Determining Drivers of Community Structuring and Denitrification Potential Within a Shallow Coastal Lake/Lagoon (Lake Ellesmere) Exposed to Anthropogenic Nutrient Deposition

by Matthew Highton

A thesis submitted for the degree of Bachelor of Biomedical Science majoring in Infection and Immunity with Honours
At the University of Otago
Dunedin
New Zealand

October 2014
Abstract

As a result of human agricultural activity, large amounts of the potent greenhouse gas nitrous oxide (N\textsubscript{2}O) are entering our atmosphere. In NZ, urine from dairy cows contributes large loads of nitrogen to the surrounding land and waterways via leaching or surface runoff. Lake Ellesmere, a coastal lake draining farmland, maintains lower nitrogen levels in its waters than inflowing rivers suggesting a nitrogen removal mechanism, present within the lake. It is likely that much of the incoming nitrogen could be emitted from the lake as N\textsubscript{2} or N\textsubscript{2}O microbial denitrification.

We investigated the microbial community structure and denitrification potential across 18 sites in Lake Ellesmere by using 16S ribosomal rRNA gene sequencing, denitrification enzyme assays and quantification of key nitrogen cycling genes, with a focus on denitrification genes, by quantitative PCR. Physical and chemical analyses (water salinity, temperature, clarity, total nitrogen, nitrate/nitrite, dissolved reactive phosphorous, ammonia, sediment organic matter, and sediment texture class classification) were also carried out simultaneously on the same sites.

Analyses identified sediment texture class and organic matter content as the strongest drivers of bacterial community composition within the lake. Water column nutrients such as Total Nitrogen and Phosphorous were only weakly correlated to community structure suggesting they functioned as weaker modifiers of the system at the time of sampling. Analysis of discrete functional populations by qPCR (e.g. nitrogen fixers, denitrifiers) demonstrated significant variation of gene copy number within the lake sites, however no significant
drivers of these populations were identified for the denitrifiers. Nitrogen fixers were found to be enriched in high silt sediments.

These results demonstrate the difficulty in analyzing these diverse, variable systems. Nevertheless they restate the importance of sediment texture class as a strong determinant for microbial community structure. Further work will need to focus on temporal variability within the system and the production of supporting evidence through controlled testing of denitrification.
Acknowledgements

Thank you first to my supervisor Dr. Sergio Morales for your much needed support throughout the year, the hours spent deciphering my unending flow of questions and the many trips to the 8th floor on the promise of free food. Thank you to the whole Morales lab team, especially Stephanie, for showing me the ropes. I would have been lost without you.

For the yarns and the steady flow of cakes, thank you to the 5th floor crew.

To Marc Schallenberg, Josie Crashaw, cheers for the awesome trip to lake Ellesmere and for providing some of the data which was essential for my project to come together,

Thanks to David Hamilton for the use of his DEA results.

To Bridget, thank you for your love and support and the long hours spent on the bus to be with me. Thank you also to Mum, Dad and the rest of the family for your support and encouragement.

Lastly thank you to the Department of Microbiology and Immunology and University of Otago for having me. It's been a great year!
# Table of contents

1 **Introduction** ................................................................................. 1  
   1.1 General overview............................................................................. 1  
   1.2 Growing agricultural demands and the fate of anthropogenic nitrogen.. 1  
   1.3 Understanding the Nitrogen cycle .................................................. 2  
      1.3.1 Nitrogen fixation: ................................................................. 2  
      1.3.2 Nitrification: ....................................................................... 3  
      1.3.3 Denitrification: ................................................................. 3  
      1.3.4 Dissimilatory Nitrate Reduction to Ammonia: ......................... 4  
      1.3.5 Anammox........................................................................... 4  
   1.4 Determinants of N₂O emission ...................................................... 4  
   1.5 16S analysis ................................................................................ 7  
      1.5.1 Role of observed diversity and community composition......... 7  
      1.5.2 Applying 16S to the study of nitrogen cycling .......................... 8  
   1.6 Functional gene abundance .......................................................... 9  
   1.7 Coastal lagoon environments........................................................ 10  
   1.8 Lake Ellesmere ........................................................................ 13  
   1.9 Aims........................................................................................... 14  

2 **Methods** .................................................................................. 16  
   2.1.1 Sample collection................................................................. 16  
   2.1.2 Sediment sampling ............................................................... 17  
   2.1.3 Water sampling .................................................................... 17  
   2.2 DNA extraction ........................................................................ 18  
   2.3 Quantitative PCR (qPCR) ........................................................... 19  
      2.3.1 *nifH* .............................................................................. 20  
      2.3.2 *nirS* .............................................................................. 20  
      2.3.3 *nosZI* ............................................................................ 20  
      2.3.4 *nosZII* ........................................................................... 21  
   2.4 16S rRNA gene amplicon sequencing ......................................... 21  
   2.5 16S analysis .............................................................................. 22  
   2.6 Denitrification enzyme assay ...................................................... 22  
   2.7 Water and sediment nutrient analysis .......................................... 23
2.8 Statistical and multivariate analysis ................................................................. 23

3 Results ................................................................................................................. 25

3.1 Physical and Chemical gradients ................................................................. 25
  3.1.1 Water column and sediment chemistry .................................................. 25
  3.1.2 Physical factors ....................................................................................... 25
  3.1.3 Correlating physical and chemical variables ........................................... 30

3.2 Microbial community structure ................................................................. 33
  3.2.1 Community composition ...................................................................... 33
  3.2.2 Alpha diversity ....................................................................................... 34

3.3 Beta diversity ................................................................................................. 36
  3.3.1 Principle component analysis (PCA) ...................................................... 36

3.4 UniFrac ......................................................................................................... 40

3.5 Community structure of organisms involved in nitrogen cycling .......... 42
  3.5.1 Potential nitrogen cycling community drivers ......................................... 45

4 Discussion ........................................................................................................ 50

4.1 State of the lake-chemical factors ............................................................ 50

4.2 State of the lake physical factors ............................................................. 52

4.3 Community structure and drivers ............................................................. 52

4.4 Determinants of nitrogen cycling community structure .......................... 55

4.5 Limitations of the experimental design ....................................................... 59

4.6 Future directions ......................................................................................... 60

4.7 Conclusions ................................................................................................. 61
List of figures

