Sedimentary DNA reveals centuries of hidden diversity in lake cyanobacterial communities

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
MARIE-EVE MONCHAMP
Master of science (MSc), Université de Montréal, Canada
Born on 28.02.1982
Citizen of Canada

Accepted on the recommendation of
PD Dr. Piet Spaak
Dr. Francesco Pomati
Prof. Dr. Irene Gregory Eaves
Prof. Dr. Jukka Jokela

2017
The major part of this thesis was conducted at the Swiss Federal Institute for Environmental Science and Technology (Eawag/ETH) in the Department of Aquatic Ecology partly in the Department of Surface Water Research and Management in Dübendorf, Switzerland. Financial support was provided by a grand from the Swiss Enlargement Contribution under project IZERZO-142165 (CyanoArchive) to Piet Spaak in the framework of the Romanian-Swiss Research Program.

Cover illustration: Photos of sediment cores collected in six peri-alpine lakes that were part of this PhD project. From front to back: (top) Lake Zurich, Hallwilersee, Lake Maggiore; (bottom) Lake Constance, Lake Lugano, Greifensee.
This thesis is based on the following papers:


Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>7</td>
</tr>
<tr>
<td>Zusammenfassung</td>
<td>9</td>
</tr>
<tr>
<td>CHAPTER I: Introduction and thesis outline</td>
<td>13</td>
</tr>
<tr>
<td>CHAPTER II: Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two peri-Alpine lakes</td>
<td>31</td>
</tr>
<tr>
<td>CHAPTER III: A century of climate change and eutrophication homogenized lake cyanobacterial communities</td>
<td>71</td>
</tr>
<tr>
<td>CHAPTER IV: Sedimentary and egg-bank DNA from 3 European lakes reveal concurrent changes in the composition and diversity of cyanobacterial and <em>Daphnia</em> communities</td>
<td>103</td>
</tr>
<tr>
<td>CHAPTER V: The dark side of lakes: diversity and distribution of nonphotosynthetic cyanobacteria</td>
<td>141</td>
</tr>
<tr>
<td>CHAPTER VI: General discussion and outlook</td>
<td>167</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>179</td>
</tr>
</tbody>
</table>
Summary

Anthropogenic global changes have affected almost all ecosystems on Earth and threatened biodiversity worldwide, but the extent of the impact of this pervasive influence is not fully understood. In aquatic ecosystems, eutrophication of lakes has led to the periodic dominance of cyanobacteria, which affects the services that lake ecosystems provide to human society (e.g., tourism, drinking water, fisheries), and sometimes directly threatens the health of animals, including humans. Many lakes in the European peri-Alpine region are regularly affected by the presence of cyanobacteria, but how rapid changes in local and regional lake conditions have contributed to determine their diversity and distribution, for example favouring toxic and bloom-forming taxa, is not well known. Peri-Alpine lakes, with their well-studied eutrophication and climate history, represent excellent study sites for reconstructing past assemblages of planktonic organisms and for investigating their diversity over long time scales of anthropogenic influence by applying molecular techniques to the sedimentary archive.

The general aim of this thesis project was to investigate the effects of environmental change, especially climate warming and alteration of nutrient regimes, over the distribution, the biodiversity, and the phylogenetic structure of lacustrine cyanobacterial communities. We used high-throughput sequencing on DNA extracted from dated sediment cores and performed phylogenetic and diversity analyses of the cyanobacterial communities over two centuries to cover the periods of pre-, mid- and post eutrophication of lakes. The first project (Chapter II) was to validate the sedimentary DNA-based reconstruction method by comparing the long-term diversity of cyanobacterial communities recovered from the sediments of two lakes (Greifensee and Lake Zurich) with the diversity assessed by microscopy in water samples collected over four decades as part of lake monitoring programmes. This study confirms the feasibility of the sedimentary DNA approach for assessing the richness and phylogenetic diversity of past communities of cyanobacteria.

The second project (Chapter III) was to investigate the biogeography of cyanobacteria and compare their community composition and phylogenetic structure across lakes. For that, the study region was expended to ten lakes across the peri-Alpine area. The results of this project show evidence for the homogenisation of lakes abiotic conditions due to climate warming and eutrophication, and reveal that over the past 150 years, the richness and
similarity of cyanobacterial communities has generally increased, particularly gaining buoyant and colonial taxa.

The presence of cyanobacteria do not only impair water quality and ecosystem services, but they also have negative effects on other levels of lake food-webs. **Chapter IV** aims at linking the diversity of *Daphnia* and cyanobacteria in lakes impacted by anthropogenic eutrophication. This study using the sedimentary *Daphnia* egg bank and cyanobacterial DNA from two peri-alpine lakes (Switzerland) and a shallow lake in the Danube Biosphere Reserve (Romania) confirms the invasion of *D. galeata* and the subsequent replacement of native *Daphnia* lineages by *D. galeata* taxa in the two Swiss lakes. The results further suggest that *D. galeata*, *D. cucullata*, and their sexual descendants appear to have been favoured by the presence of filamentous cyanobacteria at all sites. These results provide novel insights into long-term community interactions between two important plankton groups.

During our investigation of cyanobacterial diversity in the sedimentary archive (chapters II & III), we detected two newly described clades of ancestral cyanobacteria called Melainabacteria and ML635J-21. **Chapter V** presents a short literature review describing these nonphotosynthetic cyanobacteria, and explores their diversity within and across the same ten lakes investigated in chapter II, with the aim to compare their dynamics to the trends observed in cyanobacteria. This study reveals that, unlike cyanobacteria, the richness and composition of nonphotosynthetic cyanobacteria did not vary significantly over the past ~150 years, suggesting that the ladder are not governed by the same environmental factors as the photosynthetic clades.

Taken together, the results of these four studies reveal previously unreported patterns in the composition and structure of cyanobacterial communities over broad temporal and spatial scales, and highlight the effects of human-induced environmental changes on cyanobacterial assemblages in freshwater lakes that provide important ecosystem services to humans. Our results support the hypothesis that shifts in the diversity and distribution of taxa might not be reversible even after applying stringent remediation measures for the reduction of water eutrophication, and highlight the importance of increasing our understanding of the underlying mechanisms that regulate community structure to gain predictive ability for future lake management purposes.
Zusammenfassung


Die zweite Studie (Kapitel III) untersucht die Biogeographie von Cyanobakterien und vergleicht die Komposition der Gemeinschaft und die phylogenetische Struktur...
verschiedener Seen, dazu wurde das Studiengebiet auf zehn Seen in der peri-alpinen Region erweitert. Die Resultate zeigen auf, dass der Klimawandel und die Eutrophierung zur Homogenisierung der abiotischen Konditionen der Seen geführt haben und dass sich Artenreichtum, sowie die Ähnlichkeit der Gemeinschaft von Cyanobakterien, besonders schwimmende und koloniale Taxa, generell erhöht haben.


Während der Diversitäts-Analyse der Cyanobakterien in sedimentären Archiven (Kapitel II & III) sind wir auf eine neu-beschriebene Gruppe von Ur-Cyanobakterien, genannt Melainabacteria und ML635J-21, gestossen. Das **Kapitel V** präsentiert zuerst eine kurze Literaturübersicht, welche diese Ur-Cyanobakterien, die nicht Photosynthese betreiben, beschreiben und erforstcht dann ihre Diversität innerhalb und zwischen denselben zehn Seen, die für die Studie im Kapitel zwei untersucht wurden, mit dem Ziel, ihre Dynamik und Trends zur denen der Cyanobakterien zu vergleichen. Die Ergebnisse zeigen auf, dass sich über die letzten 150 Jahren, ungleich zu den Cyanobakterien, der Artenreichtum und die Zusammensetzung der Melainabacteria nicht signifikant verändert haben. Dies bedeutet, dass Melainabacteria nicht durch dieselben Umweltfaktoren wie die Kladen, die Photosynthese betreiben, geregelt werden.

von strengen Sanierungsmassnahmen für die Reduktion der Eutrophierung von Gewässern nicht reversibel sind. Dies zeigt wie wichtig unser Verständnis für die unterliegenden Mechanismen, welche die Gemeinschafts-Strukturen regulieren, ist, damit wir sinnvolle Massnahmen für ein zukünftiges See-Management generieren können.
CHAPTER I

Introduction and thesis outline

General background

Understanding how biodiversity is generated and maintained is an focal topic in ecological research that has the potential to influence the way ecosystem are managed and conserved. Since the eighteenth century, scientists have investigated the geographic distribution (that is, the biogeography) of animals and plants over time. Contrary to animals and plants diversity of microbes could only be studies by culturing environmental samples on standard media, with the disadvantage that everything that did not grow on these media could not be detected. Genetic methodologies developed more recently allowed to take a step further into the exploration of the distribution and diversity of non-culture microorganisms that were until then impossible to study. Although it is generally accepted that large organisms have biogeographies, for microbes there is this idea that ‘Everything is everywhere, but the environment selects’ (Baas Becking, 1934), which implies microorganisms do not have any large-scale patterns of geographic distribution. This is because dispersal capabilities of microorganisms are thought to be greater than the effect of past contingencies on their geographic distribution, thus, the contemporary environment is what contributes most to maintaining diversity in distinct microbial assemblages (Martiny et al., 2006).

Species distribution (in both macro- and micro-organisms) is mostly under the control of climate, as well as local abiotic (e.g., sunlight, pH, humidity) and biotic (e.g., presence of prey, predators, parasites) conditions. Species dispersal can be limited by the structure of the landscape (e.g., physical barriers like mountain ranges, or a patchy territory), and the physical ability of the organism itself also plays a role in dispersal success. The distribution and dominance of species in terrestrial and aquatic habitats can be modified by changes in local and global conditions. Human activities, especially over the last few centuries, have contributed to alter almost all ecosystems on Earth and contributed to modify the global climate (Vitousek et al., 1997a). They have had profound effects upon the global geochemical cycles (phosphorus [P], nitrogen [N], and carbon) and on the availability of nutrients on the planet (Vitousek et al., 1997b; Smith et al., 1999). As a consequence, these modifications have had impacts on biodiversity, community assembly, and dispersal of species due to the modification of landscapes and associated degradation of environmental
quality in most ecosystems since the industrial revolution (Vitousek et al., 1997a; Hunter, 2007). Intensification of land-use is an example of a humans impact that threatens local biodiversity but also contributes to biotic homogenisation (i.e., the decrease in beta diversity) across the landscape at broader scales (Gámez-Virués et al., 2015; Gossner et al., 2016). In microbial communities, environmental filtering (the process by which species compatible to local conditions survive and persist in a community while others fail (Mayfield et al., 2009)) has been identified as a major mechanism underlying community assembly (e.g., Lindström & Langenheder, 2012; Logares et al., 2013). The extent and directionality of the impacts of human activities on the biodiversity and distribution of species, via modification of the environment, vary across ecosystems, spatial scale, and taxonomic groups (Gossner et al., 2016), and are not fully known.

**Human-induced eutrophication of lakes**

Anthropogenic eutrophication, also called cultural eutrophication, is generally defined as the enrichment with nutrient as a consequence of human activities leading to excessive algal growth (Hutchinson, 1973). It has been for decades – and still is – the most widespread water-quality issue in lakes, reservoirs, and coastal marine ecosystems worldwide (Glibert et al., 2006; Smith & Schindler, 2009). Phosphorus has been identified as the main limiting factor for phytoplankton growth in freshwaters (Schindler, 1974, 1977), and the suggestion was made that mitigation of phosphorus in lakes should lead to their recovery from eutrophication (Schindler, 1974). Over the recent decades, governments of developed countries have implemented eutrophication management frameworks, such as the installation of wastewater treatment plants and the ban of phosphate in detergents. These measures focussing primarily on the control of P loading resulted in the return to near natural trophic state of some lakes (Schindler, 2012), but despite sustained efforts, cyanobacteria continue to proliferate and still constitute a major threat to water quality worldwide (Smith & Schindler, 2009; Schindler, 2012; Taranu et al., 2015). While P was managed, N has not received the same attention and has kept increasing as a consequence of human activities. In a recent meta-analysis, Elser and colleagues (2007) have reported that the enrichment of lakes with both P and N can have a strong synergistic effect on algal biomass production. Several laboratory (e.g., Berman & Chava, 1999) and experimental (Levine & Schindler, 1999) studies have highlighted the effect of N (as well as specific N-forms (Donald et al., 2011, 2013)) in promoting cyanobacterial biomass in general. Some have also found a taxa-specific response to enhanced N, favouring especially nitrogen-fixing and bloom-forming
species (Dolman et al., 2012; Donald et al., 2013), as well as an effect on cyanobacterial toxicity (Giani et al., 2005; Monchamp et al., 2014).

The consequences of eutrophication, via direct and indirect effects on lake physical and chemical conditions, are visible at multiple levels of the food-web. The most striking symptom of eutrophication is the excessive growth of phytoplankton, which often leads to a shift in phytoplankton composition in favour of the dominance of cyanobacteria (Smith & Schindler, 2009; O'Neil et al., 2012). Cyanobacteria, especially in high abundance, have negative impacts on other compartments of the food-web. It is considered poor quality food for zooplankton compared to eukaryotic algae because it lacks essential nutrients such as sterols and poly-unsaturated fatty acids (von Elert et al., 2003).

A major concern raising from the presence of cyanobacteria in lakes is the fact that many taxa can synthesize toxins and other secondary metabolites that can be harmful to humans and other animals (Carmichael et al., 1990; Carmichael, 1992; Chorus & Bartram, 1999). Cyanotoxins are classified into five main groups: hepatotoxins, neurotoxins, cytotoxins, dermatotoxins, and irritant toxins (lipopolysaccharides) (Mankiewicz et al., 2003). The hepatotoxins like microcystins are the most common in freshwaters taxa (Carmichael, 1992), and Microcystis aeruginosa is the most ubiquitous producer (Frangepal et al., 2008). Microcystins can be synthesized by other genera as well, mainly Planktothrix (order Oscillatoriales), and Anabaena (also named Dolichospermum; order Nostocales). All these taxa are often of concern for authorities, as they can accumulate to form large blooms at the water surface. The presence of toxic Microcystis strains in the diet of the large filter feeder Daphnia has been shown to have negative effects on the feeding behaviour, reproductive success, and survival in Daphnia (DeMott et al., 1991; reviewed in Lampert & Haney, 1987). In high concentrations, microcystins can be lethal to animals, including humans (Carmichael, 1994). In 1996, the accidental poisoning of the water source to a dialysis centre in Brazil led to the death of over 50 patients (Carmichael et al., 2001). Numerous cases of domestic and wild animals intoxications due to the ingestion of microcystins have been reported worldwide (Carmichael, 1992; Lambert et al., 1994; Mez et al., 1997).

Scientists have recognized that cyanobacterial bloom events have increased both in strength and in frequency over the past decades due to the combined effect of climate change and accelerated eutrophication (Paerl & Huisman, 2008; Otten et al., 2012). The main factors regulating the growth of photosynthetic cyanobacteria are temperature, light, and nutrient (mainly P and N) availability (Whitton & Potts, 2002). Because cyanobacteria in general have an optimal growth temperature higher than most eukaryotic phytoplankton
(reviewed in Visser et al., 2016), they should continue to be favoured under future climate warming scenario (Carey et al., 2012). Warming can also have indirect effects on physical lake conditions. For instance, it can extend the duration of the period of thermal stratification, and change the strength and frequency of water mixing (Livingstone, 2008; Adrian et al., 2009). These modifications in turn have consequences on the oxygen and nutrient regimes, ultimately affecting the biology of lakes. Reduced winter mixing due to warming has been shown to increase the abundance of the potentially toxic cyanobacterium Planktothrix rubescens (Posch et al., 2012). The fact that some cyanobacterial taxa have physiological traits that make them highly competitive under various environments (e.g., nitrogen fixation, regulation of cell buoyancy, adaptation to low light) increases the difficulty to determine effective measures for their management in lakes. How local and global environmental changes affect the biodiversity and the distribution of cyanobacteria are still not fully understood.

**Aims of the thesis**

The general aim of my thesis is to explore and describe the temporal trends in diversity of natural freshwater cyanobacterial communities in relation to environmental changes using DNA preserved in sediment cores in combination with long-term biological and chemical-physical lake data. Specifically, I focused on the impacts of eutrophication and climate warming on species turnover, phylogenetic diversity, composition and biogeography of freshwater cyanobacteria.

**Study system**

Lakes in the European peri-Alpine region provide important ecosystem services to human society. They are important economically as touristic attraction and for fisheries. The peri-Alpine region is densely populated where several large lakes supply drinking water for millions of people. Most peri-alpine lakes have undergone a phase of eutrophication in the twentieth century as a consequence of excessive nutrient inputs due to human activities, mainly intensive land-use and urbanisation. As a consequence of eutrophication, frequency and strength of cyanobacterial bloom events has increased, impairing ecosystem functioning and ecosystem services. Phosphorus control measures were implemented in the 1980s, which led to re-oligotrophication of lakes and the associated decrease in primary productivity.
Phosphorus control measures were implemented in the 1980s which led to the re-oligotrophication of lakes and the associated decrease in primary productivity. Nevertheless, the presence of nuisance cyanobacteria continues to impair water quality (Posch et al., 2012; Salmaso et al., 2015; Taranu et al., 2015). The studies presented in this thesis focus mainly on 10 lakes spread across the peri-Alpine region (north-eastern and north-western Plateau, and Alpine south side; Figure 1). They are relatively large (lake area between 5 and 582 km$^2$) and deep (Depth$_{\text{max}}$ between 24 and 370 m), and their present trophic status range from oligotrophic to eutrophic. The lakes geographic location, as well as the main morphological and chemical characteristics are summarized in Table 1.
Figure 1. Map of the European per阿尔卑斯 region that extends throughout France, Switzerland, Austria, Italy, and Germany. The location of the lakes studied is shown (red circles).

Table 1. List of lakes with their geographical location, main morphological characteristics, and current trophic status. $P_{\text{max}}$ refers to the maximum annual mean TP concentration measured over the first 20 m of the water column at the peak of eutrophication, and $P_{\text{recent}}$ gives the most recent available annual mean TP concentration.

<table>
<thead>
<tr>
<th>Lake name</th>
<th>Depth$_{\text{max}}$ (m)</th>
<th>Lake area (km²)</th>
<th>Volume ($10^5$ m³)</th>
<th>$P_{\text{max}}$ (µg/l)</th>
<th>$P_{\text{recent}}$ (µg/l)</th>
<th>Elevation (m a.s.l.)</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alpine north side - western Plateau</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geneva</td>
<td>309</td>
<td>582</td>
<td>89,000</td>
<td>78.69</td>
<td>13.81</td>
<td>372</td>
<td>46°27'N</td>
<td>6°32'E</td>
</tr>
<tr>
<td>Annecy</td>
<td>82</td>
<td>27.59</td>
<td>1,124</td>
<td>15.65</td>
<td>6.00</td>
<td>447</td>
<td>45°51'N</td>
<td>6°10'E</td>
</tr>
<tr>
<td><strong>Alpine north side - eastern Plateau</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constance (Upper Lake)</td>
<td>252</td>
<td>472</td>
<td>51,400</td>
<td>83.90</td>
<td>7.60</td>
<td>395</td>
<td>47°35'N</td>
<td>9°28'E</td>
</tr>
<tr>
<td>Zurich (Lower Lake)</td>
<td>137</td>
<td>67.3</td>
<td>3,300</td>
<td>67.81</td>
<td>11.89</td>
<td>406</td>
<td>47°15'N</td>
<td>8°41'E</td>
</tr>
<tr>
<td>Hallwilersee</td>
<td>47</td>
<td>10</td>
<td>280</td>
<td>213.81</td>
<td>18.27</td>
<td>449</td>
<td>47°17'N</td>
<td>8°12'E</td>
</tr>
<tr>
<td>Baldeggersee</td>
<td>66</td>
<td>5.2</td>
<td>173</td>
<td>439.96</td>
<td>23.66</td>
<td>463</td>
<td>47°11'N</td>
<td>8°15'E</td>
</tr>
<tr>
<td>Greifensee</td>
<td>33</td>
<td>8.5</td>
<td>148</td>
<td>492.43</td>
<td>40.20</td>
<td>435</td>
<td>47°21'N</td>
<td>8°41'E</td>
</tr>
<tr>
<td><strong>Alpine south side</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maggiore</td>
<td>370</td>
<td>212</td>
<td>37,500</td>
<td>31.00</td>
<td>11.00</td>
<td>194</td>
<td>45°57'N</td>
<td>08°38'E</td>
</tr>
<tr>
<td>Lugano #</td>
<td>288</td>
<td>48.9</td>
<td>5,860</td>
<td>177.25</td>
<td>16.30</td>
<td>270</td>
<td>45°58'N</td>
<td>08°57'E</td>
</tr>
<tr>
<td>Pusiano</td>
<td>24</td>
<td>5</td>
<td>69</td>
<td>197.00</td>
<td>88.00</td>
<td>259</td>
<td>45°48'N</td>
<td>09°16'E</td>
</tr>
</tbody>
</table>

# Refers to the Figino basin in Lake Lugano.
The evolution of cyanobacteria

Cyanobacteria are photosynthetic prokaryotes that are important primary producers in fresh and marine waters. They are amongst the oldest organisms on Earth, and they have played an important role in the oxygenation of the planet about 2.4 billion years ago (Summons et al., 1999; Fischer et al., 2016a). Cyanobacteria are distributed across several habitats from fresh, brackish, and salt waters, to deserts and extreme polar environments (Whitton, 2012). Cyanobacteria form a diverse phylum that has been sub-divided into four main groups based on morphological characteristics (Rippka et al., 1979). With the development of molecular techniques, however, it became clear that there were major incongruences between the traditional classification based upon morphological attributes and phylogenetic reconstructions (Komárek et al., 2014). In addition, in the last five years, scientists have sequenced the genomes of some uncultured bacteria and discovered that they were direct ancestors of cyanobacteria (Di Rienzi et al., 2013; Soo et al., 2014). These basal clades share some cyanobacterial traits like a double cell wall, but surprisingly they lack the photosystem that would allow them to perform oxygenic photosynthesis. Based on genome sequencing and phylogenetic studies, a new classification was proposed for Phylum Cyanobacteria, with three class-level lineages: Melainabacteria, Oxyphotobacteria (cyanobacteria), and ML6235J-21 (Figure 2). The function of these new cyanobacterial clades as well as their ecological role are not yet known.

Phylogenetic reconstructions using environmental DNA from sedimentary archives

Phylogenetic analyses are evolutionary reconstructions of the relatedness among DNA sequences and they can be used to assess structure and genetic diversity in natural populations. Phylogenetic diversity (PD) is defined as the sum of the length of all the branches composing a phylogenetic tree in the minimum spanning path (Faith, 1992) and can be calculated for any given subset of taxa or for the whole phylogeny. 16S rRNA markers have been used to resolve the taxonomic diversity in cyanobacterial populations (Otsuka et al., 2001), and to determine genetic diversity, evolution, and structure in phylogenetic studies (Ishida et al., 1997; Otsuka et al., 1999; Rantala et al., 2004; Moreira et al., 2013).
Figure 2. (From Fischer et al., 2016b). A phylogenetic tree showing the relationship among sequences of the 16S rRNA gene in the Phylum Cyanobacteria. The ML635J-21 (red) is the most basal clade, and the Oxyphotobacteria (photosynthetic cyanobacteria; green) and the Melainabacteria (blue) are sister-clades within Phylum Cyanobacteria.

With the recent emergence of molecular techniques such as high-throughput DNA sequencing allowing the investigation of organisms directly from the environment, the number of genetic studies of microscopic organisms has increased exponentially (Shendure & Ji, 2008). Because these technologies are still new, the genetic diversity of natural communities prior to the last twenty years remains mostly unknown. This is especially true for non-culture microorganisms that cannot be used in laboratory experiments. Exploring DNA preserved in sediments allows scientists to go back in time and study whole communities at the genetic level over long time-scales. Environmental DNA (eDNA) refers to nucleic acids isolated directly from environmental samples, either from living or dead organisms, gut contents, as well as extracellular DNA (Taberlet et al., 2012). It is used to study both micro- and macro-organisms present in various environments (e.g., water, ice, sediments). In the case of nucleic acids preserved in sedimentary records, the term sedimentary DNA (sedDNA) is used. Lacustrine sediments are rich archives that contain information about past planktonic communities. Classical palaeolimnological markers (e.g., subfossil diatoms and chironomids) have been extensively used to reconstruct the history of lakes with regards to
local environment and climatic conditions (Smol & Cumming, 2000; Millet et al., 2010; Simpson & Birks, 2012). Planktonic groups that do not leave biological remains in the form of cysts, spores, or frustules leave their DNA in sediments. This can now be investigated by applying the most recent sequencing technologies using sedimentary records. Temporal changes in the abundance and diversity of various planktonic organisms in freshwater and marine ecosystems have recently been investigated using sedDNA over time-scales of decades (e.g., Brede et al., 2009; Savichtcheva et al., 2015) to millennia (e.g., Boere et al., 2011; Lejzerowicz et al., 2013; Hou et al., 2014).

The sediments of many lakes worldwide have an annually laminated structure (Zolitschka et al., 2015), which offers unique temporal resolution for palaeolimnological studies. The formation of annual organic varves in the sediments (Figure 3) is a consequence of the combined high primary production, the calcareous nature of the sediments, and the anoxic or hypoxic conditions in the deep-water strata of some lakes. Several peri-alpine lakes show records of annual organic varves over decades or centuries (Hollander et al., 1992; Lotter et al., 1997; Naeher et al., 2013; Zolitschka et al., 2015). When the sediment structure is not affected by mixing or by bioturbation by bottom-dwelling organism, they offer perfect material for conducting highly time-resolved reconstruction studies on the past macro- and microorganism assemblages.

Box 1. Model and description of the steps of carbonaceous organic varves formation in the sediments of lakes.

**Organic varves**

- A pale layer forms in spring after the first Diatom bloom;
- In late spring and over summer, a white layer might form following calcite precipitation;
- In fall and winter, a dark layer forms due to accumulation of organic matter and minerogenic detritus from runoff.

**Figure 3.** Model of carbonaceous organic varves formation. Modified from Sturm & Lotter (1995)
Chapter I – Introduction & Thesis Outline

**Thesis outline**

In **Chapter II**, I reconstructed cyanobacterial communities over 200 years from the sedimentary archives of two pre-alpine lakes, Greifensee and Lake Zurich, and used the 16S rRNA sequences retrieved from the sediments to derive a phylogeny summarizing the diversity of cyanobacteria in the two lakes. Further, I validated the approach by comparing the genetic data for the sediments to long-term records of phytoplankton as identified by microscopy in water samples collected over the past five decades in the same two lakes. The results of this first study show that the phylogenetic reconstruction approach applied to sedimentary DNA relate well with historical data on cyanobacterial communities in the lakes in terms of richness and taxonomic composition. In the following chapters of my thesis, the same sequencing approach (with few modifications) was used to investigate the past diversity of cyanobacteria in multiple European lakes impacted by human-induced environmental changes.

In **Chapter III**, I address questions related to the biodiversity and biogeography of cyanobacteria over ~150 years in ten pre-alpine lakes affected by climate warming, as well as rapid eutrophication and re-oligotrophication. I found evidence for the homogenisation of lakes abiotic conditions due to climate warming and eutrophication. These changes, especially the increasing temperatures, appear to have facilitated the spread of buoyant and colonial taxa, which have increased in both richness and prevalence over the past ~150 years and led to the biotic homogenisation of cyanobacterial communities across the region.

In **Chapter IV**, I present a retrospective study on the genetic structure of *Daphnia* populations related to changes in cyanobacterial community composition in a relatively shallow lake located in the Danube Delta Biosphere Reserve (Romania), and two deep pre-alpine lakes (Switzerland), all of which were impacted by anthropogenic eutrophication over the twentieth century. Here, I explored the links between *Daphnia* and cyanobacterial communities using the egg banks and DNA preserved in lake sediment cores. In all lakes, *D. galeata* taxa were dominant since the onset of eutrophication, and appear to have been favoured by the presence of filamentous cyanobacteria. Interestingly, the presence of potentially toxic cyanobacteria did not appear to have an effect on the genetic structure of *Daphnia* populations.
Chapter V presents a short literature review and an exploratory study focusing on the
diversity and distribution of newly described clades of nonphotosynthetic cyanobacteria. In
this chapter, I describe the diversity of these poorly known lineages in the same ten lakes
studied in Chapter III, and I discuss their biogeography and composition in parallel with the
photosynthetic cyanobacterial communities described previously (Chapter III). The results
show that the phylogenetic similarity across lake communities did not significantly change
over time and is not dependant on geographic distance between lakes.

