Non-cyanobacterial nitrogen fixation insights in humic freshwater lakes and ponds

LEYDEN FERNANDEZ-VIDAL
The biological basis of nitrogen fixation beyond the canonical role of Cyanobacteria is not well understood in freshwater ecosystems. To address this gap in knowledge, the main objective of this thesis is to study non-cyanobacterial nitrogen fixation in freshwater lakes and ponds.

Microbial communities and diazotrophic potential were characterized by direct metagenome sequencing from seasonally stratified lakes and permafrost thaw ponds, both of which are systems featuring strong redox gradients. To quantify the nitrogen fixation process, we also adopted and applied a $^{15}$N tracer method to estimate realized diazotrophic activity in five of the studied lakes. Chemical characteristics were also measured concomitantly to link diazotroph distribution patterns to chemical features and metabolic traits in the studied freshwaters.

Exploring a 3-year metagenomic time series of a humic lake (Trout Bog), widespread and stable occurrence of $nifH$ genes were detected. This marker gene for nitrogen fixation appeared with accessory genes, validating the marker. The diazotrophic community was diverse and dynamic with contributions from Geobacter, Desulfobacterales, Methylococcales, Acidobacteria, Verrucomicrobia and Chlorobi. Accordingly, nitrogen fixation may be fueled by a variety of metabolic processes (heterotrophic sulfate/iron reducers, methylotrophs and photolithotrophs) in oxygen depleted dark waters. Interestingly the photolithotrophic Chlorobi also carried a Fe-only nitrogenase ($anfH$) recently implicated in alternative methane production. Overall, we demonstrated widespread potential for nitrogen fixation within hypolimnia in stratified humic lakes, and analyses of depth profiles also confirmed the presence of active diazotrophic communities in boreal lakes. These active nitrogen fixing communities were characterized by overall higher bacterial abundances, cellular aggregation and increased phosphorus availability as compared to communities where nitrogen fixation was not detected.

Expanding our work to include recently formed freshwater ecosystems, we characterized microbial communities in arctic thaw ponds at different ontogenetic stages. We also investigated the possible role of $anfH$ in methane production, but the abundance of this gene was not correlated with high methane concentration in the water column. $NifH$ was detected in all systems, and interestingly the hypolimnetic waters in the more established systems emerged as suitable niche for diazotrophs with $nifH$ abundances positively correlated to elevated methane concentrations. Based on this observation, we propose that nitrogen-fixing microorganisms may be important partners in complex syntrophic networks established between bacteria and archaeal methanogens.

In conclusion, the results presented across the different types of terrestrially influenced freshwater systems revealed widespread potential for nitrogen fixation within hypolimnia of humic lakes and permafrost thaw ponds. Furthermore, nitrogen fixation was confirmed at and below the redoxcline in five Finnish humic lakes. This implies that we need to look beyond phototrophic cyanobacteria to more fully understand the role of nitrogen fixation and overall nitrogen cycling in freshwater ecosystems.

**Keywords:** humic lakes, thaw ponds, non-cyanobacterial diazotrophs, nitrogen fixation, nitrogenase

**Leyden Fernandez-Vidal, Department of Ecology and Genetics, Limnology, Norbyvägen 18 D, Uppsala University, SE-75236 Uppsala, Sweden.**

© Leyden Fernandez-Vidal 2019

ISSN 1651-6214

urn:nbn:se:uu:diva-395057 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-395057)
To my family
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from Wiley.
Additional Papers

In addition to the papers included in this thesis, the author has contributed to the following papers during the PhD studies:

Contents

Introduction ................................................................................................... 11
The nitrogen cycle and N-fixation in freshwaters ........................................ 12
The nitrogenase enzyme and N-fixation biochemistry ................................ 13
Environmental factors controlling N-fixation in aquatic ecosystems .......... 15
Biogeochemical approaches to measure N-fixation ..................................... 16
Identifying N-fixers .................................................................................. 17

Scope and Objectives of the Thesis .............................................................. 20

Methods ........................................................................................................ 21
Collection of metagenomic and environmental data from humic-rich boreal lakes ....................................................................................... 21
Nitrogen fixation experiments .................................................................... 22
Collection of metagenomic and environmental data from permafrost thaw ponds .................................................................................. 23
DNA extraction and sequencing .................................................................. 24
Bioinformatics .......................................................................................... 25
Data analysis and visualization .................................................................... 26

Results and Discussion ................................................................................. 28
General findings of diazotroph abundances and taxonomy in boreal humic lakes and permafrost thaw ponds .................................................. 28
How do seasonal microbial dynamics and nutrient status affect potential diazotroph abundances and community composition in a north American humic bog lake? .............................................................. 30
How is nitrogen fixation regulated in freshwater non-cyanobacterial diazotrophs? .................................................................................. 32
Which are the main energy acquisition strategies of diazotrophs in dark hypolimnetic humic waters? .......................................................... 34
Does nitrogen fixation take place in the aphotic and anoxic waters of stratified boreal lakes? Which environmental factors may affect such nitrogen fixation? .......................................................................... 34
How does bacterial aggregation affect nitrogen fixation? ...................... 35
What is the role of anfH nitrogenase in methane production in permafrost thaw ponds? .............................................................................. 37
How do potential NCDs inhabiting thaw ponds couple metabolic processes related to methane metabolism with nitrogen fixation? ........ 39
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCD</td>
<td>Non-cyanobacterial diazotrophs</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>DNRA</td>
<td>Dissimilatory nitrate reduction to ammonium</td>
</tr>
<tr>
<td>Anammox</td>
<td>Anaerobic ammonia oxidation</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosyl triphosphate</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosyl diphosphate</td>
</tr>
<tr>
<td>16S rRNA</td>
<td>16S ribosomal RNA</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in situ hybridization</td>
</tr>
<tr>
<td>nanoSIMS</td>
<td>Nano-scale secondary ion mass spectrometry</td>
</tr>
<tr>
<td>DOC</td>
<td>Dissolved organic carbon</td>
</tr>
<tr>
<td>DIC</td>
<td>Dissolved inorganic carbon</td>
</tr>
<tr>
<td>TN</td>
<td>Total nitrogen</td>
</tr>
<tr>
<td>TP</td>
<td>Total phosphorus</td>
</tr>
<tr>
<td>TDN</td>
<td>Total dissolved nitrogen</td>
</tr>
<tr>
<td>TDP</td>
<td>Total dissolved phosphorus</td>
</tr>
<tr>
<td>ALW</td>
<td>Artificial lake water</td>
</tr>
<tr>
<td>NMDS</td>
<td>Nonmetric multidimensional scaling</td>
</tr>
<tr>
<td>MAG</td>
<td>Metagenome assembled genome</td>
</tr>
<tr>
<td>NP-MAG</td>
<td>Nitrogenase positive MAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemicals:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N, N₂</td>
<td>Nitrogen, dinitrogen</td>
</tr>
<tr>
<td>O₂</td>
<td>Dioxide</td>
</tr>
<tr>
<td>CH₄</td>
<td>Methane</td>
</tr>
<tr>
<td>NH₄</td>
<td>Ammonia</td>
</tr>
<tr>
<td>Mo</td>
<td>Molybdenum</td>
</tr>
<tr>
<td>Fe</td>
<td>Iron</td>
</tr>
<tr>
<td>P</td>
<td>Phosphorous</td>
</tr>
<tr>
<td>S</td>
<td>Sulfur</td>
</tr>
</tbody>
</table>
Introduction

Nitrogen is a crucial element for life. It is a key component of amino acids and nucleic acids and the disruption of the N cycle can bring imbalance to ecosystems. The N cycle has been widely studied in seawater, i.e., at the ocean's surface (Capone and Knapp, 2007), at deep-sea vents (Capone et al., 2006) and more recently at and below the chemocline of the central Baltic Sea (Farnelid et al., 2013). However, the N cycle and specifically nitrogen fixation are not well understood in freshwater ecosystems. Aquatic fixation of atmospheric nitrogen is considered an essential source of biologically available nitrogen in freshwater and marine environments (Finke and Seeley, 1978; Capone, 2006). Due to the crucial role of nitrogen fixation, it is important to understand how this process works in freshwater lakes. For instance, earlier work has shown that cyanobacterial nitrogen fixation is significant in a broad range of natural waters (Finke and Seeley, 1978; Hübel and Hübel, 1980; Lehtimäki et al., 1997; Goebel et al., 2007) and recent evidence has shown that other bacterial and archaeal species may also have leading roles in the marine N cycle and this also applies to fixation (Farnelid et al., 2013). In freshwater ecosystems, there are only a few studies about nitrogen fixation by non-cyanobacterial microorganisms. Further, there is only patchy information about the occurrences of nitrogen fixation at and below the redoxcline in stratified freshwater lakes (Halm et al., 2009).