Figure 1: Sampling sites and sample collection order within. .......................... 16
Figure 2: Water column chemistry per site and across Lake Ellesmere ............... 27
Figure 3: Sediment chemistry per site and across Lake Ellesmere .................... 28
Figure 4: Physical gradients per site and across Lake Ellesmere ....................... 29
Figure 5: Water nutrient concentrations in relation to Selwyn River inflow .......... 31
Figure 6: Principal component analysis (PCA) of physical and chemical parameters within Lake Ellesmere. ................................................................. 31
Figure 7: Linear regression between lake chemical gradients ............................ 33
Figure 8: Relative abundance of phyla within Lake Ellesmere ....................... 34
Figure 9: Alpha diversity measures, Lake Ellesmere ..................................... 35
Figure 10: Clustering of sites based on PCA of genus level relative abundance data and color coding based on sand percentage at each site ................. 38
Figure 11: Clustering of sites based on PCA of genus level relative abundance data and colour coding based on organic matter content and redox ........... 38
Figure 12: Clustering of sites based on PCA of phylum level relative abundance data and color coding based on sand percentage at sites .......... 39
Figure 13: Two-way cluster analysis of phyla by sample ................................. 39
Figure 14: Water nutrients show correlation to phyla level PCA components .... 40
Figure 15: Principal coordinate analysis (PCoA) at the genus level separates lake site communities based on sand percentage ................................. 41
Figure 16: Distribution of nitrogen fixation and denitrifying genes throughout Lake Ellesmere ................................................................. 44
Figure 17: Denitrification enzyme assay over 18 lake sites ............................. 45
Figure 18: PCA separates physical and chemical gradients and biological within Lake Ellesmere ................................................................. 46
Figure 19: Sediment denitrification genes correlate poorly with physical and chemical lake gradients ................................................................. 49
List of tables

Table 1: Non-parametric Spearman's correlation for lake chemical and physical factors.................................................................32

Table 2: Non-parametric Spearman’s correlation of lake alpha diversity and physicochemical parameters.................................................................35

Table 3: Non parametric analysis of PCA plot principle components ..........41

Table 4: Non parametric analysis of lake nitrogen cycling genes ...............47
1 Introduction

1.1 General overview

Greenhouse gas emissions are a significant and growing concern for the world. As a result of human activities, namely agriculture, large amounts of the potent greenhouse gas nitrous oxide ($\text{N}_2\text{O}$) are entering our atmosphere. An understanding of the processes that produce this gas and the microbes and environments involved can inform action against climate and environment change.

1.2 Growing agricultural demands and the fate of anthropogenic nitrogen

Agriculture continues to increase to support global population growth (van Beek, Meerburg, et al., 2010). Conventional agricultural methods result in large inputs of anthropogenic nitrogen onto farmland to promote crop and livestock growth. Not all of this nitrogen can be consumed by produce, inevitably leaving excess which affects local and distant environments (Syswerda, Basso, et al., 2012). Significant amounts of this nitrogen, usually in the form of nitrate, enter aquatic systems via run off and leaching where it can cause eutrophication (Swaney, Hong, et al., 2012). It has been estimated that 15 to 30 percent of watershed anthropogenic nitrogen in temperate climates is exported by riverine flow (Swaney, Hong, et al., 2012).

Different farming practices, soils, crops, livestock and climates can result in significantly different leaching rates (Bryant, Snow, et al., 2011; Syswerda,
Basso, et al., 2012). Of concern in New Zealand is our large agricultural industry’s focus on dairy. Urine from dairy cows can contribute large nitrogen loads to leaching and is seemingly a more important controlling factor for leaching than the addition of fertilizers (Decau, Simon, et al., 2004).

Human anthropogenic nitrogen inputs associated with agriculture have significantly perturbed the global nitrogen cycle with anthropogenic sources now contributing to 45 percent of yearly global fixed nitrogen. The concern is the fate of this anthropogenic nitrogen; the potential for eutrophication of aquatic environments and production of N$_2$O which is a potent greenhouse gas (Canfield, Glazer, et al., 2010).

1.3 Understanding the Nitrogen cycle

Key to understanding the fate of anthropogenic nitrogen is the process by which it flows through the environment. Nitrogen transitions through its various molecular forms primarily via microbially catalyzed redox reactions (Zehr and Ward, 2002). Flow between pools of molecules depends on molecular stability and accessibility of these molecules to microbes. An overview of the role of key nitrogen cycling processes follows.

1.3.1 Nitrogen fixation:

Most nitrogen exists in the gaseous form of N$_2$ in the atmosphere where it is largely inaccessible to the majority of the biosphere. Nitrogen fixing bacteria are able to access this huge atmospheric sink, bringing this essential life-giving element into the biosphere for all other organisms. This is an energy intensive but necessary process requiring large amounts of ATP (Alberty, 1994) which produces the product ammonia (NH$_3$).
1.3.2 Nitrification:

Nitrification is a two-step aerobic process converting NH\textsubscript{3} into nitrite and subsequently nitrate. The first step, ammonia oxidation, is carried out by a group of bacteria and archaea called ammonia oxidizing bacteria and ammonia oxidizing archaea (AOB, AOA)(Stahl and de la Torre, 2012). A second distinct group carries out the oxidation of nitrite to nitrate. Importantly this process can be a significant source of N\textsubscript{2}O, producing it as a byproduct (Khalil, Mary, et al., 2004).

1.3.3 Denitrification:

Denitrification consists of four modular reduction steps taking nitrate (NO\textsubscript{3}\textsuperscript{-}) through nitrite (NO\textsubscript{2}\textsuperscript{-}), to nitric oxide (NO), to the gaseous product nitrous oxide (N\textsubscript{2}O) which is a potent greenhouse gas. A final step converts N\textsubscript{2}O into the inert gas N\textsubscript{2}. Due to the modular nature of denitrification the final step is not always carried out meaning that N\textsubscript{2}O can be the final product of the process. Indeed, denitrification is considered to be a major N\textsubscript{2}O producing process because it produces N\textsubscript{2}O as an obligate intermediate rather than as a by product such as in nitrification (Canfield, Glazer, et al., 2010). It is also considered to be a major nitrogen sink in coastal ecosystems (Giblin, Tobias, et al., 2013) though some argue that its role as a sink has been overemphasized (Burgin and Hamilton, 2007). Denitrification is mainly a facultative anaerobic process, occurring only when oxygen is unavailable as an electron acceptor. However it can persist with a rising presence of O\textsubscript{2} as has been noted in a number of novel bacterial strains (Miyahara, Kim, et al., 2010; Takaya, Catalan-Sakairi, et al., 2003).
1.3.4 Dissimilatory Nitrate Reduction to Ammonia:

Dissimilatory nitrate reduction to ammonia (DNRA) produces ammonia from nitrate. Thus it competes with denitrification, moving nitrogen back into the cycle rather than acting as a sink. It has been suggested that DNRA may have a significant role in the fate of nitrogen in shallow sediments, perhaps being more significant than denitrification (Giblin, Tobias, et al., 2013).

1.3.5 Anammox

Ammonium (NH$_4^+$) and nitrite (NO$_2^-$) are converted directly into N$_2$ by the microbiologically mediated process of anammox (Anaerobic ammonia oxidation) (Strous, Fuerst, et al., 1999). Thus like denitrification, this process acts as a sink for nitrogen. For many years scientists predicted that such a pathway was present but it wasn’t until 1999 that Strous et al. were able to identify the responsible microbes (Strous, Fuerst, et al., 1999). This discovery was applied to wastewater treatment but it also significantly complicated our view of the global nitrogen cycle when in 2002, it was revealed to be responsible for a huge portion of nitrogen removal from the ocean (Thamdrup and Dalsgaard, 2002). That anammox was significant in the ocean was sure but debate still reigned over the relative roles of denitrification and anammox. A recent study shows that variation in the observed anammox vs denitrification rate in the ocean is likely due to differences in organic matter concentration at sampled sites (Babbin, Keil, et al., 2014).