Finally, in Chapter VI I conclude the thesis by interpreting the main results in the context
of the literature, and I provide an outlook with future perspectives for the field of sedimentary
DNA to help gain predictive ability over the fate of lake ecosystems regarding their diversity
and the presence of nuisance plankton taxa.
Literature cited Chapter I


Ishida T, Yokota A, Junta S (1997) Phylogenetic relationships from of filamentous rRNA


Savichtcheva O, Debros D, Perga ME et al. (2015) Effects of nutrients and warming on
Planktothrix dynamics and diversity: A palaeolimnological view based on sedimentary DNA and RNA. *Freshwater Biology*, **60**, 31–49.


Whitton BA, Potts M (eds.) (2002) *The ecology of Cyanobacteria: Their diversity in time and

CHAPTER II

Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two peri-Alpine lakes

Marie-Eve Monchamp\(^{1,2}\), Jean-Claude Walser\(^3\), Francesco Pomati\(^{1,2}\) and Piet Spaak\(^{1,2}\)

\(^1\) Eawag, Swiss Federal Institute of Aquatic Science and Technology, CH-8600 Dübendorf, Switzerland;

\(^2\) Institute of Integrative Biology, ETH Zürich, CH-8092 Zurich, Switzerland;

\(^3\) Genetic Diversity Centre (GDC), Department of Environmental Systems Science, ETH Zürich, CH-8092 Zurich, Switzerland.

*Journal of Applied and Environmental Microbiology*, doi 10.1128/AEM.02174-16

The definitive version is available at http://aem.asm.org/
Abstract

We reconstructed cyanobacterial community structure and phylogeny using DNA that was isolated from layers of stratified sediments spanning 200 years of lake history in the peri-Alpine lakes Greifensee and Lake Zurich (Switzerland). Community analysis based on amplification and sequencing of a 400-nucleotide (nt)-long 16S rRNA fragment specific to *Cyanobacteria* revealed operational taxonomic units (OTUs) capturing the whole phylum, including representatives of a newly characterized clade termed *Melainabacteria*, which shares common ancestry with *Cyanobacteria* and has not been previously described in lakes. The reconstruction of cyanobacterial richness and phylogenetic structure was validated using a data set consisting of 40 years of pelagic microscopic counts from each lake. We identified the OTUs assigned to common taxa known to be present in Greifensee and Lake Zurich and found a strong and significant relationship (adjusted $R^2 = 0.89; P < 0.001$) between pelagic species richness in water and OTU richness in the sediments. The water-sediment richness relationship varied between cyanobacterial orders, indicating that the richness of Chroococcales and Synechococcales may be underestimated by microscopy. PCR detection of the microcystin synthetase gene *mcyA* confirmed the presence of potentially toxic cyanobacterial taxa over recent years in Greifensee and throughout the last century in Lake Zurich. The approach presented in this study demonstrates that it is possible to reconstruct past pelagic cyanobacterial communities in lakes where the integrity of the sedimentary archive is well preserved and to explore changes in phylogenetic and functional diversity over decade-to-century timescales.
**Introduction**

Understanding patterns in species diversity across space and time is a fundamental topic in ecological research that has the potential to influence how ecosystems are conserved and managed. In lake ecosystems, eutrophication and warming have favored algal growth and triggered dramatic changes in phytoplankton community composition, such as the dominance of cyanobacteria and the increasing frequency of bloom formation, events that can harm aquatic species and humans (Carey *et al.*, 2012; Paerl & Paul, 2012; Rigosi *et al.*, 2014). Proliferation of cyanobacteria perturbs the physical and chemical environment, affecting pelagic and benthic communities. This can drive changes in the food web structure and in ecosystem function (reviewed in Sukenik *et al.*, 2015). Several cyanobacterial taxa are known to synthesize toxic metabolites that impair important ecosystem services as well as water quality (Chorus & Bartram, 1999). Although freshwater *Cyanobacteria* have been extensively studied over the past few decades, little is known about the long-term patterns of diversity and richness of this phylum in natural assemblages. One of the main limitations to our understanding of ecological systems is the lack of long-term data. Because of the recent development of molecular tools, the genetic diversity and phylogenetic structure of planktonic communities prior to the last 20 years are mostly unknown. Especially because there have been important changes in the trophic status of lakes over the last century, it is important to explore the diversity of communities in recent history. This will help understand the changes in lake ecosystems and gain predictive ability. We will address this in these two lakes in a forthcoming paper. Here, we present the validation of our approach for reconstructing past cyanobacterial communities using DNA archived in lake sediments.

The microfossils of various planktonic organisms (e.g., diatom frustules, chrysophyte cysts, cladocerans, and chironomids) have been traditionally used in paleolimnology to reconstruct the past trajectory of lakes (Smol & Cumming, 2000; Leavitt & Hodgson, 2001). However, because only few cyanobacteria produce resting cysts (akinetes), their past diversity and abundance of planktonic cyanobacteria cannot be investigated using fossil remains. Algal pigments have also been used to study the past dynamics of phytoplankton communities from sedimentary archives (Leavitt & Hodgson, 2001). In a recent large scale study in northern temperate subarctic lakes, sedimentary pigments revealed that cyanobacterial abundance has increased over the past 200 years relative to other phytoplankton taxa (Taranu *et al.*, 2015). Another recent study, combining pigment and DNA analysis of lake sediment cores collected in western Quebec (Canada), showed an increase in cyanobacterial abundance over the past 30 years in lakes located in both protected and nonprotected areas (Pal *et al.*, 2015). Although pigments are useful to investigate temporal
changes in the abundance of major phytoplankton groups with different pigment profiles, they do not provide information about species richness and diversity within communities (Leavitt & Hodgson, 2001).

In combination with newly developed genetic tools, environmental DNA (eDNA) studies open up the possibility of investigating present and past diversity at finer scales. In recent years, several studies have reported the successful isolation of DNA from environmental samples (Pedersen et al., 2014). Sedimentary DNA (sedDNA) and sedimentary ancient DNA (sedaDNA) isolated from marine and freshwater sediment cores are increasingly being used to investigate the long-term dynamics of various planktonic taxa, such as diatoms (Epp et al., 2009; Stoof-Leichsenring et al., 2012, 2015), copepods (Bissett et al., 2005), cladocerans (Hairston et al., 2001), protists (Capo et al., 2015), and viruses (Coolen, 2011). Recent sedaDNA-based studies were applied to investigate diversity changes in phytoplankton (Hou et al., 2014) and cyanobacterial communities (Fernandez-Carazo et al., 2013) during the Holocene. The abundance and diversity of natural populations of cyanobacteria have been investigated in sedimentary archives of peri-alpine lakes using quantitative PCR and cloning (Savichtcheva et al., 2011, 2015; Domaizon et al., 2013). Sedimentary DNA has also been used to investigate the past distribution and diversity of potentially toxic Microcystis in Lake Erie (Rinta-Kanto et al., 2009), and the saxitoxin-producing Cylindrospermopsis raciborskii in a subtropical lagoon (Martínez de la Escalera et al., 2014).

Several studies have described the composition of past cyanobacterial populations, but the phylogenetic diversity of whole cyanobacterial communities has, to our knowledge, never been assessed over timescales of decades to centuries using sedimentary archives. Knowing the composition and phylogenetic relatedness of cyanobacterial communities over a long period of time may allow us to infer the mechanisms of ecological change and forecast future trends (Emerson and Gillespie, 2008; Moreira et al., 2013). In this study, DNA preserved in lake sediments was used to reconstruct cyanobacterial phylogenetic diversity from two peri-Alpine lakes, Greifensee and Lake Zurich, over the last 200 years. The phytoplankton communities of these lakes have been monitored over the last five decades, and the presence of potentially toxic cyanobacteria has been documented in both. We amplified a 400-nucleotide (nt)-long fragment of the V3 and V4 variable regions of the 16S rRNA gene using cyanobacterium-specific primers on samples taken from sediment cores. We also used a PCR-based approach to reconstruct the history of potentially microcystin-producing cyanobacteria over the last century. Finally, we validated our sequencing approach by comparing the cyanobacterial richness estimated from the sedimentary archives.
to independent data sets of species richness that were quantified by microscopy over the last 40 years from water samples collected in these lakes.

Figure 1. Map of Greifensee and Lake Zurich showing the sampling sites. The insert shows the location of the lakes within Switzerland. (Maps created with ESRI ArcMap version 10.3.1 using Swisstopo data.)

Materials and Methods

Study sites
Greifensee (47°20′N, 8°40′E) and Lake Zurich (47°15′N 8°41′E) are two natural peri-Alpine lakes located near the city of Zurich in northeastern Switzerland (Fig. 1). The lakes were chosen on the basis of the availability of long-term data on phytoplankton communities, and because recent sediments of both lakes are characterized by the formation of annual varves (Fig. 2). These consist of a pale summer layer and a dark winter layer (Zolitschka, 2007) that allow high temporal resolution dating. Greifensee (surface area 8.45 km², max. depth 32 m) is monomictic, with one complete mixing event occurring every winter. It is currently classified as eutrophic (average phosphorus of 52 µg/liter in 2015) and has anoxic deep water layers over summer when the lake is strongly thermally stratified (from June to December). Lake Zurich is mesotrophic (average phosphorus of 14.8 µg/liter in 2010), and with a surface area
of 65 km² and a maximum depth of 136 meters, it is one of the largest European peri-Alpine lakes. The lake is divided into two basins by a natural dam and the focus of this study is on the lower lake, located near the city of Zurich. The lake is considered to be monomictic or dimictic, but the increase of strength of thermal stratification in the last few decades as a consequence of climate warming impedes complete mixing of the water column (Anneville et al., 2004; Posch et al., 2012).

**Long-term phytoplankton data**

The phytoplankton community in Greifensee has been monitored over the last five decades by the Swiss Federal Institute of Aquatic Science and Technology (Eawag). In the present study, we used a data set consisting of species composition and counts as measured by microscopy from integrated water samples collected over the upper 20 meters with a Schröder sampler (Bürgi et al., 2003) every month from 1974 to 2010. For Lake Zurich, we used a dataset consisting of phytoplankton samples collected by the Zurich Drinking Water Company (Wasserversorgung Zürich - WVZ) at 14 discrete depths over the water column at monthly sampling intervals from 1976 to 2010. In both lakes, phytoplankton identification and cell counts were performed using the Utermöhl method (Lund et al., 1958). Taxonomic affiliation of species was harmonized over the entire data sets according to (Pomati et al., 2015) and species that were not found in at five time points were excluded. This was done to remove taxa that are potentially the result of misidentification or inconsistency in taxonomic identification through time.

**Sediment sampling**

Three sediment cores of 63 mm diameter and approximately one-meter length were collected in 2013 using a gravity corer in the deepest part of Greifensee (32 m) and at 98-m depth in the center of Lower Lake Zurich. The cores were sealed and stored in a vertical position in a dark room at 4°C until processing within the following months. One core from Greifensee and two cores from Lake Zurich were opened longitudinally and photographed in a room where no DNA work or PCR amplification had been performed before. The opened cores were visually inspected to identify disturbances that may affect the temporal reconstruction. One half of each core was used as a reference for counting the annual laminations in the sediments. For Greifensee, the reference half-core was also used for radiometric measurements to build an age model (Fig. S1). The other half of the Greifensee core and one half of each Lake Zurich core were used for DNA isolation. In the case of Lake
Zurich, we used the second core later on to collect a few additional samples between the years 1800 and 1935 in order to complete the time series. All cores were processed at different times to avoid cross contaminations. In addition, older sediments were processed independently from recent sediments to minimize contamination.

Figure 2. Photographs of oxidized varved sediments showing the depth profile of the upper 40 centimeters in a sediment core from lakes Greifensee (A) and Zurich (B). Arrows indicate the depths at which the sedDNA samples were collected and are identified by the corresponding year.
**Sediment core chronology**

Varves were formed from the 1930s onwards in Greifensee (Hollander et al., 1992) and from the early 1900s onwards in Lake Zurich (Naeher et al., 2013; Fig. 2). The undisturbed varves indicate the absence of benthic bioturbation or sediment mixing (Hollander et al., 1992). In order to select layers of sediments corresponding to specific years over the recent history of the lakes, we built a high-resolution dating profile by varve counting. The Lake Zurich sediment core profiles were aligned with a recent high-resolution age model of a core taken at the same location (A. Gilli, ETH Zürich, personal communication). In Greifensee, we performed radiometric analysis ($^{210}$Pb and $^{137}$Cs; Krishnaswamy et al., 1971; Figure S1) to corroborate visual varve counting. Briefly, 20 samples were collected from the half core at 1-cm intervals and were freeze-dried. The sediments were then homogenized, and 3 to 6 g of sediments was used for radiometric measurements on a high-purity germanium (HPGe) well detector (gamma spectrometer) at Eawag facilities. The upper 1.5 cm of each core (corresponding approximately to the years 2009-2013) was discarded from the richness comparison because of possible mixing in the surface sediments likely caused during core collection and handling.

**Sedimentary DNA isolation**

Based on the age model, we collected sediment strata that corresponded to 1 or 2 years spanning the period for which annual varves were visible (Fig. 2). We collected additional samples from older sediments between the years ~1840 to 1930 in Greifensee and between ~1800 to 1900 in Lake Zurich for a total of 20 time points between ~1840 and 2012 in Greifensee, and 23 time points between ~1800 and 2009 in Lake Zurich. Sediment samples for DNA analysis were taken from the center of the cores while carefully avoiding the sediments in contact with the plastic liner to prevent contamination. Samples were immediately transferred to the DNA clean facility dedicated to eDNA work at Eawag in Dübendorf. Samples were either immediately processed, or kept at −20°C until DNA isolation. Strict ancient DNA work protocols were followed in order to prevent contamination with modern DNA (following protocols from Deiner et al., 2014). All instruments used for sedDNA isolation were placed in a laminar flow hood under a UV lamp a minimum of 20 minutes prior to use. From each stratum, sedDNA was extracted from approximately 1 g of wet sediments using the PowerSoil DNA Isolation Kit (MoBio Laboratories, Inc., CA, USA) according to the manufacturer's manual. The extractions were performed in batches of seven samples, and one negative extraction control (which were treated in the same way as the
samples except it did not contain sediments) was added for every seven samples. Each sedDNA extract was quantified using a Qubit (1.0) Fluorometer (Thermo Fisher Scientific) following the manufacturer protocol for the double-stranded DNA high sensitivity Assay (dsDNA HS). The genomic sedDNA was inspected for degradation on agarose gels dyed with SYBR Safe DNA Gel Stain. A PCR amplification test was also performed on all sedDNA extracts using published cyanobacteria-specific primers amplifying a ~800-nt-long fragment of the 16S rRNA gene (Neilan et al., 1997; Jungblut et al., 2005) to assess the level of DNA preservation in the sediments.

**DNA preparation and amplification**

Library preparation for Illumina high-throughput sequencing (HTS) was performed following the optimized protocol of the Genetic Diversity Centre (GDC; ETH Zürich, Switzerland). The protocol uses a two-step PCR approach. A first PCR was performed using previously published cyanobacterium-specific primers CYA359-F, 5′-GGGGAATYTTCCGCAATGGG-3′ and CYA784-R, 5′-ACTACWGGGGTATCTAATCCC-3′ (Nübel et al., 1997). These amplify an approximately 400-nt-long fragment of the V3-V4 regions of the 16S rRNA gene. The primers were modified for Illumina sequencing by adding overhanging adaptors (Table S1) and inserting up to three random nucleotides between the adaptor and the primer sequence to increase cluster complexity on the flow-cell. All modified primer pairs were pooled for the PCR amplifications to avoid amplification bias. The specificity of the primers was verified *in silico* against the Greengenes database (DeSantis et al., 2006) to ensure the primer target a wide range of taxa all over the *Cyanobacteria* phylum, which was the most important characteristic considered.

Sedimentary DNA samples (including negative extraction controls) were amplified in two separate PCR reactions of 20 µl each containing 10× Roche FastStart PCR buffer (Roche, Inc., Basel, Switzerland), 25 mM MgCl₂, 0.2 mM deoxynucleoside triphosphate (dNTP) mix, 0.2 mM each primer (Microsynth, Balgach, Switzerland), and 0.05 U of Faststart Taq polymerase (Roche). The volume of the template DNA used in each reaction varied between 1 and 4 µl depending on the sedDNA concentration. The thermal cycler PCR program included a first denaturation step at 95°C for 4 min, followed by 30 cycles at 95°C for 20 s, annealing at 62°C for 30 s, extension at 72°C for 60 s and a final extension step at 72°C for 5 min. PCR products were purified with 0.8× Agencourt AMPure XP beads (Beckman Coulter, Nyon, Switzerland) and resuspended in 20 µl of 10 mM Tris. The purified product was quantified with a Qubit (1.0) Fluorometer.
Depending on the yield, which varied between <0.2 and 7 ng/µl, 2 to 15 µl of purified PCR product was used as template for a second round of low-cycle PCR to add the Illumina Nextera XT index kit with 1× Kapa Hifi HotStart Readymix (Kapa Biosystems, MA, USA). This second PCR was performed with an initial denaturation step at 95°C for 3 min, followed by 8 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s, with a final extension step at 72°C for 5 min. Samples were purified once more with 0.8× AMPure beads and the final product was quantified on a 7500 Fast real-time PCR system (ABI) using the KAPA library quantification kit (Illumina).

**High-throughput amplicon sequencing**
The indexed samples were pooled at equimolar concentrations into one library. One negative PCR control was also added to the pool. Paired-end 2 × 300 bp sequencing was performed on an Illumina MiSeq (software v.2.5.0.5; Illumina Inc.) at the Genetic Diversity Centre (GDC), ETH Zürich. A total of 6,445,171 raw reads were obtained from 39 samples of Greifensee, and 6,192,583 raw reads were obtained from 43 samples from Lake Zurich.

**Data processing**
The sequencing data were quality controlled and processed using an in-house workflow developed at the GDC (see Fig. S2 in the supplemental material). A data quality check of the raw reads was done using FastQC v0.11.2 (Schmieder & Edwards, 2011)). First, ambiguous and low-quality nucleotides (false-discovery rate [q value], < 10) at the end of the reads were removed to improve overlap recognition and error correction. Next, forward and reverse reads were merged into amplicons using USEARCH (v8.0.1623_i86linux64; Edgar, 2013) allowing a minimum overlap of 15 and minimum merging length of 300 nucleotides. Cutadapt v1. 5 (Martin, 2011) was used to trim full-length forward and reverse primer site of the merged reads. Mismatches were not allowed, except for wildcards to compensate for wobble bases. Last, quality filtering and amplicon size selection was performed using Prinseq-lite v0.20.4 (Schmieder & Edwards, 2011). The quality filtering, merging, primer trimming and the size selection steps removed about 30% of the data. The negative control had less than 0.14% of reads compared to the average of the other samples and was therefore removed.

The primer-trimmed, quality-filtered, and size-selected amplicons were clustered into operational taxonomic units (OTUs) using the UPARSE workflow (Edgar, 2013). The clustering was based on a minimum identity of 97% sequence similarity level, an abundance size threshold of 5. *De novo* and reference based chimera filtering were also applied. All the
steps and parameter applied are detailed in Fig. S2 in the supplemental material and (Schmieder & Edwards, 2011).

**Taxonomic assignments**

The taxonomic assignment of OTUs was done following UTAX with our own database constructed from the Greengenes database (DeSantis et al., 2006) with the addition of few decoy sequences (see Fig. S2 in the supplemental material) http://www.drive5.com/usearch/manual/utax_user_train.html. The assignment was done with a confidence threshold of 0.85. OTUs that were assigned to chloroplasts or heterotrophic bacteria were discarded. All except for two photosynthetic cyanobacteria OTUs from Greifensee were assigned to the order (and class) level with high confidence. At the family level, 51 OTUs were assigned, and 27 OTUs were identified to the genus level. Similarly, in Lake Zurich, only one photosynthetic cyanobacteria OTU could not be assigned to the order and class levels. Sixty-two OTUs were assigned a family name, and 31 were assigned to the genus level. None of the OTUs could be assigned a species name with high confidence.

**Diversity analyses**

All subsequent data analysis steps were carried out with R software v3.0.1 (R Core Team, 2013) and the support of various libraries and packages for R and from Bioconductor (http://www.bioconductor.org). First, the OTU count table, including the taxonomic assignment alongside with the OTU FASTA sequences and a project-specific metafile, were imported in R using the Bioconductor package phyloseq (see workflow in Fig. S2 in the supplemental material) (McMurdie & Holmes, 2013). For the comparison with species richness estimated from microscopy data, only OTUs assigned to photosynthetic cyanobacteria were used. To account for differences in sequencing depth among samples (see Fig. S3 in the supplemental material), the abundance of sequencing reads (cyanobacteria only) for each sample was rarefied to an even sampling depth (10,789 in Greifensee and 10,857 in Lake Zurich) using the rarefy_even_depth function in phyloseq.

The OTU composition of duplicated samples was compared by Permutational Multivariate Analysis of Variance Using Distance Matrices (PERMANOVA) using the vegan R package (Oksanen et al., 2013). The adonis function was run using the OTU table transformed into a matrix of Jaccard distances based on the incidence of OTUs. The results of the PERMANOVA confirmed our assumption that the sample replicates were not significantly different from one another. For a comparison of richness estimated from water
and sediments, linear models were run using R on the rarefied sedDNA samples and the annual estimates of pelagic species richness in water.

**Comparison of species richness using microscopy and sedDNA**

Samples for microscopy were collected monthly between 1975 and 2010 in both lakes. However, in Greifensee, some samples were not collected; therefore, the number of months sampled varies between 8 and 12 per year. To correct for the bias introduced by uneven sampling effort, we estimated the annual species and genus richness by randomly selecting eight time points per year. This random selection was made for 1,001 permutations for each year, and the median was used as our richness estimate. Even though the sampling effort was constant \((n = 12)\) for Lake Zurich over the entire sampling period, the same randomization procedure with permutations to calculate cyanobacterial richness over eight sampling dates per year was performed in order to have comparable estimates of annual richness in both lakes. To compare annual richness in pelagic and sediment samples, the mean annual richness in water was calculated by averaging values of three pelagic samples (i.e., year before, year, and year after) around the corresponding year in sedDNA samples. This last step was done to correct for potential inaccuracies in the chronology of the sediment samples.

We used linear models to explain OTU richness using species richness in water and lake identity as predictors at 12 time-points over the years for which both sequencing and microscopy data were available (between 1975 and 2010). We tested for the significance of these explanatory variables using Akaike Information Criterion corrected for small sample size (AICc). A delta AICc value \(\geq 2\) was considered to be a significant difference.

**Phylogenetic analysis of the OTUs**

To verify the accuracy of the taxonomic assignment based on UTAX, we used the OTU FASTA sequences from both lakes to construct a phylogeny. Sequences were aligned in the software Geneious (Kearse et al., 2012) using the Geneious multiple alignment tool with the default settings. *Chloroflexus aurantiacus* was added as out-group, and 13 cyanobacteria reference sequences from GenBank or CyanoBase (http://genome.kazusa.or.jp/cyanobase) were also added. The alignment was used to build a tree based on Bayesian inference (Fig. 3) under 10,000,000 generations using MrBAYES v3.2 (Ronquist et al., 2012). The first 25,000,000 generations were excluded at the burn-in step and the tree’s standard deviation of split frequencies was below 0.05.
Detection of mcyA genes in the sediments

To confirm the presence of cyanobacteria that can potentially produce microcystin, we used the published primers mcyA-Cd 1R and mcyA-Cd 1F (Hisbergues et al., 2003). These amplify an ~300-nt-long amplicon in the mcyA condensation domain of microcystin-producing strains of the genera Planktothrix, Microcystis, and Anabaena. PCR was carried out in a final volume of 50 µl containing 1× PCR Buffer I (Applied Biosystems), 2.4 U AmpliTaq Gold DNA Polymerase (Applied Biosystems), 0.5 mM of MgCl₂ (in addition to the 1.5 mM contained in buffer), 0.2 mM of each dNTP, 0.2 µM of both primers, and 4.8 µg of bovine serum albumin (BSA, GeneON, Germany). The PCR program consisted of a first step of polymerase activation at 95°C for 10 min, followed by 35 cycles at 95°C for 15 s, annealing at 56°C for 30 s, and extension at 72°C for 45 s, and a final extension step at 72°C for 5 min. The PCR products were visualized on a 2% agarose gel stained with SYBR Safe DNA gel stain. PCR products were purified using the Illustra GFX PCR DNA and gel bands purification kit (GE Healthcare, Little Chalfont, UK) and directly sequenced (Microsynth, Balgach, Switzerland) to verify the specificity of the PCR product.

Accession numbers

The sequence reads obtained in this study were deposited in the European Nucleotide Archive (ENA) under project number PRJEB13044 (http://www.ebi.ac.uk/ena). Unique mcyA sequences from this study have been deposited in GenBank under accession numbers KX437768 and KX437769.

Results

Cyanobacterial phylogenetic diversity reconstructed from sediments

A total of 163 OTUs spanning the phylum Cyanobacteria were recovered from the sediments of Greifensee and Lake Zurich (Fig. 3). In Greifensee, 78 OTUs were obtained from sequencing the 39 sedDNA samples (20 time points from ~1840 to 2012), while 85 OTUs were obtained in the 41 Lake Zurich sedDNA samples (23 time points from ~1800 to 2009). Most OTUs were assigned to photosynthetic cyanobacteria, 74 % (58 OTUs) in Greifensee and 81% (69 OTUs) in Lake Zurich (Fig. 3). About 25% (20 OTUs) in Greifensee and 19% (16 OTUs) in Lake Zurich were classified as Melainabacteria or ML635J-21, deep-branching
groups of nonphotosynthetic cyanobacteria. Interestingly, the two lakes displayed contrasting communities of nonphotosynthetic cyanobacteria, with only 7 (24%) shared OTUs (Fig. 3). Among photosynthetic cyanobacteria, 25 OTUs were unique to Greifensee, and 36 OTUs to Lake Zurich (Fig. 3). Thirty-five percent (33 OTUs) of the 95 photosynthetic cyanobacteria OTUs detected in this study were shared (i.e., identified in at least one sample from both lakes).

In order to have a comparable estimate of OTU and species richness, we calculated the annual richness of cyanobacterial orders. The OTU richness within each cyanobacterial order estimated from the sedDNA samples was similar within and between the lakes over the six time points spanning the last 40 years (Fig. 4A and B). The order Chroococcales accounted for most of the annual richness in all samples (40 to 56%). Synechococcales accounted for 13 to 26% of the annual OTU richness, whereas Nostocales, Oscillatoriales, and Pseudanabaenales each accounted for less than 20% of the annual richness estimates in all samples.

Cyanobacterial diversity in water

A total of 42 cyanobacterial species in 26 genera were identified in the pelagic samples from Greifensee between 1974 and 2010 (n = 481), and 37 species belonging to 21 genera in Lake Zurich between 1976 and 2010 (n = 420). Nineteen genera (68%) were found in both lakes. Only 2 genera (7%) were unique to Lake Zurich, whereas 7 genera (25%) were only present in Greifensee. About half of the species (53%) found in Greifensee were unique to that lake, whereas 15 out of 37 species (40.5%) composing the Lake Zurich data set were solely found in that lake. Of all species listed in the two data sets, 35% were found in both lakes, which is a similar proportion as that for the shared OTUs in the sediments (Fig. 3).

The annual species richness within each cyanobacterial order varied over the years (Fig. 4C and D). In Greifensee, the orders Chroococcales and Synechococcales generally comprised the highest proportion of species annual richness (mean = 32% in both groups), whereas in Lake Zurich, Chroococcales and Pseudanabaenales constituted, on average, 50% of all species each year. Synechococcales, Nostocales, and Oscillatoriales never individually comprised more than 25% of the annual species richness.
Figure 3. Phylogenetic tree of cyanobacterial OTUs based on Bayesian posterior probabilities. All 163 cyanobacteria 16S rRNA OTU sequences from the sediment samples of Greifensee and Lake Zurich were used to build the phylogeny. *Chloroflexus aurantiacus* was used as an outgroup and 13 additional sequences obtained from GenBank and Cyanobase were added as references. Values at nodes indicate posterior probabilities calculated from 7,500,000 trees. The shapes encode the samples from the two lakes, and the colors for the cyanobacterial order, with the nonphotosynthetic lineages of *Melainabacteria* and ML635J-21 grouped under the grey color. The 5 most abundant OTUs in Greifensee and the 3 most abundant OTUs in Lake Zurich (as presented in Fig. 4A-B) are identified with a star.
Most abundant taxa in sediments

Five OTUs predominated (in terms of abundance of sequencing reads) in Greifensee over the six time-points between 1974 and 2010 (Fig. 5A), accounting for approximately 96% of the total number of reads. All others each contributed less than 1% of the reads in that lake. In Lake Zurich, three OTUs were most abundant over the six time points between 1975 and 2010. They accounted for >80% of the reads, and all of the other OTUs each contributed less than 4% of the reads. The first two (Synechococcus spp.) and third (Anabaena sp.) most abundant OTUs in Lake Zurich were also abundant in Greifensee.