New techniques have revealed the importance of N-fixing freshwater microorganisms in the global N cycle. First, only specific lineages of prokaryotes can fix nitrogen. Second, nitrogen fixation by bacteria and archaea can lead to NH$_4^+$ accumulation that can be toxic for fish and other aquatic life, and also lead to eutrophication and high oxygen demand in freshwater reservoirs (Siripong and Rittmann, 2007). In summary, nitrogen fixation is an overlooked process in freshwaters ecosystems, crucial to the maintenance and health of aquatic life.
The nitrogen cycle and N-fixation in freshwaters

Simplified, the N cycle in aquatic ecosystems consists of five pathways: nitrification, denitrification, anaerobic (or aerobic) ammonium oxidation, dissimilatory nitrate reduction to ammonium (DNRA), and nitrogen fixation (Canfield et al., 2010). Figure 1 illustrates the main chemical transformations of nitrogen in stratified lakes.1

Nitrification is the oxidation of NH₄⁺ to NO₃⁻ in three steps under aerobic conditions (Cébron et al., 2003; Canfield et al., 2010). In contrast, denitrification is the reduction of NO₃⁻ to N₂, mainly mediated by a diverse group of anaerobic bacteria. Denitrification in lakes is expected to occur in the water column and in some lakes can be more significant in the sediments (Seitzinger, 1988). The anaerobic ammonia oxidizing bacteria (anammox), which include Planctomycetes members, can couple NH₄⁺ oxidation to NO₂⁻ reduction (van Niftrik and Jetten, 2012). An important characteristic of the anammox bacteria is the versatility of the adaptation machinery. Accordingly, they are widespread in different natural systems (Zhu et al., 2015).

DNRA or dissimilatory NO₃⁻ reduction to NH₄⁺ is reported to mainly take place in sediments and wetlands. The inhibition of this biological process can lead to the increase of NO₂⁻ concentration in the water column of aquatic systems, affecting other N cycle pathways and the life of freshwater organisms in general (Kelso et al., 1997). Finally, the process of biological nitrogen fixation is reported to occur in sediments, as well as, in the water column of freshwater ecosystems. With this pathway, some bacteria and archaea (known as diazotrophs) transform atmospheric nitrogen to NH₄⁺, which can then be incorporated into more complex organic molecules and sustain the rest of the food web (Canfield et al., 2010).

---

1 Stratified lakes: In summer, the temperature (and density) differences between upper and lower water layers become more distinct in temperate lakes. Deep lakes generally become physically stratified into three identifiable layers, known as the epilimnion, metalimnion, and hypolimnion.
The focus of this PhD thesis is the biologically driven nitrogen fixation in freshwaters. The bacteria and archaea that can fix nitrogen have in common that they all carry the nitrogenase gene \((\text{nifH})\). N-fixers can catalyze the fixation by using the nitrogenase enzyme, whose multiple subunits are encoded by the genes \(\text{nifH}, \text{nifD}, \) and \(\text{nifK}\) (Hu and Ribbe, 2011).

The nitrogenase enzyme and N-fixation biochemistry

The nitrogenase operon structure is highly conserved throughout the taxonomic spectrum of diazotrophs. However, this functional guild can be divided into three different groups considering the phylogenetic history. The first group carries the common Mo-Fe nitrogenase, which is present in the Proteobacteria and Cyanobacteria phyla. The second group includes a broad range of organisms that does not require oxygen for growth while carrying anaerobic Mo-Fe nitrogenase (i.e. *Clostridia*, acetogenic bacteria, and several methanogens). The third group consists of a limited subset of diazotrophs, with alternative nitrogenases, including the Mo-independent \(\text{anf}\) and \(\text{vnf}\) genes. The product of the \(\text{anf}\) gene is a protein wherein the cofactor Mo is replaced by iron, whereas the \(\text{vnf}\) gene code for a vanadium nitrogenase protein (Raymond et al., 2004).
Two metalloproteins form the molecular structure of the most abundant nitrogenase in nature (molybdenum-iron nitrogenase): the dinitrogenase heterotetramer or molybdenum-iron (MoFe)4 protein (the product of *nifD* and *nifK* genes) and the dinitrogenase reductase homodimer (the product of the *nifH* gene) (Figure 2). The site of substrate reduction is contained in the dinitrogenase heterotetramer, typically in the subunit coded by the *nifD* gene. The electron donor to MoFe protein is the dinitrogenase reductase, also known as the nitrogenase Fe protein. The electron transfer step is carried out by the [Fe4S4] cluster within the dinitrogenase reductase redox-active center (Halbleib and Ludden, 2000; Hu and Ribbe, 2011).

**Figure 2.** Molybdenum iron nitrogenase structure and biochemical activity. Fdxr is the abbreviation for reduced ferredoxin and Fdxo is the abbreviation for oxidized ferredoxin. The figure was adapted from Esteves-Ferreira et al. (2017).

Compared to other metabolic reactions, the ATP (adenosyl triphosphate) consumption by the nitrogenase to reduce one molecule of N₂ is high (16 ATPs). Thus, nitrogen fixation is a metabolically expensive process which accordingly requires a precise regulation. For instance, the nitrogenase transcription is known to be downregulated by NH₄⁺ and O₂. Furthermore, post-translationally, the enzyme is inactivated by ADP-ribosylation of the Fe protein under energy limiting conditions or sufficient nitrogen supply (Halbleib and Ludden, 2000; Hu and Ribbe, 2011; Yang et al., 2011). Interestingly, in a sediment core incubation experiment, nitrogenase activity did not respond to the addition of “fixed” N. This fact suggests that the regulation of the nitrogen-fixing enzyme in nature is more complicated than some biochemical models previously have proposed (Halbleib and Ludden, 2000).
Environmental factors controlling N-fixation in aquatic ecosystems

Biological nitrogen fixation occurs in all aquatic ecosystems. The microbially mediated nitrogen fixation responds to the availability of phosphorus and nitrogen in lakes. A recent study of benthic N fixation in oligotrophic northern lakes (Gettel et al., 2013), pointed to organic matter as a potential limiting factor for heterotrophic N-fixers, whereas P is typically necessary for some filamentous autotrophic microorganisms (i.e., heterocyst of N₂ fixing Cyanobacteria).

The worldwide use of NH₄⁺ and other nitrogen compounds as fertilizer can lead to eutrophication in oligotrophic lakes (Canfield et al., 2010). Also, eutrophication is exacerbated by increasing global warming, making this complex phenomenon harder to manage (Dokulil, 2014). However, not only direct N addition leads to nitrogen surplus, as also P levels can contribute, at least indirectly, to inputs of this nutrient. The N:P ratio is crucial in controlling N-fixer communities in the water column where small N inputs can enhance nitrogen fixation and cyanobacterial growth. Accordingly, low P levels can depress diazotroph blooms. When N is limited, diazotrophs activate the highly energy demanding process of nitrogen fixation. If the supply of nitrogen is sufficient, they may use less expensive metabolic pathways to sustain growth and this may have repercussions on the functioning of the entire ecosystem.

Nevertheless, N-fixer microorganisms need P to carry on the nitrogen fixation, possibly because of the considerable ATP consumption required by the nitrogenase (Esteves-Ferreira et al., 2017). In oligotrophic lakes characterized by low nutrient concentrations, it is more likely that diazotroph ecology follows Liebig’s law of the minimum. This law states that in environments where multiple nutrients are in low concentrations, only one will affect the growth of the organism, and for diazotrophs this nutrient often seems to be phosphorous (Warsi and Dykhuizen, 2017). There is another hypothesis related to symbiotic N-fixers, which demand P for building and maintaining the symbiosis. The exact mechanisms of action for such P modulation, as well as, the role of non-cyanobacterial fixation in the water column of lakes and ponds require further research (Vitousek et al., 2002).

An early study, suggested that heterotrophic bacteria were unable to fix N in the surface of lakes due to the deleterious effect of O₂ over nitrogenase enzyme (Howarth et al., 1988). Other factors that modulate nitrogen fixation by plankton include light, depth of the mixing zone and grazing (Vitousek et al., 2002). For instance, autotrophic Cyanobacteria need light to fuel nitrogen fixation, but the photosynthetic process produces oxygen besides energy and the nitrogenase enzyme need to be protected from oxygen exposition. Thus, cyanobacterial diazotrophs may rely on stored carbon as an energy source (under low light intensity) or they may form heterocyst to couple nitrogen fixation
with high-intensity light conditions and an active photosynthesis (Gettel et al., 2013).