1.4 Determinants of N$_2$O emission

The nitrogen cycle is made up of a number of reactions with a varying ability to introduce or remove nitrogen in various forms. Anthropogenic nitrogen has a
number of potential paths to follow. Of interest are the determinants of this flow path, especially with respect to the production of N\textsubscript{2}O.

The ever increasing level of N\textsubscript{2}O emissions is an important driver for ecological research. Measuring N\textsubscript{2}O emissions is not enough. To make a potential impact researchers have looked into the determinants of N\textsubscript{2}O emissions in many environments.

The availability of substrates for N\textsubscript{2}O producing processes is a clear determining factor for its production. It was shown in a meta-analysis by Pin\textadChoice{a}-Ochoa and A\textadChoice{lvarez-Cobelas}} (2006) that dissolved oxygen and nitrate levels were significantly linked to denitrification rates with 70 percent of the variability in denitrification between aquatic environments being determined by nitrate levels. Although there is a definite link between the amount of nitrogen entering a system and N\textsubscript{2}O production, it is important to remember not all denitrification produces N\textsubscript{2}O as the end product (Jones, Stres, et al., 2008; Sanford, Wagner, et al., 2012), therefore the denitrifying microbes present and their denitrifying potential must be considered.

Our ability to identify denitrifiers in the environment may still be limited by our understanding of the diversity of denitrifying genes and the microbes carrying them, as new denitrification gene variants continue to be discovered along with new organisms containing them (Green, Prakash, et al., 2010; Sonia Henry, Baudoin, et al., 2004; Jones, Graf, et al., 2013; Jones, Welsh, et al., 2011). The nitrous oxide reductase enzyme encoded by the nosZ gene is of particular interest because it catalyses the reduction of N\textsubscript{2}O to N\textsubscript{2}, hence its activity defines the emission of the greenhouse gas (N\textsubscript{2}O) versus the harmless nitrogen sink (N\textsubscript{2}). In an important recent discovery it was shown that we may
have been grossly underestimating the N₂O reducing community with the discovery of a new separate clade of N₂O reducing genes which haven’t yet been targeted in studies (Jones, Graf, et al., 2013). Furthermore, the new clade was often found to be present in the absence of other denitrification genes (nirK, nirS), providing evidence that non-denitrifying N₂O reducers may be important in mitigating N₂O emissions (Sanford, Wagner, et al., 2012).

Conditions within an ecosystem such as O₂ concentration, pH and carbon availability are also key determinants of N₂O production. Consider the effect of O₂; it has been demonstrated that denitrification can occur in the presence of O₂, but that under increasing aerobic conditions the amount of N₂O relative to N₂ produced is higher (Rassamee, Sattayatewa, et al., 2011). This has been attributed to the sensitivity of nitrous oxide reductase (the enzyme catalyzing the conversion of N₂O to N₂) to O₂ (Bonin, Gilewicz, et al., 1989). This represents a significant problem for waste water treatment where aerobic and anaerobic conditions must be applied to promote nitrification and subsequently denitrification. Removing all the dissolved oxygen from waste is difficult, so production of N₂O via denitrification may be promoted. Novel aerobic denitrifiers may provide a solution to this problem (Miyahara, Kim, et al., 2010; Takaya, Catalan-Sakairi, et al., 2003).

Linking determinants of N₂O emission with biogeochemistry and community changes is important for understanding the mechanism by which N₂O is produced. This is a difficult process for in situ studies. Simplified, controlled environments called micro and mesocosms allow for better linking of a variable to an outcome. Predictions can later be tested and applied to more complex environments. For example Harter et al., (2014) were able to link the
application of biochar to soils directly with changes in the proportion of nitrate and nitrous oxide reducers and a reduction in N$_2$O emission. When we do wish to examine the role of community changes in situ, 16S rRNA gene sequencing becomes the tool of choice.

1.5 16S analysis

16S analysis provides a powerful tool to assess the microbial diversity within an environment. It allows assessment of the community composition, richness (number of distinct species) and evenness (relative abundance of distinct species). Phylogeny can be constructed directly from the comparison of 16S sequences (Clementino, Vieira, et al., 2008) or from fingerprints constructed from differential 16S restriction patterns and separation by electrophoresis (Galand, Bourrain, et al., 2012).

1.5.1 Role of observed diversity and community composition

Microbial diversity and community composition are thought to have an important role in ecosystem function but the exact impact of diversity and composition in different systems is still unclear (Danovaro and Pusceddu, 2007). Though there are many studies that examine ecological diversity/composition using a variety of techniques, few seem to be able to comment on the significance of the observed diversity/composition (Cavigelli and Robertson, 2001; Laurent Philippot, Spor, et al., 2013). Cavigelli and Robertson, (2001) make a case for the effects of diversity and composition on N$_2$O reduction; they propose that due to the observed differences in reduction activity between different taxa (in response to different O$_2$ levels or other variables), species composition could have a significant effect on an
environment’s ability to reduce N\textsubscript{2}O. Unfortunately they are unable to clearly link observed reductive abilities between taxa and community level reduction. This could have been due to biases introduced by the current inability to culture all the denitrifiers from a community. More recently, Philippot et al., (2013) were able to confirm a link between diversity and function of the denitrifying community. They found that even a small reduction in the biodiversity of a community (25% reduction) could have a profound effect on decreasing the level of denitrification. Furthermore, they showed that the effect was amplified upon addition of wheat residues to microcosms. This is an important finding as it suggests that redundancy (ability of multiple species to carry out the same essential function) is not as powerful as was thought in a community, especially for a phylogenetically diverse group such as denitrifiers and that environmental factors could significantly amplify the effects of diversity loss. The implication for an environment acting as a nitrogen sink is that reductions in biodiversity could lead to a reduction in their ability to carry out this ecosystem service.

1.5.2 Applying 16S to the study of nitrogen cycling

It is difficult to directly apply 16S based methods for the study of denitrifying bacterial diversity and function because it is such a phylogenetically diverse group (Cheneby, Philippot, et al., 2000; L Philippot, 2005; Santoro, Boehm, et al., 2006). Hence function does not group with phylogeny in this case and culture based methods need to be used in conjunction with 16S analysis to confirm function. Analysis of functional denitrification genes is more useful for direct study of these groups, while 16S analyses provide a view of overall community structure, which may be important for the functioning of specific communities.
1.6 Functional gene abundance

An understanding of the enzymes and the genes encoding denitrification has allowed analysis of its evolutionary basis and the quantification of denitrifiers in the environment (Sonia Henry, Baudoin, et al., 2004; López-Gutiérrez, Henry, et al., 2004; Petersen, Blazewicz, et al., 2012; Smith, Nedwell, et al., 2007). Similarly, the quantification of other nitrogen cycling genes allows the identification of other functional groups of microbes. Some studies suggest community assemblages may be based more upon functional genes rather than species identity, implying that analysis of these functional genes is of even more importance than pinpointing taxonomic identity (Burke, Steinberg, et al., 2011).