Figure 4. (Top) Proportions of cyanobacterial OTU richness within each order recovered from the sediments over the 6 years investigated between 1975 to 2010 in Greifensee (A) and Lake Zurich (B). (Bottom) Proportions of annual species richness (microscopic observations) within each order estimated from pelagic samples at the same time points in Greifensee (C) and Lake Zurich (D).
Most abundant taxa in water

The annual species abundances (cells/liter) were calculated from the time series of pelagic samples for the 6 years investigated between 1975 and 2010. In Greifensee Anathece bachmannii and Aphanocapsa delicatissima (Synechococcales) were the two most abundant species, followed by Aphanothece sp. (Chroococcales) (Fig. 5C). Planktothrix rubescens (Oscillatoriales) accounted for more than 80% of cell counts at all time-points in Lake Zurich, and Aphanothece sp. (Synechococcales) and Aphanizomenon flos-aquae (Nostocales), were the second and third most abundant species, respectively (Fig. 5D).

Figure 5. (Top) Proportions of reads of the 10 most abundant OTUs over the 6 years investigated between 1975 to 2010 in Greifensee (A) and Lake Zurich (B). (Bottom) Proportion of annual cell counts in the pelagic samples of the 10 most abundant species at the same time points in Greifensee (C) and Lake Zurich (D).
Comparison of cyanobacterial richness in sediments and water

The annual cyanobacterial richness varied in each lake, both in the sediments and the pelagic samples. Between 12 and 37 OTUs were found in the sedDNA samples from Greifensee, whereas in the microscopy data sets, the estimated annual richness varied between 6 and 20 species. In Lake Zurich, the annual richness recovered from the sediments ranged from 12 to 52 OTUs, whereas the estimated annual richness in water varied between 7 and 25 species. There was a strong and significant correlation between the annual species richness in water and the OTU richness in the sediment from the corresponding years ($r = 0.81$, $n = 12$). The linear model (Fig. 6) shows a strong and significant positive relationship between pelagic species richness and OTU richness in sediments in both lakes, with the intercepts differing between lakes (adjusted $R^2 = 0.89$; $n = 12$; $P < 0.001$).

![Figure 6](image)

**Figure 6.** Linear model showing the relationship between annual cyanobacterial richness in the water column and OTU richness in the sediments of Greifensee and Lake Zurich. The linear model including pelagic species richness and lake identity as factors was highly significant ($p < 0.001$, adjusted $R^2 = 0.89$, $n = 12$ between 1975 and 2010). The colored lines show the linear fit of the model in respective lakes (Greifensee: $y = 1.20x + 7.64$; Lake Zurich: $y = 1.20x + 20.52$) and the gray dashed line is the 1:1 relationship.
In another linear model, we used genus richness to explain OTU richness in sediments. This was to verify that the relationship between species and OTUs was not caused by biases in fine taxonomic classification. The best linear model explaining OTU richness included both genus richness and lake identity as explanatory variables (adjusted $R^2 = 0.81$; $n = 12$; $P < 0.001$; Fig. S4 in the supplemental material).

To better elucidate the differences between the richness of cyanobacteria estimated from the sediments and water samples, we used linear models on the annual richness of species grouped by order (Fig. 7). The water-sediment richness relationship was significant in Chroococcales, Oscillatoriales, Pseudanabaenales, and Synechococcales ($P < 0.02$). Lake identity was a significant factor in explaining OTU richness in all orders except for Pseudanabaenales. The linear model was not significant for the order Nostocales ($P = 0.09$).

**Figure 7.** Linear models showing the relationship of annual OTU richness estimated from the sediments and annual species richness estimated in water with samples grouped by order in Greifensee and Lake Zurich. The colors indicate the cyanobacterial order, and the gray dashed line represents the 1:1 line. The water-sediments richness relationship was significant in all orders ($P < 0.02$) except Nostocales ($P = 0.09$), and lake identity was a significant explanatory variable in the Chroococcales, Oscillatoriales, and Synechococcales models ($P < 0.01$).
Detection of potentially toxic cyanobacteria in the sediments

Because potentially microcystin-producing cyanobacteria such as *Anabaena* sp., *Microcystis* sp., and *Planktothrix rubescens* have been present in Greifensee and Lake Zurich for several decades, we screened the sequencing data for OTUs assigned to potentially toxic cyanobacterial taxa and verified the presence of *mcyA* genes in the sedDNA samples by PCR. The conservative taxonomic assignment, based on a high confidence threshold (85%), did not allow us to determine OTUs to a species level, but we identified OTUs corresponding to the aforementioned genera. In Greifensee, two OTUs assigned to genus *Microcystis* were found sporadically between the 1930s and the 1970s, and either one or both OTUs were detected in all eight samples between 1984 and 2012. In Lake Zurich, a single OTU reference sequence that is assigned to *Planktothrix* sp. was found in a majority of samples between ~1800 and 2009.

PCR amplification of *mcyA* genes confirmed the presence of potentially microcystin-producing cyanobacteria in 4 of 15 sediment layers from Greifensee and in 13 out of 19 layers in Lake Zurich (Table 1). Interestingly, in Greifensee, only the most recent sedDNA samples (years 2006, 2009, 2011 and 2012) tested positive for the presence of the *mcyA* gene. Direct sequencing of the amplicon revealed that the genes found in Greifensee were all related to the same single species of *Microcystis*. In Lake Zurich *mcyA* genes were detected in sediments between the years 1912 and 1962, and in samples between the years 1993 and 2010 (Table 1). A single *mcyA* sequence related to *Planktothrix rubescens* was present in 12 out of 13 samples from Lake Zurich (the sequencing of the sample dated to 1922 failed).

Discussion

Reconstruction of cyanobacterial phylogenetic diversity from sediments

Our work shows that it is possible to study cyanobacterial communities by sequencing DNA from lake sediment cores. We successfully sequenced amplicons recovered from DNA archived in the sediments of two lakes over the last two centuries, and we were able to validate the data with two independent time series, consisting of 40 years of phytoplankton microscopic identification from the same lakes. We also reconstructed the history of potentially microcystin-producing cyanobacteria over the last century. Our results are consistent with the historical information describing the cyanobacterial community composition in the two lakes.
One of the main limitations in sedimentary DNA studies is the degradation of DNA over time (Anderson-Carpenter et al., 2011). However, the cold and anoxic/hypoxic conditions at the bottom of the two deep and stratified lakes studied here are ideal for DNA preservation (Coolen et al., 2004). We first verified the quality of the DNA preserved over the past 200 years in the sediments of Greifensee and Lake Zurich by amplifying an 800-nt-long fragment of the 16S rRNA gene. This test confirmed the possibility to sequence a shorter DNA fragment of ~400 bp.

Table 1. PCR amplification results for mcyA genes in sediment samples from Greifensee and Lake Zurich.

<table>
<thead>
<tr>
<th>Sample year(s)</th>
<th>Greifensee</th>
<th>Zurich</th>
</tr>
</thead>
<tbody>
<tr>
<td>1840–1845</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1901</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1903</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1912</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>1920</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>1922</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>1931–1933</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1934</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>1935</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>1943</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>1950</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>1954–1957</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1961–1964</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>1968</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1975</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1982</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1984</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1993</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>1997</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>2003</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>2006</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2009</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>2010</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>2011</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>2012</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

*a The presence or absence of the gene in a given year(s) is indicated by the + or – sign.*
Another important limitation of cyanobacterial investigations using sequencing technologies is due to the lack of exhaustive and well-curated reference databases, which limits the taxonomic assignment of OTUs. While the existing reference databases (DeSantis et al., 2006; Pruesse et al., 2007; Cole et al., 2009) are well developed for microbial 16S rRNA analysis, the coverage of cyanobacteria, especially the freshwater taxa, needs to be improved. In this study, however, sequencing of a relatively large DNA fragment (400 nt) allowed us to use the OTUs reference sequences in a more detailed phylogeny analysis based on Bayesian inferences (Fig. 3). This opens up the possibility to investigate the cyanobacterial phylogenetic diversity and community structure.

Because the diversity of whole natural cyanobacterial communities had never been assessed using HTS technologies on sedimentary records, we had no clear expectations regarding the recovery efficiency of our approach. However, based on the time series of pelagic observations, we had some prior knowledge of the cyanobacterial community composition over the past 5 decades in the two lakes. The sequencing of circa 40 sedDNA samples per lake spanning 200 years yielded a similarly high diversity covering all major clades of cyanobacteria in the two lakes (Fig. 3). Even though the PCR primers were thought to be cyanobacterium-specific, they co-amplified chloroplasts and heterotrophic bacteria. This co-amplification did not have an impact on community reconstruction because the Illumina sequencing run produced millions of amplicons, which were sufficient for an optimal coverage of the cyanobacterial diversity in most samples (Fig. S3 in the supplemental material). In general, older samples contained less DNA, which lead to a lower number of amplicons sequenced. This problem can be solved by pooling DNA extracts, and increasing the template DNA for PCR amplifications or by pooling PCRs. In this study, we were mainly interested in the diversity of recent samples (i.e., between 1975 and 2010) for comparison with pelagic samples; therefore, we did not attempt to optimize the coverage in the older sample.

Interestingly, sequencing revealed the presence of unexpected deep-branching groups of cyanobacteria termed *Melainabacteria* and ML635J-21 in the sediments of the two lakes (Fig. 3). Recent evidence from whole-genome sequencing confirmed that *Melainabacteria* constitute a class within the *Cyanobacteria* phylum because both groups share common ancestral traits, such as the cell envelope structure and the presence of putative circadian rhythms (Soo et al., 2014). Taxa within this group have been detected in various environments, including groundwater, drinking water and wastewater treatment plants (Di Rienzi et al., 2013; Soo et al., 2014), terrestrial plants, and animal guts (McGorum et al.,
2015). To our knowledge, their presence has not been previously reported in lakes. The sequencing of these clades from the sediments of our two lakes shows that the cyanobacterium-specific primers used can target taxa over the whole phylum. From our results, we cannot conclude whether these nonphotosynthetic cyanobacteria live in the water column or colonize the sediments, but the contrasting diversity in the two lakes (Fig. 3) may reflect the adaptation of these taxa to specific local conditions. More eDNA studies using HTS technologies could help elucidate the ecological roles and the sensitivity of these clades to environmental changes.

**Comparison of sediment and pelagic samples**

The strong and significant relationship observed between annual cyanobacterial richness estimated at 12 time points from sedDNA and from pelagic samples between the mid-1970s and 2010 in the two lakes (Fig. 6) reinforces the validity of our reconstruction approach. The greater cyanobacterial richness observed in the sediments of both lakes compared to the microscopic estimates in the pelagic samples (Fig. 6) can be partially explained by differences in the detection limit of the two methods. Other studies have shown that diversity estimation based on morphology generally underestimate the true diversity of cyanobacteria, which emerges from genetic methods (Wilson *et al.*, 2005; Miller & McMahon, 2011). Our results suggest that the richness of *Chroococcales* and *Synechococcales* was widely underestimated in the microscopy data compared to the genetic data from sedDNA (Fig. 7). This is probably because several taxa within the two groups are unicellular picocyanobacteria (<2 µm diameter), which are difficult to classify based on morphology (Hayes *et al.*, 2007; Dvořák *et al.*, 2014).

With the sequencing approach, we were able to verify the presence of known potentially toxic cyanobacterial taxa that have been observed in the lakes, like *M. aeruginosa* and *P. rubescens*. Regular blooms of *M. aeruginosa* have been reported in Greifensee over the past 15 years (Eawag, unpublished), and the phytoplankton community of Lake Zurich has been largely dominated by *P. rubescens* over the last 3 decades (Posch *et al.*, 2012). Our data confirmed the presence of two OTUs that are assigned to *Microcystis* species in Greifensee and a single OTU sequence that is related to *P. rubescens* in Lake Zurich. The detection of the mcyA genes started at the same time that there was an increase in the abundance of one of the OTUs assigned to *Microcystis* sp. in Greifensee. Although our results show that the number of sequencing reads do not reflect the number of cells counted in pelagic samples (Fig. 5), it is likely that an increase in the relative OTU abundance reflects a change in the
pelagic community, in this particular case, the dominance of a microcystin-producing *M. aeruginosa* genotype.

In Lake Zurich, the detection of an OTU assimilated to *Planktothrix rubescens* was supported by the amplification of the *mcyA* gene related to the same taxa. The latter finding is consistent with early reports of the presence of *P. rubescens* (formerly *Oscillatoria rubescens*) forming large reddish blooms known as the Burgundy blood phenomenon at the surface of Lake Zurich in 1897 (Zülig, 1981; Liechti, 1994), and with a recent study showing that a single genotype of *P. rubescens* constituted almost 100% of the lake’s population over almost 30 years (1980 to 2008) (Ostermaier *et al.*, 2012).

The relative abundances of sequencing reads did not match the relative annual species densities estimated in the water by microscopy (Fig. 5). Several methodological and biological explanations for this result can be hypothesized. First, PCR and high-throughput sequencing of bacterial 16S rRNA genes introduce biases that can lead to inaccurate population data (Kennedy *et al.*, 2014). Also, the number of 16S rRNA gene copies per cell is known to vary among cyanobacterial taxa (Schirrmeister *et al.*, 2012). Traditional microscopy methods are also not free of biases as plankton identifications and cell estimations can vary greatly from one person to another. Rare phytoplankton taxa have been shown to be severely underestimated by traditional sampling methods (Rodríguez-Ramos *et al.*, 2013), and it is known that many cyanobacterial taxa are impossible to distinguish on the basis of their morphology only (Hayes *et al.*, 2007). Furthermore, local phenomena in the lake, such as the presence of grazers, buoyant cells, and water currents affect sedimentation rate of plankton and may influence the proportion of phytoplankton cells that can be found in the sediments. For the above-mentioned reasons and others, abundance data should be interpreted with extreme caution in sedDNA studies. Nonetheless, the high richness recovered in this study, and the strong relationship observed between independent data sets of cyanobacterial community composition (sediments versus water) confirm that the approach for reconstructing past diversity was successful in both Greifensee and Lake Zurich.

**Conclusions**

This study presents a validated approach to characterize into the past composition of cyanobacterial communities archived in lake sediments. Our results demonstrate that amplicon sequencing of a relatively large DNA fragment is useful for investigating richness and phylogenetic relatedness of cyanobacteria in lakes where the sediments are
undisturbed. Our approach applied to varved sediments will allow us to explore phylogenetic diversity and community assembly of cyanobacteria over centuries with high temporal resolution. In a forthcoming paper, we investigate the impact of human-induced environmental changes on cyanobacterial phylogenetic diversity and community structure. The ability to recover and sequence important functional genes, like those underpinning the production of secondary metabolites, will assist us in studying the factors that favored toxic cyanobacterial taxa. This approach can, in principle, be extended to other planktonic organisms to help address ecological questions, such as those related to eutrophication, climate change, colonization processes, and invasive species, which are all relevant to the assessment and management of ecosystem processes and services.

Acknowledgements

Sequencing data were generated at the Genetic Diversity Centre of ETH Zürich. Greifensee phytoplankton data was provided by Eawag, and Lake Zurich data was provided by Wasserversorgung Zürich (WVZ). We thank A. Gilli, P. Turko, A. Lück, and A. Zwyssig for their help with sampling. We also thank S. Kobel and M. Thali for their help in the lab, N. Dubois (Eawag) for offering technician time and facilities for radioisotope analyses, and A. Gilli for providing an age model of Lake Zurich sediments. Finally, we thank H. Hartikainen, S. Pichon, C. Tellenbach, and M. Thomas for helpful discussions, and the three anonymous reviewers for their comments, which improved the manuscript.

Funding

This work was supported by the Swiss Enlargement Contribution under project IZERZO-142165 (CyanoArchive) to Piet Spaak in the framework of the Romanian-Swiss Research Program.
Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two peri-Alpine lakes
**Table S1.** Description of all primers used in the study. The random nucleotides added at the 5’end of the primers are in **bold**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Adapter sequence</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYA359F_nex0</td>
<td>TCGTCGGCAGCGTCAGATGTGTAAGAGACAG</td>
<td>GGGGAATTTCCGCAATGGG</td>
</tr>
<tr>
<td>CYA784R_nex0</td>
<td>GTCTCGTGAGGCTCGAGATGTGTAAGAGACAG</td>
<td>GACTACWGGGTATCTAATCCC</td>
</tr>
<tr>
<td>CYA359F_nex1</td>
<td>TCGTCGGCAGCGTCAGATGTGTAAGAGACAG</td>
<td>NGGGGAATTTCCGCAATGGG</td>
</tr>
<tr>
<td>CYA784R_nex1</td>
<td>GTCTCGTGAGGCTCGAGATGTGTAAGAGACAG</td>
<td>NGACTACWGGGTATCTAATCCC</td>
</tr>
<tr>
<td>CYA359F_nex2</td>
<td>TCGTCGGCAGCGTCAGATGTGTAAGAGACAG</td>
<td>NGGGGAATTTCCGCAATGGG</td>
</tr>
<tr>
<td>CYA784R_nex2</td>
<td>GTCTCGTGAGGCTCGAGATGTGTAAGAGACAG</td>
<td>NGACTACWGGGTATCTAATCCC</td>
</tr>
<tr>
<td>CYA359F_nex3</td>
<td>TCGTCGGCAGCGTCAGATGTGTAAGAGACAG</td>
<td>NGGGGAATTTCCGCAATGGG</td>
</tr>
<tr>
<td>CYA784R_nex3</td>
<td>GTCTCGTGAGGCTCGAGATGTGTAAGAGACAG</td>
<td>NGACTACWGGGTATCTAATCCC</td>
</tr>
</tbody>
</table>
Figure S1. Depth-age model of the sediments of Greifensee showing caesium ($^{137}\text{Cs}$) and lead ($^{210}\text{Pb}$) profiles over the 21 upper centimetres of the sediment core. The $^{137}\text{Cs}$ profile (bars) shows two expected peaks resulting from the Chernobyl nuclear reactor accident (1986), and as a result of nuclear tests (1963). The excess $^{210}\text{Pb}$-based sedimentation rate for recent sediments in Greifensee was constant, as shown by the linear radionuclides decrease ($R^2=0.90$).
Chapter II – Supplementary Material

**Supplementary Material**

Raw Reads MiSeq PE300 → Quality Control

Paired read sample → A

**B1** End Trimming

- PRINSEQ-lite (v0.20.4)
  - min read length: 200nt
  - trim ambiguous nucleotides
  - trim low quality threshold: 10
  - trim quality window size: 5

 merges reads → B2

**B2** Merge Reads

- usearch (v8.0.1623_il68lunix64)
  - min overlap: 15
  - max overlap: 200
  - max mismatch density: 0.25

**C** Primer Trimming

- cutadapt (v1.5)
  - overlap, full length
  - error rate: 0
  - wildcards allowed

**D** Quality Filtering

- PRINSEQ-lite (v0.20.4)
  - sequence size range: 370-410
  - min quality mean: 20
  - max number of ambiguous nucleotides: 1
  - complexity filter: dust
  - complexity threshold: 10

**E** OTU Clustering

- UPARSE Workflow
  - (usearch v8.0.1623_il68lunix64)
    1. Sort reads (sortbylength)
    2. De-replicate reads (derep_fulllength)
    3. Abundance sorting (sortbysize = 5)
    4. OTU clustering (cluster_otus)
    5. Chimera removal (uchime_ref)

**OTUs**

Sample: Lake Zurich
- Number of OTUs: 954
- Reads mapped: 4,116,248
- Reads not mapped: 161,855

Sample: Greifensee
- Number of OTUs: 952
- Reads mapped: 4,238,057
- Reads not mapped: 174,419
Figure S2. Summary of the bioinformatics workflow applied to the paired-end sequences obtained from the Illumina® MiSeq Platform.
Figure S3. Rarefaction curves for all samples in Greifensee (A) and Lake Zurich (B); and for samples between 1975 and 2010 in Greifensee (C) and Lake Zurich (D).
Figure S4. Linear model showing the relationship between cyanobacterial genus richness in the water column and OTU (97% similarity) richness in the sediments of Greifensee (1975-2006) and Lake Zurich (1982-2006). The best model (based on AICc) includes both genus richness and lake identity as explanatory variables for the number of OTUs in the sediments (Adjusted $R^2 = 0.85$, $p < 0.001$, $n = 12$). The colored lines show the linear fits of the model in respective lakes. The grey dashed line represents the 1:1 relationship.
Literature cited Chapter II


Boere AC, Rijpstra WIC, de Lange GJ, Malinverno E, Sinninghe Damsté JS, Coolen MJL (2011) Exploring preserved fossil dinoflagellate and haptophyte DNA signatures to infer ecological and environmental changes during deposition of sapropel S1 in the eastern Mediterranean. Paleoceanography, 26, 1–16.


R Core Team (2013) *R: A language and environment for statistical computing*. Vienna, Austria.


Rinta-Kanto JM, Saxton MA, DeBruyn JM et al. (2009) The diversity and distribution of toxigenic Microcystis spp. in present day and archived pelagic and sediment samples from Lake Erie. *Harmful Algae, 8*, 385–394.


CHAPTER III

A century of climate change and eutrophication homogenized lake cyanobacterial communities

Marie-Eve Monchamp\textsuperscript{1, 2}, Piet Spaak\textsuperscript{1, 2}, Isabelle Domaizon\textsuperscript{3}, Nathalie Dubois\textsuperscript{4}, Damien Bouffard\textsuperscript{4}, Francesco Pomati\textsuperscript{1, 2}

\textsuperscript{1} Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, 8600 Dübendorf, Switzerland;

\textsuperscript{2} Swiss Federal Institute of Technology (ETH) Zürich, Institute of Integrative Biology, 8092 Zürich Switzerland

\textsuperscript{3} INRA, Université de Savoie Mont Blanc, UMR CARRTEL, 74200, Thonon-les-bains, France

\textsuperscript{4} Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Surface Waters Research & Management, 8600 Dübendorf, Switzerland;

\textsuperscript{5} Swiss Federal Institute of Technology (ETH) Zürich, Department of Earth Sciences, 8092 Zürich, Switzerland

\textit{Manuscript submitted to Nature Ecology & Evolution}
Abstract

Human impacts on biodiversity are well recognized, but uncertainties remain regarding patterns of diversity change at different spatial and temporal scales (González-Orozco et al., 2016; Vellend et al., 2016). Changes in microbial assemblages are, in particular, not well understood, partly due to the lack of community composition data over relevant scales of space and time (Jones et al., 2012). Here, we investigated biodiversity patterns in cyanobacterial assemblages over a century of eutrophication and climate change, by sequencing DNA preserved in the sediments of 10 lakes of the European peri-alpine region. We found species losses and gains at the lake scale, while species richness increased at the regional scale over the past ~100 years. Our data show a clear signal for homogenization of cyanobacterial communities (beta-diversity loss), with the composition and phylogenetic structure of assemblages becoming more similar across sites in the most recent decades, as have the general environmental conditions in and around the lakes. We attribute patterns of change in community composition to raised temperatures affecting the strength of the thermal stratification and, as a consequence, nutrient fluctuations, which favoured cyanobacterial taxa able to regulate buoyancy (Carey et al., 2012; Posch et al., 2012). Our results reinforce previous reports of human-induced homogenization of natural communities, and reveal how potentially toxic and bloom-forming cyanobacteria have widened their geographic distribution in the European temperate region.

Main text

Anthropogenic-induced alterations of ecosystems have contributed to reshape biodiversity of natural communities globally (Vitousek et al., 1997; Newbold et al., 2015). Patterns of responses in local and regional biodiversity are regulated by mechanisms acting at different spatial and temporal scales (Vellend, 2010). Recent controversial debates on the accuracy of meta-analyses reporting biodiversity patterns have highlighted the sensitivity of the choice of the spatial and temporal scales, or geographical bias, for the final assessment of biodiversity loss or gain (McGill et al., 2015; González-Orozco et al., 2016; Vellend et al., 2016). However, beta-diversity loss over space or time (i.e. previously differentiated assemblages increasingly resembling one another) has been consistently reported (Dornelas et al., 2014; Magurran et al., 2015; McGill et al., 2015), supporting the hypothesis that human impacts can reduce environmental heterogeneity and thus increase homogenization of natural communities (Gámez-Virués et al., 2015; Magurran et al., 2015), with potential impacts on
ecosystems functioning and services (Cardinale et al., 2012). Yet, this hypothesis remains untested in aquatic microbial assemblages, particularly photosynthetic microorganisms, whose biodiversity patterns are mostly unknown at large spatial and temporal scales.

Aquatic ecosystems are among the most sensitive to anthropogenic activities and biodiversity loss (Adrian et al., 2009). In lakes, combined warming and eutrophication (i.e. pollution caused by excessive discharge of nutrients) have favoured the dominance of cyanobacteria over eukaryotic phytoplankton (Paerl & Huisman, 2008; Sukenik et al., 2015) due to eco-physiological adaptations such as buoyancy regulation, fast growth rate, adaptation to low light levels, and ability to produce dormant cells (Carey et al., 2012; Rigosi et al., 2014). Cyanobacterial blooms create unfavourable conditions for other phytoplankton (e.g. low light conditions), as well as for organisms of higher trophic levels such as zooplankton and fish (e.g. anoxia) (Vonlanthen et al., 2012; Sukenik et al., 2015). Over the past few decades, the frequency and severity of cyanobacterial-dominated communities have increased in lakes and reservoirs worldwide (Downing et al., 2001), despite remediation measures applied at the regional and international scale (Carey et al., 2012; Taranu et al., 2015). Cyanobacterial blooms are often dominated by toxic species, and there is a global concern that climate change is promoting the geographic expansion of some potentially harmful taxa (Sinha et al., 2012; Salmaso et al., 2015a, 2015b) enhancing their dominance due to a combined effect of increasing nutrient loads (Carey et al., 2012; Rigosi et al., 2014).

The lack of long-term historical data (both biotic and abiotic), however, limits our understanding on how anthropogenically-induced changes have affected lake cyanobacterial taxa distribution and community composition over broad spatial and temporal scales. Studies in the European peri-alpine lakes have reported new occurrence of potentially harmful cyanobacteria such as Planktothrix rubescens (Ernst et al., 2009) and Dolichospermum lemmermannii (Salmaso et al., 2015b) over the last few decades that could be attributed to changes in local nutrient concentrations and a warmer climate. Yet, the relative influence of these environmental factors on the structure (diversity and composition) of cyanobacterial communities remains unclear.

In this study, we reconstructed trends in cyanobacterial community structure across 10 lakes of the European peri-alpine region (see the map in SI Fig. S1), and related diversity and compositional patterns to known long-term environmental drivers such as temperature and nutrients (Fig. 1). The average annual air temperature in the investigated region increased over the past ~150 years (SI Fig. S2a), with an intensification of the trend since the 1980s. Increased air temperature raised surface water temperatures of lakes and led to a
stronger (Fig. 1a) and longer thermal stratification (SI Fig. S2b). Over the twentieth century, peri-alpine lakes also underwent evident shifts in phosphorus and nitrogen levels (Fig. 1b-c). The severity of change in nutrient loads varied from lake to lake, but the main shifts occurred roughly simultaneously across the whole peri-alpine region. Total phosphorus (TP) increased rapidly around the 1950s to reach a maximum in the late 1960s – early 1970s (Fig. 1b). Mitigation of phosphorus discharges into lakes determined a rapid decrease of this nutrient, leading to the return of most lakes to TP concentrations that are similar to pre-eutrophication. On the contrary, nitrogen inputs have not been controlled and the concentrations of this nutrient has increased over the twentieth century (Glibert et al., 2006; Otten et al., 2012) (Fig. 1c).

We found that, as a consequence of climate change, eutrophication, and re-oligotrophication, the general environmental conditions across the study lakes have homogenized. The difference in mean annual air temperatures between north and south of the Alps decreased significantly over the last century (Fig. 1d). The annual mean TP concentrations across the lakes became also more uniform in the two most recent decades compared to the range of concentrations measured in earlier times (i.e. between 1950 and 1999; Fig. 1e). TP concentrations at present are equivalent to mesotrophic to eutrophic conditions in most lakes (Table S1, Fig. 1b and 1e). Conversely, the lakes display a wide range of NO$_3$-N concentrations, which have remained relatively constant since the maximum of eutrophication (Fig. 1f).