The presence of benthic N-fixers seems to be an early indicator of eutrophication in high-latitude oligotrophic lakes. Such benthic N-fixers include autotrophic and heterotrophic bacteria where the latter require a source of labile carbon and are thus highly dependent on N/C availability to conduct nitrogen fixation (Gettel et al., 2013).

Biogeochemical approaches to measure N-fixation

An early method used for measuring nitrogen fixation was the C$_2$H$_2$ reduction assay. This can be used as an indirect approach to calculate nitrogen fixation rates. The nitrogenase metabolizes C$_2$H$_2$ as a substrate instead of N$_2$, and its reduction rate is a sensitive proxy for nitrogenase activity. The advantages of this assay include easy application in field studies and simple experimental equipment. As the C$_2$H$_2$ reduction assay is not a direct measurement of nitrogen fixation (Montoya et al., 1996), it has nowadays been largely substituted by isotopic-labelling techniques.

There are two natural and stable isotopic forms of nitrogen, the most abundant, $^{14}$N (99.6337%), and the rare isotope, $^{15}$N (0.3663%). $^{15}$N isotope can be used in three kinds of studies. The first type of approach involves the estimation of the natural abundances and the isotope fractionations in aquatic systems (Robinson, 2001; Canfield et al., 2005). The second type of studies consists of adding various $^{15}$N labeled compounds (depending on the N transformation under investigation) as a tracer, to know in which reaction the N compounds are involved. The third and rather specialized isotopic method described in this section, known as nanoSIMS, allows tracing $^{15}$N labeled compound incorporation into the cell (Marchant et al., 2016).

Natural abundances and isotope fractionations: Microbially-mediated assimilatory nitrate reduction, nitrogen fixation, denitrification, and nitrification result in distinct fractionations ($^{15}$N/$^{14}$N) in aquatic systems. The study of the natural variations in $^{15}$N can provide information about the intensity of these N cycle processes in an ecosystem (Robinson, 2001; Canfield et al., 2005). In paleolimnology, the natural $^{15}$N abundances constitute a way to assess changes in the global nitrogen cycling through geologic time (i.e., using lake sediment records) (Canfield et al., 2010; McLauchlan et al., 2013).

Enrichment analysis of $^{15}$N labeled chemical forms: The introduction of isotopic enrichment analysis helped to solve the problem of the coupling of many chemical reactions involved in the N cycle. For example, $^{15}$N added as N$_2$ can be monitored in diazotrophs and used to measure the movement of $^{15}$N into the particulate matter (this approach is explained in more detail in the Method section) (Montoya et al., 1996; Grokopf et al., 2012).
**Nano-scale secondary ion mass spectrometry:** Nano-scale secondary ion mass spectrometry, also known as nanoSIMS, constitutes a variation of isotopic tracer technique, which allows the observation of labeled compounds incorporated by microorganisms at the single-cell level (Marchant *et al.*, 2016).

The combination of biogeochemical and microbiological methods constitutes a step-in advance in the general knowledge about the N cycle and fixation (Grokopf *et al.*, 2012).

**Identifying N-fixers**

In the early microbiological era, it was only possible to study N transformation in microorganisms that had been cultured (Figure 3). Subsequently, the introduction of PCR (Polymerase Chain Reaction) and later qPCR (quantitative PCR) techniques allowed the amplification and quantification of genes of interest (involved in nitrogen fixation, i.e., *nifH*) directly from environmental samples. More recently, metagenomic studies have been exploited for estimating microbial composition and abundances in natural environments while the parallel development of metatranscriptomics has enabled inferences about genes activated under specific physicochemical conditions (Ravin *et al.*, 2015).

The marine nitrogen fixation process has been widely studied using various combinations of molecular biology and isotopic techniques (Grokopf *et al.*, 2012). Grokopf *et al.* (2012), analyzed data from the Atlantic Ocean, and suggested that earlier nitrogen fixation rates had been underestimated. The authors based their assumption on a combined approach, in which they both quantified the abundance of clades of diazotrophs and estimated rates with $^{15}$N tracers. They used TaqMan assays (a highly specific qPCR) to target the presence of nitrogenase reductase (*nifH*) gene to quantify the relative abundance of the diazotrophic phylotypes. As a result, they proposed a more accurate method to measure $^{15}$N rates. The new approach (Mohr *et al.*, 2010; Grokopf *et al.*, 2012) helps to solve the problem of N$_2$ gas bubbles, which leads to underestimation using the method previously described by Montoya and co-workers (see section above) (Montoya *et al.*, 1996). Unfortunately, only a few scientific studies refer to inland water diazotrophs, even though they are crucial for keeping an optimal status on freshwater lakes (Halm *et al.*, 2009).
Figure 3. Timeline in microbiological science. This figure lists the main scientific events from the early microbiology to current times, which also reflects the main techniques used to study N-fixing microorganisms.

Advances in molecular biology techniques, such as PCR and qPCR, have revealed previously unknown diversity of uncultivated microorganisms capable of fixing N. For instance, in marine environments, new heterotrophic, non-cyanobacterial prokaryotes were discovered by nifH PCR amplicon datasets. This new group of heterotrophic diazotrophs seems to be ubiquitous and can grow in high-latitude, dark, cold or coastal marine waters (Bombar et al., 2016).

To further elucidate the organisms engaged in this process, the use of nanoSIMS and fluorescence in situ hybridization (FISH) or halogen in situ hybridization (HISH) can be of great use. By the combination of nanoSIMS and FISH, the identity of a microorganism is linked to its activity and the approach can be applied directly to environmental samples (i.e., by measuring the incorporation of stable isotopes into the population) (Marchant et al., 2016). One example is the study conducted by Halm and coworkers (Halm et al., 2009). They used meromictic Lake Cadagno as a model to analyze nitrogen cycling processes in stratified water bodies. Lake Cadagno is also characterized by a high salinity of the monimolimnion. 16S rRNA gene sequencing revealed that the microbial community at the chemocline was dominated by Chlorobium clathratiforme. They also targeted nifH genes expressed as mRNA, showing that N-fixers were more active right at the chemocline. Furthermore, by in situ
hybridization-secondary ion mass spectroscopy, they assured the diazotrophic character of *Chlorobium*.

Using metagenomics studies involving freshwater communities, Linz and coworkers (2018) identified a group of overlooked diazotrophs inhabiting Trout Bog Lake, a humic rich freshwater system. Compared to PCR and qPCR approaches, metagenomics studies avoid pitfalls and limitations introduced by primer coverage. Metagenomic and metatranscriptomic analyses have also been applied successfully to assess a N-fixer community in the Amazon river plume (Hilton *et al.*, 2015). With this analysis, it was possible to reveal new diazotrophic populations, untargeted by the commonly used PCR-primers. Interestingly, they found that highly abundant transcripts in the diazotroph metatranscriptomes were annotated as non-coding regions. This fact opens more questions about the role of non-coding or unannotated coding sequences in the regulation of nitrogen fixation.

The literature reviewed here clearly illustrates the importance of studying N transformation and particularly nitrogen fixation in freshwater ecosystems. Identifying aquatic N-fixers and their metabolic traits is crucial for designing novel strategies to face the new challenges brought on by the global climate change (Canfield *et al.*, 2010). However, still many questions remain unsolved about microbially mediated N transformations taking place in lakes, rivers and ponds. For instance, the diazotrophic community that inhabits water bodies of lakes is poorly characterized, as well as the specific hotspots for enhanced nitrogen fixation. For instance, anoxic bottom waters in stratified lakes are enriched in phosphorous released from sediment (Marsden, 1989; Dokulil, 2014), making these water masses a suitable and overlooked niche for diazotrophs. Further studies are needed to outline details on metabolic and regulatory capabilities of non-cyanobacterial N-fixers in freshwater ecosystems.
Scope and Objectives of the Thesis

The broader aim of this PhD thesis was to identify environmental factors controlling nitrogen fixation and potential diazotrophs in carbon-rich freshwater ecosystems from humic lakes to permafrost thaw ponds. We tested four hypotheses: (i) anaerobic photolithotrophs, iron-reducers and methanotrophs are significant nitrogen-fixing organisms in humic lakes and permafrost thaw ponds, (ii) such nitrogen fixers are found in hypoxic micro-habitats, (iii) strong redox gradients are hotspots of enhanced nitrogen fixation activity or may affect diazotroph abundances and (iv) deep waters with limited solar radiation exposure are conducive for nitrogen fixation.