Quantification of gene abundance and hence microbial abundance relies on quantitative PCR (qPCR). This technology became cheaper in the early 2000s and since then has seen widespread use (Ginzinger, 2002). In denitrification studies, qPCR is used to amplify genes encoding a number of metallo-enzymes; nitrate reductases, nitrite reductases, nitric oxide reductases and nitrous oxide reductases (Philippot, 2002). Multiple enzymes may catalyze the same denitrification step. For instance two distinct nitrite reduction genes encode functionally similar but structurally different nitrite reductase proteins (Silvestrini, Galeotti, et al., 1989; Ye, Fries, et al., 1993). Therefore when analyzing denitrification via functional gene analysis, all functionally redundant genes catalyzing a step must be targeted to capture the full community (Sonia Henry, Baudoin, et al., 2004). As illustrated before in section 1.4, if some of the genes encoding denitrification remain unknown and unclassified this type of analysis becomes less useful (Jones, Graf, et al., 2013).
One important consideration is that denitrification genes only give a measure of
denitrification potential because denitrification genes are usually facultative.
Analysis of mRNA can fill this gap by showing us what is actually occurring in
an environment. Though this technique has been limited by our ability to extract
intact RNA from environmental samples. This highlights the importance of
studies focusing on the further development of DNA/RNA extraction methods
for environmental samples (Griffiths, Whiteley, et al., 2000; Paulin, Nicolaisen,
et al., 2013a; Saleh-Lakha, Shannon, et al., 2011). Co-extraction protocols are
even more desirable as they allow for the comparison calculation of a
transcription rate.

1.7 Coastal lagoon environments

Coastal lagoons are shallow bodies of water with limited attachment to the
open ocean. These are highly variable environments with inter and intra lagoon
variability being determined by a range of natural and human factors. Indeed
the level of attachment to the ocean itself is a major defining factor for the
lagoon environment (Duck and da Silva, 2012). Coastal lagoons are highly
productive and ideally situated for settlement. Because of this they have been
exploited by humans as harbours, food sources and dumping grounds.

Much of the research on these environments has focused on their diverse
macrofauna (Anger, Spivak, et al., 1994; Carvalho, Pereira, et al., 2011;
Pereira, Carvalho, et al., 2012; Sfriso and Marcomini, 1997; Taylor, Nixon, et
al., 1995; Tyler, McGlathery, et al., 2001). Considering the threat of increasing
anthropogenic nitrogen levels, the susceptibility of these environments to
eutrophication, their prevalence (coastal lagoons cover 13% of the coastline
worldwide (Pérez-Ruzafa and Marcos, 2012) and their usefulness, there needs to be a focus on how these systems react to increased nitrogen load. With the large sediment surface area relative to water volume in shallow coastal lagoons and the potential of the microbiota to act as a nitrogen sink, emphasis needs to be placed on the benthic microbial community.

Current studies have shown that this microbial population is essential for the recycling of nutrients to sustain primary production and the movement of detritus to higher trophic levels (Danovaro and Pusceddu, 2007; Manini, Fiordelmondo, et al., 2003). Benthic microbiota are able to carry out denitrification and may allow these environments to act as a sink for nitrogen.

Linking micro and macrofauna is also of importance because macrofauna can have an impact on microbial ability to carry out denitrification. For instance the bioturbation effects by chironomids and oligochaetes are positively correlated with denitrification in sediments because of their ability to mobilise nitrate, making it available for denitrifiers (Svensson and Leonardson, 1996; Svensson, Enrich-Prast, et al., 2001).

Studies using direct amplification and/or fingerprinting of 16S samples from lagoon environments have been carried out and comparisons between coastal, eutrophic and salt/fresh water lagoons have been made, however these analyses tend to focus on the water column rather than sediment (Benlloch, Rodríguez-Valera, et al., 1995; Clementino, Vieira, et al., 2008; Ghai, Hernandez, et al., 2012). Water column studies have demonstrated diverse microbiota occupying a range of environments. Benlloch et al., (1995) showed that the eutrophic Prevost lagoon had a more diverse microbial population than
Arcachon lagoon on the French Atlantic coast. This is surprising considering that diversity is generally associated with more healthy environments.

A more recent study carried out on large coastal Mediterranean lagoons demonstrated that anthropogenic inputs can have a powerful effect on microbial community composition and diversity. The study used 16S and metagenomic analyses to compare the microbial populations between two lagoons and microbiota from similar environments. Mar Menor, a salty lagoon that has some contact with the Mediterranean showed similar microbial composition to other similarly saline aquatic environments with high levels of Alphaproteobacteria dominating. Interestingly it was observed that at higher taxonomic levels community composition was generally maintained between Mar Menor and similarly saline environments. It was only at lower taxonomic levels that major differences were observed. Lake Albufera, a hypertrophic freshwater lagoon did not share similar microbial composition to non-eutrophic freshwater environments, containing as expected in a eutrophic system, increased levels of cyanobacteria. More surprising were the reduced levels of Actinobacteria and common freshwater Betaproteobacteria, which owing to their small size and large surface area were suspected to have lost their competitive edge in the high nutrient environment (Ghai, Hernandez, et al., 2012).

Studies of lagoon sediments have also demonstrated diverse microbial communities and again linked anthropogenic inputs, microbial community composition and diversity (Borin, Brusetti, et al., 2009; Tsuboi, Amemiya, et al., 2013). Even within the single albeit large environment of the Venice lagoon, microbial community differences and a reduction in diversity were observed
between sites of heavy anthropogenic input and less impacted sites (Borin, Brusetti, et al., 2009).

The significance of benthic microbial diversity loss specific to coastal lagoons has only been tentatively explored. Danovaro and Pusceddu (2007) make the broad conclusion that more diverse lagoons have higher functional performance but the specific consequences for function of this diversity loss need further attention, especially with respect to denitrification. Linking 16S sequencing to functional gene analyses could reveal some insight in this area.

A google scholar search for the keywords quantitative PCR, coastal lagoon, nitrogen cycle produces 5370 results, however no articles directly pertaining to examining nitrogen cycling in coastal lagoons by function gene qPCR were found, suggesting such studies are non existisitant or very rare. In any case this powerful technology needs to be applied to these environments. Quantification of functional genes has been very useful for understanding other aquatic environments. For instance it has shown the importance of archaea in nitrification (Mincer, Church, et al., 2007) and demonstrated how nitrogen cycling processes change across the nitrate gradient of an estuary (Dong, Smith, et al., 2009a). In the latter study biogeochemical rates were related to gene abundances. Smith and Osborn suggest that such studies are important because they allow the linking of genetic potential of the community to environmental factors that control a process (Smith and Osborn, 2009).