To study compositional and phylogenetic changes of cyanobacterial communities across lakes and over time, we used DNA extracted from sediment cores (sedDNA) collected in the 10 lakes (Methods). SedDNA has been successfully used to elucidate changes in freshwater pelagic community composition of various plankton forms, such as microbial eukaryotes (Capo et al., 2016), diatoms (Stoof-Leichsenring et al., 2015), and cyanobacteria (Domaizon et al., 2013; Savichtcheva et al., 2015; Monchamp et al., 2016). We reconstructed patterns in richness and community similarity by amplifying and sequencing part of the cyanobacterial 16S rDNA from sediment layers, and studied the dynamics of operational taxonomic units (OTUs), which represent groups of sequences with a minimum similarity of 97% to each other, within and between lakes (Methods). We have demonstrated earlier that our reconstruction method provides an accurate description of cyanobacterial community composition, which is highly correlated to patterns obtained by historical microscopic observations of water samples (Monchamp et al., 2016).
Figure 1. History of environmental conditions in the 10 peri-alpine lakes. (a) Normalized Schmidt Stability Index (SSI) values plotted against time showing the general increase in water column stability over the second half of the twentieth century. The dashed line shows the average SSI value over all lakes and time points. (b) Annual average total phosphorus concentrations plotted on a log scale (TP, [µg/L]): the dashed lines show total phosphorus concentration reconstructions based on Diatoms assemblages in the sediments. (c) Annual average nitrate concentrations in the water column plotted on a log scale (NO$_3$-N, [mg/L]). (d) Pairwise difference between the mean annual air temperature across the three meteorological stations around the Alps (N; Alpine north side - eastern Plateau, NW; Alpine north side - western Plateau, and S; Alpine south side). Significant linear fit in each time-series (NW-S: R$^2$ adj. = 0.156, p = 8.08e-12, N - S: R$^2$ adj. = 0.0949, p = 1.10e-10) is indicated with a full line. (e) Range of TP and (f) NO$_3$-N concentrations during the eutrophication (1950-1974; 1975-1999) and post-eutrophication (2000-2015) periods. Lake names: Lugano (LUG), Pusiano (PUS), Maggiore (MAG), Constance (BOD), Greifensee (GRE), Zurich (ZRH), Hallwilersee (HAL), Baldeggersee (BAL), Annecy (ANN), Geneva (GEN).

We found that the cyanobacterial community composition varied over time and across lakes (ranging from 9 to 52 OTUs, average = 25 ± 10.5 OTUs per lake) (Fig. 2a). Individual lake change in OTU richness over time was generally insignificant (at the p=0.05 level), with three exceptions: lakes Zurich, Lugano, and Hallwilersee that showed a significant increase. Overall, there was a significant temporal trend towards a slight increasing OTU richness across all lakes (R$^2$= 0.18, p = 0.000132; n = 76; Fig. 2a). Species accumulation curves
confirmed this trend across all sampled lakes (Fig. 2b), suggesting a generalised gain of species at the peri-alpine regional scale over the last century. A similar increase in taxa richness has been observed in meta-analysis studies across sites for communities that have been affected by anthropogenic disturbance, particularly post-disturbance successional changes in community composition (Vellend et al., 2013, 2016; Dornelas et al., 2014).

While reports of successful reconstructions of lake community richness and composition based on long DNA markers exist (Coolen et al., 2006; Boere et al., 2009; Hou et al., 2014), possible biases due to DNA degradation with sediment age are of concern in palaeo-genetic studies (Boere et al., 2011). We carefully inspected the dataset for the identification of patterns suggesting preferential DNA degradation (e.g. disappearance of OTUs in specific clades with sediment age), and we did not detect any significant bias. For example, cyanobacteria forming dormant cells (akinetes) were not particularly more represented or more diverse in older layers compared to more recent sediments. Our reconstructed data compare well with historical records of several of these well-studied lakes (Monchamp et al., 2016), reporting, for example, the presence of the cyanobacterium *P. rubescens* in Lake Zurich, Baldeggersee, and Hallwilersee over the last two centuries (Züllig, 1982; Liechti, 1994). In lake Zurich, Pomati et al. have also shown an increase in phytoplankton richness over the past 30 years possibly due to reduction in phosphorus inputs combined with climate warming-mediated enhancement of resource heterogeneity over the water column (Pomati et al., 2012). Joint effects of nutrients and climate have more generally been suspected to drive phytoplankton richness across lakes with different trophic status, and to favour certain cyanobacterial taxa or groups based on their eco-physiological traits (Rigosi et al., 2014; Özkan et al., 2016).

Temperature can have a direct metabolic effect on phytoplankton, but also indirect effects mediated by changes in the physical (strong thermal stratification leading to increased stability of the water column) and chemical lake state (nutrient re-circulation due to mixing events). For example, the strength of thermal stratification in lake Zurich has increased by ~20% over the last three decades due to warming, and this impaired the frequency and magnitude of lake annual mixing (Livingstone, 2003; Posch et al., 2012). Weaker mixing contributed to a reduction in the annual phytoplankton turnover in the latter lake (Matthews & Pomati, 2012) due to limited winter re-setting of the phytoplankton community, and favoured the dominance of the harmful cyanobacterium *P. rubescens* (order Oscillatoriales) (Posch et al., 2012). We used linear mixed-effect models to study the potential environmental drivers of cyanobacterial OTU richness change in our dataset (Methods). The normalized maximal annual Schmidt Stability Index and annual average TP concentrations were the only two
significant and interacting variables ($p = 0.002$, $n = 30$) predicting natural log-transformed OTU richness ($\text{marginal } R^2 = 0.614$; Fig. 2c). The fixed (SSI and TP) effects explained 61.4% of the variation in OTU richness, with no detectable additional effect of the random factor (lake). The remaining explanatory variables (NO$_3$-N, and pairwise interactions) were not statistically significant ($p > 0.2$).

Figure 2. Changes in OTU richness and phylogenetic structure. (a) Temporal plot of rarefied OTU richness in the 10 lakes (see Figure 1 caption for full lake names). Each coloured line shows the lake-specific significant (full lines) or non-significant (dashed line) relationship between OTU richness and time. The black line shows the overall fit for all lakes combined. (b) Species accumulation curves showing the average regional OTU richness estimated for a given number of lakes sampled at different time periods. The 95% confidence intervals for each curve are shown for reference in SI Fig. S3. (c) Ln-transformed cyanobacterial OTU richness relative to normalized annual maximal Schmidt Stability Index values and ln-transformed mean annual TP concentrations across all lakes over the last century. (d) Violin plots of pairwise un-weighted Unifrac similarities estimated across lake communities at each decade. The horizontal line at 0.43 represents the mean value of pairwise similarity across all samples.
The observed change in community structure has increased the average similarity between lake communities (estimated at each decade based on un-weighted Unifrac distance matrices – Methods) since the 1940s (Fig. 2d). This suggests a general homogenization of the phylogenetic community structure across lakes (i.e. loss of beta-diversity), particularly over the few most recent decades. PERMANOVA (Methods) confirmed that the groups of communities sampled at each time period have significantly different centroids ($p = 0.01$).

**Figure 3.** Proportion of rare and common OTUs across all lakes. The samples were grouped either by 10 or 25 years in order to have the highest possible number of lakes at each period without replication of lakes within each group. A simplified model depicting the intensity of change in temperature (increased warming, red) and phosphorus concentrations (eutrophication, darker green) in lakes is shown above the graphs for interpretation of patterns.
The proportion of weakly spatially distributed OTUs among lakes ("rare"; found in < 25% of lakes at a given time period) decreased over the last century. Conversely, the proportion of OTUs common in multiple lakes (found in > 75% of lakes at a given time period) increased about 4-fold between the 1950s and the 1990s, after having remained relatively stable until the period of maximum eutrophication in the late 1960s – early 1970s (Fig. 3). Our data suggest that changes in lake chemistry and physics might have promoted the dispersion of taxa that best fit the generalised contemporary environmental conditions.

The increase in similarity among cyanobacterial communities across lakes was in fact only driven by certain phylogenetic and taxa groups. Only three out of five cyanobacterial orders showed an increase in richness over the past century: Chroococcales, Nostocales, and Oscillatoriales (Fig. 4). This result indicates a higher gain in filamentous and colonial taxa able to regulate buoyancy, relative to other cyanobacteria, across peri-alpine lakes. This observation is consistent with previous studies in lakes of the same region reporting phytoplankton compositional shifts in favour of colony-forming taxa due to climate warming and changes in nutrient concentrations (Anneville et al., 2004; Salmaso, 2010; Posch et al., 2012; Gallina et al., 2013; Savichtcheva et al., 2015). The ability to regulate cells buoyancy allows Chroococcales, Nostocales and Oscillatoriales to float and sink periodically to access light (at the surface) and nutrients (below the photic zone) during strong and extended water stratification (Reynolds et al., 1987; Carey et al., 2012).

Taxa from the above orders, which comprise bloom forming and toxic organisms, particularly species of the genus Microcystis (Chroococcales), *D. lemmermannii* (Nostocales), and *P. rubescens* (Oscillatoriales), increased in prevalence and geographic range across all lakes and over time (Fig. 5). Previous studies have reported the increasing occurrence of blooms of *P. rubescens* in recently re-oligotrophied lakes around the European Alps (Ernst et al., 2009). The competitive advantage of this cyanobacterium is thought to rest in its adaptation to low light conditions, and its ability to migrate vertically over the water column to harvest nutrients (Walsby, 2005). The invasion by *D. lemmermannii*, in several deep southern subalpine lakes was also recently reported (Salmaso et al., 2015a, 2015b). In our study, we found two *Dolichospermum* spp. in the sediments of several northern peri-alpine lakes in the 1800s and during the twentieth century (Fig. 5). In the southern lakes, however, we detected *Dolichospermum* spp. only after the 1980s. Interestingly, these OTUs were found in sediments of all lakes, except in the oligotrophic Lake Annecy, suggesting that *Dolichospermum* species are able to colonise a wide range of lakes when N and P are not limiting.
The interaction between climate warming and nutrient fluctuations played an important role in our study system. It appears to have favoured taxa able to thrive in environments with a more stable water column and where limiting nutrients were on a declining trend (after they had been enriched). For instance, colony-forming and buoyancy-regulating species are favoured under such conditions (Rigosi et al., 2014). Evidence in the literature supports the hypothesis that reduction in spatial environmental heterogeneity filters species traits and determines homogenization in the composition of natural communities across sites (Gámez-Virués et al., 2015; Magurran et al., 2015). Coloniality, by increasing the species realised size, is also a grazing-defence trait (Litchman & Klausmeier, 2008). We cannot exclude that top-down control form changing zooplankton communities might have played a role in our detected trends in cyanobacterial community composition. Recent experimental evidence suggests that grazers (i.e. *Daphnia*) can significantly affect the local and regional diversity and community composition of phytoplankton (Birtel & Matthews, 2016). We have, however, no direct evidence from our data to test this hypothesis.

Our results reveal previously unreported patterns of richness and community composition in cyanobacterial assemblages over spatial and temporal scales that are relevant to both lake ecology and environmental policy. Biotic homogenization has been shown to occur in various ecosystems, including marine fish communities, as a consequence of warming (Magurran et al., 2015), and across microbial decomposers, plants, and animals in grassland communities as a consequence of land-use intensification (Gossner et al., 2016), independent of trends observed in local species richness. We show prime evidence that this process is affecting also lake cyanobacterial communities, whose homogenization could have caused the generalised problems of bloom forming and toxic cyanobacteria in the European peri-alpine region over the past decades (Jacquet et al., 2005; Ernst et al., 2009; Posch et al., 2012; Salmaso et al., 2015b). The patterns presented here are spatial- and temporal-scale dependent, and more knowledge is necessary about the state and history of cyanobacterial diversity and community composition on a European or continental scale. The loss of spatial beta-diversity via the increase in biotic homogenization related to regional changes in climate and nutrient loading signals a general problem in freshwater ecosystems that might not be reversible by applying traditional lake management schemes focused solely on phosphorus reduction.
Figure 4. OTU richness change within cyanobacterial orders. The solid black lines indicate significant fits of the generalized additive model (GAM) at a $p < 0.05$ level, and the shaded areas around the curves show the confidence intervals at 95%.
Chapter III – Beta-diversity loss in cyanobacterial communities

<table>
<thead>
<tr>
<th>Species</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synechococcus</td>
<td>15</td>
</tr>
<tr>
<td>Synechococcus</td>
<td>63</td>
</tr>
<tr>
<td>Synechococcus</td>
<td>1</td>
</tr>
<tr>
<td>Synechococcus</td>
<td>1254</td>
</tr>
<tr>
<td>Nostoc</td>
<td>354</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1307</td>
</tr>
<tr>
<td>Nostoc</td>
<td>500</td>
</tr>
<tr>
<td>Nostoc</td>
<td>860</td>
</tr>
<tr>
<td>Nostoc</td>
<td>788</td>
</tr>
<tr>
<td>Nostoc</td>
<td>871</td>
</tr>
<tr>
<td>Nostoc</td>
<td>11416</td>
</tr>
<tr>
<td>Nostoc</td>
<td>549</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1457</td>
</tr>
<tr>
<td>Nostoc</td>
<td>40</td>
</tr>
<tr>
<td>Nostoc</td>
<td>220</td>
</tr>
<tr>
<td>Nostoc</td>
<td>152</td>
</tr>
<tr>
<td>Nostoc</td>
<td>557</td>
</tr>
<tr>
<td>Nostoc</td>
<td>50</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1031</td>
</tr>
<tr>
<td>Nostoc</td>
<td>123</td>
</tr>
<tr>
<td>Nostoc</td>
<td>211</td>
</tr>
<tr>
<td>Nostoc</td>
<td>310</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1258</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1388</td>
</tr>
<tr>
<td>Nostoc</td>
<td>2152</td>
</tr>
<tr>
<td>Nostoc</td>
<td>7181</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1806</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1566</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1718</td>
</tr>
<tr>
<td>Nostoc</td>
<td>2334</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1034</td>
</tr>
<tr>
<td>Nostoc</td>
<td>742</td>
</tr>
<tr>
<td>Nostoc</td>
<td>2741</td>
</tr>
<tr>
<td>Nostoc</td>
<td>7101</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1046</td>
</tr>
<tr>
<td>Nostoc</td>
<td>2169</td>
</tr>
<tr>
<td>Nostoc</td>
<td>946</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1743</td>
</tr>
<tr>
<td>Nostoc</td>
<td>2609</td>
</tr>
<tr>
<td>Nostoc</td>
<td>815</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1354</td>
</tr>
<tr>
<td>Nostoc</td>
<td>587</td>
</tr>
<tr>
<td>Nostoc</td>
<td>631</td>
</tr>
<tr>
<td>Nostoc</td>
<td>625</td>
</tr>
<tr>
<td>Nostoc</td>
<td>430</td>
</tr>
<tr>
<td>Nostoc</td>
<td>407</td>
</tr>
<tr>
<td>Nostoc</td>
<td>118</td>
</tr>
<tr>
<td>Nostoc</td>
<td>58</td>
</tr>
<tr>
<td>Nostoc</td>
<td>175</td>
</tr>
<tr>
<td>Nostoc</td>
<td>154</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1274</td>
</tr>
<tr>
<td>Nostoc</td>
<td>143</td>
</tr>
<tr>
<td>Nostoc</td>
<td>152</td>
</tr>
<tr>
<td>Nostoc</td>
<td>104</td>
</tr>
<tr>
<td>Nostoc</td>
<td>1516</td>
</tr>
<tr>
<td>Nostoc</td>
<td>149</td>
</tr>
</tbody>
</table>

Prevalence
Figure 5. Prevalence of OTUs in all lakes between the 1900s and 2015. The absence of an OTU at a given time period is depicted by a white box, while the gradient of red illustrates the proportion of lakes where each taxon was found at each given time period (prevalence). Hierarchical clustering based on Euclidian distances revealed five clusters of taxa with different temporal patterns. Cluster I comprises five OTUs assigned to unicellular and non-bloom forming *Synechococcus* spp. (order Synechococcales) that were highly prevalent over all the studied time period (Cluster I). Cluster II contains taxa that were relatively rare (i.e. isolated in certain lakes) until the 1990s when they became common in most lakes investigated, including genera from Chroococcales (such as *Microcystis*) and Oscillatoriales (such as *P. OTU #344*, 100% similarity to *P. rubescens* previously isolated from Lake Zurich (Beard et al., 1999)). Cluster III, which contains about 43% of all OTUs identified in this study, is characterised by rare OTUs that were never high in prevalence across our study lakes. Some of them were present in a few lakes prior to the 1970s, and were not found in recent decades. Cluster IV is a group of taxa with relatively low prevalence over the twentieth century, but that are more common during the eutrophication peak. Finally, cluster V contains OTUs that mostly increased in prevalence after the period between 1950 and 1974, and that remained common thereafter. This cluster was characterised by the presence of a majority of Chroococcales, Nostocales, and Oscillatoriales including genera such as *Microcystis* and the known invasive *Dolichospermum* (OTU #614; 99% similarity to *D. lemmermannii* LN871456) previously isolated from a southern peri-alpine lake (Salmaso et al., 2015b).
Methods

Long-term physico-chemical characteristics of the lakes and environmental change. Seven lakes north, and three lakes south of the European Alps were sampled between 2009 and 2014 (see the map in SI Fig. S1). The lakes were selected on the basis of their geographical location, their eutrophication history, and because they cover a wide gradient of trophic levels and morphological characteristics (SI Table S1). Water samples from these lakes have been collected and analysed as part of multiple sampling programmes by various governmental and institutional instances over the last two to six decades (See SI Table S2 for details about sampling and source of monitoring data). Total phosphorus (TP) and nitrate (NO$_3$-N) concentrations were generally determined on a monthly basis. From these time-series, we averaged TP and NO$_3$-N integrated over the 20 uppermost meters, or over whole water when the depth-resolved data was not available (Lake Pusiano). We used publicly available homogenized monthly air temperature data (Begert et al., 2005) recorded at three meteorological stations (see map in SI Fig. S1): Zurich/Fluntern (N; Alpine north side – eastern Plateau), Lugano (S; Alpine south side), and Geneva (NW; Alpine north side – western Plateau) to show the increase in air temperature between the years 1860 and 2015.

Water stability index. We used the Schmidt Stability Index (SSI; as defined by (Schmidt, 1928)) to calculate the strength of the water column thermal stratification in all the lakes, with the exception of Lake Pusiano where no depth-resolved temperature profiles were available. We used the maximal annual index value obtained for each lake in every year where a minimum of 3 profiles had been collected in summer (between July and September; n = 312). To account for possible biases related to uneven number of water column temperature profiles sampled across the years, the depth-resolved water stability profiles were bootstrapped 1000 times by choosing 3 random points during the summer period and compared to hourly-resolved one-dimensional hydrodynamic model recently published on Lake Geneva (Schwefel et al., 2016). The estimated error in the maximal water stability index calculation was around 8% over a period of 30.

Sediment core sampling and dating. Two types of gravity corer (UWITEC-63 or Eawag-63/S corer) were used to collect sediment cores of approximately 1 m in length and 63 mm in diameter in 10 peri-alpine lakes between 2009 and 2015. The cores from Greifensee, Hallwilersee, Baldeggersee, Lake Zurich, Lake Constance, Lake Lugano, Lake Maggiore and Lake Pusiano were transported to Eawag facilities in Dübendorf where they were stored in
upright position in the dark at 4°C until analysis. The cores from Lake Geneva and Lake Annecy were preserved and processed similarly at INRA (Thonon-Les-Bains, France).

The cores from Greifensee and Lake Zurich were previously dated based on varve counting, in addition to radiometric methods for the Greifensee core (\(^{210}\)Pb, \(^{226}\)Ra, \(^{137}\)Cs) (Monchamp et al., 2016). All other cores were dated based on both varve counting (when applicable) and radiometric methods (SI Fig. S4). A sediment core from each lake was opened longitudinally, photographed, and described. A half core from each lake was subsampled at 1-cm intervals for radionuclide measurements on a high-purity germanium (HPGe) well detector (gamma spectrometer, Canberra Industries Inc.) at the Eawag (Switzerland) for lakes Hallwilersee, Baldeggersee, Constance, Lugano, Maggiore, and Pusiano) and at the Modane underground laboratory (LSM environmental radioactivity facility, France) for lakes Annecy and Geneva.

**Molecular analysis.** The sediment samples dedicated to sedimentary DNA (sedDNA) analysis were collected either from a half of the reference core, or from a second half-core that was correlated to the reference core based on lithological tie points and lamina counting performed on the reference and working cores. Sediment cores were opened and processed in a room where no DNA work had been done before, following procedures to reduce contamination with foreign DNA detailed in (Monchamp et al., 2016). Samples were transferred to a clean laboratory, where sedDNA was extracted from approximately 1g of sediments using the PowerSoil DNA Isolation kit (MoBio Laboratories, Inc., CA, USA) for all lakes except Geneva and Annecy, for which the UltraClean Soil DNA Isolation Kit from the same company was used, according to the manufacturer’s instructions. We followed strict clean-laboratory procedures and used negative controls at various analytical steps to rule-out contamination by external DNA. Samples were extracted in two replicates that were subsequently combined to reduce heterogeneity. The sequencing library preparation was done following the MetaFast protocol developed by Fasteris (Geneva, Switzerland) using cyanobacteria-specific primers (Nübel et al., 1997) that included a 10-12 nt-long tag for identifying the samples and for creating size variation in the amplicon (SI Table S3).

PCR reactions in a final volume of 40 µL contained 1 × AmpliTaq Gold Buffer I (Thermofisher Scientific), 0.2 mM of each dNTP, 0.5 × MgCl₂, 2.4U AmpliTaq Gold polymerase, 0.2 µg/µL BSA, and 0.2 µM each primer (Microsynth, Balgach, Switzerland) and 3 µL template sedDNA at a concentration of 4 ng/µL (or lower, in cases when the sedDNA was insufficient). The PCR program on the thermal cycler included a first denaturation step at 95°C 10 min, 35 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 45 s, followed by a last
extension step at 72°C 5 min. PCR products were purified using the Illustra GFX PCR DNA and gel band purification kit (GE Healthcare, Chicago, Illinois, USA) and eluted in 30 µL of TE buffer (10 mM). Purified PCR products were quantified with a Qubit Fluorometer, and 111 samples were pooled in equimolar concentration into one multiplex library. Negative samples (i.e. PCR reactions without DNA) were added to the sequencing run to control for contamination. The library was sent to the Fasteris sequencing company (Geneva, Switzerland) for the adapter ligation and sequencing on a MiSeq Illumina platform.

**Sequence data processing.** For the quality control and processing of the sequencing data, we followed an in-house workflow developed at the Genomic Diversity Centre (GDC; ETH Zürich, adapted from (Monchamp et al., 2016)). A total of 10,609,311 quality-checked, primer-trimmed and cleaned sequences from 112 samples were clustered in 2,826 operational taxonomic units (OTUs) using the UPPARSE workflow (Edgar, 2013) at a minimum sequence similarity of 97% and a read abundance threshold of 5. Four samples had less than 10,000 reads each, and were therefore removed from the dataset. Finally, we filtered out OTUs that were not assigned to cyanobacteria (confidence threshold of 0.89), and removed 13 more OTUs that were found in less than three samples over the whole data set. This was done to reduce biases introduced by very rare taxa or possible sequencing errors. The seven negative samples that were included in the sequencing run each comprised less than 1% of the reads compared to the samples average, and were therefore removed. For the OTU richness estimation, the OTU counts were rarefied to account for differences in sequencing depth. In total, the filtered dataset comprised 143 OTUs assigned to photosynthetic cyanobacteria distributed across 108 samples (down to 133 OTUs in 76 samples after rarefaction).

The taxonomic assignment of OTUs was based upon a GreenGenes database (McDonald et al., 2012) supplemented with few decoy sequences belonging to prokaryotes other than cyanobacteria (see (Monchamp et al., 2016)). Of all the photosynthetic cyanobacteria OTUs found in this study, only four were not assigned at the order level (based on a confidence threshold of 0.89). Eighteen OTUs were not assigned to the family level, sixty more were not assigned to genus level, and none were assigned a species name with high confidence.

**Phylogenetic diversity estimation.** The cyanobacterial OTUs reference sequences were aligned with PyNAST (Caporaso et al., 2010) and used to infer a phylogeny based on maximum-likelihood in FastTree (Price et al., 2010). All diversity analyses were performed with the software R v3.3.2 (R Core Team, 2013). The ‘phyloseq’ package in Bioconductor
was used to import, visualise, and subset sequence data, as well as for estimating richness and community similarity, which we respectively defined as the number of OTUs in a lake at a given point in time, and the compositional dissimilarity between sites. Community dissimilarity was estimated using both taxonomic-based (Jaccard) (Faith et al., 1987) and phylogenetic-based (Unifrac) distances (Lozupone & Knight, 2005). The patterns were comparable between the two; therefore we present and discuss only the analyses based on phylogenetic distances. We estimated OTU richness in the meta-community formed by all lakes at each given point in time using randomisation in the function ‘specaccum’ in the ‘vegan’ (Oksanen et al., 2013) package for R. The rarefied OTU counts were used to estimate richness and community dissimilarity.

**Compositional turnover.** We compared the phylogenetic dissimilarity among communities of lakes using the function ‘adonis’ in the ‘vegan’ package by permutational multivariate analysis of variance using distance matrices (PERMANOVA). The samples were split in different time periods to cover the timespan of the study. The number of samples prior to the 1940s was insufficient for calculating pairwise dissimilarity across communities, and therefore these samples were excluded from the analysis. To understand changes in cyanobacterial community composition across sites, we studied the prevalence and distribution of shared and specific OTUs across our 10 lakes over time. We calculated the proportion of lakes where each OTU was found at each given time period (prevalence). The colour-coded image map of taxa prevalence was produced using CIMminer (Weinstein et al., 1997) and taxa were clustered for pattern detection by calculating Euclidean distances based on prevalence and the dendrogram constructed by average linkage method. Rare OTUs were defined as the OTUs found in less than 25% of the lakes at a given time period, and the common OTUs were defined as the ones present in at least 75% of the lakes at a given time period.

**Statistical models.** Our objective was to compare and quantify the strength of the relationship between cyanobacteria and environmental drivers (strength of water thermal stratification, TP, and NO₃-N) over the last century in the 10 peri-alpine lakes. The annual average values of NO₃-N and TP of the year before, the year, and the year after the approximate year corresponding to a sedDNA sample was calculated for each observation (i.e. for a sample dated to year 1990, the mean values of years 1989, 1990, and 1991 were averaged and used in the model). This averaging over a 3-year period was done to account for uncertainties related to the sediments dating. The abiotic variables were weakly cross-correlated (r < 0.55). OTU richness was used as response variables in individual models, and a lake random effect was evaluated. We applied a natural logarithm transformation to the
response variable, and verified that the residuals were normally distributed. We used linear mixed-effect models (LMMs) from the ‘lme4’ package for R based on the 30 observations for which OTU richness, depth-resolved water temperature, TP, and NO$_3$-N data were available (Lake Annecy and Lake Pusiano were therefore excluded from the model). The best model was selected based on the comparison of the nested models using parametric bootstrap methods in the ‘pbkrtest’ package for R.

For calculating community dissimilarities, and for reporting the richness and prevalence of OTUs over time across all sites, the sedDNA samples were grouped by time periods of either ten or twenty-five years, depending on data availability. The lakes chemical data were pooled in a similar way (twenty-five years periods) to reflect temporal changes in the overall nutrient concentrations across lakes (Fig. 1e-f).

**Data availability.** The Illumina sequences have been submitted to the European Nucleotide Archive (ENA) under project number PRJEB21329.
Acknowledgements

High-throughput sequencing data was produced at Fasteris (Geneva). We thank J-C. Walser (GDC; ETH Zürich) for bioinformatics support, and B. Müller for helping with lake chemical data acquisition. We thank M. Thali, A. Lück, A. Zwyssig, C. Chardon, A. Lami, S. Gerli, H. Penson, and C. Ouellet-Plamondon for technical assistance, as well as M. Lavrieux for help with the sediment dating. We are grateful to H. Hartikainen, C. Tellenbach, M.K. Thomas, K. Räsänen, and R. Ptacnik for intellectual feedback and fruitful discussion. Coring, dating and DNA extractions for lakes Annecy and Geneva were performed in the context of the “REPLAY” program funded by EC2CO INSU, and the Iper Retro program funded by ANR-VULNS-005 (France). Physical-chemical data of Greifensee was produced by AWEL (Canton Zurich); Wasserversorgung Zürich (WVZ) for Lake Zurich; Lake Constance Water Supply for Lake Constance; Abteilung für Umwelt (AfU) Kanton Aargau (A. Stöckli) for Hallwilersee; Eawag/Kanton Luzern for Baldeggersee, F. Lepori for Lake Lugano, M. Manca (CNR Institute of Ecosystem Study, Italy) for Lake Maggiore; the Observatory on Alpine Lakes (SOERE OLA), CIPEL, SILA, and OLA-IS developed by Eco-Informatics ORE INRA for lakes Annecy and Geneva.

Funding

This work was supported by the Swiss Enlargement Contribution, project IZERZ0 – 142165, “CyanoArchive” to Piet Spaak, in the framework of the Romanian-Swiss Research Programme.

