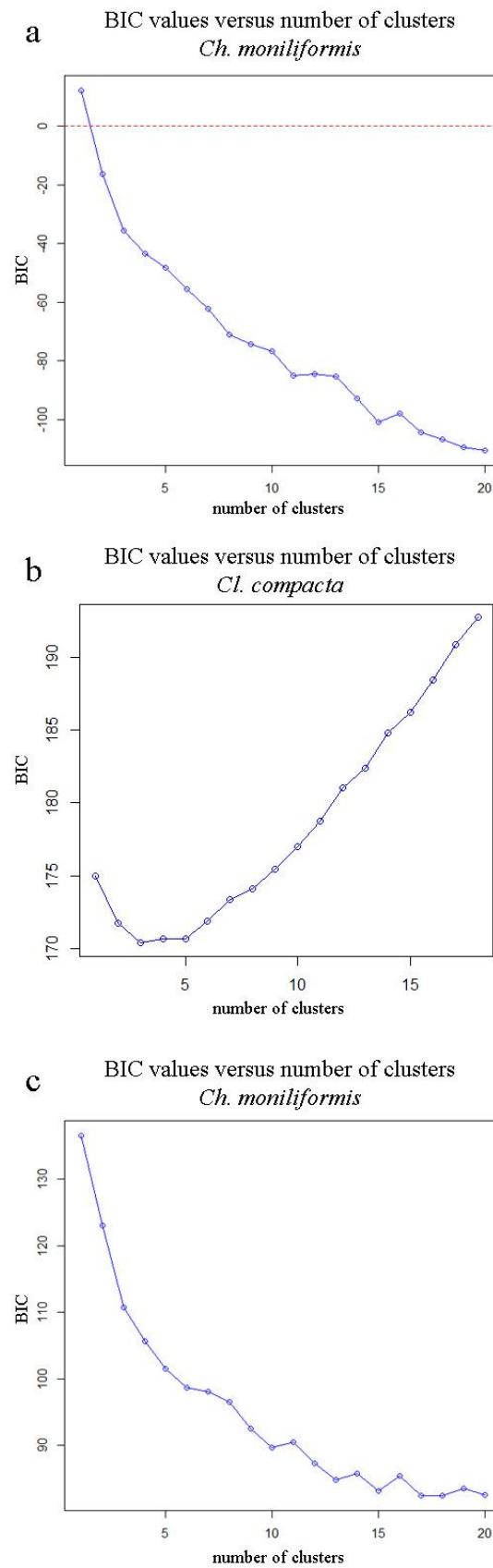
Appendix 4.9. Bayesian information criterion (BIC) values estimated for different number of clusters using DAPC. a = *Chaetomorpha moniliformis*, b = *Cladophora compacta*, c = *Cladophora kursanovii*
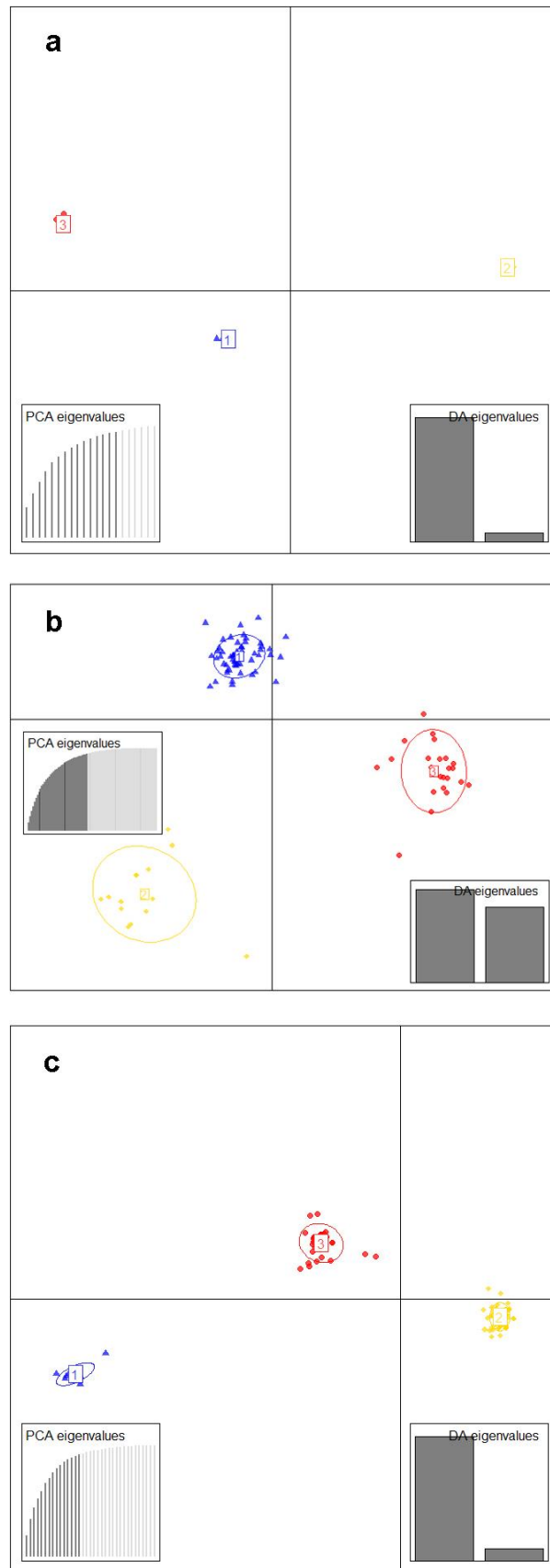
Appendix 4.10. DAPC scatterplots for the species analyzed using three predefined clusters (see text for details). PCA and DA eigenvalues used are shown. a = *Chaetomorpha moniliformis*, b = *Cladophora compacta*, c = *Cladophora kursanovii*
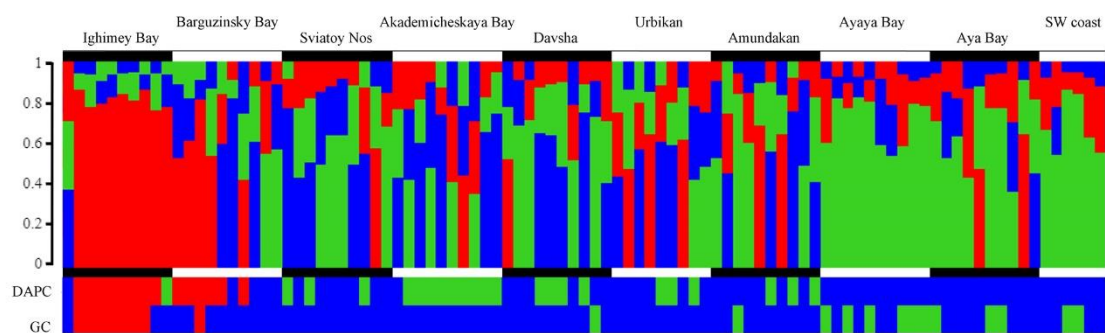
Appendix 4.11. Results of cluster analysis of *Cladophora kursanovii* in STRUCTURE using a K = 3 and comparison with DAPC and GC which also have three clusters.