1.8 Lake Ellesmere

Lake Ellesmere (Te Waihora) is a large shallow brackish water lagoon, situated on the east coast of New Zealand, just south of Banks Peninsula (Fig. 1-see
section 2.1.1). Human intervention has caused a dramatic change in the state of this lake. The control of opening and closing of the lagoon to the sea allowed much of the lake to be reclaimed as farmland. Similarly much of the surrounding Canterbury plains have been developed into dairy pastures. Leaching of nitrogenous compounds from these farms directly into the lake or through rivers and groundwaters has resulted in severe eutrophication, undermining the value of this lake (Hamil and Schallenberg, 2013).

Like other lagoons, much of the study of this system has focused on its diverse mega or macrofauna (D. Jellyman, 2011; D. J. Jellyman, 2001; Mitchell, Hamilton, et al., 1988). Recent studies and reports have examined the hypereutrophic status of this lake, focusing on lake chemistry and water quality management (Hamil and Schallenberg, 2013; Hughey and Taylor, 2009; Schallenberg, Larned, et al., 2010). The results of such studies demonstrate that nutrient loading from freshwater tributaries into the lake is large and highly variable over years and seasons. Total nitrogen and phosphorous levels entering the lake peak in winter and diminish during the summer while wind driven sediment resuspension may drive considerable nutrient input from the sediments into the lake waters during summer (Hamil and Schallenberg, 2013). Lake opening and closing appears to only have minor effects on lake water quality (M. Schallenberg, Larned, et al., 2010).

1.9 Aims

This study was carried out in response to the observation that mean total nitrogen (TN) levels are approximately two times greater in Lake Ellesmere’s tributaries than in the lake (Hamil and Schallenberg, 2013). We posed the
question: what is the fate of nitrogen entering this system? It was suspected, given the ecologically widespread role of denitrification as a nitrogen sink and high levels of inflowing dissolved inorganic nitrogen (DIN) in the form of nitrate, that this process might be responsible for maintaining the comparatively lower lake nitrogen levels. Moreover if this process was active it might be carrying out a considerable ecosystem service or disservice based on the identity of the end product: inert N$_2$ gas or N$_2$O, a powerful greenhouse gas.

To address this hypothesis we aimed to determine

1) The structure and distribution of benthic microbial denitrification and nitrogen fixing communities within lake Ellesmere sediments using qPCR of three denitrification genes (nirS, nosZI and nosZII) and one nitrogen fixation gene (nifH).

2) The overall microbial community structure using 16S rRNA gene sequencing.

Comparison of these measures to the lake’s chemical/physical and denitrification activity analyses would identify the state and drivers of denitrification within Lake Ellesmere. It was hypothesized that as agriculturally derived nitrogen flowed into the lake it would produce a concentration gradient that would determine the state of overall and denitrifying communities within the lake sediments.
2 Methods

2.1.1 Sample collection

Sediment and bottom water samples were collected on the 9th, 10th and 11th of April, 2014 from 18 sites (Fig 1.) throughout Lake Ellesmere along with concomitant measurements of sediment redox potential, lake depth, water clarity, bottom water temperature, top water temperature, top water salinity and bottom water salinity.

Figure 1: Sampling sites and sample collection order within Lake Ellesmere. (Top panel) sampling locations, (Bottom panel) sampling order over day one (black), two (blue) and three (red).
2.1.2 Sediment sampling

Sediment samples (four replicate sediment samples per site [total of 72 samples]) were collected using a weighted perspex 75mm diameter core tube. Intact sediment cores were retrieved and the first 4cm of the sediment column were extruded, collected in a plastic container and homogenized by mixing with a metal spoon. A 2mL subsample of homogenized sediment was transferred to a 2mL microfuge tube. All replicates were processed and stored separately without compositing. Sediment redox was measured from the last homogenised sediment core using a Schott millivolt meter with a redox probe. Measurements were recorded once the voltage reading stabilized. Sediment samples were stored on dry ice in the field and at -80°C in the lab until processed for microbial community nucleic acid extraction.

Additional sediment sampling for analyses of texture, organic matter and denitrification enzyme assays (DEA) were carried out by collecting three 4cm cores from each site and pooling.

2.1.3 Water sampling

Water clarity was measured at each site using a Secchi disk. To measure lake depth, the Secchi disk was lowered to the lake floor. Water samples were collected using a messenger activated van Dorn water sampler submerged to just above the lake floor at each lake site. One 50mL water sample was collected from the sampler into a 50mL falcon tube for denitrification enzyme assay (DEA) analyses, and two additional 50mL samples were collected for water nutrient analysis.
2.2 DNA extraction

Total community DNA was extracted from all 72 samples following the methods of Paulin et al., (2013). One gram of 0.5mm silica beads, 0.75g of 0.1mm silica beads, 0.5g of sediment, 50µL of 10mg/mL salmon sperm DNA (Invitrogen, Carlsbad, CA, USA) and 500µL each of cetyltrimethyl ammonium bromide (CTAB) buffer (1 volume of 10% (w/v) CTAB in 0.7 M NaCl mixed with 1 volume of 240 mM potassium phosphate buffer, pH 8.0) and phenol chloroform isoamyl alcohol 25:24:1 (PCI, ACROS Organics, Geel, Belgium) was added to a 2ml sterile O-ring tube. Two 15 second intervals of bead beating at 1750rpm were carried out to lyse microbial cells within the sediment. The resulting slurry was centrifuged at 10,000 rpm for 10 minutes to separate out nucleic acid containing aqueous phase from the phenol phase. The supernatant was recovered (between 400-750µL) into a 1.5mL Eppendorf tube. An equal volume of PCI was added and then the tube was centrifuged again at 13,000rpm for 10 minutes. The resulting supernatant was recovered into 1.5mL Eppondorfs (no more than 400µL per tube) and twice the volume of 20% polyethylene glycol 6000 (PEG) was added to precipitate out the nucleic acids for 2 hours at 4ºC. Following incubation the precipitate was pelleted by centrifugation for 10 minutes at 13,000rpm. The supernatant was discarded and replaced with 70% ice-cold ethanol to wash the pellet. After a further 10 minute centrifugation at 13,000rpm, the ethanol was removed by pipette, nucleic acid pellet left to air dry and then resuspended in 50µL of sterile water. Extractions were then incubated at 37ºC with 4µL of 20Unit/µL RNase I (Ambion, Austin, Texas, USA) for 30 minutes in a 50µL solution of 0.2M NaCl to degrade co-extracted RNA. Samples were stored at -20ºC.
DNA was quantified and assessed for purity using a Nanodrop 1000 (ThermoScientific, Wilmington DE, USA). All DNA concentrations were normalized to 5ng/µL and 10ng/µL for qPCR and 16S rRNA gene sequencing respectively.