Author contributions: M-E.M., P.S. and F.P. designed the study. I.D. and N.D. contributed materials/analysis tools/samples. M-E.M. and I.D. collected data, and all the analyses were carried out by M-E.M. and D.B.. M-E.M. and F.P. wrote the manuscript, which was revised and edited by all authors.
A century of climate change and eutrophication homogenized lake cyanobacterial communities
**Supplementary Table S1.** List of lakes (corresponding to the abbreviations in Figures 1 and 2), geographical location, main morphological characteristics, and current trophic status. \( P_{\text{max}} \) refers to the maximum annual mean TP concentration measured over the first 20 m of the water column at the peak of eutrophication, and \( P_{\text{recent}} \) gives the most recent available annual mean TP concentration.

<table>
<thead>
<tr>
<th>Lake name</th>
<th>Lake ID</th>
<th>Max depth (m)</th>
<th>Lake area (km²)</th>
<th>Volume ( \times 10^5 )m³</th>
<th>( P_{\text{max}} ) (µg/l)</th>
<th>( P_{\text{recent}} ) (µg/l)</th>
<th>Elevation (m a.s.l.)</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Number of samples (range of sedDNA time-series in yrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>West of the Alps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geneva</td>
<td>GEN</td>
<td>309</td>
<td>582</td>
<td>89 000</td>
<td>78.69</td>
<td>13.81</td>
<td>372</td>
<td>46°27'N</td>
<td>6°32'E</td>
<td>12 (1872-2007)</td>
</tr>
<tr>
<td>Annecy</td>
<td>ANN</td>
<td>82</td>
<td>27.59</td>
<td>1 124.5</td>
<td>15.65</td>
<td>6.00</td>
<td>447</td>
<td>45°51'N</td>
<td>6°10'E</td>
<td>10 (1913-1996)</td>
</tr>
<tr>
<td><strong>North of the Alps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constance (Upper Lake)</td>
<td>BOD</td>
<td>252</td>
<td>472</td>
<td>51 400</td>
<td>83.90</td>
<td>7.60</td>
<td>395</td>
<td>47°35'N</td>
<td>9°28'E</td>
<td>11 (1901-2014)</td>
</tr>
<tr>
<td>Zurich (Lower Lake)</td>
<td>ZRH</td>
<td>137</td>
<td>67.3</td>
<td>3 300</td>
<td>67.81</td>
<td>11.89</td>
<td>406</td>
<td>47°15'N</td>
<td>8°41'E</td>
<td>12 (1890-2009)</td>
</tr>
<tr>
<td>Hallwilersee</td>
<td>HAL</td>
<td>47</td>
<td>10</td>
<td>280</td>
<td>213.81</td>
<td>18.27</td>
<td>449</td>
<td>47°17'N</td>
<td>8°12'E</td>
<td>12 (1798-2014)</td>
</tr>
<tr>
<td>Baldeggersee</td>
<td>BAL</td>
<td>66</td>
<td>5.2</td>
<td>173</td>
<td>439.96</td>
<td>23.66</td>
<td>483</td>
<td>47°11'N</td>
<td>8°15'E</td>
<td>10 (1878-2013)</td>
</tr>
<tr>
<td>Greifensee</td>
<td>GRE</td>
<td>33</td>
<td>8.5</td>
<td>148</td>
<td>492.43</td>
<td>40.20</td>
<td>435</td>
<td>47°21'N</td>
<td>8°41'E</td>
<td>11 (1797-2010)</td>
</tr>
<tr>
<td><strong>South of the Alps</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maggiore</td>
<td>MAG</td>
<td>370</td>
<td>212</td>
<td>37 500</td>
<td>31.00</td>
<td>11.00</td>
<td>194</td>
<td>45°57'N</td>
<td>08°38'E</td>
<td>11 (1908-2014)</td>
</tr>
<tr>
<td>Lugano #</td>
<td>LUG</td>
<td>288</td>
<td>48.9</td>
<td>5 860</td>
<td>177.25</td>
<td>16.30</td>
<td>270</td>
<td>45°58'N</td>
<td>08°57'E</td>
<td>11 (1900-2014)</td>
</tr>
<tr>
<td>Posina</td>
<td>PUS</td>
<td>24</td>
<td>5</td>
<td>69</td>
<td>197.00</td>
<td>88.00</td>
<td>259</td>
<td>45°48'N</td>
<td>09°16'E</td>
<td>11 (1883-2014)</td>
</tr>
</tbody>
</table>

# Refers to the Figino basin in Lake Lugano.
**Supplementary Table S2.** Source of phosphorus and nitrogen data and length of the available time series in the 10 lakes.

<table>
<thead>
<tr>
<th>Lake</th>
<th>Data collection or reference</th>
<th>TP</th>
<th>TP20m</th>
<th>NO3</th>
<th>Time-series P (from - to)</th>
<th>Time-series N (from - to)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annecy</td>
<td>Observatory on Alpine Lakes (SOERE OLA), CISALB, CIPEL, SILA, and OLA-13 developed by Eco-Informatics ORE INRA</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1966-2010</td>
<td>1966-1961/2004-2010</td>
</tr>
<tr>
<td>Baldeggersee</td>
<td>Eawag/Kanton Luzern</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1957-2011</td>
<td>1957-2011</td>
</tr>
<tr>
<td>Geneva</td>
<td>Observatory on Alpine Lakes (SOERE OLA), CISALB, CIPEL, SILA, and OLA-13 developed by Eco-Informatics ORE INRA</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1957-2010</td>
<td>1958-2010</td>
</tr>
<tr>
<td>Greifensee</td>
<td>Amt für Amt, Wasser, Energie und Luft (AWEL), Canton Zürich</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1951-2010</td>
<td>1947-2010</td>
</tr>
<tr>
<td>Hallwilersee</td>
<td>Abteilung für Umwelt (AUF) Kanton Aargau</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1969-2010</td>
<td>1997-2010</td>
</tr>
<tr>
<td>Lugano</td>
<td>Federal Office for the Environment (FOEN)</td>
<td></td>
<td>X</td>
<td></td>
<td>1983-2006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CIPAIS - Commission Internazionale per la Protezione delle acque Iaio – Svizzera, SUPSI - University of Applied Sciences and Arts of Italian Switzerland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maggiore</td>
<td>CNR Institute of Ecosystem Study, Italy</td>
<td>X</td>
<td></td>
<td></td>
<td>1972-2014</td>
<td></td>
</tr>
<tr>
<td>Zurich</td>
<td>Wasserversorgung Zürich (WVZ)</td>
<td>X</td>
<td>X</td>
<td></td>
<td>1972-2014</td>
<td>1976-2008</td>
</tr>
</tbody>
</table>

TP: Total phosphorus over the whole water column
TP20m: Total Phosphorus over the 20 upper meters of the water column
NO3: Nitrate over the 20 upper meters of the water column
### Supplementary Table S3. List of all samples, primers sequences and tags sequences used in this study.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Forward primer</th>
<th>Tag forward primer</th>
<th>Reverse primer sequence</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUG6_1</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R1</td>
</tr>
<tr>
<td>LUG6_2</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R2</td>
</tr>
<tr>
<td>LUG6_3</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R3</td>
</tr>
<tr>
<td>LUG6_4</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R4</td>
</tr>
<tr>
<td>LUG6_5</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R5</td>
</tr>
<tr>
<td>LUG6_6</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R6</td>
</tr>
<tr>
<td>LUG6_7</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R7</td>
</tr>
<tr>
<td>LUG6_8</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R8</td>
</tr>
<tr>
<td>LUG6_9</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R9</td>
</tr>
<tr>
<td>LUG6_10</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R10</td>
</tr>
<tr>
<td>LUG6_11</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R11</td>
</tr>
<tr>
<td>LUG6_12</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R12</td>
</tr>
<tr>
<td>LUG6_13</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R13</td>
</tr>
<tr>
<td>LUG6_14</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R14</td>
</tr>
<tr>
<td>LUG6_15</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R15</td>
</tr>
<tr>
<td>LUG6_16</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R16</td>
</tr>
<tr>
<td>LUG6_17</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R17</td>
</tr>
<tr>
<td>LUG6_18</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R18</td>
</tr>
<tr>
<td>LUG6_19</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R19</td>
</tr>
<tr>
<td>LUG6_20</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R20</td>
</tr>
<tr>
<td>LUG6_21</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R21</td>
</tr>
<tr>
<td>LUG6_22</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R22</td>
</tr>
<tr>
<td>LUG6_23</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R23</td>
</tr>
<tr>
<td>LUG6_24</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R24</td>
</tr>
<tr>
<td>LUG6_25</td>
<td>F1</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R25</td>
</tr>
<tr>
<td>MAGI_1</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R1</td>
</tr>
<tr>
<td>MAGI_2</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R2</td>
</tr>
<tr>
<td>MAGI_3</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R3</td>
</tr>
<tr>
<td>MAGI_4</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R4</td>
</tr>
<tr>
<td>MAGI_5</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R5</td>
</tr>
<tr>
<td>MAGI_6</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R6</td>
</tr>
<tr>
<td>MAGI_7</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R7</td>
</tr>
<tr>
<td>MAGI_8</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R8</td>
</tr>
<tr>
<td>MAGI_9</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R9</td>
</tr>
<tr>
<td>MAGI_10</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R10</td>
</tr>
<tr>
<td>MAGI_11</td>
<td>F6</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R11</td>
</tr>
<tr>
<td>ANN_3</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R2</td>
</tr>
<tr>
<td>ANN_4</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R3</td>
</tr>
<tr>
<td>ANN_5</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R4</td>
</tr>
<tr>
<td>ANN_6</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R5</td>
</tr>
<tr>
<td>ANN_7</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R6</td>
</tr>
<tr>
<td>ANN_8</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R7</td>
</tr>
<tr>
<td>ANN_9</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R8</td>
</tr>
<tr>
<td>ANN_10</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R9</td>
</tr>
<tr>
<td>ANN_11</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R10</td>
</tr>
<tr>
<td>ANN_12</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R11</td>
</tr>
<tr>
<td>ANN_13</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R12</td>
</tr>
<tr>
<td>ANN_14</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R13</td>
</tr>
<tr>
<td>ANN_15</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R14</td>
</tr>
<tr>
<td>ANN_16</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R15</td>
</tr>
<tr>
<td>ANN_17</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R16</td>
</tr>
<tr>
<td>ANN_18</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R17</td>
</tr>
<tr>
<td>ANN_19</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R18</td>
</tr>
<tr>
<td>ANN_20</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R19</td>
</tr>
<tr>
<td>ANN_21</td>
<td>F7</td>
<td>NNAAGAAGTAC</td>
<td>GGAGGTAATTTCGCGAAAGG</td>
<td>R20</td>
</tr>
</tbody>
</table>
### (Supplementary Table S3 continued)

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Name forward</th>
<th>Tag forward primer</th>
<th>Forward primer sequence</th>
<th>Name reverse</th>
<th>Tag reverse primer</th>
<th>Reverse primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEM_1</td>
<td>F7</td>
<td>NNNTCTCGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R11</td>
<td>NNNNGAGCTGAA</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_3</td>
<td>F7</td>
<td>NNNTCTCGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R12</td>
<td>NNNNGGCTGAA</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_4</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R3</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_10</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R2</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_11</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R3</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_15</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R4</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_22</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R5</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_27</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R6</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_29</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R7</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_32</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R8</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_35</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R9</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_42</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R10</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>LEM_47</td>
<td>F8</td>
<td>NNNCATCGCTGTA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R11</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_18B</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R1</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_27</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R2</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_29</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R3</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_39</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R4</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_43B</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R5</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_49</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R6</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_72</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R7</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_74</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R8</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_84</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R9</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>G13_78</td>
<td>F9</td>
<td>NNNNCACACCTCAAG</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R10</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z19_25</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R1</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z19_6</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R2</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_1</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R3</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_4</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R4</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_5</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R5</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_6</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R6</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_9</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R7</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_11</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R8</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_12</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R9</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z18_15</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R10</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
<tr>
<td>Z19_3</td>
<td>F10</td>
<td>NNNNTCTTGTGGCA</td>
<td>GGGGAATTCCCCCAATGGG</td>
<td>R12</td>
<td>NNCTCTGAC</td>
<td>GACTACWGGGGTATCTAATCCC</td>
</tr>
</tbody>
</table>
Supplementary Figure S1. Map of the study area showing the location of the 10 lakes around the European Alpine mountain range. The location of the meteorological stations where air temperatures were recorded is indicated by the black squares (north-eastern Plateau, north-western Plateau, and Alpine south side).
Supplementary Figure S2. (a) Annual average air temperatures at three meteorological stations (N; Alpine north side - eastern Plateau, NW; Alpine north side - western Plateau, and S; Alpine south side; see location on Fig. S1). (b) Duration of the period of annual thermal stratification in the nine peri-alpine lakes between 1950 and 2015. The significant linear fit ($R^2_{adj.} = 0.18, p = 9.84e^{-15}$) over all time-points across all lakes is indicated with a black line. Lake Pusiano is not included in this analysis because no depth-resolved water temperature data were available. We evaluated the duration of the period of thermal stratification based on a minimum of 8 profiles per year. We defined thermal stratification as the period during which the temperature difference between the surface waters and the 50 meters-depth water layer (or the bottom layer) exceeded 3°C.
Supplementary Figure S3. Species accumulation curves showing the average regional OTU richness estimated for a given number of lakes sampled at different time periods from the 1880s to the 2010s. The number of lakes within each decade is variable, and the curves reflect the increase in the number of new OTUs found when sampling additional lakes. For each time period, the sequence of lakes is bootstrapped 100 times and the shaded area around each curve shows the confidence intervals at 95%.
Supplementary Figure S4. Caesium-137 downcore profiles. The peaks marking the fallout after the Chernobyl accident (1986) and the maximum fallout due to nuclear weapon tests (1963) are marked in the plots. For Lake Zurich, the dating was based on annual varve counts, which are visible in the sediments since the early 1900s.
Literature cited Chapter III


R Core Team (2013) R: A language and environment for statistical computing. Vienna, Austria.


CHAPTER IV

Sedimentary and egg-bank DNA from 3 European lakes reveal concurrent changes in the composition and diversity of cyanobacterial and *Daphnia* communities

Marie-Eve Monchamp*1,2, Ioana Enache1,3,4, Patrick Turko1, Francesco Pomati1,2, Geta Rîșnoveanu4, Piet Spaak1,2

1 Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Dübendorf, Switzerland;

2 Swiss Federal Institute of Technology (ETH) Zürich, Institute of Integrative Biology, Switzerland;

3 Institute of Biology Bucharest, Department of Ecology, Taxonomy and Nature Conservation, Romanian Academy, Bucharest, Romania;

4 University of Bucharest, Faculty of Biology, Department of Systems Ecology and Sustainability, Doctoral School in Ecology, Bucharest, Romania

Published online in *Hydrobiologia*, doi: 10.1007/s10750-017-3247-7

The definitive version is available at http://rdcu.be/tn2E
Abstract

Eutrophication generally favours the growth of cyanobacteria over eukaryotic green algae in freshwater lakes. Cyanobacteria constitute a poor food source for the waterflea Daphnia, an important primary consumer of phytoplankton in lakes. While it is known that some Daphnia species are adapted to eutrophic conditions and can cope with cyanobacteria in their diet, it is less known whether cyanobacterial community composition can influence Daphnia population structure in lakes. We studied the variation in genetic diversity of Daphnia resting eggs and cyanobacterial DNA preserved in sediment cores from three European lakes impacted by eutrophication. Our retrospective analysis confirms that D. galeata invaded the two pre-alpine lakes around the middle of the twentieth century, hybridized with and became dominant over D. longispina. This coincides with the presence in all lakes and the increase in the proportion of colonial and filamentous cyanobacteria in the pre-alpine lakes. The recent re-oligotrophication of the lakes did not reverse the cyanobacterial and Daphnia assemblages to their pre-eutrophication composition and genetic structure, suggesting that both changed irreversibly due to anthropogenic influence on the ecosystems. Genetic analyses applied to lake sedimentary archives have the potential to unveil how different compartments of the food-web co-vary in a changing environment.
Introduction

A majority of lakes have been impacted by human-induced eutrophication (i.e., the over-enrichment by nutrients) over the past century (Schindler 2012; Smith 2003), with negative consequences on water quality and on ecosystem processes (Smith et al. 2006; Sukenik et al. 2015). In Europe, most lakes underwent a major phase of eutrophication around the middle of the twentieth century that was in most cases followed by a re-oligotrophication phase due to targeted reduction of phosphorus loading (Anneville et al. 2005; Vonlanthen et al. 2012). Re-oligotrophication generally led to improved water quality and ecosystem functioning by alleviating symptoms of eutrophication, such as deep-water oxygen depletion that affects fish survival, and low autotroph to herbivores energy transfer efficiency characteristic of cyanobacteria-dominated phytoplankton communities (Bürgi et al. 2003; Müller et al. 2014; von Ellert et al. 2003). The long-term effects of eutrophication followed by re-oligotrophication on the genetic structure of plankton communities, however, remain largely unknown. One of the first long-term (100 yr) population genetic structure assessments of the impact of eutrophication showed that the composition and genetic diversity of populations of the waterflea Daphnia underwent permanent changes attributable to eutrophication (Brede et al. 2009). Similarly, long-term changes (over 100 yr) in diversity of Synechococcus (picocyanobacteria) populations in a deep pre-alpine lake have been linked to the lake’s trophic status, and revealed a similar phylogenetic composition during the periods of pre-eutrophication and re-oligotrophication (Domaizon et al. 2013).

Daphnia (Crustacea: Anomopoda) is ecologically important because of its dual role as a primary consumer of phytoplankton and an important food for fish (Lampert 2006), thus occupying a central position in many lentic food webs. Daphnia dominates the zooplankton community and often reaches high densities in many lakes and ponds worldwide. Under harsh conditions (e.g., food limitation, cooling), Daphnia reproduces sexually, producing haploid sexual eggs. These eggs are contained in a sclerotized sac (ephippium), which often sink to the sediments, from which they may hatch the following spring or be buried. Buried eggs may preserve their DNA intact and remain viable for decades (e.g., Turko et al. 2016) or even centuries (Frisch et al. 2014), making it possible to use them to investigate changes in the genetic structure of Daphnia populations over long time-scales (Brede et al. 2009; Kerfoot et al. 1999; Turko et al. 2016).

The Daphnia longispina-galeata-cucullata complex is one of the most prevalent in European lakes (Schwenk and Spaak 1995; Petrusak et al. 2008a). Both interspecies hybridisation and introgression within the complex are known, and members are differentially
adapted to environmental conditions (Flößner 2000; Petrusek et al. 2008b; Schwenk et al. 2000). Due to this local adaptation to a variety of conditions, combined with the excellent dispersal ability of ephippia (Bohonak and Jenkins 2003), environmental change has the potential to dramatically alter the ranges of Daphnia species. Human-induced eutrophication of lakes has been identified as a main cause of biological invasions by species of the genus Daphnia in Europe and elsewhere (Moest et al. 2015; Brede et al. 2009; Weider et al. 1997). One documented example is the invasion by Daphnia galeata of several European pre-alpine lakes during the period of maximal eutrophication in the 1960s and 1970s. This major event was followed by hybridisation with native species and subsequent introgression, leading to drastic changes in the genetic composition of local Daphnia populations (Alric et al. 2016; Brede et al. 2009; Rellstab et al. 2011).

One main consequence of the eutrophication process is that it favours primary productivity of phytoplankton, especially of cyanobacteria (Paerl et al. 2001). Changes in environmental conditions (such as fluctuations in nutrient concentrations) influence cyanobacterial abundance, but also its taxonomic composition in lakes (Rigosi et al. 2014). Warming and over-enrichment of nutrients are important factors facilitating the global expansion of cyanobacteria, especially “invasive” taxa in the orders Chroococcales and Nostocales (Reviewed in Sukenik et al. 2015). The latter two groups comprise taxa that have eco-physiological traits, like buoyancy regulation, that allow them to thrive in light depleted and highly competitive environments (Carey et al. 2012). While some of these range shifts in response to environmental change have been documented, little information is available about the long-term genotypic population structure changes of Daphnia populations and its consequences for the food-web.

In this study, we seek to trace the genetic histories of several Daphnia populations during and after human-induced eutrophication, a period that was characterized by the dominance of cyanobacteria in the phytoplankton community. Cyanobacteria have long been known to be a poor food source for zooplankton in general (Lampert 1987) for various reasons. They lack essential nutrients needed by zooplankton (von Elert et al. 2003), large colonies interfere with the feeding apparatus of Daphnia, and although smaller filaments or colonies pieces can be ingested, they might be poorly digested or assimilated. Further, some strains of cyanobacteria can produce toxic compounds, which can affect Daphnia reproduction and survival (Drugă et al. 2016; Rohrlack et al. 1999). The microcystins form the most common group of cyanotoxins observed in temperate lakes (Chorus and Bartram 1999). They are synthesized mainly by colonial and bloom-forming cyanobacteria of the
genera *Microcystis*, *Planktothrix*, and *Dolichospermum* (formerly called *Anabaena*), which often dominate phytoplankton communities in productive lakes. Thus, knowing about the cyanobacterial species composition and the presence of microcystins-producing strains in natural lake communities is important for understanding the interaction between *Daphnia* and cyanobacteria.

*Daphnia* – cyanobacteria interactions have been extensively discussed in the literature, but much of the findings have been contradictory (Ger et al. 2016; Lemaire et al. 2012). Genotype × genotype interactions might explain some of the contradicting results reported, as it appears that certain *Daphnia* genotypes are better adapted than others to local conditions (including cyanobacterial presence). For example *D. cucullata* have been shown to have a competitive advantage over other *Daphnia* in the presence of filamentous cyanobacteria (Gliwicz and Lampert 1990), which contribute to their success in eutrophic lakes. Lemaire et al (2012) found that the reaction of daphnids to cyanobacteria depended not only on the origin of the *Daphnia* clone, but also on the lake or pond from which the cyanobacteria strain originated. The physiological sensitivity of zooplankton taxa to microcystins has also been shown to differ (DeMott et al. 2001). These results suggest that *Daphnia* are capable of evolving adaptive mechanisms to cope with cyanobacteria. This suggestion has been experimentally confirmed using the “resurrection ecology” (Kerfoot et al. 1999) approach on the dormant egg bank (Hairston et al. 2001; Hairston et al. 1999; Jiang et al. 2016).

Palaeolimnological reconstructions are useful to track compositional (species) and genetic changes in plankton assemblages. *Daphnia* populations have been successfully reconstructed from lakes sedimentary egg bank (Ishida et al. 2012; Limburg and Weider 2002; Turko et al. 2016) over long time scales. Similarly, DNA preserved in lacustrine sediments (sedimentary DNA) was also used to track long-term changes in the diversity of a broad range of pelagic organisms, such as cyanobacteria (e.g., Monchamp et al. 2016; Savichtcheva et al. 2011), eukaryotic algae (Hou et al. 2014; Stoof-Leichsenring et al. 2012), protists (Capo et al. 2016), and zooplankton (Bissett et al. 2005). Here, using sedimentary and egg-bank DNA, we aimed at identifying trends in taxonomic and genetic diversity of *Daphnia* and cyanobacteria, and at determining whether shifts in community composition happened at the same time at two levels of the food web. For this, we reconstructed 1) the taxonomic and genetic diversity in *Daphnia* populations from the sedimentary egg bank using microsatellites, and 2) the taxonomic and phylogenetic diversity of cyanobacteria (including
potentially toxic strains) by sequencing DNA preserved in sediment cores of three lakes dating back ~60 to 80 years.

**Material and Methods**

**Lake characteristics**

A summary of the main characteristics of the three lakes is available in Online Resource 1 in the Supporting Information. The lakes were chosen because the structure of their sedimentary archive showed potential for dating. In addition, the two pre-alpine lakes have been well studied and the history of eutrophication of these lakes is known. For the third lake located in Eastern Europe, no long-term data was available, and little is known about the timing and the strength of the human-mediated eutrophication.

Lake Gorgova (long. 45.15, lat. 29.18; surface area 13.8 km$^2$) is located in the Danube Delta (Romania), which is the end point of the Danube River and its connection to the Black Sea. The Delta comprises of hundreds of lakes, all of which are interconnected to different degrees to each other or to the main branches of the Danube River (Oosterberg et al. 2000). Gorgova is a relatively shallow lake (maximum depth ~ 3 m), but because of its connection with the main river, its water level fluctuates seasonally. The lake is completely mixed several times per year due to the dominant winds in the region and the shallow depth of the lake. With a concentration of total phosphorus (TP) of 26 µg/l measured over the water column in spring 2015, the lake classifies as meso-eutrophic (OECD 1982). The lake floor is covered with high densities of the submerged macrophytes *Potamogeton* sp. (Oosterberg et al. 2000; This study).

Lake Greifensee (long. 47.34, lat. 8.68) is located 20 km east of the city of Zurich, north of the Swiss Alps. It has a surface area of 8.5 km$^2$ and a maximum depth of 33 m. The lake is eutrophic (annual mean TP concentration around 40 µg/l) and is completely mixed once every winter. The phosphorus concentration in the lake has been monitored on a monthly basis over the past six decades, and the lake is known for presenting regular blooms of *Microcystis* species (Monchamp et al. 2016).

Lake Hallwil (Hallwilersee; long. 47.28, lat. 8.21) is another pre-alpine lake located about 65 km west of Greifensee. Its size (10 km$^2$) and maximum depth (46 m) are similar to Greifensee, but the lake is meso-eutrophic (mean annual TP concentration around 18 µg/l). Hallwilersee also differs from Greifensee with respect to water mixing: because of the
presence of relatively high hills around the lake, the winds are not sufficient to trigger natural mixing of the lake, even though the spring and autumn thermal conditions are adequate (Liechtli 1994). The chemical characteristics of Hallwilersee have been monitored for five decades, and the presence of the potentially harmful cyanobacterium *Planktothrix rubescens* in the lake has been reported since the 1880s (Züllig 1982).

Both pre-alpine lakes underwent a major phase of eutrophication around the mid twentieth century, caused by massive inputs of phosphorus to the lakes as a consequence of the increasing agriculture and urban development activities in the watershed from the 1950s onward (Liechtli 1994). Governmental efforts to control pollution by phosphorus (installation of wastewater treatment plants, ban of phosphorus in detergents) helped reduce phosphorus concentrations in most lakes of the region since the end of the 1980s. In addition to these measures, the authorities of Hallwilersee started to mechanically aerate the bottom of the lake in 1986 (Stöckli 2010) to help reduce the release of phosphorus from the sediments. Although there are no long-term records of the water chemistry in Lake Gorgova, it is known that the general water quality in the Danube Delta has severely degraded over the past few decades as a consequence of increased pollution, both locally and along the Danube River (reviewed in Oosterberg et al. 2000). The human-induced changes in the Danube Delta wetlands over the second half of the twentieth century (e.g., man-made channelization, fisheries) and the nutrient-polluted Danube river waters that feed the Delta have accelerated eutrophication (Cristofor et al. 1993). Overall in the Delta, the period of maximal eutrophication was recorded during the second half of the 1980s (Vădineanu et al. 1992), when concentrations of total reactive phosphorus were as high as 200 µg/l in some lakes (Rișnoveanu et al., 2004). After 1990, the economic collapse in the Danube river basin helped reduce phosphorus concentration in both the Danube river waters and the lakes in the Delta.

**Sediment sampling, description, and dating**

In Lake Gorgova, three sediment cores of 24 to 34 cm in length were collected from the centre of the lake in May 2015 using a gravity corer equipped with PVC tubes of 63 mm in diameter. The cores were sealed and kept in a cool and dark place until they were transported back to Eawag (Dübendorf, Switzerland) where they were stored at 4°C in a vertical position until opening. Sediment cores (63 mm diameter) were collected from the centre of lakes Greifensee in June 2013 and Hallwilersee in 2012 and 2014 using an
These sediment cores were preserved as previously described until they were opened and all cores were processed within few weeks after sampling.

The sediment cores were sliced longitudinally using large brass blades. The sediments were photographed immediately, allowed to oxidize at 4°C in the dark for 24 hours to improve contrasts, and photographed again. The longest core from Lake Gorgova (34 cm) was used for sediment dating. Briefly, the sediments layers were sampled at 1-cm intervals in one half of the sediment core using sterile blades. The samples were transferred to plastic boxes and were stored at −20°C overnight before they were freeze-dried for 48 hours. The dried samples were then crushed, and all coarse particles (e.g., mollusc shells) were removed. The homogenized sediments were weighted into plastic tubes for gamma spectrometry measurements (\(^{7}\)Be, \(^{137}\)Cs, \(^{210}\)Pb) on a high-purity germanium (HPGe) well detector (Canberra Industries). The \(^{137}\)Cs and \(^{210}\)Pb isotopes are useful to accurately date sediments back to the 1950s and to estimate sedimentation rate. Because of its very short half-life (~53 days), \(^{7}\)Be can be used to verify the presence of the surface sediments in the core (up to ~1 year). For the two pre-alpine lakes, the sediment cores were previously dated based on \(^{137}\)Cs, \(^{210}\)Pb, and varve counts (Monchamp et al. in prep; Monchamp et al. 2016).