Furthermore, the manuscripts included in this thesis address the following specific questions:

- How do seasonal microbial dynamics and nutrient status affect potential diazotroph abundances and community composition in a humic bog lake? Paper I
- How is nitrogen fixation regulated in freshwater non-cyanobacterial diazotrophs? Which are the main energy acquisition strategies of diazotrophs in dark hypolimnetic humic waters? Paper II
- Which are the non-cyanobacterial community members involved in nitrogen fixation? Does nitrogen fixation take place in the aphotic and anoxic waters of stratified boreal lakes? Which environmental factors may affect such nitrogen fixation? How does bacterial aggregation affect nitrogen fixation? Paper III
- What is the role of anfH nitrogenase in methane production in permafrost thaw ponds? How do potential non-cyanobacterial diazotrophs in this type of ecosystem couple metabolic processes related to methane metabolism with nitrogen fixation? Paper IV
Collection of metagenomic and environmental data from humic-rich boreal lakes

(Paper I, Paper II, Paper III)

We collected metagenomic and environmental data from six different boreal humic lakes, one from North America (Wisconsin, USA), and the other five located in the Evo Forest in southern Finland (Figure 4).

The metagenomic time series from the north American Trout Bog Lake (46° 02’ 2” N and 89° 41’ 10” W) was used in Paper I and Paper II and included multiple years of sampling. In Paper I, I analyzed metagenomic data from a seasonal time series, in which sampling occurred at approximately weekly intervals during open water season and primarily during the summer stratified period from 2007 to 2009. The samples were composite samples from epilimnion and hypolimnion, respectively. Meanwhile, Trout Bog metagenome assembled genomes explored in Paper II, were assembled using mainly hypolimnetic DNA data collected on seasonal experiments over nine years (2005-2013). Trout Bog Lake is studied at the North Temperate Lakes Long Term Ecological Research Network (LTER) site; and feeds data to the Global Lake Ecological Observatory Network (GLEON). Temperature and dissolved oxygen were measured throughout the three years sampling campaign (Paper I). In addition, during the sampling campaign of 2008 dissolved organic carbon [DOC], dissolved inorganic carbon [DIC], total nitrogen [TN], total phosphorus [TP], total dissolved nitrogen [TDN] and total dissolved phosphorus [TDP]) were also measured (Linz et al., 2018) in Paper I. The average water temperature in the hypolimnion was approximately 5°C during the collection period, suboxic conditions prevailed and light penetration in the lake was limited to a shallow upper photic layer (Baines et al., 2000).

In Paper III, data from five boreal lakes: Åliven Mustajärvi, Keskinen Rajajärvi, Mekkojärvi, Valkeakotinen and Yliven Rajajärvi were used as a model to study nitrogen fixation. The lakes are located on an area of 48 km² in southern Finland (61° 10’ to 61° 13’ N and 25° 5’ to 25° 12’ E). The lakes were sampled during August 2015 at different depths across the lake water column. The samples were taken from epilimnion, metalimnion and hypolimnion, respectively, following a gradient of DOC, oxygen and nutrient concentrations. Oxygen concentration and temperature were measured in situ using an oxygen
probe, YSI 55 model combined temperature and oxygen probe (Yellow Springs Instruments, Yellow Springs, Ohio, USA). NH$_4^+$, NO$_3^-$, PO$_4^{3-}$, SO$_4^{2-}$ were measured by the laboratory facility at the Limnology Department, Uppsala University. Finally, cell counts and morphology were analyzed following a previously described protocol (Osman et al., 2017), but using a fluorescence imaging system Leica AF6000. For bacterial cell counts and morphology (particle size measurement), we used ImageJ (Version 1.51 23 April 2018) (Schneider et al., 2012).

Nitrogen fixation experiments
(Paper III)

Along with environmental parameters, nitrogen fixation was measured at different depths during the sampling campaign for Paper III. N$_2$ fixation rates were measured by our adaptation of the method described by Mohr et al. (2010). Mainly, we changed the artificial seawater protocol for our own artificial lake water protocol (see below).

Before the incubations, we prepared artificial lake water (ALW) as described by Ricão Canelhas et al. (2016). As we needed to handle samples which contained oxygen-sensitive microorganisms, N$_2$ gas was bubbled into the ALW to remove the oxygen. Then, in an anaerobic chamber, 60 ml of ALW was transferred into septum-capped glass bottles until overflow, and 1 ml of $^{15}$N$_2$ gas (98 at%, Cambridge isotope laboratories) was injected to the bottle. The bottles were sealed, shaken vigorously until the bubble disappeared, and kept overnight in an orbital reciprocating shaker. The $^{15}$N$_2$ enriched and anoxic ALW was transferred from the glass bottle to a syringe and transported to the field in a portable anaerobic chamber for setting up of the incubation experiment.

For incubations, the lake water was transferred to 1.2L septum-capped poly-carbonate bottles until overflow (two incubations and one time zero (t0) sample). Aliquots containing 50 ml of $^{15}$N$_2$-enriched ALW were injected to the incubation bottles, which were mixed. Then incubation bottles and one control (without adding $^{15}$N$_2$-enriched ALW) were incubated in situ for 24h at the same depth where the lake water was collected, while t0 sample was filtered immediately after the addition of $^{15}$N$_2$-enriched water. The incubation samples and controls were filtered similarly after 24h. Filters were sent to the University of Jyväskylä in Finland for isotope analysis using mass spectrometer gas chromatograph (MS-GC). The subsequent calculations were done as described in Montoya and coworkers (Montoya et al., 1996).
Collection of metagenomic and environmental data from permafrost thaw ponds
(Paper IV)

The DNA and environmental data used in Paper IV were collected near the shore of Hudson Bay, subarctic Canada, in August 2014 (Figure 4 A, C). The study site has sporadic permafrost (covering < 10% of the total area) and consists of a palsa bog site called SAS2. As the palsas thaw, ponds emerge and can eventually be up to 4m deep (in SAS2 study site). In Paper IV, we analyzed physicochemical and DNA sequencing data corresponding to eleven Canadian thaw ponds from the SAS2 area (one pond sample of the original 12 ponds failed during the DNA sequencing process). The ponds represent three different stages of thaw pond succession; emerging ponds, middle-age or developing ponds (approximately two to five years) and old ponds (40-50 years).
Middle-age and old ponds were sampled across a vertical gradient, which allowed us to collect data from different limnological layers (epilimnion, metalimnion, and in the old ponds also hypolimnion). A detailed description of the geographic location and characteristic of the ponds can be found in a previous publication (Wurzbacher et al., 2017).

Nutrients (total organic carbon (TOC), NH$_4^+$, NO$_2^-$, NO$_3^-$, total N (TN), PO$_4^{3-}$, SO$_4^{2-}$) were measured in the laboratory facility at the Department of Ecology and Genetics/Limnology (Uppsala University) using standard methods. Iron concentrations were measured as described in Viollier and coworkers previous publication (Viollier et al., 2000) (Paper III and Paper IV). CH$_4$ and CO$_2$ were analyzed as described by Kankaala and coauthors (2006) with the exception that instead of N$_2$, room air was used to extract the gas from water (Paper III and Paper IV).

**DNA extraction and sequencing**

Paper I and Paper II: Trout Bog Lake DNA extraction and sequencing were performed as previously described (Bendall et al., 2016), by the Department of Energy Joint Genome Institute (DOE JGI) (Walnut Creek, CA, USA). Samples were sequenced on the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA), except for four libraries that were sequenced using the Illumina TruSeq protocol on the Illumina GAIIx platform; all samples were sequenced using paired ends with read lengths of 150 base pairs, see (Linz et al., 2018). Raw sequences and MAGs (metagenome assembled genomes) are publicly available at Project JGI-ID: 416375 (JGI-IMG database).

Paper III: Samples from five Finnish lakes were collected by pressure-filtration and sealed cartridges were subsequently frozen in liquid nitrogen. DNA was extracted using the Mobio PowerSoil kit (Mobio Inc., Carlsbad, CA, USA). DNA concentrations were measured using PicoGreen (Invitrogen, Paisley, UK) and the DNA was fragmented using the Covaris E220 ultrasonicator. Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala Biomedical Centre (Uppsala University, Sweden). Libraries were prepared from 20ng of DNA/fragmented DNA using the ThruPLEX DNA-seq Prep Kit (R400427, Rubicon Genomics) according to the manufacturer’s protocol (QAM108-003). In more details, sequencing was performed with NovaSeq, paired-end 150bp read length, v1 sequencing chemistry.