2.3 Quantitative PCR (qPCR)

Quantitative PCR (qPCR) was performed on all 72 samples (18 sites x 3 biological replicates) separately. Triplicate technical reactions (minimum) were performed for each sample using modified conditions detailed below. All reactions contained 2µL 5ng/µL DNA, 5µL of fast SYBR Green Master Mix (Applied Biosystems, Foster City, California, USA), 0.5µL of 10µM relevant forward primer, 0.5µL of 10µM relevant reverse primer and 2µL ddH2O. pGEM-T easy (Promega, Madison, Wisconsin, USA) cloned template standards were included in every run to allow absolute quantification of templates. Reaction wells were prepared using an automated VERSA liquid handling robot (Aurora, Vancouver, B.C., Canada). The robot added 8µL of hand prepared qPCR mix to each well followed by 2µL of 5ng/µL DNA. Amplification was performed in 384 well plates (Applied Biosystems) on the ViiA 7 real time qPCR machine (Applied Biosystems) using the following thermal cycling protocol. The plate was held at 95°C for 10 minutess followed by 40 amplification cycles of 96°C for 3 seconds, 58.5°C for 3 seconds and 68°C for 30 seconds. To ensure complete strand elongation the plate was held for one 5 minute cycle at 72°C. A melt curve was then produced holding the plate at 96°C for 15 seconds, 50°C for 1 minute and then ramping up to 96°C while data was collected. Specific genes and their respective transcripts targeted in this study were: nifH,
nitrogenase gene; nosZ, nitrous oxide reductase gene; nirS, nitrite reductase gene

2.3.1 nifH

*nifH* qPCR was carried out according to the parameters described above and using:

1) 10µM *nifH* forward and reverse primers (Rösch and Bothe, 2005; Yergeau, Kang, et al., 2007).

2) Cloned *nifH* standard templates from copies 1x10⁵ to 1x10⁰.

2.3.2 nirS

*nirS* qPCR amplification was carried out according to the parameters described above and using:

1) 10µM *nirS* forward and reverse primers (Throbäck, Enwall, et al., 2004; Yergeau, Kang, et al., 2007).

2) Cloned *nirS* standard templates from copies 1x10⁵ to 1x10⁰

2.3.3 nosZI

*nosZI* qPCR amplification was carried out according to parameters described above and using:

1) 10µM *nosZI* forward and reverse primers were used (S Henry, Bru, et al., 2006).

2) Cloned *nosZI* standard templates from copies 1x10⁶ to 1x10¹
2.3.4 *nosZII*

*nosZII* qPCR amplification was carried out according to parameters described above but with the following alterations:

1) 20µM *nosZII* forward and revers primers were used (Jones, Graf, et al., 2013a).

2) Cloned *nosZII* standard templates from copies 1x10⁶ to 1x10¹.

3) Luminaris Low ROX Colour HiGreen master mix (Thermo Scientific) was used in place of SYBR Green Master Mix.

4) Amplification was run using the following cycling procedure, modified from Jones et al., (2013):

qPCR was initialized at 50°C for 2 minutes and then 95°C for 10 minutes followed by one amplification cycle of 95°C for 15 seconds, primer annealing at 60°C for 30 seconds and elongation at 72°C. The amplification step was repeated a further 5 times but the annealing temperature was reduced by 1°C for each subsequent cycle. Forty-four cycles of amplification with annealing at 54°C were carried out. Finally, a melt curve was generated by holding the reaction at 95°C for 15 seconds, 50°C for 1 minute and then ramping up to 95°C while data was collected.

2.4 16S rRNA gene amplicon sequencing

10µL aliquots of 10ng/µL for all 72 Ellesmere replicates were placed in a 96 well qPCR plate and desiccated using a Vacuum desiccator jar and desiccation beads (Bel-Art, Wayne, NJ, USA). The plate was sealed using a qPCR compatible optical cover for mailing. 16S rRNA gene amplification and
amplicon sequencing on the Illumina MiSeq platform were performed at the Department of Energy Argonne National Laboratories (USA) following the Earth Microbiome Project standard protocol (Caporaso, Lauber, et al., 2012)

2.5 16S analysis

16S sequence data was analysed using the Quantitative Insights Into Molecular Ecology (QIIME) 1.7.0 workflow (Caporaso, Kuczynski, et al., 2010). Data was first demultiplexed. Sequences were classified to an operational taxonomic unit (OTU) using an open reference strategy based set of reference sequences from the greengenes database. The sequence pool was then subsampled to a depth of 17400 sequences to eliminate biases in the depth of sampling. Weighted and unweighted Unifrac distances were calculated and analysed by principle coordinate analysis. Uncertainty in these principle component plots was measured by repeated resampling (jackknifing). OTUs were classified to 6 taxa levels (kingdom, phylum, class, order, family, genus) and plots produced to describe relative abundance of taxa by site. An alpha diversity analysis and rarefaction was carried out to quantify microbial richness within the 72 lake samples and assess the effect of our chosen sampling depth (17400 sequences) on the observed richness. Shannon diversity was also calculated for the 72 lake samples.

2.6 Denitrification enzyme assay

Sediment denitrification enzyme activity was measured using the acetylene block assay (Bruesewitz, Hamilton, et al., 2011) on sediment and water samples. Analysis was carried out by David Hamilton at the Environmental Research Institute, University of Waikato, Hamilton. Briefly, four different
treatments (control, 10mg/L nitrate, 12mg/L glucose, additional nitrate and glucose) were applied to homogenized lake water (15mL) and sediments (15mL) from the 18 Ellesmere sites in the presence of acetylene, to block the conversion of N₂O to N₂. 8mL emissions samples were taken every hour over 6 hours and measured using a Varian CP 3800 gas chromatograph with an ECD detector.

2.7 Water and sediment nutrient analysis

Water nutrient, sediment texture and organic matter analysis was carried out by Josie Crawshaw (MSc student, University of Otago, Department of Zoology). Briefly, dissolved nutrients (NO₂⁻/NO₃⁻, DRP, NH₃) were measured in 25mL filtered water samples using colorimetry as described in Schallenberg and Burns, (2004). Total nutrients (TN, TP) were measured from an unfiltered lake water sample using the same method. Sediment organic matter content was measured as the weight of dry sediment burnt off at 450°C. Sediment porosity was calculated as the percentage of total weight contributed by wet weight of the sediment samples. Sediment percentage attributable to different grain sizes; sand (63-2000 µm), silt (2-63 µm), clay (0-2 µm) was measured using a Mastersizer 2000 laser diffraction particle size analyzer (Malvern Instruments, Malvern, Worcestershire, UK) after sediments were treated with hydrogen peroxide to remove organic matter.