The recent sediments (~75 years) of Greifensee and Hallwilersee are characterized by the presence of annual laminations (varves) that formed as a consequence of high primary productivity and oxygen depletion in the deep-water layers in the two lakes. The age of the sediments is relatively easy to determine based upon these annual varves that are composed of a pale layer corresponding to spring/summer, and a dark layer that forms during winter (Zolitschka 2007).

**Sediment core preparation for ephippia extraction**

After dating, we used our core photographs to apply the age model to the cores from which we retrieved *Daphnia* eggs. For lakes Greifensee and Hallwilersee, where annual varves are visible, we sampled the sediment cores by slicing 1 or 2-year sections with acid-washed zinc blades. For Lake Gorgova, the temporal resolution provided by the age model was lower; therefore we cut layers of 1-cm, which generally corresponds to a 2 to 3-year period. In Greifensee, four time periods were chosen (1956-57, 1966-67, 1974-75, and 2003-04) to capture the time of intense cultural eutrophication, which resulted in rapid *Daphnia* community composition changes (Brede et al. 2009), and one more recent time period after re-oligotrophication. Similarly, we sampled the Hallwilersee sediments for the years 1950, 1960, 1970, and 2000. In Lake Gorgova, we did not have prior knowledge on the exact timing or the extent of the eutrophication and re-oligotrophication periods, therefore we chose
to sample sediments layers (sixteen in total) at high temporal resolution spanning between the mid-1950s and 2015. These sediment sections were passed through a clean 150μm brass sieve, and the remnants were examined under a stereomicroscope at 20× magnification to pick out the *Daphnia* ephippia. The ephippia were counted, individually placed in a drop of water, and opened with flame-sterilized dissection needles. The eggs were removed and transferred into 200 μl tubes containing 25 μl alkaline lysis buffer for DNA extraction as described below. In case an egg broke during handling, only the broken egg was kept, as we considered the unbroken egg to probably be contaminated. Although the two eggs from each ephippium have the same mother and cannot be considered fully independent, we chose to analyze both, for two reasons. First, as our goal was reconstruct the population as it existed in nature, we sampled the egg bank as fully as possible. All of the eggs would have contributed to the next generation, had they hatched. Second, as the eggs are produced sexually, each of them should represent different genotypes. The DNA was extracted from these eggs following the HoTShot protocol (Montero-Pau et al. 2008) and stored at −20°C until use. The eggs were genotyped using microsatellites as detailed below.

**Microsatellite genotyping and analysis**

For assessing the genetic diversity of the *Daphnia* egg banks in the three lakes, we amplified a panel of 8 microsatellite markers in an optimized multiplex protocol (DaB10/14, Dp512, SwiD1, SwiD10, SwiD12, SwiD14, SwiD4, and SwiD5; Brede et al. 2006). Because there was low genetic variation in Lake Gorgova, we added another eight microsatellite markers to the aforementioned ones in order to obtain greater resolution for taxa in the *D. longispina* complex. The markers were SwiD2, SwiD6, SwiD15, DaB17/17, Dp519, Dp281NB, Dgm105, and Dgm109 (Brede et al. 2006).

The microsatellite markers were analyzed on an ABI 3130 XL sequencer (Applied Biosystems), and the fragments length were inferred using the software STRand (Toonen and Hughes 2001) version 2.4.59 (http://www.vgl.ucdavis.edu/STRand) and binned into integer alleles using the R package ‘MsatAllele’ (Alberto 2009). The binned allele data were used for the determination of species identity via comparison with a panel of reference genotypes, which consisted of 57 *D. galeata*, 30 *D. longispina*, 31 *D. cucullata*, and 49 *D. galeata × longispina* hybrids that were collected in several lakes covering the northern and southern European peri-alpine area (Möst 2013). We supplemented this reference dataset by producing 50 random hybrids of *D. cucullata* with each of the other parental species, using the R package ‘adegenet’. The real and simulated reference genotypes were used as a training set to derive a clustering procedure using Discriminant Analysis of Principal
Components (DAPC, Jombart 2008). The derived discriminant functions were then applied to the samples to obtain a posterior probability of membership in each of the 6 modeled clusters (3 parental species and 3 hybrids). Clustering was based on a >0.9 probability. Complex hybrid classes (e.g., a hybrid in which at least one parent is also a hybrid) were inferred when membership probability was evenly split between two of the modeled clusters.

**Sedimentary DNA isolation for cyanobacterial analyses**

For cyanobacterial high-throughput amplicon sequencing, we used sediments collected at 1-cm intervals down the Gorgova core. Because of the laminated structure (presence of annual varves) of the sediments in Greifensee and Hallwilersee, and based on the existing high-resolution age-depth models, we were able to sub-sample the cores at specific years with the goal to compare communities dated to similar time periods as the *Daphnia* populations. In both cores, we chose samples covering the periods of pre-, mid- and post-eutrophication. While sub-sampling the cores, we carefully avoided the sediments in contact with the plastic layer to reduce the risks of contamination with foreign DNA. Sediment samples were immediately transferred into sterile tubes, and carried to the Eawag clean lab facility for DNA extraction and further analyses. We performed the DNA extractions in batches of seven samples, plus a negative control (which contained all the reagents, but did not contain sediments). We followed strict environmental DNA work protocols at all steps to reduce the risks of contamination with external DNA. For Lake Gorgova, DNA extracts collected at each 1-cm layer were pooled in a total of eleven samples, each comprising 2 to 4 layers, to obtain a similar temporal resolution as used in the *Daphnia* dataset.

The forward and reverse PCR primers (Monchamp et al. 2016) specific to cyanobacteria were tagged with a 8 to 12 random nucleotides (nts) barcode, and used for amplifying a ~400-nt-long section of the cyanobacterial 16S rDNA gene. The PCR reactions were performed in a final volume of 40 µl containing ~12 ng template DNA (Monchamp et al. in prep). The PCR products were purified with the Illustra GFX PCR DNA and gel band purification kit (GE Healthcare, Chicago, USA), quantified using a Qubit fluorometer (Thermo Fisher Scientific) and pooled at equimolar concentration. The final library was sent to Fasteris (Geneva, Switzerland) for paired-end 2 × 250 bp sequencing on a MiSeq platform (Illumina Inc.).

**Sequence data processing and taxonomic determination of cyanobacterial operational taxonomic units (OTUs)**
The workflow used for data quality control and processing is fully detailed in Monchamp et al. (2016). After primer trimming, quality filtering, size selection, and removal of samples containing less than 4,000 reads, a total of 2,177,896 sequences were obtained distributed in 26 samples (9 from each Lake Gorgova and Hallwilersee, 8 from Greifensee) and 2 negative controls. The negative controls each contained less than 0.01% of the average number of reads in the samples; therefore they were excluded. Cleaned amplicons were clustered into OTUs at the 0.03 similarity level and an abundance size threshold of 5. The taxonomic assignment was done following UTAX (http://www.drive5.com/usearch/manual/utax_user_train.html). Only OTUs confidently assigned to photosynthetic cyanobacteria (based upon a confidence threshold of 0.85) were retained for analysis. The final dataset comprised 942’812 sequence reads recovered from sediments dating back to years between 1928 and 2015.

**Detection of potentially microcystin-producing cyanobacteria**

For the detection of potentially microcystin-producing cyanobacteria in the sediments of Lake Gorgova and Hallwilersee, we followed the method described in (Monchamp et al. 2016) previously applied to the Greifensee sediment core. Briefly, a ~400-nt-long fragment of the *mcyA* gene specific to the genera *Anabaena*, *Microcystis*, and *Planktothrix* (Hisbergues et al. 2003) was amplified and the product was visualized on an agarose gel. The products were purified and directly Sanger-sequenced (Microsynth, Balgach, Switzerland). Each chromatogram was visually inspected, and sequences were aligned in Muscle (Edgar 2004). Unique sequences were used in a BLAST search against the GenBank genetic sequence database to determine the closest relative. We complemented our dataset with the published results from a previous study on the Greifensee sediments (Monchamp et al. 2016).

**Genetic and phylogenetic diversity analyses of Daphnia and cyanobacteria**

The *Daphnia* genotypic and the cyanobacterial OTU data were analysed using the software R (R Core Team 2013) with the help of various packages. *Daphnia* genetic differentiation between subpopulations separated in time was quantified using Jost’s D (Jost 2008) as implemented in the package ‘mmod’ (Winter 2012) using our microsatellite markers. These statistics were used to create a genetic distance matrix for each lake, which was used in a non-metric multidimensional scaling (NMDS) analysis in the package ‘MASS’ to visualize the genetic divergence between *Daphnia* populations from different time periods in individual lakes. We chose to classify the OTUs into 5 orders for visualizing the composition of cyanobacteria in our lakes, as the taxonomic resolution was low for some OTUs. For all
cyanobacterial analyses, we used the OTU counts rarefied at the level of the sample with the lowest number of sequencing reads (4,338). The principal component analysis (PCA) is based on the proportion of *Daphnia* taxa and the proportion of cyanobacterial orders recovered in each sample using standardized community values. Beta diversity between cyanobacterial communities of individual lakes over time was estimated using the package ‘phyloseq’ in Bioconductor (McMurdie and Holmes 2013). Similarly to the *Daphnia* analyses, two distance matrices between cyanobacterial OTUs based on phylogenetic relatedness (Unifrac distance; Lozupone and Knight 2005) and on taxonomic dissimilarities (Jaccard distance; Faith et al. 1987) were created in the package ‘phyloseq’. The NMDS was applied to the phylogenetic distance matrix to visualize temporal beta diversity (i.e., within-lake dissimilarities over time).

**Results**

**Sediment age-depth models**

The $^{137}$Cs activity was generally low in Lake Gorgova, which makes difficult to accurately estimate years for this sediment core (Online Resource 2 in the Supporting Information). As a consequence, the dates used in the subsequent analyses and figures for Lake Gorgova should be interpreted with caution. $^{137}$Cs activity was detected in the lower section of the Gorgova core (34 cm), indicating that these sediments date back to the mid-1950s at the earliest, when $^{137}$Cs from atmospheric deposition was first detectable in the northern hemisphere. The unsupported $^{210}$Pb activity and the $^{137}$Cs profile also indicate the presence of some bioturbation in that lake. We identified two peaks attributed to the fallout following the Chernobyl accident in 1986 (11.5 cm), and to the maximum radionuclide fallout caused by the atmospheric nuclear weapon tests in 1963 (21.5 cm). A third peak of $^{137}$Cs was measured at 7.5 cm depth in the Gorgova core. Based on the shape of the peak and after comparison with the lead profile, we concluded that the latter caesium peak probably originated from terrestrial input from the catchment rather than from atmospheric deposition. In Greifensee (Monchamp et al. 2016) and Hallwilersee (Monchamp et al. unpublished data), the two well-resolved $^{137}$Cs peaks of 1986 and 1963 (Online Resource 2 in the Supporting Information) correlate well with the model based on annual varve counts. Only minor stratigraphic disturbances were visible within the Gorgova core section, but a high density of mollusc shells fragments were however found in the sediments between 19 and 34 cm-depth. Most layers from 18 cm-depth to the surface contained low densities of shell
fragments compared to the older sediments (Online Resource 3 in the Supporting Information).

The sedimentation rates were calculated based on $^{137}$Cs in all lakes (Table 1) and validated with the $^{210}$Pb measures. The highest sedimentation rate was measured in Lake Gorgova (0.43 cm/yr) between 1963 and 1986. The sedimentation rate in Greifensee was relatively stable throughout the period of eutrophication and re-oligotrophication (between 0.30 – 0.34 cm/yr), while in Hallwilersee, the sedimentation rate was higher in the recent sediment layers (yr 1986-2014).

<table>
<thead>
<tr>
<th>Lake</th>
<th>Period</th>
<th>Sedimentation rate (cm/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorgova</td>
<td>1986-2015</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>1963-1986</td>
<td>0.43</td>
</tr>
<tr>
<td>Greifensee</td>
<td>1986-2013</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>1963-1986</td>
<td>0.34</td>
</tr>
<tr>
<td>Hallwilersee</td>
<td>1986-2014</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>1963-1986</td>
<td>0.20</td>
</tr>
</tbody>
</table>

**Table 1.** Sedimentation rates calculated based on $^{137}$Cs in the three lakes.

**Compositional change in Daphnia and cyanobacterial assemblages**

The analysis of the time-series reconstructed over the past ~60 years from sedimentary archives revealed changes in the genetic diversity of *Daphnia* (egg bank) and the taxonomic composition of cyanobacteria (sedimentary DNA) in the three lakes studied. The abundance and composition of *Daphnia* genotypes varied over time in all the lakes. Overall, we identified 8 *Daphnia* taxa (species and hybrid classes) in Lake Gorgova, whereas only 5 were found in each of the pre-alpine lakes. Only few eggs were found in the Gorgova sediment layers dated to the 1950s and the early 1960s ($n<4$ per 1-cm layer). The majority were assigned to *D. cucullata*, *D. galeata*, and their hybrids, and a couple of *D. galeata × cucullata*. *D. longispina × cucullata* hybrids were found in the oldest sample (Fig. 1). Between 9 and 26 eggs were found in each of the other sediment layers of Lake Gorgova. *D. galeata* and their hybrid descendants were generally predominant in the upper half of the core (from ~1980s up to 2014; Fig. 1). The number of eggs retrieved from each sediment layer (each spanning one year) was generally higher in Greifensee (data not available for Hallwilersee), and the
number of successfully genotyped eggs ranged from 17 to 47 in the latter two lakes. Eggs assigned to pure *D. longispina* were only found in samples dated back to the 1950s and the 1960s in both pre-alpine lakes (Fig. 1). The *D. galeata × longispina* hybrid was generally the predominant genotype in the sediments of Hallwilersee and Greifensee, except during the 1970s when *D. galeata* represented the highest proportion of the egg bank in both lakes (Fig. 1).

![Figure 1. Proportion of *Daphnia* genotypes in populations reconstructed from ephippial eggs over 7 decades from lake sediment cores. The gradient of green reflects the timing and intensity of the main period of human-induced eutrophication in each lake. *Daphnia* taxa legend key: **cuc**, *D. cucullata*; **gal**, *D. galeata*; **gal.gal × lon**, *D. galeata × longispina*; **gal × cuc**, *D. galeata × cucullata*; **gal ×cuc .lon × cuc**, *D. galeata × cucullata . D. longispina × cucullata*; **gal × lon**, *D. galeata × longispina*; **lon × cuc**, *D. longispina × cucullata*; **lon**, *D. longispina*; **unidentified**, unidentified.](image)
Figure 2. Proportion of rarefied sequence reads assigned to each cyanobacterial order in all samples at the corresponding approximate year reflecting changes in community composition over time in the three lakes. The gradient of green identifies the main period of human-induced eutrophication between the 1950s and the 1980s.
Similarly to the *Daphnia* assemblages, the cyanobacterial community composition varied over time in the three lakes. The proportion represented by each cyanobacterial order (based on rarefied 16S rDNA sequence reads) was similar in the two pre-alpine lakes, whereas the reconstructed communities in Gorgova were different from the pre-alpine lakes (Fig. 2). Filamentous and buoyant taxa in the order Nostocales dominated the community in all samples from Lake Gorgova, whereas small single-celled taxa in the order Synechococcales were predominant in all samples from Greifensee and Hallwilersee. The proportion occupied by OTUs assigned to Nostocales increased in the 1960s and again in the 1990s in both pre-alpine lakes, especially in Greifensee (Fig. 2). In Gorgova, we observe an increase in the proportion of Chroococcales, which only represented a small percentage of the OTUs until the mid-1970s, and accounted for up to 25% of the sequence reads after the 1980s. At the same time in that lake, we observed a decline in the proportion of Nostocales, which were partly replaced by Synechococcales. Because the Synechococcales sequences were highly dominant in all samples in the two pre-alpine, we were unable to visualize the temporal variation in the other cyanobacterial groups, which contain several toxic and bloom-forming taxa. For better visualization of the temporal changes in the potentially harmful cyanobacterial groups, we plotted the proportions of sequence reads within each order, but this time excluding reads assigned to Synechococcales (see Online Resource 4 in the Supporting Information). This plot reveals that the increased in the proportion of OTUs assigned to Nostocales was followed by a decrease in Hallwilersee after the 1980s, and that the Nostocales OTUs were partly replaced by sequences assigned to taxa in the order Oscillatoriales (such as *Planktothrix* sp. and *Phormidium* sp.). In Greifensee, the proportion of Nostocales OTUs were the second-most dominant after the Synechococcales, with the exception of one time point in ~1950 when Chroococcales were slightly more important (Online Resource 4 in the Supporting Information).

The PCA correlation biplot showed distinct patterns in the distribution of the five cyanobacterial orders and the *Daphnia* taxa between the Danube Delta lake and the two pre-alpine lakes, with sites well separated in the ordination space (Fig. 3). The first axis (PC1) was mainly described by the presence of various *D. galeata* hybrid classes and a high proportion of the cyanobacterial orders Nostocales, Synechococcales, and Chroococcales, and explained 41.1% of total variance. The second axis (PC2, 15.4% of variance explained) was mainly related to Oscillatoriales and Pseudanabaenales, pure *D. galeata* as well as unidentified *Daphnia* taxa. There was a general trend towards a shift in community composition over time along the vertical axis (PC2) in Lake Gorgova. In Greifensee, the community shifted from a more *D. longispina*-dominated community to a *D. galeata x D.*
longispina-dominated community. All samples from the pre-alpine lakes were characterized by the presence of Synechococcales (Fig. 3).

**Figure 3.** Principal component analysis (PCA) representation based on the proportion of the five main cyanobacterial orders (bloom-forming groups in bold) and the *Daphnia* community composition in the three lakes over the past 6-7 decades. Ellipses show 85% confidence intervals. The cyanobacterial 16S rDNA sequence data and the *Daphnia* genotypic composition analysed at various depths in the sediment cores were grouped and approximate years are used as label coloured by lake. *Daphnia* taxa legend key: cuc, *D. cucullata*; gal, *D. galeata*; gal.gal × lon, *D. galeata × longispina*; gal × cuc, *D. galeata × cucullata*; gal × cuc.lon × cuc, *D. galeata × cucullata*; gal × lon, *D. galeata × longispina*; lon × cuc, *D. longispina × cucullata*; lon, *D. longispina*; unidentified, unidentified *Daphnia* taxa.
Temporal evolution of the genetic divergence in *Daphnia* and cyanobacterial communities

The historical trends in genetic differentiation among *Daphnia* populations (Jost’s *D*) and genetic dissimilarity among cyanobacterial communities (Unifrac) in individual lakes are reported in Fig. 4. The stress values for all NMDS configurations based on Jost’s *D* are lower than 0.03, and are comprised between 0.069 and 0.12 for the NMDS analyses based on Unifrac distances.

In Lake Gorgova, the genetic differentiation between consecutive *Daphnia* populations was greater in the older samples (Fig. 4a). The position of the samples in the plot indicates that turnover was lower between consecutive samples since the early 1990s. The temporal patterns in genetic distance based on allelic frequency across *Daphnia* populations reconstructed from the egg banks were similar in the two pre-alpine lakes over the past fifty years. The cyanobacterial phylogenetic diversity in the older samples (before the onset of the main eutrophication event) in Lake Gorgova was the most divergent from all the other communities in that lake (Fig. 4b). In Greifensee, the turnover was lower between consecutive communities of 1961, 1976 and 1981, which correspond to the period of maximal eutrophication in that lake. During and after re-oligotrophication of the lake (in the 1990s and 2000s) the *Daphnia* genetic diversity did not return to the original pre-eutrophication structure. The cyanobacterial community from 2010, however, was more similar to that of the pre-eutrophication period (1943 and 1950) compared to the ones corresponding to the mid-eutrophication period. In Hallwilersee, the cyanobacterial communities were characterized by a relatively high divergence between 1928 and 1980 and a considerably lower turnover post-eutrophication (1994 to 2014). These patterns in temporal phylogenetic dissimilarity among communities were also reflected in the individual-lake NMDS representations based on Jaccard distances (i.e., taxonomic diversity based on incidence of OTUs; see Online Resource 5 in the Supporting Information).

Detection of potentially microcystin-producing cyanobacteria

The *mcyA* gene was detected by PCR in sedimentary DNA samples collected in Gorgova and Greifensee, confirming the presence of potentially toxic cyanobacteria at these sites. In Lake Gorgova, the gene was amplified in twenty-eight out of thirty sediment layers (Fig. 5). Sanger sequencing revealed that all amplicons correspond to a single *mcyA* sequence assimilated to *Microcystis* sp. In Greifensee, the *mcyA* gene was found in only four of the thirteen layers tested spanning the last eight decades. The gene was only found in the top
section of the core corresponding to the period between 2000 and 2013. Surprisingly, the \textit{mcyA} sequences in Lake Gorgova and Greifensee were all identical, and were identified as a \textit{Microcystis} sp. (GenBank accession number KX437769.1; This analysis on the Greifensee sediments was done as part of a previous study, see Monchamp et al. 2016). In Hallwilersee, the gene was not detected in any of the nine samples tested spanning between the 1930s and the 2010s. Overall in lakes where the \textit{mcyA} gene was detected, we found no temporal genetic variation in the composition of the potentially microcystin-producing cyanobacteria.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Nonmetric multidimensional scaling (NMDS) representations based on (\textbf{a}) \textit{Daphnia} genetic distances based on allele frequency, and (\textbf{b}) cyanobacterial phylogenetic dissimilarity (un-weighted Unifrac distances based on incidence of OTUs). The full line between consecutive samples shows the temporal trend in each \textit{Daphnia} or cyanobacterial community, and samples are labelled using approximate years. The samples corresponding to the period of maximal eutrophication in each lake are highlighted in green.}
\end{figure}
Figure 5. Results of PCR amplifications for the detection of mcyA genes in the sediment cores from the three lakes. Each box represents a sample corresponding to an approximate year over the past nine decades. Samples were grouped in decades to simplify visual representation. Samples where mcyA genes were detected are marked in red, whereas a green square means the gene was not detected. The mcyA genes data from the Lake Greifensee sediment core originate from a previous publication (Monchamp et al. 2016; see Chapter II).

Discussion

Here we show that changes in the genetic diversity of Daphnia and the phylogenetic diversity of cyanobacteria, as reconstructed from the sediments of three natural communities over the past 6 to 8 decades, occurred at comparable temporal scales and time periods. Daphnia and cyanobacterial community composition appeared to respond to human-mediated eutrophication. Interestingly, the original composition and genetic structure of the two plankton groups were not restored after re-oligotrophication of the lakes.

Generally, the taxonomic diversity of Daphnia eggs was higher in Lake Gorgova compared to the two pre-alpine lakes. This is mainly because D. cucullata is present in many sediment layers and hybridized with both D. galeata and D. longispina. D. cucullata is a species typical found in eutrophic lakes (Benzie 2005; Flößner 2000). We found that pure D. cucullata and D. galeata were predominant in the eutrophic Lake Gorgova, where filamentous cyanobacteria (affiliated mainly to the order Nostocales) were always found in high proportion. These findings are consistent with previous studies having reported that D. galeata thrives in meso- to eutrophic systems, and that D. cucullata has a competitive advantage in eutrophic conditions (Benzie 2005; Flößner 2000) and in the presence of filamentous cyanobacteria (DeMott et al. 2001; Gliwicz and Lampert 1990).
In a pilot study done in 2002, Cremer et al. (2004) reconstructed the diatom assemblages from sediment cores collected in Lake Gorgova and from four other lakes located in the Danube Delta. Based on the time-series of diatom composition, it was determined that the lakes were meso- to eutrophic throughout the period covered by the study (~1920s to 2002), and that most likely, the onset of eutrophication took place in much earlier times in lakes of the Delta. This is not surprising, as the Delta is the end point of the Danube River which has a vast drainage basin and flows through ten countries. Our findings both on the *Daphnia* and cyanobacterial taxonomic composition in the sediments of Lake Gorgova are consistent with the conclusion of Cremer and colleagues. A few signs in the Gorgova sediment core suggest a drastic change in the ecology of the lake around the 1970s, such as a change in sediment structure and the decline in density of mollusc remains between 18 and 19 cm-depth. These changes most likely reflect rapid modifications of lake conditions (e.g., hypoxia, reduction in the macrophyte coverage) in the late 1960s and early 1970s, period which corresponds to the onset of human-mediated eutrophication in the Danube Delta (Rîșnoveanu et al., 2004). At the same time, we observed an increase in the proportion of OTUs assigned to Chroococcales, a group comprising of large colonial taxa that are generally favoured under high nutrient conditions (Sukenik et al. 2015). Interestingly, after the peak of eutrophication in Lake Gorgova (according to the literature on the Danube Delta lakes), there was a slight increase in the proportion of Synechococcales, which could also be an indicator of change in local lake conditions (e.g., nutrient availability). The relative contribution of picocyanobacteria to the total phytoplankton community has been shown to increase with re-oligotrophication (Agawin et al. 2000) due to their low phosphorus requirements. In large and deep water bodies such as pre-alpine lakes, picocyanobacteria are known to account for the majority of the picophytoplankton (reviewed in Callieri 2007).

Our reconstructed time-series of *Daphnia* populations reflect the species composition reported in previous studies over the past few decades based on pelagic surveys of several lakes located north and south of the European Alps (Keller et al. 2008), and on the egg banks of Greifensee (Brede et al. 2009) and Hallwilersee (Turko et al. 2016). To our knowledge, this constitutes a first report of the *Daphnia* and cyanobacteria genetic diversity in Lake Gorgova. The invasion of *D. galeata* and the dominance of *D. galeata × longispina* hybrids in the pre-alpine lakes Greifensee and Constance has previously been linked to human-induced eutrophication (Brede et al. 2009). Our results suggest that the *Daphnia* community in Hallwilersee was similarly affected by local environmental changes leading to the replacement of *D. longispina* by *D. galeata* and their hybrids.
The composition of the *Daphnia* resting egg bank in Gorgova was different from that of the two pre-alpine lakes, but interestingly, *D. galeata* and their hybrids were also generally dominant in that lake. We did not find eggs belonging to *D. longispina* in Gorgova, but hybrids and backcrosses (mostly *D. galeata × cucullata*, *D. longispina × cucullata* and *D. galeata × longispina*) signalled the historical presence of *D. longispina* in the lake. Sampling older sediment layers could reveal an invasion and the subsequent replacement of native taxa by *D. galeata* related to changes in environmental conditions.

The cyanobacterial diversity in the two pre-alpine lakes is known from long-term monitoring data (Bürgi et al. 2003; Bürgi and Jolidon 1998) and from sedimentary DNA reconstructions (Monchamp et al. 2016; Monchamp et al., in prep). The diversity reconstruction over the past few decades in Greifensee confirms the presence of bloom-forming and potentially toxic taxa in the genera *Microcystis* (order Chroococcales) and *Dolichospermum* (*Anabaena*; order Nostocales). In Hallwilersee, we report the presence of Nostocales, as well as OTUs related to Oscillatoriales, including the potentially harmful cyanobacterium *Planktothrix rubescens*. This particular taxon is known for regularly forming metalimnetic blooms in Hallwilersee (Liechtl 1994) and other pre-alpine lakes north (Posch et al. 2012) and south (D’Alelio et al. 2011) of the Alps. To our knowledge, there are no published reports on the composition of the cyanobacterial assemblage in Lake Gorgova; therefore we are unable to compare our results with other studies. Our compositional data recovered from sedimentary DNA sequencing confirms historical records in Greifensee and Hallwilersee, but as it has been shown before (Monchamp et al. 2016), the number of reads cannot directly be matched to microscopy counts due to biases related to both methods.

Although it is difficult to establish a direct link between the time-series of *Daphnia* and cyanobacteria reconstructed from natural archives, we can speculate about the temporal trends observed in the composition and diversity of the two plankton assemblages. Interestingly, concomitantly to the major turnover in the *Daphnia* communities in Greifensee and Hallwilersee, we observed an increase in the proportion of filamentous and bloom-forming cyanobacteria assimilated to order Nostocales in both lakes, and related to Oscillatoriales over the past three decades in Hallwilersee. Filamentous cyanobacteria were generally predominant in the sediments of Lake Gorgova over the timespan reflected by our sediment core. There was, however, an increase in the proportion of other colony-forming cyanobacteria (in the order Chroococcales) over the recent few decades. The *Daphnia* populations of yr 2003-2004 in Greifensee and yr 2000 in Hallwilersee were similar in composition and did not return to the pre-eutrophication population structure, probably as a
consequence to the competitive advantage of *D. galeata* in a changing environment due to its broad ecological niche (Weider 1993). *D. galeata* is also known for investing more in sexual reproduction compared to other species, and to be able to hybridize with almost all other species in the *D. longispina* complex (Taylor et al. 2005).