Paper IV: DNA extraction was done by the same procedure as in Paper III. Sequencing was performed using Illumina HiSeq 2500 running in paired-end 2 × 150 bp mode at the Science for Life Laboratory SNP/SEQ facility hosted by Uppsala University, Sweden. The sequences have been deposited in NCBI SRA under accession number: PRJNA517163.
Bioinformatics

To calculate the abundances of nitrogenase genes in the individual samples in Paper III and Paper IV, we first worked with the unassembled sequence data. Initially, raw reads were trimmed using Sickle (Joshi and Fass, 2011) in Paper IV and Trimmomatic (Bolger et al., 2014) in Paper III. Secondly, forward and reverse reads were merged using PEAR (Zhang et al., 2014). After merging the reads, Prinseq was used for quality control (Schmieder and Edwards, 2011) and fastq to fasta transformation.

Paper I: To measure \textit{nifH} abundances, we used merged reads where paired-end sequencing reads were merged with FLASH v1.0.3 (Magoč and Salzberg, 2011). Later we used two different approaches to select and count reads matching nitrogenase sequences. The first was to compare all reads to genomes in JGI-IMG, including MAGs from Trout Bog Lake and other N-fixer genomes. The second was to compare all reads to a database of \textit{nifH} sequences from ubiquitous bacteria and archaea microorganisms (Gaby and Buckley, 2014). In both approaches, we used BLAST searches to recruit reads and to enumerate how many matched an annotated nitrogenase sequence.

Paper II: We analyzed MAGs containing nitrogenase sequences corresponding to the Project JGI-ID: 416375 from JGI-IMG database. Trout Bog MAGs were assembled by Bendall and coworkers (2016) and we used PhyloPhlan (Segata et al., 2013) for taxonomical annotation. For functional assignment we used Pfam (Finn et al., 2014), KO, and COG-based annotations (all these annotations were included in the IMG genome records).

Paper III: DNA samples containing raw reads from five different lakes and collected at different depths were used in this analysis. For the genome assembly, our preferred strategy was to generate the MAGs gathering all the DNA sequences obtained per lake instead of subdividing by depth of sampling. Metagenomic assembly was performed using MEGAHIT (Li et al., 2016). The contigs were mapped with BBMap (Bushnell, 2014) and binned into MAGs with Metabat (Kang et al., 2015). CheckM (Parks et al., 2015) was used for quality assessment. We used Prokka (Seemann, 2014) for the functional assignment of the contigs and PhyloPhlan (Segata et al., 2013) for taxonomical annotation of MAGs containing \textit{nifH} genes. Then we prepared a database using all nitrogenase sequences identified from the previously generated nitrogenase positive MAGs (NP-MAGs). Then we compared all reads to our custom database of \textit{nifH} sequences. In both approaches, we used BLASTx (Altschul et al., 1990) searches to recruit reads matching as a nitrogenase gene fragment. The nitrogenase gene abundances were then normalized by sample library size (number of reads). The raw reads from the samples collected from the five lakes at different depths were used as inputs for the homology searches.

Paper IV: We prepared two custom nitrogenase gene databases, one contains Mo-iron nitrogenase and the second contain only-iron nitrogenase genes.
To build these databases, we gathered nitrogenase gene sequences (classified by Prokka (Seemann, 2014) annotation system) using as input contigs assembled by pulling together the DNA sequencing data collected from the eleven thaw ponds. After performing the BLASTx searches (Altschul et al., 1990), we counted the amount of read fragments which match a nitrogenase gene. Thus, relative nitrogenase gene abundances were estimated by normalizing nitrogenase gene count per sample library size. Sample library size was calculated as the number of prokaryotic sequences reads, identified by MEGAN (Huson et al., 2016), detected in the merged library.

The prokaryotic (raw) reads were identified using MEGAN v6.10.5 (Huson et al., 2016). DIAMOND (Buchfink et al., 2014) was used to compare all metagenomic DNA sequencing reads (merged reads) against a DIAMOND-NR database (February 2018). Then, each sample was analyzed using MEGAN6 "daa2rma" tool that performs taxonomic and functional classification of all reads. For the taxonomical assignment, we used NCBI database (updated in 2017), and for functional assignment SEED database (Overbeek et al., 2014) (updated in 2015).

NP-MAGs were reconstructed using the same method described for Paper III, and we also used the same annotation functional system for the MAGs.

Data analysis and visualization

Paper I: All the statistical analyses were performed in R version 3.3.2 (2016-10-31) (R Core Team, 2016).

For data visualization, we used the package ggplot2 (Wickham, 2011). R scripts used to generate the main figures and statistical analysis are available at https://github.com/microbioinformatic/Trout-Bog.

Paper II: Easyfig (Sullivan et al., 2011) was the software utilized to draw and reconstruct nitrogenase operon structures. ITOL was used for tree visualization (Letunic and Bork, 2016).

Paper III: Data analysis and graphics were generated using R version 3.4.4 (2018-03-15) (R Core Team, 2016). To understand the relationship between environmental data, particles size, cell abundances and the nitrogen fixation process, we used the Wilcoxon rank-sum test. We also applied Spearman correlation analysis to study the strength of the relationships among average particle size and the environmental data. Also, Spearman analysis was applied to evaluate the correlation among bacterial counts and the environmental data. ggplot2 (Wickham, 2017) and package Phyloseq (McMurdie and Holmes, 2013) were used to draw the graphics.

Paper IV: Multiple correlation analysis (Spearman’s correlation) was used to compare nitrogenase gene abundances against environmental data. To draw the monotonic function that best fits the association between nitrogenase gene abundances and methane concentration, we used ggplot2 (R version 3.4.4
(2018-03-15) (Wickham, 2017) and the method “loess.” A permutational analysis of variance (PERMANOVA) test was used to assess differences in abundances among ponds, ages, and layers. The R package Phyloseq (McMurdie and Holmes, 2013) was used to draw the graphics regarding nifH phylotype abundances. Nonmetric multidimensional scaling (NMDS) analysis of diazotroph diversity was performed using the function metaMDS and Bray-Curtis distances with the function envfit (999 permutations) from the vegan package (Oksanen et al., 2017).
Results and Discussion

General findings of diazotroph abundances and taxonomy in boreal humic lakes and permafrost thaw ponds

This thesis highlights the relevance of non-cyanobacterial nitrogen fixation by mapping potential diazotrophs across different boreal freshwater habitats from humic lakes to permafrost thaw ponds. We found similar features among the NCD (non-cyanobacterial diazotroph) taxonomy and behavior across the different environments, even though they were collected from distinct freshwater systems (lakes and ponds), climatic zones (subarctic and boreal), and continents (North America and Europe).

For instance, in Paper III potential active diazotrophs were identified analyzing the presence and abundances of different nifH phylotypes in samples with a positive nitrogen uptake by biomass. Active nitrogen fixation was detected in all the suboxic hypolimnetic waters overlying sediments (bottom samples) (see the example of Mekkojärvi in Figure 5A). We found in three out of five lakes that nifH phylotype abundances and diversity were higher in samples with an active nitrogen fixation compared to samples in which nitrogen fixation was not detected. NP-MAGs (nitrogenase positive MAGs) identified as Deltaproteobacteria, Gammaproteobacteria and Verrucomicrobia were found in samples with a positive nitrogen uptake in all five lakes. Chlorobi nifH phylotypes were also associated with an active nitrogen fixation process in four out of five lakes. Some NP-MAGs were exclusively detected in one lake, as was the case of nifH-holding Methanomicrobia, a group which was found only in Alinen Mustajärvi. In Paper I and II, we identified similar taxa, i.e., Chlorobi, Geobacter, Desulfo bacterales, M ethylococcales, and Acidobacteria, in the hypolimnion (suboxic conditions) of Trout Bog Lake (Figure 5B). More specifically, nitrogenase sequences identified as Proteobacteria were abundant in May, while from June to the end of October, chlorobial nitrogenase reads were dominant.
In permafrost thaw ponds (Paper IV), diazotrophic community composition was different in samples grouped by pond age and depth layer and to a lesser extent also among the individual ponds (Figure 5C). *Proteobacteria* (mainly *Deltaproteobacteria*) and *Chlorobi* were the most abundant nifH phytypes in the hypolimnion of the old ponds. In contrast, in samples from epilimnion and metalimnion diazotrophs related to *Alphaproteobacteria* and *Gammaproteobacteria* (mainly *Methylococcaceae*) were in higher proportion compared to *Chlorobi* and *Deltaproteobacteria*.

*Cyanobacteria*-related reads contributed negligibly in the hypolimnion of Trout Bog Lake (Paper I). In the lake data analyzed in Paper III, *Cyanobacteria* related reads were not detected in samples with active nitrogen fixation in any of the five Finnish lakes. In permafrost thaw ponds only one out of 11
ponds analyzed contained reads related to *Cyanobacteria* (B1-B2). Thus, in humic lakes and carbon-rich permafrost thaw ponds, non-cyanobacterial diazotrophs seem to dominate the potential nitrogen-fixing communities inhabiting these boreal and subarctic freshwaters.