2.8 Statistical and multivariate analysis

Analyses were performed using JMP 11.2 (SAS institute. Cary, NC, United States of America). Four types of analysis were performed to explore relationships between measured variables. Principal component analysis was
used to screen for potential correlations within the lake data set based on chemical and physical factors, relative microbial abundances at the genus and phyla levels and microbial community measures such as richness, shannon diversity and nitrogen cycling gene copy number. Similarly two-way cluster analyses were used on phyla level relative abundance data to identify community drivers. Correlations identified in these types of analysis were confirmed by non-parametric analyses (Spearman's correlations) and linear regressions.

Extrapolation of lake parameters onto map bases was carried out in Surfer 11 (Golden software, Colorado, USA) using the Kriging gridding method.
3 Results

3.1 Physical and Chemical gradients

Chemical and physical gradients within Lake Ellesmere sediment and water column were analysed over 18 sites (Fig. 2). Data provided by Josie Crawshaw, Department of Zoology, University of Otago.

3.1.1 Water column and sediment chemistry

With the exception of dissolved reactive phosphorus (DRP), ammonia (NH₃) and Redox, water column and sediment nutrient concentrations were homogenous with large changes only observed within one or two sites (Fig. 2). In particular, total nitrogen (TN), total phosphorous (TP) and sediment organic matter levels ranged from 2,016 to 2,846µg/L, 198 to 256µg/L and 0.7 to 6.5% with outlier sites at 1175µg/L, 98µg/L and 21.7% respectively (Fig. 2-3). Lake salinity (surface and bottom) was high (mean, 9.9ppt) and only a weak gradient was observed. Lower salinity was observed at the northern end of the lake (where most of the lake’s freshwater tributaries are located) with higher levels at the southern end of the lake (where the controlled opening to the ocean is located) (Fig. 2F-N). Site E14 showed exceptionally low salinity levels (Fig. 2F). All redox measurements made were negative (Fig. 3B, D).

3.1.2 Physical factors

Physical factors (depth, sediment texture) formed gradients from the centre of the lake to the perimeter with silt, clay levels and depth values highest at the lake centre. Sand levels following the opposite trend (Fig. 4).
Figure 2: Water column chemistry per site (left column) and across (right column) Lake Ellesmere. Water samples from 18 (E01 to E18) sites around Lake Ellesmere were analysed for total nitrogen (TN) (A,H), total phosphorous (TP) (B,I), dissolved reactive phosphorous (DRP) (C,J), total combined nitrite and nitrate (NO$_2^-$/NO$_3^-$) (D,K), ammonia (NH$_3$) (E,L) using colorimetry as described in section 2.7. On site conductivity measurements were made on top and bottom water samples to estimate top (F,M) and bottom (G,H) salinity. Map figures H to N represent 18 data points extrapolated by Kriging gridding in Surfer 11 (Golden software, Colorado, USA).
**Figure 3**: Sediment chemistry per site (left column) and across (right column) Lake Ellesmere. Organic matter (as percentage of total dry sediment) (A,C) and redox (B,D) was measured at 18 sites around Lake Ellesmere. Organic matter was measured using methods detailed in section 2.7. Redox was measured on site from a pooled mixture of four 4cm sediment cores using an oxidation reduction potential (ORP) meter. Map figures C and D represent 18 data points extrapolated by Kriging gridding.
Figure 4: Physical gradients per site (left column) and across (right column) Lake Ellesmere. Percent clay (A,H), silt (B,I) and sand (D,K) were measured in sediment samples from 18 lake sites by the methods described in section 2.7. Lake depth was (D,K) measured on site using a secci disk. Map figures H to K represent 18 data points extrapolated by Kriging griding.
3.1.3 Correlating physical and chemical variables

To determine the influence of Selwyn river inflow on measured lake parameters, distance from the Selwyn river mouth (E2) to all sites was determined and correlated to chemical parameters. TN, DRP and TP levels were lower at more distant site decreasing 1.4, 2.3 and 1.2 fold form E2 to the most distant site (Fig. 5).

Principal component analysis (PCA) was used to search for further relationships between physical and chemical variables within the lake (Fig. 6). Loadings demonstrated that physical characteristics tended to associate with the first principal component, explaining 35.8% of the variability between the sites while chemical changes were mainly associated with the second component, accounting for 18.6% of the variability (Fig. 6). This segregation between components suggests that the water column nutrients TN, TP and water clarity (secchi depth) are unrelated to the major physical parameters (sediment texture and depth). Other measures of water (DRP, NO$_2^-$/NO$_3^-$, NH$_3$, Salinity) and sediment chemistry (redox, organic matter content) however did show potential relationships to the physical component.

Non-parametric analysis and linear regression was used to further explore the relationships suggested by the PCA. It was found that sand percentage correlated with a host of other lake variables but most strongly with silt, clay, depth, bottom salinity, porosity and organic matter (Table 1). Total combined nitrite/nitrate (NO$_2^-$/NO$_3^-$) levels were strongly correlated to lake salinity but this trend was not supported by linear regressions (Fig. 7B). TP and TN were not strongly correlated, but a linear regression demonstrated a strong association between these chemical factors (Fig. 7A).
Figure 5: Water nutrient concentrations in relation to Selwyn River inflow. Linear regression of DRP, TP and TN \( \mu g/L \) by distance from lake site E2 at the mouth of the Selwyn River. Data points represent comparisons between one off measurements from each of 17 lake sites. Site E17 was excluded from these analyses as an outlier. Shaded bars (blue, green, red) indicate confidence of fit.

Figure 6: Principal component analysis (PCA) of physical and chemical parameters within Lake Ellesmere. (Left panel) Clustering of sites based on lake water chemistry (TN, TP, NH\(_3\), NO\(_2^-\)/NO\(_3^-\), DRP, bottom salinity), sediment chemistry (organic matter content, redox) and physical factors (depth, sand%, silt%, clay%, secci depth, porosity, bottom temperature, distance from site E2, distance from site E16). (Right panel) Factor loadings for inputted variables. Axis percentage values indicate the % of variability explained by that axis.
Table 1: Non-parametric Spearman's correlation for lake chemical and physical factors.