The fact that we did not detect the *mcyA* gene in any samples from Hallwilersee was surprising because the lake’s phytoplankton assemblage is known to often comprise a relatively high biomass of *Planktothrix rubescens* (Liechti 1994; Züllig 1982), a taxon known to include microcystin-producing strains (Ostermaier and Kurmayer 2010). Our results do not allow us to rule out the possibility that toxic *P. rubescens* genotypes were present in Hallwilersee. Although we were able to detect *P. rubescens* in general, toxic strains may have made up a small part of the population and therefore remained below the detection limit in the sediment layers sampled. As shown in a previous study (Monchamp et al. 2016), the number of sequences derived from sedimentary archives does not always correlate with the abundance of cells in water estimated based on microscopic surveys, probably due to the limitations associated with both methods. An unlikely explanation for the low recovery of *Planktothrix rubescens* in the sediments could be related to technical aspects of the method, for example low specificity of the PCR primers used or DNA degradation with sediment age. It has however been shown in a previous study using the same approach that a broad range of representatives of the phylum Cyanobacteria, including *Planktothrix* species, could be recovered from the sediments of pre-alpine lakes over centuries (Monchamp et al. 2016, Savichtcheva et al. 2011, 2015). In this study, we confirm the presence of an OTU confidently assigned to *P. rubescens* in Hallwilersee in several sediment samples since the early 1880s, as well as in Lake Gorgova since the 1960s.

**Methodological approach and limitations**

One of the main concerns in palaeogenetic studies is related to DNA degradation. The fate of DNA in the sediments depends on various factors, such as temperature, pressure, oxygen concentration, microbial nuclease activity, and adsorption to particles (Boere et al. 2011; Coolen and Overmann 1998; Panieri et al. 2010). The cold, dark, and hypoxic to anoxic conditions prevailing at the bottom of the two pre-alpine lakes are favourable for DNA preservation. In the shallow lake Gorgova, the conditions are less ideal (warmer temperatures, mixing). Nevertheless, we were able to recover good quality DNA in high concentration, and amplify 400 nt-long fragments of both the 16S rRNA and the *mcyA* genes over the relatively recent sediments studied (less than a century old). The sequencing depth
and the number of cyanobacterial OTUs in the samples isolated from all lakes were comparable, and we did not detect any patterns that could be attributed to diagenetic processes (e.g. the loss of OTUs with sediment age). We carefully inspected the OTU table and found no evidence for preferential DNA preservation across the cyanobacterial groups.

Many papers have been published that reconstruct past *Daphnia* populations by genotyping resting eggs from the sediment (Brede et al. 2009; Orsini et al. 2013). Due to the reproductive mode of *Daphnia*, cyclic parthenogenesis, resting eggs are the outcome of sexual reproduction, thus they reflect the genes present in the population at the moment of production. *Daphnia* clones that do not reproduce sexually will not be found in the sedimentary archive, but because the final persistence of a *Daphnia* population is depending on the resting eggs (recovery from catastrophes like strong winters, high temperature or food scarcity), the resting egg bank is a good ecological and evolutionary representation of past populations.

We were able to date the sediment with a yearly resolution in Greifensee and Hallwilersee due to the presence of varves and the lack of mixing in the surface sediments. Lake Gorgova sediments represented a more challenging archive to date, because of multiple mixing events every year, and the shallowness of this lake. Nevertheless, based on $^{137}$Cs and $^{210}$Pb activity, we were able to construct an approximate, but robust age model. Our sedimentation rate calculations in the sediments of Lake Gorgova (see Table 1) based on the lead and caesium profiles were in good agreement with previous observations from Cremer et al. (2004), who measured a sedimentation rate of 0.47 cm/year between 1986 and 2002 and of 0.37cm/year between 1963 and 1986 in a core collected in the same lake. We suspected that the upper section of our core, which consisted of relatively aerated and loose sediments, was compressed during transportation of the core to our lab in Switzerland, which could have contributed to underestimate recent sedimentation rates. Nevertheless, the detection of $^7$Be in the uppermost sediment layer confirmed that the surface sediments were present and we believe that our dating of Lake Gorgova sediments is sufficiently accurate for a coarse reconstruction of past temporal patterns in this lake.
Conclusion

In this paper, we report about the diversity of the grazer *Daphnia* and important primary producers, cyanobacteria, as reconstructed from sediments cores collected in three lakes impacted by human-induced eutrophication over the past century. Our results reveal co-occurring changes in *Daphnia* and cyanobacterial composition and genetic diversity, most likely as a consequence of human-induced eutrophication. Although it is not possible in this study to infer a causal link between the cyanobacteria and the *Daphnia* communities, our diversity analyses reveal that *D. galeata* hybrids and filamentous and potentially bloom-forming cyanobacteria appear to have been favoured over the past 6-8 decades in the three eutrophied lakes studied, and importantly, the *Daphnia* and cyanobacteria assemblages have not returned to the pre-eutrophication composition after the re-oligotrophication phase in the two pre-alpine lakes. Long-term retrospective studies like ours have the potential to help identify the factors governing community assembly of freshwater plankton, and to inform us on how different compartments of the food-web are linked and co-vary in a changing environment.

Acknowledgements

High-throughput sequencing was performed at Fasteris (Geneva, Switzerland). We thank Jean-Claude Walser (Genetic Diversity Centre, ETH Zürich) for bioinformatics support, and Marco Thali (Eawag) for his help with the sequencing library preparation. We very much appreciate the help of Adrian Cacencu, who was our guide in the Danube Delta, and who helped with Lake Gorgova sampling. We also thank Alois Zwyssig (Eawag) for his help with sampling, Nathalie Dubois and Alfred Lück (Eawag) for sampling and sediment dating, Marcin Dziuba, Esther Keller, and Aglaia Pârvu for their help in the lab. Finally, we thank the two anonymous reviewers for their feedback which greatly improved the manuscript.

This work was supported by the Swiss Enlargement Contribution; project IZERZ0 – 142165, “CyanoArchive” to Piet Spaak, in the framework of the Romanian-Swiss Research Programme.

This work was supported by the Swiss Enlargement Contribution; project IZERZ0 – 142165, “CyanoArchive” to Piet Spaak, in the framework of the Romanian-Swiss Research Programme.
Sedimentary and egg-bank DNA from 3 European lakes reveal concurrent changes in the composition and diversity of cyanobacterial and *Daphnia* communities
Online Resource 1. Summary of the lakes physical, morphological, and chemical characteristics.

<table>
<thead>
<tr>
<th>Lake</th>
<th>Gorgova</th>
<th>Greifensee</th>
<th>Hallwilersee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation (m.a.s.l.)</td>
<td>1</td>
<td>435</td>
<td>449</td>
</tr>
<tr>
<td>Maximal depth (m)</td>
<td>~3</td>
<td>33</td>
<td>46</td>
</tr>
<tr>
<td>Lake surface area (km²)</td>
<td>13.8</td>
<td>8.5</td>
<td>10</td>
</tr>
<tr>
<td>TP (µg/L) *</td>
<td>26</td>
<td>40</td>
<td>18</td>
</tr>
<tr>
<td>TP Max (µg/L) #</td>
<td>NA</td>
<td>492.43</td>
<td>213.81</td>
</tr>
<tr>
<td>Presence of submerged macrophytes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Water mixing</td>
<td>polymictic</td>
<td>monomictic</td>
<td>Monomictic or incomplete mixing (Hypolimnion mechanically aerated since 1986)</td>
</tr>
</tbody>
</table>

* Refers to the most recent available phosphorus concentrations measured over the water column
# Refers to the maximum annual mean phosphorus concentration recorded over the water column
Online Resource 2. Profile of lead-210 and Caesium-137 activity (per gram of dry sediments) in the sediments of lakes Gorgova, Greifensee, and Hallwilersee. The Cs-137 fallout after the Chernobyl accident of 1986, and the maximum fallout after the nuclear tests in 1963 are identified. The Greifensee plot is based on radionuclide data from Monchamp et al. 2016, and the Hallwilersee plot is based on data from Monchamp et al. (in preparation, see Chapter III).
Online Resource 3. Profile of the mollusc shell fragments counted per 1-cm layer in the Gorgova sediment core.
Online Resource 4. Proportion of rarefied sequence reads assigned to each cyanobacterial order excluding order Synechococcales in all samples at the corresponding approximate year reflecting changes in community composition over time in the three lakes. The gradient of green identifies reflects the intensity of the main period of human-induced eutrophication in each lake.
Online Resource 5. Nonmetric multidimensional scaling (NMDS) representations based on taxonomic dissimilarity (Jaccard distances, based on OTU incidence) across cyanobacterial communities within each lake over time. The line connects consecutive sample to show the temporal trajectory of the cyanobacterial communities composition.
Literature cited Chapter IV


Frisch, D., P. K. Morton, P. R. Chowdhury, B. W. Culver, J. K. Colbourne, L. J. Weider & P.


CHAPTER V

The dark side of lakes: diversity and distribution of nonphotosynthetic cyanobacteria

Marie-Eve Monchamp\textsuperscript{1,2}, Piet Spaak\textsuperscript{1,2}, and Francesco Pomati\textsuperscript{1,2}

\textsuperscript{1} Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Dübendorf, Switzerland;

\textsuperscript{2} Swiss Federal Institute of Technology (ETH) Zürich, Institute of Integrative Biology, Zürich, Switzerland;

\textit{Manuscript in preparation for Frontiers in Microbiology}
Abstract

Freshwater cyanobacteria are the focus of many studies as they are infamous for their potential impacts on water quality, aquatic ecosystem functioning and services. A group of nonphotosynthetic bacteria has recently been re-classified as part of the Cyanobacteria phylum based on whole-genome sequencing. These ancestral cyanobacteria have been mostly found in aphotic environments, in sediments, water, and animal guts. The ecological role of these ancient clades of cyanobacteria and their community structure across aquatic ecosystems or over time are currently unknown. In this study, we report for the first time about the long-term diversity and distribution of two clades of nonphotosynthetic cyanobacteria recovered by sequencing a fragment of the 16S rDNA gene form sediment cores of ten lakes in the European peri-Alpine region. We describe and discuss their composition and phylogenetic diversity over the past ~100 years in parallel with patterns observed in the phylogenetic diversity of photosynthetic cyanobacteria recovered from the same samples. Unlike photosynthetic cyanobacteria, the richness of environmentally associated nonphotosynthetic cyanobacteria did not significantly change in lake sediments over the last century, and the across-lake community similarity in nonphotosynthetic clades could not be explained by temporal succession. Interestingly, the distance-decay of pairwise community phylogenetic similarity (based on Unifrac) as a function of geographic distance was not significant in both the photosynthetic and nonphotosynthetic groups, suggesting no limit of dispersal at the regional scale. This study is one of the earliest reporting the diversity of nonphotosynthetic cyanobacteria in the environment and contributes to a better understanding of their niche and ecological function in nature.
Introduction

Cyanobacteria constitute a highly diverse group of gram-negative prokaryotes that colonize a wide range of environments, from desert crusts to fresh and marine waters, and from the tropics to the poles (Whitton & Potts, 2002). They have played a crucial role in modifying the Earth’s atmosphere through the process of oxygenic photosynthesis, which enabled the creation of life in more complex forms (Fischer et al., 2016a). Cyanobacteria have been studied for decades, and their diversity has been described both morphologically (Rippka et al., 1979; Komárek et al., 2014) and genetically (Shih et al., 2013). Aquatic cyanobacteria are often considered a nuisance in freshwaters and marine coastal ecosystems as they can form large blooms at the surface of water replete with nutrients (mainly nitrogen and phosphorus), impairing water quality and disrupting food-webs. Because of their negative consequences on water and general ecosystem quality, they are often used as indicators of pollution caused by human activities.

With the recent development of molecular techniques allowing the investigation of non-cultivable organisms directly in nature, scientists have unveiled an unexpected diversity of cyanobacteria – as well as other groups of microorganisms – in many ecosystems e.g., (Sogin et al., 2006; Di Rienzi et al., 2013; Soo et al., 2014; Meola et al., 2015). Genome sequencing recently revealed important information about a clade of nonphotosynthetic prokaryotes related to cyanobacteria. The name Melainabacteria was proposed (Di Rienzi et al., 2013) because several representatives of this group have been found in aphotic environments. They were first thought to constitute a sister-phylum of cyanobacteria (Di Rienzi et al., 2013), but more recent genomic information confirmed the position of Melainabacteria as sister-clade of photosynthetic cyanobacteria part of the same phylum (Soo et al., 2014). On the basis of this new genomic evidence, a re-classification was proposed for phylum Cyanobacteria, with the class-level lineages Oxyphotobacteria (all cyanobacteria capable of photosynthesis) and Melainabacteria, as well as a third possible class called ML635J-21 (Soo et al., 2014). The ML635J-21 is the most basal lineage and forms a paraphyletic group which is the ancestor of both Melainabacteria and Oxyphotobacteria (Fischer et al., 2016b). For ease of discussion, we hereafter refer to photosynthetic cyanobacteria as class Oxyphotobacteria, and we use the nomenclature proposed by Soo et al. (2014) for the class Melainabacteria, wherein the following orders are included: Gastranaerophilales (YS2), Obscuribacterales (mle1-12), Caenarcaniphilales (ACD20) and Vampirovibrionales (SM1D11).
Because there are only a hand-full of Melainabacteria genomes and no ML635J-21 sequenced genomes, the metabolism, functions, and ecological role of these ancestral organisms are not fully known. Many representatives of Melainabacteria and ML635J-21 (McDonald et al., 2012; Quast et al., 2013) have been found in aphotic environments, such as ground water (Di Rienzi et al., 2013), marine and lacustrine sediments (Ley et al., 2005), animal and human faeces (Soo et al., 2014) and guts (Ley et al., 2005), consistent with the assumption that they are nonphotosynthetic organisms. A list of nonphotosynthetic cyanobacteria representatives with their associated environments is presented in Table 1. After diverging from Melainabacteria, the Oxyphotobacteria developed oxygenic photosynthesis around 2.4–2.35 billion years ago based on molecular clock estimates (see Shih et al., 2016) and geological data (Fischer et al., 2016a).

**Table 1.** List of Melainabacteria and ML635J-21 representatives with the type of environment where they were found.

<table>
<thead>
<tr>
<th>Class</th>
<th>Order</th>
<th>Environment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melainabacteria</td>
<td>Gastranaerophilaes</td>
<td>Human faecal samples</td>
<td>Qin et al., 2010; Soo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Gastranaerophilaes</td>
<td>Koala faecal samples</td>
<td>Soo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Gastranaerophilaes</td>
<td>Soil sample</td>
<td>McGorum et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Obscuribacterales</td>
<td>Activated sludge from a batch (aerobic) reactor performing enhanced biological phosphorus removal (EBPR)</td>
<td>Soo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Obscuribacterales</td>
<td>Plant washing; Soil sample; Horse ileal content;</td>
<td>McGorum et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Obscuribacterales</td>
<td>Aquifer</td>
<td>Wrighton et al., 2012; Di Rienzi et al., 2013</td>
</tr>
<tr>
<td></td>
<td>Caenarcaniphilaes</td>
<td>Anaerobic reactor treating a high-strength organic wastewater</td>
<td>Soo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Vampirovibrionales</td>
<td><em>Chlorella vulgaris</em> culture</td>
<td>Soo et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Vampirovibrionales</td>
<td><em>Chlorella</em> in commercial ponds</td>
<td>Ganuza et al., 2016</td>
</tr>
<tr>
<td>ML635J-21</td>
<td>NA</td>
<td>Plant washing; Soil sample; Marine and lacustrine sediments</td>
<td>McGorum et al., 2015</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>Meromictic lake</td>
<td>Gies et al., 2014</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>Marine surface water samples (uncultured marine bacteria)</td>
<td>e.g., Venter et al., 2004a</td>
</tr>
</tbody>
</table>
All sequenced genomes of Melainabacteria to date confirm that they completely lack the photosynthesis apparatus, supporting the hypothesis that the acquisition of the photosystem in Oxyphotobacteria happened after divergence from the nonphotosynthetic Melainabacteria (Soo et al., 2014). All representatives of Melainabacteria (based on 11 sequenced genomes) are proposed to be able to perform fermentation and some are thought to be capable of anaerobic as well as aerobic respiration (Di Rienzi et al., 2013; Soo et al., 2014). Some Melainabacteria in the orders, Vampirovibrionales, and Obscuribacterales are also capable of aerobic respiration. This ability would have been acquired by lateral gene transfer after the evolution of oxygenic photosynthesis in the sister-clade Oxyphotobacteria (Shih et al., 2016). The Gastranaerophilales are found in human and other animal guts and although their exact role is unknown, they might have a beneficial effect for their hosts by aiding digestion and as a source of vitamins B and K (Di Rienzi et al., 2013). Some Gastranaerophilales are flagellated, but other representatives appear to have only a subset of genes necessary to encode a functional flagellum (Soo et al., 2014). The Gastranaerophilales and Caenarcaniphilales appear to lack the genes for anaerobic and aerobic respiration, suggesting that they only support their metabolism by fermentation (Soo et al., 2014).

Vampirovibrio chlorellavorus (previously assigned to genus Bdellovibrio) is a predatory bacteria that was originally described in 1972 and was first suggested to be a member of the Deltaproteobacteria (Gromov & Mamkaeva, 1972). It was recently isolated from a 36-old co-culture sample lyophilised cells of Chlorella vulgaris (NCBI 11383; Coder & Starr, 1978), and whole genome sequencing confirmed the position of Vampirovibrio in the class Melainabacteria (order Vampirovibrionales) part of phylum Cyanobacteria (Soo et al., 2015). Vampirovibrio chlorellavorus predated on Chlorella cells, and consistent with photosynthetic cyanobacteria as well as other class of Melainabacteria, it lacks the genes for carbon fixation and photosynthesis (Soo et al., 2014).

The aim of this study is twofold: First, to study the diversity of nonphotosynthetic cyanobacteria by investigating lake sedimentary archives, and second, to investigate changes in their composition and prevalence over time and to discuss patterns in parallel with previously observed diversity and community structure of photosynthetic cyanobacterial in the same lakes as a comparison (Monchamp et al. submitted). For this, we used a temporal dataset consisting of 16S rDNA sequences obtained in samples collected in sediment cores from ten European peri-Alpine lakes. Because we were only interested in the diversity of environmentally associated nonphotosynthetic cyanobacteria, we present the composition of all groups of nonphotosynthetic cyanobacteria, but exclude the gut-associated Gastranaerophilales order from the diversity analyses.
Materials and Methods

Data collection

We used the high-throughput MiSeq data from Monchamp et al. (submitted) (Chapter III of this thesis) consisting of 16S rDNA cyanobacterial sequences obtained from sequencing of 107 sedimentary DNA samples spanning over ~150 years of sedimentary archives from 10 lakes located around the Alps. In our analysis of cyanobacterial sequences, we found that our dataset comprised of a large proportion of 16S rRNA genes that were classified as part of the Cyanobacteria phylum, but did not belong to Oxyphotobacteria. Instead, they were highly related to the Melainabacteria and ML635J-21 groups (based on a confidence threshold of 0.89) in the Greengenes database (DeSantis et al., 2006). Here, we use this subset of sequence data for investigating the alpha and beta diversity of nonphotosynthetic cyanobacteria in the ten lakes. All details related to sediment sampling, sample processing, DNA extraction, and PCR amplification are described in detail in our previous work (Chapter II and III of this thesis). Paired-end sequencing (2 × 250 bp) was performed on an Illumina MiSeq at the sequencing company Fasteris (Geneva, Switzerland).

Sequence data processing

Briefly, the raw 16S rDNA sequences were quality controlled using the workflow developed at the Genetic Diversity Centre ETH Zürich (GDC; Monchamp et al., 2016). After the quality filtering, primer trimming and size-selection steps, the amplicons were clustered into operational taxonomic units (OTUs) following the UPARSE workflow (Edgar, 2013) based on an abundance threshold of 5, and a minimum sequence similarity of 97%. The taxonomic assignment of OTUs was done based on the Greengenes database (DeSantis et al., 2006) that we supplemented with a few decoy bacterial sequences (Monchamp et al., 2016) using PyNAST (McDonald et al., 2012). The confidence threshold was set at 0.85. The aligned sequences were imported in FastTree (Price et al., 2010) to infer a phylogeny based on maximum-likelihood.

The reference OTUs FASTA sequences, the tree file, and the taxonomic assignment file were imported in the software R (R Core Team, 2013) version 3.3.2, using the package ‘phyloseq’ in Bioconductor (McMurdie & Holmes, 2013). We then subsetted the communities to retain only the nonphotosynthetic assigned to class Melainabacteria (two
Vampirovibrionales, six Obscuribacterales, and twenty-two Gastranaerophilales), twenty-nine were assigned to the deepest-branching clade ML635J-21, and only four nonphotosynthetic cyanobacterial OTUs were not assigned to a class. OTUs that did not appear in a minimum of two samples over the whole dataset were excluded to reduce the biases associated to rare taxa and to possible sequencing errors. The subsetted dataset of 16S rDNA sequences was used for inferring a phylogeny based on maximum-likelihood with bootstrap analysis (100 replicates) using the alignment program MAFFT (Katoh & Standley, 2013). The tree was visualised and annotated with the online program iTOL (Letunic & Bork, 2016). Relevant physiological traits attributed to these clades based on genome sequencing (Di Rienzi et al., 2013; Soo et al., 2014) were added to the tree for references (Fig. 1). It should be noted that this list of traits is man extrapolation of the information based on genome sequencing of a small subset of representatives in the Melainabacteria, thus the definitive presence or absence of traits in the OTUs found in the present study cannot be confirmed or infirmed based on our study.

**Compositional and phylogenetic diversity analyses**

For the diversity analyses, we used all OTUs assigned to ML635J-21, as well as the Melainabacteria representatives that are generally associated with environmental samples (i.e., the orders Vampirovibrionales (SM1D11) and Obscuribacterales (mle1-12)). We excluded the order Gastranaerophilales that is associated with gut samples as it was considered irrelevant to the present study aiming at describing the diversity of nonphotosynthetic cyanobacterial lake communities. Our dataset did not include representatives of the Caenarcaniphilales order.

Using this reduced dataset, we reproduced some of the analyses that were previously applied to Oxyphotobacteria to assess their diversity and community composition within and across the 10 peri-Alpine lakes over ~150 years (Monchamp et al., submitted; Chapter III). To be able to estimate indices alpha and beta diversity that can be compared across samples, we rarefied the samples to even sequencing depth. In order to keep the maximum number of samples, we rarefied to 201 reads, which was enough to cover a high percentage of OTU richness in the majority of samples (Rarefaction curves are shown in Fig. S1 of the supplementary materials). We used linear functions to determine the significance of variation in the log-transformed rarefied OTU richness over time within the lakes. Hierarchical clustering of taxa for pattern detection was performed by calculating Euclidian distances on OTU prevalence in the lakes over time. The dendrogram, was constructed by average
linkage method and the colour-coded image map produced in CIMminer (Weinstein et al., 1997).

For calculating beta diversity between communities across all lakes and between time periods, the rarefied samples were grouped into 10-years blocks. A distance matrix based on phylogenetic dissimilarity (Unifrac distances) across all samples at each period was calculated. The ‘adonis’ function in the ‘vegan’ package (Oksanen et al., 2013) was used to apply PERMANOVA (Anderson, 2001) to verify temporal and spatial (lake and region) effects on the dissimilarity between groups. If a lake sediment core was sampled at two depths that were grouped in the same period of time (e.g., yr. 1992 and yr. 1997), the pairwise distance between the two samples was excluded from the dissimilarity estimation (to remove the internal turnover effect within lakes). We used the package ‘Imap’ (version 1.32; https://CRAN.R-project.org/package=Imap) for producing a matrix of pairwise geographic distances between lakes (Table S1 in the supplementary material). This distance matrix was used for assessing the distance-decay relationship for all community phylogenies in each decade over the twentieth century. For comparison with the Oxyphotobacteria communities, we calculated the distance-decay relationship in the same manner using the sequence data produced in Chapter III. The data was insufficient to estimate the distance-decay relationship between photosynthetic cyanobacterial communities in the 1900s, 1910s and 1020s, and between nonphotosynthetic cyanobacterial communities in the 1900s and the 2010s. Because the assumption of independence is violated when using multiple pairwise comparisons, the significance of the distance-decay curves at each time period was assessed by a Mantel (Mantel, 1967) permutational test (with 999 repetitions) between distance matrices using the package ‘ade4’ for R.
Results

Composition and phylogenetic diversity of nonphotosynthetic cyanobacteria

Our rarefied dataset of Melainabacteria and ML635J-21 comprised of sixty-three OTUs distributed in sixty-six samples. The maximum-likelihood tree of all OTUs 16S rDNA reference sequences recovered from the sediments of the ten peri-Alpine lakes is shown in Figure 1. The two lineages ML635J-21 and Melainabacteria are well supported by the bootstrap values, and the tree topology reflects previously published phylogenies inferred from 16S rDNA sequences. A list of traits associated to some nonphotosynthetic cyanobacteria (based on genome sequencing and existing literature) is shown (Fig. 1).

Richness change over time

All the following analyses are based on the reduced data set where the gut-associated lineage Gastranaerophilales, not considered relevant for the present environmental survey, was excluded. The subsetted rarefied data comprised of forty taxa in sixty-five samples. The majority of OTUs (30) were assigned to class ML635J-21, and the Melainabacteria OTUs mostly belonged to the Obscuribacterales order (5 OTUs). A single OTU was assigned to Vampirovibrionales, and four OTUs were not assigned to a class. In contrast to the Oxyphotobacteria, which increased in richness across the 10 lakes over the twentieth century (see Chapter III), the richness of environmentally associated nonphotosynthetic OTUs did not increase significantly over the last century (n = 65, p > 0.3; Fig. 2). Lake Annecy was the only exception where the richness of nonphotosynthetic cyanobacteria increased significantly over time (R^2 = 0.86, p = 0.003).

Temporal and regional beta diversity change

Between-lake phylogenetic similarity (based on Unifrac distances) in communities of nonphotosynthetic cyanobacteria was lower than the average similarity estimated in photosynthetic cyanobacterial communities (Fig. 3). The similarity estimated between lakes at each decades between the 1940s and 1970s was stable before it slightly decreased from the 1980s onward (Fig. 3). However, PERMANOVA did not support a temporal effect over the nonphotosynthetic cyanobacterial assemblages over the time span covered by the analysis (i.e., 1940s–2000s; Table 2). This contrasts with the PERMANOVA results obtained for communities of Oxyphotobacteria, where a temporal succession effect was significant.
Interestingly, while a regional (north–south) effect was observed for the photosynthetic communities, no significant north–south discrimination of community similarity was detected in nonphotosynthetic cyanobacteria. PERMANOVA supported a lake effect in phylogenetic community similarity in both groups (Table 2).

**Table 2.** Test of the effects of temporal and spatial factors on community phylogenetic dissimilarities of nonphotosynthetic and photosynthetic cyanobacteria determined by PERMANOVA*. p-values are based on 9,999 permutations, bold face indicates statistical significance (p < 0.05), DF; degree of freedom. Samples dated to years between the 1940s and the 2010s were used.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Nonphotosynthetic cyanobacteria</th>
<th>Photosynthetic cyanobacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF</td>
<td>Pseudo-F</td>
</tr>
<tr>
<td>Temporal (decades)</td>
<td>8</td>
<td>1.2031</td>
</tr>
<tr>
<td>Regional (north-south)</td>
<td>1</td>
<td>1.8719</td>
</tr>
<tr>
<td>Spatial (lake)</td>
<td>9</td>
<td>3.6927</td>
</tr>
</tbody>
</table>

*PERMANOVA: Permutational Multivariate Analysis of Variance using Distance Matrices.

**Temporal changes in OTU prevalence**

The average linkage clustering highlights five groups (Fig. 4). Cluster I comprises of four OTUs that were prevalent in the lakes at all time periods. Representatives of this cluster are all affiliated to class ML635J-21. Clusters II and III show a large number of OTUs that appear to be randomly distributed and that were never found in high prevalence across the ten lakes. Cluster IV highlights a group of OTUs that were generally more prevalent in older times compared to the last 3–4 decades. Most of the latter OTUs were also found in recent sediments, but they were less common across the region. They are mainly representatives of the ML635J-21 class, and one OTU is assigned to Obscuribacterales. Cluster V is composed of 2 ML635J-21 representatives, and one unidentified OTU, which is most likely associated with class ML635J-21 considering its position in the phylogeny (see OTU 114 in Fig. 1). Interestingly, OTUs in Cluster V show the opposite pattern observed in Cluster IV: they became more common across the region since the 1940s (Fig. 4).
Distance-decay in cyanobacterial community similarity

Because there was a significant lake effect, but not a regional effect over beta diversity (i.e., across lake dissimilarity) in nonphotosynthetic cyanobacteria, we wanted to find out whether the geographic distance between lakes explained the patterns in pairwise dissimilarity across communities. The community dissimilarity values obtained for each pairwise set of samples plotted against geographic distance revealed no significant distance-decay curve at all time periods in nonphotosynthetic cyanobacterial communities (Fig. 5a). Interestingly, the relationship was also non-significant in photosynthetic cyanobacterial communities (Fig. 5b).