Next I aimed to address specific questions about the main genomic and metabolic features of NCDs in the boreal carbon-rich lakes and ponds analyzed.

**How do seasonal microbial dynamics and nutrient status affect potential diazotroph abundances and community composition in a north American humic bog lake?**

(Paper I)

Using Trout Bog (Wisconsin, US) as a model system to address this question, we analyzed data from a three-year time series from two different water masses: 1) epilimnion and 2) hypolimnion. Potential diazotrophs in the epilimnion and hypolimnion showed similar maximum abundances (~0.030% nitrogenase genes per sample) during 2008 (Figure 6A), and the relative abundances were not significantly affected by oxygen concentration. Similar behavior was observed during 2007 and 2009, except for one epilimnetic sample collected on 27 July 2009.

The presence of potential diazotrophs seemed to be favored by low TN:TP (<20) (Figure 6B). In marine environments, earlier work has shown that nitrogen fixers seem to respond negatively to an average particulate N:P ratio > 22 (Martiny *et al.*, 2014). Other work has also demonstrated that active nitrogen fixation prevails in N-rich marine water if P is also present and ambient N:P is low. One explanation for this could be that nitrogen fixation is used to stabilize and maintain an ideal intracellular redox state (similar to carbon fixation pathway). This finding hints to a decoupling of nitrogen fixation from the exclusive function of satisfying the cellular N demand (Bombar *et al.*, 2016).
Figure 6. Abundances of potential diazotrophs related to nitrogen and phosphorus availability during 2008 summer season in Trout Bog Lake. (A) Comparison between abundances of potential diazotrophs in the epilimnion and hypolimnion (No DNA data from hypolimnion was collected on 30 April 2008 and 24 June 2008). (B) TN:TP ratio (measured in part per billion) compared with nitrogenase gene abundance data from epilimnion and hypolimnion.
How is nitrogen fixation regulated in freshwater non-cyanobacterial diazotrophs?

(Paper II)

We analyzed nifH gene operons in MAGs associated with Chlorobi, Geobacter, Desulfobacterales, Methylococcales, and Acidobacteria. By this analysis, we identified the nifH gene operon features co-regulated with the nitrogenase (nifH) which is the canonical gene marker for diazotrophs. For the iron-molybdenum cofactor synthesis, additional genes nifX, nifE, nifN, nifB and nifV may be necessary (Chen et al., 2001) and were usually accompanied by the nifHDK cluster in the MAGs analyzed here. However, in Methylococcales related genomes, these genes were carried in a different operon than nifHDK (Figure 7). In addition to Methylococcales, accessory genes nifQ, nifZ, and nifW also appeared in betaproteobacterial diazotroph genomes but were located near a cluster formed by nifHDK, nifB, nifE and nifN. Other nitrogen fixation accessory genes included nifA and nifL, both associated as regulators; affecting the nitrogenase gene expression.

Also, we observed coding regions for PII, DraG, and DraT (Figure 7), with draG/draT genes presence being more usual among nifH positive MAGs. PII family proteins are directly involved in N metabolism pathways regulation (Conroy et al., 2007) while DraG and DraT proteins are specific modifiers of the nitrogenase enzyme (Oetjen and Reinhold-Hurek, 2009). Thus, DraG and DraT proteins may represent a post-translational regulation system for nitrogen fixation in many diazotrophs.

In Betaproteobacteria and Verrucomicrobia, we accordingly observed the presence of molybdenum transport genes (modAB) in the neighborhood of the nifHDK cluster (Figure 7). For all other potential diazotrophs identified in our dataset, modAB were also present, but located elsewhere on their reconstructed genomes. Thus, molybdenum availability seemed crucial to regulate Mo-nitrogenase expression and activity (Chen et al., 2001).

Retrotransposable elements flanking the nifHDK genes were also identified in several freshwater populations (Figure 7). This finding may be related to an extensive horizontal gene transfer. Surprisingly, even though the hypolimnion of Trout Bog Lake rapidly becomes anoxic after stratifying in the spring (Linz et al., 2017), we observed the presence of many genes involved in the protection of bacterial cells from oxidative stress and reactive oxygen species.
Figure 7. Alignment of a nitrogenase gene clusters of 15 MAGs corresponding to potential nitrogen fixers. Coding region sequences (CDS) are colored according to GenBank and KO annotation, with orange representing unknown or unclassified genes. In the gene color legend is visualizing the general function attributed to the CDS.
Which are the main energy acquisition strategies of diazotrophs in dark hypolimnetic humic waters?

(Paper II)

Functional inferences were based on extensive shotgun metagenomics data mainly from Trout Bog Lake hypolimnion (a bottom layer, usually suboxic and affected by low light penetration).

By analyzing NP-MAGs, we inferred functions linked to anoxygenic photosynthesis in the phototrophic *Chlorobi*. An alternative unique antenna system, the chlorosome, allows energy to be obtained under low-irradiance and anoxic/suboxic conditions (Oostergetel *et al.*, 2010).

Furthermore, the presence of *rnf* genes in *Methylococcales* (Figure 7) and *Desulfobacterales* suggested that a different energy-generating metabolism may drive the costly nitrogen fixation in the absence of photosynthesis in these NCD (Tremblay *et al.*, 2013).

A complete pathway to carry out sulfate reduction was detected in *Desulfobacterales*. The use of alternative terminal acceptors for the electron transport chain is crucial for bacteria to mobilize energy from reduced organic compounds in oxygen-depleted waters efficiently (Wang *et al.*, 2017).

Does nitrogen fixation take place in the aphotic and anoxic waters of stratified boreal lakes? Which environmental factors may affect such nitrogen fixation?

(Paper III)

We used five stained lakes (average range watercolor 133-388 mg platinum⁻¹) from Evo Forest in Finland as models for studying nitrogen fixation. In all lakes, irradiance and dissolved oxygen were low in the hypolimnetic waters. Nitrogen fixation occurred independent of oxygen concentration (Figure 8, correlogram). However, nitrogen fixation was detected in the bottom hypolimnetic samples in all five lakes. Low temperature, higher phosphate and ammonia concentration were detected in samples with an active nitrogen fixation (Figure 8). Thus, in these freshwater ecosystems, when nitrogen fixation is active, ammonia production may exceed concomitant consumption. Low temperature as a regulator of nitrogen fixation could also reflect the effect of a different environmental factor correlated with the temperature such as in phosphate gradients and depth (see Figure 8, correlogram).
Figure 8. The boxplots represent the relation between the occurrence of nitrogen fixation with the temperature, phosphate and ammonia concentration. The correlogram depicts a multiple correlation analysis (Spearman method) calculated using the chemical data, the ratio [Fe/P0₄³⁻] was also included in the analysis (ratio variable), as well as the nitrogen fixation rates.

How does bacterial aggregation affect nitrogen fixation? (Paper III)

Environmental conditions in the five humic lakes analyzed in Finland affected the ecology of N fixing organisms. For instance, the presence of high molecular weight humic substances, insoluble Fe(III) and either autochthonous (i.e., algal polysaccharides) or allochthonous (i.e., cellulose) biopolymers may facilitate particle formation in freshwater lakes (Linz et al., 2018; He et al., 2019). Bacteria may attach and form aggregates surrounding such particles in order to obtain nutrients and energy necessary for fueling biological processes,
such as nitrogen fixation, especially in aphotic zones (Linz et al., 2018; He et al., 2019). Using samples collected from five Finnish lakes, we estimated particle size from DAPI stained microscopic images (Figure 9). We found that particle size was significantly larger in the samples with active nitrogen fixation. This property is influenced by i) the cell size and ii) bacterial aggregation. Thus, the presence of big cells may affect our perceptions of the formation of bacterial aggregates as we can spot big cells along with bacterial aggregates in the microscopic images (Figure 9). Average particle size increased at greater depth in cold waters in all the studied boreal humic lakes. Also, a moderate negative correlation (rho= -0.62) was estimated between the average particle size and Fe(III) concentration in the water column, in four out of five lakes. This observation supports the hypothesis about the possible role of insoluble Fe(III) and humic substances in sustaining heterotrophic diazotrophs contributing to particle formation and bacterial aggregation.

Figure 9. Microscopic images showing DAPI stained cells from Alinen Mustajärvi. Sample depth is also included.
What is the role of \textit{anfH} nitrogenase in methane production in permafrost thaw ponds?