The relationship across all pairs of communities at each decade was also tested using compositional similarity (Jaccard distances based on presence-absence of OTUs). Interestingly, there was no significant relationship between compositional similarity in nonphotosynthetic cyanobacteria and geographic distance (Fig. S2a in the supplementary materials), but the Mantel test confirmed that the distance-decay relationship was significant at five time periods (1950s, 1970s, 1990s, 2000s, and 2010s) in photosynthetic cyanobacterial communities (Fig. S2b in the supplementary materials).

Discussion

In this study, we report about the long-term composition and phylogenetic diversity of recently discovered clades of nonphotosynthetic cyanobacteria in ten European freshwater lakes. Little is known about the diversity and distribution of the basal clades Melainabacteria and ML635J-21 in nature. The class ML635J-21 is the least understood since no genome was ever sequenced. The ecological function and the niche occupied by these ancestral clades of nonphotosynthetic cyanobacteria is unknown to date. Our results constitute a first attempt to describe their diversity across broad regional and temporal scales, and to investigate their distribution in parallel to the photosynthetic cyanobacterial lineage as a reference.

The number of nonphotosynthetic cyanobacteria OTUs recovered from the sedimentary archive of peri-Alpine lakes was much lower in comparison to the richness of Oxyphotobacteria determined previously. Most of the OTUs were assigned to the deepest-branching paraphyletic group of nonphotosynthetic cyanobacteria, called ML635J-21, and to the Melainabacteria order Gastranaerophilaes. The latter OTUs were not retained for analysis because of their association mostly with guts and faeces samples (Di Rienzi et al.,
2013; Quast et al., 2013). Further, their diversity and distribution in our samples appear to be random, suggesting they might their presence in lakes might be the result of punctual events, for example via run-off, wastewater, or the presence of wild animals and livestock in the lake surroundings.

Unlike Oxyphotobacteria, the richness of environmentally associated nonphotosynthetic cyanobacteria did not significantly increase over time, with the exception of Lake Annecy. The exact niche(s) occupied by Melainabacteria and ML635J-21 taxa is not yet determined, but because of their nonphotosynthetic nature, they most likely live deep in the water column below the photic zone. As a consequence, they are less likely to be able to disperse across lakes in comparison with Oxyphotobacteria that mostly live in the upper water strata of lakes.

Because of the lack of genomic data for representatives of ML635J-21, we do not know about their ecology and their metabolic functions. In this study, we found members of this clade at every depth of the sediment cores and in all the peri-Alpine lakes investigated. While it could be hypothesized that these organisms could be living in the sediments rather than in the sediment / water interface or in the water column, this is unlikely considering that their presence has been reported in various liquid environments, like surface sea water (Venter et al., 2004), termite guts, microbial mats (Quast et al., 2013) and in various compartments of a meromictic lake (Gies et al., 2014). As direct ancestors of Melainabacteria and Oxyphotobacteria, they are most probably aquatic life forms, but more research would be needed to verify if they remain metabolically active in the sediments. More exploratory studies and genome sequencing could help shed light on the ecology of nonphotosynthetic cyanobacteria in general.

Similar to patterns observed in Oxyphotobacteria, there was a significant lake effect on the phylogenetic distance estimated between each pair of nonphotosynthetic cyanobacterial communities. This could indicate that local lake conditions are the most important drivers of community structure in this group, or could be signalling dispersal limitation across lakes. Our results, however, do not suggest distance-decay in phylogenetic community similarity in neither photosynthetic or nonphotosynthetic cyanobacterial communities, and the Mantel test reveals only a sporadic (in only half of the time periods investigated) significant decrease in compositional similarity in Oxyphotobacteria over geographic distance (Fig. S2 in the supplementary materials). Our results suggest that the taxa composition of Oxyphotobacteria is sometimes more dissimilar between lakes separated by greater geographic distance, but that the community phylogenetic structure is not modified significantly by the taxa composition. A phylogenetic study of the cyanobacterium *Chroococcidiopsis* found in hot and
cold deserts showed that distance-decay did not explain their distribution on a global scale, and that *Chroococcidiopsis* variants found in hot and cold deserts were evolutionarily distinct (Bahl *et al.*, 2011).

Our results suggest that nonphotosynthetic cyanobacteria communities are not shaped by the same drivers as Oxyphotobacteria, and that their composition is mostly lake-dependant. Extending the survey of nonphotosynthetic cyanobacterial diversity to other climatic regions could reveal patterns due to environmental selection that may have played a major role in the establishment of these clades in different regions. This, and the fact that their diversity and richness did not vary a lot over the last century, could indicate that they have been present in the lakes over long evolutionary time. In contrast, PERMANOVA suggests that Oxyphotobacteria phylogenetic diversity is under both local (lake) and regional (north–south) selection and appear to be more responsive to environmental changes as seen in our previous study (Monchamp *et al.*, submitted) and in others (e.g., (Anneville *et al.*, 2005; O'Neil *et al.*, 2012).

**Conclusion**

This study represents an initial exploratory survey of the composition and diversity of unknown clades of nonphotosynthetic cyanobacteria, reconstructed from lake sediments over time and across lakes of the peri-Alpine region. Our results highlight the rich diversity of these mysterious organisms that might play important roles in lake biogeochemistry. High-throughput sequencing of environmental DNA has the potential to illuminate the diversity of understudied microbial groups over large spatial and temporal scales. In combination with genome sequencing, diversity surveys will help gain understanding on the evolution, the function, and the ecological role of such microbes, like nonphotosynthetic cyanobacteria, in natural environments.
Acknowledgements

High-throughput sequencing was performed at Fasteris sequencing company in Geneva (Switzerland). We thank Jean-Claude Walser (Genetic Diversity Centre, ETH Zürich) for bioinformatics support, and Marco Thali (Eawag) for his help with the sequencing library preparation. We also thank Alois Zwyssig, Alfred Lück, Adrian Gilli, Patrick Turko, Nathalie Dubois, and Andrea Lami for help in the field, and Isabelle Domaizon for providing sedimentary DNA samples from lakes Geneva and Annecy.
Figure 1. A maximum-likelihood phylogenetic tree based on all OTU reference sequences assigned to non-photosynthetic cyanobacteria found in this study, with *Microcystis aeruginosa* as outgroup. Bootstrap values > 0.50 are shown. The clade in red was assigned to class ML635J-21, and taxa highlighted in blue are representatives of class Melainabacteria, which splits into three orders: Gastranaerophilales, Obscuribacterales, Vampirovibrionales (the fourth order Caenarcaniphilales was not present). A trait table next to the phylogeny indicates whether representatives of the group have been found to possess (full circle), or not (empty circle) a given physiological trait. No circle signifies missing information.
Figure 2. Natural log-transformed OTU richness of nonphotosynthetic cyanobacteria associated with environmental samples (ML635J-21, Obscuribacterales, and Vampirovibrionales) over time in the ten lakes. The dashed lines show the non-significant linear fit for each lake, and the full line (Lake Annecy only) shows the significant linear fit for that lake ($R^2 = 0.86, p = 0.003$). The overall relationship over all observations ($n = 65$) is not significant at the $p = 0.05$ level (black dashed line).
Figure 3. Phylogenetic dissimilarities among lake communities of non-photosynthetic cyanobacterial at each decade from the 1940s to the 2000s. The horizontal dashed lines, representing the mean dissimilarity value across all observations for each group of cyanobacteria (black for photosynthetic cyanobacteria (see Monchamp et al. in prep; Chapter III); PC, dark grey for nonphotosynthetic cyanobacteria; N-PC), are shown for reference.
Figure 4. Colour-coded map showing the prevalence (i.e., the proportion of lakes where and OTU was found at a given time period) of ML635J-21, Obscuribacterales, and Vampirovibrionales in the lakes between the 1870s and the 2010s. Hierarchical clustering was based on Euclidian distances based on prevalence of OTUs and revealed five main clusters: I – OTUs always prevalent in a majority of lakes; II – group of OTUs not common nor isolated relatively stable over time; III – rare and isolated OTUs; IV – OTUs mostly decreasing in prevalence over time; and V – OTUs increasing in prevalence over time. The absence of an OTU at a given time period is depicted by a white box, and the gradient of colour is proportional to the prevalence of an OTU in the group of lakes sampled at a given time period.
Figure 5. Distance-decay plot showing the log-transformed pairwise phylogenetic distances (based on UniFrac) calculated among all pairwise communities of (a) nonphotosynthetic cyanobacteria and (b) photosynthetic cyanobacteria at each decade between the 1900s to the 2010s. Each point in the plot represents the pairwise log-transformed phylogenetic distance between communities plotted against the log-transformed geographic distance. A geographic distance of 0 signifies that the pairwise phylogenetic dissimilarity was calculated between samples from the same lake corresponding to the same time period). The two vertical dashed lines mark 50 km and 130 km distances for reference. The absolute geographic distances between lakes (in km) are shown in Table S1 in the supplementary materials.
The dark side of lakes: diversity and distribution of nonphotosynthetic cyanobacteria in sedimentary archives

Table S1. Pairwise geographic distances between lakes (kilometres).

<table>
<thead>
<tr>
<th></th>
<th>LUG</th>
<th>BAL</th>
<th>BOD</th>
<th>HAL</th>
<th>PUS</th>
<th>MAG</th>
<th>ANN</th>
<th>GEN</th>
<th>GRE</th>
<th>ZRH</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUG</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAL</td>
<td>147</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOD</td>
<td>183</td>
<td>101</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAL</td>
<td>157</td>
<td>10</td>
<td>101</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PUS</td>
<td>22</td>
<td>169</td>
<td>200</td>
<td>179</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAG</td>
<td>26</td>
<td>142</td>
<td>192</td>
<td>151</td>
<td>41</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN</td>
<td>217</td>
<td>220</td>
<td>317</td>
<td>222</td>
<td>229</td>
<td>191</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEN</td>
<td>195</td>
<td>156</td>
<td>256</td>
<td>157</td>
<td>212</td>
<td>171</td>
<td>72</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRE</td>
<td>155</td>
<td>36</td>
<td>65</td>
<td>37</td>
<td>175</td>
<td>156</td>
<td>255</td>
<td>192</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>ZRH</td>
<td>144</td>
<td>32</td>
<td>70</td>
<td>36</td>
<td>165</td>
<td>145</td>
<td>248</td>
<td>187</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure S1. Rarefaction plot (read count limit set at 2000 for visualisation) of all the nonphotosynthetic cyanobacteria samples from the ten lakes. The sequencing coverage was variable among samples and across lakes. There is no visible effect of time on the sequencing depth that would indicate a loss of richness over time due to DNA degradation with sediment age. Based on this plot, we decided to rarefy the samples to 200 reads in order to retain as many samples as possible for the analysis and considering that the asymptote is reached in most samples at this sequencing depth.
Figure S2. Distance-decay plot showing the log-transformed pairwise taxonomic distances (based on Jaccard similarity on incidence of OTUs) calculated among all communities of (a) nonphotosynthetic cyanobacteria and (b) photosynthetic cyanobacteria at each decade between the 1900s to the 2010s. Each point in the plot represents the pairwise log-transformed compositional distance (Unifrac) between communities plotted against log-transformed geographic distance. Geographic distance of 0 signifies that the pairwise phylogenetic dissimilarity was calculated between samples from the same lake. The relationships were all non-significant in nonphotosynthetic cyanobacteria. In photosynthetic cyanobacterial communities, however, the distance-decay relationship using OTU composition was significant at five time points: 1950s, 1970s, 1990s, 2000s, and 2010s. There were not enough samples in the 1900s, 1910s and 1920s to estimate the distance-decay relationship between photosynthetic cyanobacterial communities. The two vertical dashed lines mark distances of 50 km and 130 km for reference.
Literature cited Chapter V


R Core Team (2013) *R: A language and environment for statistical computing*. Vienna, Austria.


CHAPTER VI

General discussion and outlook

Over the twentieth century and especially during the last few decades humans have imposed tremendous stress on almost all biomes on Earth. Aquatic ecosystems have been particularly affected by rapid shifts in nutrient concentrations (through eutrophication and oligotrophication) and both by direct and indirect consequences of climate change. As a consequence, noticeable changes happened in lakes, especially the increase of cyanobacterial abundance, with negative effects on ecosystem food-webs and water quality. Because of restoration measurements, several lakes returned to P concentrations similar to the pre-eutrophic levels, but nevertheless, some potentially toxic cyanobacterial taxa continue to dominate the phytoplankton assemblages in such lakes on a regular basis (Jeppesen et al. 2002; Jacquet et al. 2005; Posch et al. 2012; Taranu et al. 2015), supporting the hypothesis that management of lakes based on P management only is insufficient for reversing the negative effects of eutrophication (Schindler 2006).

Over the recent decades, a number of sedDNA-based studies reported about compositional changes in cyanobacterial populations related to modifications of lake conditions (e.g., Domaizon et al. 2013; Savichtcheva et al. 2015). However, these studies were constrained to a few sites and were only focused on cyanobacterial populations, thus lacking an overview of the dynamics of local and regional cyanobacterial metacommunity and its reaction to environmental change. The general aim of this thesis was to fill this gap by investigating how regional (climate) and local (nutrient fluctuations) changes have affected the taxonomic and phylogenetic composition of communities, as well as the distribution of cyanobacteria over a long time scale extending throughout the periods of pre-, mid-, and post-eutrophication (including the re-oligotrophication phase), across the broad peri-Alpine region. This could be achieved by using an environmental DNA approach applied to sediment cores to reconstruct communities over two centuries.

One of the most important results emerging from this thesis is that planktonic communities – both in photosynthetic cyanobacteria and Daphnia – changed over time, but did not return to the pre-disturbance taxonomic and phylogenetic composition after the phase of re-oligotrophication (Chapters II, III, and IV). This finding is consistent with other studies based on long-term observations and palaeo-reconstructions. For instance in peri-Alpine
lakes, recent studies have highlighted that shifts in the genetic structure of *Daphnia* populations (Brede *et al.* 2009; Moest 2013) and in cyanobacterial composition (Jacquet *et al.* 2005; Posch *et al.* 2012) could not simply be reversed by phosphorus reduction in lakes. Similarly, we found that bloom-forming cyanobacterial taxa have spread throughout the peri-Alpine area since the eutrophication of lakes and have become more prevalent in the region over the last few decades, intensifying potential problems due to the presence of harmful cyanobacteria (Chapter III). The favoured taxa were mostly members of the orders Oscillatoriales, Chroococcales, and Nostocales, all known to comprise potentially invasive taxa (Sukenik *et al.* 2015). A number of recent studies have reported that the spread of Nostocales taxa from tropical and sub-tropical latitudes to temperate regions seems to be facilitated by global warming (Carey *et al.* 2012; Sukenik *et al.* 2012).

The successful colonisation of these taxa throughout the peri-Alpine region appears from my study to be a consequence of abiotic homogenisation of lake conditions due to change in climate and land-use (Chapter III). Most likely, climate warming, by affecting lake physics (e.g., strength and duration of the period of thermal stratification and chemistry (e.g., alteration of nutrient cycles), has contributed to favour large and buoyant taxa that thrive under contemporary lake conditions (a stable and strongly stratified water column), leading to the regional homogenisation of community composition. Our results are in line with the results of a multi-lake survey reporting a change in the phytoplankton composition in recently oligotrophied peri-Alpine lakes at the end of the 1980s, most likely as a response to synchronous changes in P concentrations and in climatic conditions across lakes of the region (Anneville *et al.* 2005). Recent research in terrestrial ecosystems also revealed that landscape simplification due to land-use intensification led to biotic homogenisation at multiple levels of the grasslands food-web (Gámez-Virués *et al.* 2015; Gossner *et al.* 2016). As climate change scenarios predict future global warming (Houghton *et al.* 1990), further research in aquatic systems is necessary to understand the direct and indirect consequences of warming on community assembly in cyanobacteria.

**Strengths and limitations of the phylogenetic reconstruction approach applied to sedimentary DNA**

Throughout this thesis, I use state-of-the-art molecular tools to investigate long-term patterns in the diversity and phylogenetic structure of cyanobacterial communities in several European lakes. Currently, it is still unclear how well cyanobacterial diversity is archived and preserved over long time scales in the sediments of lakes. Because of the lack of long-term
biotic and abiotic lake data, it is rarely possible to compare results obtained by sedDNA-based reconstructions with historical lake data. For this reason, I first aimed at validating our sedDNA approach by comparing sedDNA reconstructions to long-term records from well-studied lakes (Chapter II). Because DNA molecules degrade in the environment, DNA sequencing from sediment samples has some limitations. Although DNA is well-preserved in cold environments protected from sunlight, this does not completely prevent DNA damage from happening. Nucleic acids in the environment accumulate chemical changes and degrade over time, causing fragmentation of the DNA molecules and impairing the structure and recovery of the molecules (Shapiro & Hofreiter 2012). I chose peri-Alpine lakes as study sites because they offer ideal conditions for the preservation of sedDNA, i.e., the low oxygen, cold temperature, lack of light, and the general absence of mixing of the surface sediments by benthic fauna due to hypoxic conditions (Coolen et al. 2004; Boere et al. 2011). We found that the OTU richness was highly related to the estimated number of species in water, but that abundance of sequence reads in sedDNA samples did not always reflect the estimated number of cells in water (Chapter II). This could partly be due to biases related to the methodological approach, but most likely it is a consequence of lake-specific processes (e.g., mixing, currents, thermal stratification) and taxon-specific traits (e.g., buoyant cells) that influence sedimentation processes. More studies focusing on sedimentation processes in lakes (e.g., by installing sediment traps) could help identifying the limitations of sedDNA-based studies for investigating patterns of planktonic diversity relevant for ecological studies.

There are also biases associated with the laboratory methods used, from DNA isolation to PCR and high-throughput sequencing. Although the yields can vary between extraction methods used, the commercial kits used in this project has been shown to produce reproducible bacterial community structure analysis (Vishnivetskaya et al. 2014). Apart from DNA degradation and biases related to molecular methods, another major concern in sedDNA and ancient DNA studies is the contamination of samples by contemporary DNA, which can lead to artifactual results. To reduce the risks of contamination, older sediments were handled separately from more recent samples, and all samples were processed in small batches of eight samples. Further, all pre-PCR molecular analyses were done in a clean laboratory where strict ancient DNA work protocols were followed (e.g., Willerslev & Cooper 2005; Pedersen et al. 2014).

Another main limitation when studying cyanobacterial diversity in general is the under-representation of freshwater cyanobacteria in publicly available DNA reference databases, which restrains the depth of the taxonomic assignment of amplicons or OTUs. In this thesis
project, we used a relatively large DNA marker (400-nucleotide-long), considered a suitable size (considering potential DNA degradation) for sedDNA studies over long geological timescales (Coolen & Gibson 2009). This allowed us to use the 16S rDNA reference sequences for inferring detailed phylogenies (Chapters II to V). By favouring a phylogenetic approach using a relatively long DNA marker, we were able to partly overcome the limitation related to the lack of exhaustive cyanobacterial databases. For example, we constructed a phylogeny with all of the 16S rDNA sequences from this study assigned to Nostocales and reference sequences from public databases (Fig. 1), which led to the identification of an OTU highly related to Dolichospermum lemmermannii that had recently been isolated from southern peri-Alpine lakes (Salmaso et al. 2015a). This analysis confirmed the presence of D. lemmermannii in 9 out of 10 lakes, and allowed us to trace back the colonisation history of this cyanobacterium across the region (Chapter III and Fig. 2 of current chapter). Similarly in Chapter II, we were able to assess the presence since the early 1800s in lake Zurich of a single OTU related to Planktothrix rubescens previously isolated from the same lake (Fig. 3 of Chapter II).

Figure 1. A phylogenetic tree based maximum-likelihood showing the relationship among 16S rDNA sequences assigned to Dolichospermum species from this study and other reference sequences of Dolichospermum/Anabaena isolated from southern peri-Alpine lakes and from cultures, with Microcystis aeruginosa as outgroup. The analysis was bootstrapped and bootstrap values >60% (out of 100 replicates) are shown at the respective node as blue circles (circle size is proportional to the bootstrap value). OTU 614 was assigned to D. lemmermannii, and OTU 4 to D. solitaria based on this phylogeny.
Figure 2. Grid plot showing the prevalence of OTU 614 (*Dolichospermum lemmermannii*) in the sediment samples collected from the 10 peri-Alpine lakes. OTU 614 was first found in the samples from the 1910s in Lake Geneva (north-western Plateau; N-W), followed by lakes from the northern Plateau (N) over the next few decades. *D. lemmermannii* sequences were latest detected in the sediments of the southern peri-Alpine lakes (S) from the 1980s (Lugano) and the 1990s (Maggiore and Pusiano) onward.

Implications and outlook

In summary, the dataset obtained in this thesis project is one of the few to comprise phylogenetic and biogeographic information on whole communities of lake plankton (mainly phylum Cyanobacteria, but also *Daphnia* communities) over centuries and across a large geographic region. This thesis contributes important information about the past and contemporary cyanobacterial diversity, with the potential to help predict the future distribution patterns of potentially harmful cyanobacteria under expected global change scenarios. Cyanobacterial blooms will continue to occur and might even increase in frequency and severity (Paerl & Paul 2012). So far, taxa in the orders Oscillatoriales, Chroococcales, and Nostocales with specific traits like buoyancy-regulation and bloom-formation have been found to be favoured under changing environments (Carey *et al.* 2012; Rigosi *et al.* 2014). Opening the sedimentary archive of lakes for investigating the global biodiversity and
biogeography of cyanobacteria will deepen our understanding of the ecology of invasive taxa, and allow to detect invasions at an early stage before consequences become visible.

**Next step towards the understanding patterns of local community assembly**

A metacommunity is defined as ‘a set of local communities that are linked by dispersal of multiple potentially interacting species’ (Leibold *et al.* 2004). Local community assembly within the metacommunity depends simultaneously on various factors that are either stochastic (i.e., dispersal, ecological drift, historical priority effects) or more deterministic (e.g., selection imposed by environmental abiotic conditions, species interactions) (Stegen *et al.* 2012).

The mean pairwise distance (MPD) is a measure of phylogenetic diversity that is often used to assess the phylogenetic relatedness between taxa within a community, and it could be a starting point to investigate community assembly processes. It is calculated as the distance between each pair of taxa, based on the tree branch lengths extracted from a given phylogenetic tree. This measure is used to quantify the overall clustering or overdispersion of taxa on a tree. To interpret and compare the community structure across phylogenies in a metacommunity, it is necessary to examine the degree of deviation of the MPD values obtained from the observed community to a randomized null model obtained with species drawn from the regional species pool (i.e., all the species in a defined metacommunity) (Webb 2000; Webb *et al.* 2002; Graham & Fine 2008; Stegen *et al.* 2012). In communities primarily driven by neutral processes, the phylogenetic composition should not differ significantly from the one derived from a random community assembly. If deterministic processes govern community assembly over stochastic events, the community phylogenetic diversity should show some degree of structure. There is a growing number of studies investigating phylogenetic structure in microbial communities using a null model approach (Horner-Devine & Bohannan 2006; Emerson & Gillespie 2008; Stegen *et al.* 2012), but studies applied specifically to natural freshwater cyanobacterial assemblages are, to the best of our knowledge, inexistent.
Figure 3. Time series of the standardized effect size of mean pairwise distances (SES MPD) values calculated for each local community. The calculation was based on each community subtree phylogenetic distances tested against the random communities (999 permutations) with randomization of the species at the tip of the phylogenetic tree while species richness in maintained (Hardy & Senterre, 2007; Erickson et al., 2014; Swenson, 2014). The samples outside of the area delimited by the black dotted lines show significant community structure ($p < 0.05$), with negative values signalling phylogenetic clustering, and positive values signalling phylogenetic overdispersion.

Although the studies presented in this thesis were not designed to investigate the influence of such mechanisms of community assembly, the 16S rDNA sequence dataset generated is ideal to investigate patterns in phylogenetic structure of the communities across space and time. This next step could lead to the formulation of new hypotheses to be tested experimentally about the drivers of community assembly in freshwater cyanobacteria. Preliminary analyses of MPD in the reconstructed cyanobacterial communities from the 10 peri-Alpine lakes studied do not show clear evidence for the influence of deterministic processes over cyanobacterial community assembly, but rather suggest that communities are mostly random (Fig. 3). It appears, however, that there was a greater effect of deterministic processes leading to community overdispersion during the recent-most few
decades. This pattern probably reflects the increase in richness of buoyant and colonial taxa since the 1980s observed in Chapter III, but this is not enough to conclude whether the overdispersion was driven by environmental changes like warming and nutrient fluctuation.

Extending the approach used throughout this thesis and investigating patterns in community phylogenetic structure in lakes over several continents could help to disentangle the relative influence of climate and local environmental change over community assembly in freshwater cyanobacteria. Such data would be useful to verify hypotheses on species biogeography, such as the Baas Becking ‘Everything is everywhere’ hypothesis (Baas Becking 1934). Ultimately, the approach used in this thesis has the potential to illuminate global biogeographic patterns of cyanobacteria and other plankton forms that could help determine the underlying factors generating and maintaining biodiversity in lakes.
Literature cited Chapter VI


Boere AC, Rijpstra WIC, de Lange GJ et al. (2011) Exploring preserved fossil dinoflagellate and haptophyte DNA signatures to infer ecological and environmental changes during deposition of sapropel S1 in the eastern Mediterranean. *Paleoceanography, 26*, 1–16.


Acknowledgements

First, I would like to thank my supervisor Piet Spaak for selecting me in the first place, and for believing in me during this entire PhD project. I appreciated the liberty I was given to design my own project; this is a great gift that not all PhD students get.

Second, I want to thank my co-supervisor Francesco Pomati for his generous feedback on my work, for encouraging creative thinking, and for always pushing me to get out of my comfort zone. He helped me believe in myself and in my ideas and thought me the importance of becoming an independent researcher.

I am also grateful to Irene Gregory-Eaves for being the external examiner for my PhD examination, and Jukka Jokela for accepting to be on the jury as well. I further want to express my gratitude to Nathalie Dubois, Isabelle Domaizon, and Jean-Claude Walser for fruitful collaborations on various projects. I thank them for their support and feedback, and I hope we will collaborate again in the near future.

I want to thank all of the former and present members of the Pomati and Spaak groups for many fruitful discussions over the years. Especially, I thank Christoph Tellenbach for his kindness and the continuous support he offered me during my PhD, and Mridul Thomas who helped me in many ways to achieve my scientific and personal goals.

A special thanks to Marco Thali, Alfred Lück, Alois Zwissig, Irene Bruner, Esther Keller, Hannele Penson, Pravin Ganesanndamoorthy, Daniel Steiner, Andrea Lami, Adrian Gilli, and Patrick Turko for the technical support. It would not have been possible to complete this PhD project without your contribution in the lab and in the field.

I also thank the GDC team Aria Minder, Silvia Kobel, Stefan Zoller, Niklaus Zemp, and Jean-Claude Walser who were always ready to answer questions and provide great bioinformatics support. Thanks to Alice Dennis for her help with phylogenetic analyses, Rosi Siber for the help with ArcGIS, Marlène Lavrieux for help with sediment dating and for interesting discussions, and my officemates Hanna Hartikainen, Elvira Mächler, and Inês Fontes for the fun times, the helpful scientific discussions, as well as the sometimes well-needed moral support.

I want to say many thanks to all my past and present colleagues at Eawag and in the Eco department for making this place an amazing work environment. Especially (but not
exclusively) Mat Seymour, Marie Leys, Anja Bürkli, Marta Reyes, Markus Möst, Kirstin Kopp, Kirsten Klappert, Eric Harvey, Katja Räsännen, Tiina Salo, Eduardo Martin, Laura Langeloh, Ioana Enache, Iasmina Moza, Anita Narwani, Isabelle Gounand, Kristy Deiner, Lukas de Ventura, and Gioia Matheson. A special Merci to Silvana Käser for her endless patience teaching me Schwiiizertüütsch, and to Yogi Arianne Maniglia for teaching me yoga and for offering support in stressful times.

Last but not least, I thank my family who always believe in me no matter what I do, and Claudéric for his endless patience and his support until the very end. This PhD was a long journey and your presence made everything better.