(Paper IV)

Nitrogenases are present in three different forms, where the main nitrogenase gene carried by all nitrogen fixers is the molybdenum-iron nitrogenase (\textit{nifH}). Some nitrogen fixers also possess a vanadium nitrogenase gene (\textit{vnfH}) as an alternative nitrogen-fixing enzyme, or an iron-only nitrogenase (\textit{anfH}). In contrast to the other nitrogenases, the iron-only nitrogenase, AnfH, can transform N\textsubscript{2}, CO\textsubscript{2}, and H\textsubscript{2} into NH\textsubscript{3} and CH\textsubscript{4} in a single enzymatic step (Zheng \textit{et al.}, 2018).

Fe-only nitrogenase genes \textit{anfH}, \textit{anfD}, \textit{anfG} and \textit{anfK} were detected in MAGs identified as \textit{Chlorobi}. Although we found fragments of this alternative nitrogenase in contigs classified as \textit{Cyanobacteria}, \textit{Chloroflexi} and \textit{Deltaproteobacteria}, we did not detect the minimal required set of genes to code for an active iron-only nitrogenase enzyme (i.e., \textit{anfHDKG}). Maybe due to MAGs incompleteness within these three taxonomical groups or that the putative \textit{anfH} fragment could be a false positive caused by the high percentage of identity among \textit{anfH} and \textit{nifH} nitrogenase genes.

Iron-only nitrogenase (\textit{anfHDKG}) was weakly correlated with sulfate (Figure 10A). Sulfate is known to inhibit microbial uptake of molybdate, an element essential for the activity of the NifH (molybdenum-iron) nitrogenase. This may explain the high potential for nitrogen fixation utilizing AnfH in sulfate-rich ponds. Thus, the mere presence and observed relative abundance of \textit{anfHKDG} genes did not explain the variation in methane concentrations in the thaw ponds. Instead, a different nitrogenase gene cluster, \textit{nifHDK} that is not related to methane production, showed a strong positive correlation with the concentration of CH\textsubscript{4} in the water column (Figure 10B).
Figure 10. Nitrogenase gene presence and abundance in the thaw pond environment
A) Spearman's rank correlogram. The scale represents Spearman's correlation coefficient values. A cross mark means that the p-value of the analysis after correcting for multiple testing was not significant. B) The upper panel represents the trophic status per pond. Phosphate concentration was used to assign the ponds into different trophic classes according to the ranges determined by guidelines published by the Canadian Council of Ministers of the Environment (2004). Points closed in dashed-squares are the samples marked in the graphic placed below. The bottom panel illustrates the monotonic positive association between CH₄ concentrations and molybdenum-iron relative abundances of nitrogenase genes per thaw pond sample (dotted curve). Colors distinguish the different ages and the shapes distinguish layer of sampling. Gray-shading represents ± 95% confidence intervals.
How do potential NCDs inhabiting thaw ponds couple metabolic processes related to methane metabolism with nitrogen fixation?

(Paper IV)

To study the potential links between diazotroph abundances (nifH genes) and methane production beyond the mere necessity of anoxic conditions, we tested two hypotheses; (i) archaeal methanogens carry nitrogen fixation genes, and (ii) archaeal methanogens and potential diazotrophs may establish a syntrophic or otherwise mutualistic relationship.

Nitrogenase phylotype corresponding to methanogenic Euryarcheota, was scarce in our samples, but was strongly positively correlated with methane concentration (rho= 0.81, p-value < 0.001). Thus, a possible explanation for the increased CH₄ and NH₄⁺ concentrations in the bottom samples collected from mature ponds may be the association with the simultaneous presence of nifH genes and methanogenic pathways in a single taxon. This may be in conjunction with neither methane nor ammonia being oxidized efficiently because of oxygen shortage (Jansson and Taş, 2014). However, we did not completely discard the second hypothesis, as we detected other potential archaeal methanogens in our dataset, which do not possess nitrogen fixation genes.

The relative abundances of nitrogenase genes (nifH) classified as Chlorobi, Deltaproteobacteria, and Elusimicrobia were also highly correlated with methane concentration. Furthermore, in the reconstructed nitrogenase positive MAGs, we identified genes related to syntrophic propionate and butyrate oxidation (Chlorobi, Elusimicrobia and Deltaproteobacteria), an alternative Wood–Ljungdahl (WL) pathway and direct interspecies electron transfer (DIET) (Deltaproteobacteria) (Gan et al., 2012; Zhuang et al., 2014; Sikora et al., 2017; Cai et al., 2018). Thus, N-fixing microorganisms may be important partners in complex syntrophic networks established between bacteria and archaeal methanogens, indirectly linking nitrogen fixation with methane production.
Conclusions and Perspectives

Abundant and diverse communities of non-cyanobacterial microorganisms seem to have the potential to perform nitrogen fixation in the hypolimnetic and dark waters of humic lakes. In permafrost thaw ponds we also found potential diazotrophs from similar lineages as those encountered in the boreal humic lakes. Overall, the majority of the potential nitrogen fixers in the boreal systems analyzed in this thesis are anaerobic photolithotrophs, iron-reducers, sulfate reducers and methanotrophs.

From a seasonal experiment in Trout Bog Lake (2007-2009) we detected that diazotrophs were less abundant during the spring period and the N-fixer community was enriched in phyla, such as Chlorobi, Proteobacteria and Acidobacteria. To the best of our knowledge, there are few environmental studies, i.e., (Halm et al., 2009), that have demonstrated diazotrophic potential in Chlorobi. The results presented in this thesis further support that Chlorobi may play a significant role in diazotrophic humic lake communities, based on the nitrogenase gene abundances and chlorobial nifH gene detection in samples with active nitrogen fixation.

By analyzing the genomic features of microorganisms carrying nifH genes, we detected several alternatives to fuel nitrogen fixation in Trout Bog Lake. For instance, MAGs associated with Chlorobi possessed genes for anoxygenic photosynthesis, and in Deltaproteobacteria, genes linked to humic substance reduction and extracellular electron transfer were present. Carbon fixation genes were found in several MAGs containing nifH genes. In contrast, we observed differences in the energy metabolism capabilities regarding terminal electron acceptors and electron transport proteins within the diazotrophic community in Trout Bog Lake. The presence of Rnf coding regions was characteristic of Alpha- and Betaproteobacteria associated MAGs. Rnf complex protein could enhance nitrogenase enzyme activity by supplying electrons and contributing to ferredoxine reduction (Tremblay et al., 2013).

Active diazotrophic communities were detected in five Finnish boreal lakes and phosphate availability and cold temperature were linked to the positive nitrogen fixation rates. Also, the formation of bacterial aggregates seemed to promote nitrogen uptake by the biomass, suggesting that particle enrichment via bacterial aggregation may enhance nitrogen fixation in the oxygen minimum and aphotic zones in humic lakes.

In this thesis project, I also studied the diversity and metabolic functioning of potential nitrogen-fixing microorganisms in subarctic systems experiencing
climate change. Diazotrophs made up a significant part of the microbial communities in the thaw ponds and this functional guild seemed to be controlled by depth, age and chemical characteristics of the ponds. I also found some support for a role of N-fixing microorganisms (i.e., Chlorobi, Deltaproteobacteria) as partners in complex syntrophic networks established between bacteria and archaeal methanogens, indirectly linking nitrogen fixation with methane production.

The advent of novel approaches such as metagenomics and metatranscriptomics in combination with more accurate isotopic labeling techniques promise discoveries related to the N cycle pathways and the microorganisms involved. The results presented in this thesis showed that the diazotrophic capabilities of prokaryotes not related to Cyanobacteria (also known as non-cyanobacterial diazotrophs (NCD) or heterotrophic diazotrophs) have been overlooked in freshwater boreal lakes and ponds. Nowadays, the technological advantages make the gathering of large-scale genomic data more accessible than ever before and this extends to metabolic characteristics and niche capabilities of the diazotrophs, as well as, their contribution to substantial N inputs in freshwater ecosystems. Furthermore, the general findings discussed here would be relevant for nutrient dynamics and microbial ecology of N-fixers in other aquatic and terrestrial habitats.
Sammanfattning på Svenska

Kväve är ett näringsämne av central betydelse för allt liv och är ofta en begränsande tillväxtfaktor i olika akvatiska system. Trots långtgående forskning om tillförsel och omsättning av kväve i sötvatten och havsvatten är det många faktorer och processer som vi fortfarande inte har särskilt god kunskap om. Ett sådant exempel är kvävefixering i sötvatten. Detta är globalt en väldigt betydelsefull process som omvandlar svårtillgänglig kvävgas (N₂) till mer biologiskt reaktiva former som sedan ligger till grund för ekosystemens samlade produktion av ny biomassa. Tidigare har forskare antagit att kvävefixeringen främst utförs av fotosyntetiserande cyanobakterier, men nya rön har ifrågasatt denna etablerade sanning och påvisat att även många andra bakterier och mikroorganismer kan vara inblandade i denna process. I denna avhandling undersöckes vilken betydelse denna typ av tidigare förbisedda kvävefixerande mikroorganismer har i sjöar andra sötvatten, vilka de är och hur de skaffar energi för tillväxt och aktivitet.

Forskningen har främst baserats på direkt sekvensering och analys av mikroorganismernas samlade arvsmassa, så kallade “metagenom”, i skiktade sjöar och de nya dammar som bildas när permafrost tinar i det arktiska landskapet. I dessa sötvatten kan man hitta skarpa och relativt stabila gradierter i syrgashalt och redoxförhållanden, och den storskalig analyser av metagenomen som vi genomför gör det möjligt att beskriva egenskaper både hos det samlade mikrobiella samhället och de specifika mikroorganismer som bär på de gener som behövs för att fixera kvävgas. För att komplettera dessa analyser anpassades och användes även inkubationsexperiment med stabila kväveisotoper som gjorde det möjligt att också direkt mätta och jämföra kvävefixeringsprocessens hastighet i olika vattenmassor. För att identifiera miljöfaktorer som påverkar och styr kvävefixeringen och de kvävefixerande mikroorganismernas utbredning och sammansättning analyserades även vattenmassornas kemiska egenskaper, och då särskilt med avseende på syrgashalt, ammonium, nitrat, sulfat och andra redox-aktiva föreningar.

I en av avhandlingens experiment studerades det mikrobiella samhällets sammansättning och förändringar i en välstuderad humös sjö i norra Wisconsin (Trout Bog). Genom att följa förändringar i metagenomet under en treårs-period kunde både säsongs-och årstidsvariation i kvävefixerande grupper av mikroorganismer beskrivas. Dessa studier påvisade en stabil och utbredd förekomst av genen nifH som kan användas som robust genetisk markör för kvävefixerande organismer. Det faktum att dessa genetiska markörer direkt kunde...
kopplas till många andra gener som är inblandade i kvävemetabolism är ett ytterligare kvitto på att detta är en bra markör. Kvävefixerarna som hittades i sjön uppvisade stor mångfald, med samhällen som inkluderade *Sulfurimonas, Geobacter, Desulfobacterales, Methylococcales, Acidobacteria, Verrucomicrobia* och *Chlorobi* (Gröva svavelbakterier). De olika gruppernas kvantitativa betydelse varierade och sammantaget uppvisar dessa kvävefixerare stor variation i metabolism, där tillväxt och den energimässigt kostsamma kvävefixeringen verkar kunna drivas av såväl organiska föreningar kopplat till sulfat eller järnreduktion, metabolt utnyttjande av 1-kolsföreningar eller fotosyntes kopplat till litotrofi utan produktion av syre som exempelvis återfinns hos *Chlorobi*. Den sistnämnda populationen bär även på gener för ett extra system för kvävefixering som kan koppla denna process till fototrof metabolproduktion.

Kvävefixerare var särskilt vanliga och utbredda i den studerade sjöns syrefattiga och mörka djupvatten, och detta kunde senare bekräftas i uppföljande studier av djupprofiler från en större urval boreala sjöar med variierande humusinnehåll. Liknande kvävefixerande arter hittades, och i dessa sjöar be- stämdes även hastigheten på kvävefixeringsprocessen, vilket bekräftade att dessa organismer inte bara har de genetiska förutsättningarna för kvävefixering, utan även omsätter detta i praktiken och utför processen direkt i dessa vatten. Betydande kvävefixeringen kunde framför allt kopplas till prover med förhöjd bakteriehalt och synbar aggregering av celler jämfört med de prover där ingen kvävefixering kunde påvisas, men även hög fosfortillgänglighet verkade ha en positiv inverkan på kvävefixeringen.

Förutsättningarna för mikrobiell kvävefixering kan förändras under ekosystemets utveckling och åldrande. För att studera detta utförde vi liknande studier av kvävefixerare i de små vattensamlingar som bildas under permafrostens snabb tillbakagång. Dammar i olika utvecklingsstadijer provtogs med avseende på kvävefixerare och det mikrobiella samhällets sammansättning kopplades till en beskrivande analys av de starka kemiska gradienter i tex syrgashalt och metangas som återfinns i dessa system. Markörer för kvävefixerare (nifH) påvisades i samtliga dammar och deras förekomst var särskilt talrik i djupvatten från de mer ”åldrade” systemen. I dessa vatten var kvävefixerarens antal positivt associerade med förhöjd metan-halt, vilket tyder på att metan, direkt eller indirekt genom syntrofiska associationer, skulle kunna fungera som en energikälla för kvävefixerare.

Sammanfattningsvis påvisades utbredd kapacitet för kvävefixering i de talrika mindre sjöar och dammar som är utmärkande för vår klimatzon. Många av dessa sjöar kännetecknas av hög halt humusämnen från omgivande skogsområden och vätskemeler, men detta verkar inte ha någon hämmande effekt på denna process. Kvävefixerarna uppvisade en stor mångfald både i artrikedom och i metabola egenskaper. Aktiv kvävefixering hittades särskilt i direkt anslutning till skarpa syrgas- eller redoxgradienter. Mina studier understryker att kvävefixering i sötvattenssystem inte kan förstås och studeras enbart utifrån
vår etablerade världsbild där cyanobakterier är ansvariga för denna process. Vi måste även rikta blicken ner mot mörka och syrefria vatten på sjöarnas botten och ta hänsyn till de mikroorganismer som anpassats till ett liv under dessa betingelser.
Acknowledgements

I would like to thank my main supervisor Prof. Stefan Bertilsson for giving me the opportunity to do my second PhD here at the Limnology Department and for providing guidance and support for this thesis. I feel so lucky to have you as my supervisor, this second PhD could not go smoothly without the enormous supports from you.

Thank you, Prof. Sari Peura for co-supervising me and the irreplaceable practical help with field experiments, papers and thesis writing. Also, for all your advices about maternity. I am happy for what you have achieved in these five years in your professional life. You’re a very talented and an excellent person.

Christoffer and Janne thank you for your expertise in the HPLC and chemical analysis. Janne and Alina I learnt a lot during our teaching activities and course preparations. Thank you, Peter, for helping me to improve the “kappa” and for your guidance to improve my work and presentation skills. Thank you also to Anna B. for giving me the opportunity to teach at BMC and EBC. Anna S, Sebastian, Eva, Lars, Silke, Gesa and other professors in the department for your expertise in Limnology and all the feedback during Monday seminars.

Omneya, you are a good friend and thank you for all the trips and conversations we shared together. Also, about the cooking recipes and dinners in Gottsunda.

Jingying, many thanks for your great company and friendship, your patience during our trips and all your support in and outside work.

I would like also to thank:

- Martha, Christoffer, Moritz, Carla and co-supervisor Lasse Riemann for the assistance during the field work, a crucial step for the culmination of this thesis.
- Karolina, Jovana, Annika and Roger for the assistance in the lab and during Monday seminars broadcasting.
– Old and new PhD crew at Limnology (Matilda, Anna N, Anastasija, Ana Isabel, Annika, Lorena, Fernando, Javier, Bianka, Marloes, Fabian, Máte, Karolina, Theresa, Karla, Simone, Rhiannon, Roger, Blaize, Jovana, Yinghua, Maria, Jingying, Monica, Baolin, Martin, Torsten, Inga, Friederike) for all the effort to make a better PhD life in our department
– All the postdoc, researchers and master students that form or formed part of Limno for keeping a good working environment.
– Also, the administrators, Ulla, Marie, Jenny, Eva N, Linn.
– Prof Sophie K for giving me the opportunity to teach in the R course and to share this experience with Giulia, Luis and Marloes.

Many thanks to the co-authors of the papers including in this thesis (Alexandra, Alexander, Björn, Katherine, Moritz, Sanni, Sari and Stefan) for all your hard work and providing insightful feedback to improve the quality of the manuscripts.

Y para terminar me gustaría agradecer a mi familia y amigos en Cuba por no olvidarme durante esta larga ausencia y por su paciencia cuando el tiempo no me permitía responder a sus mensajes. En especial, quiero agradecer a mi esposo Klaudi, por todo su apoyo y a mi hijo Marcel por su bondad sin limites.
References


Goebel, N.L., Edwards, C.A., Church, M.J., and Zehr, J.P. (2007) Modeled contributions of three types of diazotrophs to nitrogen fixation at Station ALOHA. *ISME J.*


R Core Team (2016) R: A Language and Environment for Statistical Computing.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 1867

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-395057