| Variable 1       | Variable 2               | Spearman ρ* | Prob>|ρ||** |
|------------------|--------------------------|-------------|-----|-----|
| Sand %           | Silt %                   | -0.9794     | <.0001 |     |
| Sand %           | Clay %                   | -0.9463     | <.0001 |     |
| Sand %           | Depth (m)                | -0.8762     | <.0001 |     |
| Sand %           | porosity                 | -0.8636     | <.0001 |     |
| Sand %           | Organic Matter %         | -0.8221     | <.0001 |     |
| NO2-/NO3- ug/L   | Surface Salinity (ppt)   | -0.7548     | <.0001 |     |
| NO2-/NO3- ug/L   | Surface Conductivity (m/s) | -0.7433   | <.0001 |     |
| Sand %           | Bottom Salinity (ppt)    | -0.7069     | <.0001 |     |
| Sand %           | Surface Salinity (ppt)   | -0.697      | <.0001 |     |
| Sand %           | Bottom Conductivity (m/s) | -0.697     | <.0001 |     |
| Distance from E2 (km) | DRP ug/L               | -0.6904     | <.0001 |     |
| Distance from E2 (km) | TN ug/L                | -0.6883     | <.0001 |     |
| Sand %           | Surface Conductivity (m/s) | -0.6729   | <.0001 |     |
| Depth (m)        | Surface Conductivity (m/s) | 0.6139     | <.0001 |     |
| Clay %           | Surface Conductivity (m/s) | 0.6149     | <.0001 |     |
| Clay %           | Bottom Conductivity (m/s) | 0.6256     | <.0001 |     |
| Depth (m)        | Surface Salinity (ppt)   | 0.636       | <.0001 |     |
| Clay %           | Surface Salinity (ppt)   | 0.6371      | <.0001 |     |
| Clay %           | Bottom Salinity (ppt)    | 0.6677      | <.0001 |     |
| Silt %           | Surface Conductivity (m/s) | 0.677      | <.0001 |     |
| Silt %           | Surface Salinity (ppt)   | 0.698       | <.0001 |     |
| Depth (m)        | Bottom Conductivity (m/s) | 0.7167     | <.0001 |     |
| Silt %           | Bottom Conductivity (m/s) | 0.7332     | <.0001 |     |
| Silt %           | Bottom Salinity (ppt)    | 0.7379      | <.0001 |     |
| Organic Matter % | Depth (m)                | 0.7477      | <.0001 |     |
| Bottom Salinity (ppt) | Surface Conductivity (m/s) | 0.7536     | <.0001 |     |
| Bottom Conductivity (m/s) | Surface Conductivity (m/s) | 0.7588     | <.0001 |     |
| Bottom Salinity (ppt) | Surface Salinity (ppt)   | 0.7599      | <.0001 |     |
| Bottom Conductivity (m/s) | Surface Salinity (ppt)   | 0.7631      | <.0001 |     |
| Silt %           | Organic Matter %         | 0.7642      | <.0001 |     |
| Depth (m)        | Bottom Salinity (ppt)    | 0.7709      | <.0001 |     |
| porosity         | Depth (m)                | 0.8099      | <.0001 |     |
| porosity         | Silt %                   | 0.8099      | <.0001 |     |
| Silt %           | Depth (m)                | 0.8184      | <.0001 |     |
| Silt %           | Clay %                   | 0.8865      | <.0001 |     |
| Clay %           | Depth (m)                | 0.8885      | <.0001 |     |
| Bottom Temp (°C) | Surface Temp (°C)        | 0.9041      | <.0001 |     |
| Clay %           | Organic Matter %         | 0.9276      | <.0001 |     |
| porosity         | Clay %                   | 0.9514      | <.0001 |     |
| Bottom Salinity (ppt) | Bottom Conductivity (m/s) | 0.9762     | <.0001 |     |
| porosity         | Organic Matter %         | 0.9803      | <.0001 |     |
| Surface Salinity (ppt) | Surface Conductivity (m/s) | 0.9974      | <.0001 |     |

* Spearman's p cut-off at <0.6
** significance cut-off at p>0.05
3.2 Microbial community structure

A total of 17400 sequences were analysed per site. Analyses of the microbial community were separated into community composition, alpha and beta diversity.

3.2.1 Community composition

Microbial community structure saw significant variation within and between lake sites but had a consistent overall composition within the lake. Proteobacterial and unidentified bacterial species dominated the microbial community, making up on average 34 and 29% of all microbial species within a site (Fig. 8). Cyanobacteria/chloroplast, Bacteroidetes, Actinobacteria and Acidobacteria also made up large proportions of the community contributing on average 13, 8, 5 and 3% respectively (Fig. 8). One major observation was the enrichment (5% of all sequences in that site) of sequences assigned to the Firmicutes within site E2 compared to a lake average (1%) (Fig. 8).
3.2.2  Alpha diversity

Mean microbial richness changed significantly over the 18 lake sites (Fig. 9A, p<0.0001, One way ANOVA) ranging from 2005 microbial OTUs detected near the mouth of the Selwyn river and 4144 in the centre west of the lake (Fig. 9). These changes in richness were strongly related to the changes in sediment texture across the lake (p=0.72, 0.63, -0.7 for silt, clay, sand). As the sediment changed from sandy to silty texture, richness levels increased. Similarly, other chemical and physical changes that correlated to the texture gradient (depth, bottom salinity, organic matter), correlated to changes in richness (table 2).

Shannon diversity, a measure that accounts for richness and species evenness, recapitulated the distribution and associations displayed by richness but the strength of correlations to physical and chemical variables were reduced (Table 2).

Figure 8: Relative abundance of phyla within Lake Ellesmere. Phylum level relative abundance was generated using Qiime (see section 2.5). Bars represent the mean percentage of the microbial community attributable to different phyla within 4 site replicates from each of 18 lake sites. Colours (see key, right) indicate phylum level microbial identity.
Figure 9: Alpha diversity measures, Lake Ellesmere. Average richness (A,C) and Shannon diversity index(B,D) over 18 lake sites was calculated using Qiime. Error bars indicate one standard error from the mean of 4 site samples (n=4), p<0.0001, one way ANOVA. Map figures C to D represent 18 mean data points (n=4) extrapolated by Kriging gridding in Surfer 11 (Golden software, Colorado, USA).

Table 2: Non-parametric Spearman’s correlation of lake alpha diversity and physicochemical parameters.

<p>| Variable 1          | Variable 2                  | Spearman p | Prob&gt;||p| |
|---------------------|-----------------------------|------------|------|
| Richness            | Surface Temp (°C)           | 0.2555     | 0.0303|
| Richness            | Surface Conductivity (m/s)  | 0.4322     | 0.0001|
| Richness            | Surface Salinity (ppt)      | 0.4408     | 0.0001|
| Richness            | Bottom Temp (°C)            | 0.2964     | 0.0115|
| Richness            | Bottom Conductivity (m/s)   | 0.5415     | &lt;.0001|
| Richness            | Bottom Salinity (ppt)       | 0.531      | &lt;.0001|
| Richness            | Redox (mV)                  | -0.3651    | 0.0016|
| Richness            | Secchi (cm)                 | 0.2812     | 0.0167|
| Richness            | Depth (m)                   | 0.6019     | &lt;.0001|
| Richness            | Organic Matter %            | 0.4583     | &lt;.0001|
| Richness            | TN ug/L                     | -0.1435    | 0.2292|
| Richness            | TP ug/L                     | 0.3028     | 0.0097|
| Richness            | NO2-/NO3- ug/L              | -0.0529    | 0.6587|
| Richness            | DRP ug/L                    | -0.1158    | 0.3327|</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Richness NH3 ug/L</th>
<th>0.2987</th>
<th>0.0108</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richness Clay %</td>
<td>0.6287</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Richness Silt %</td>
<td>0.7198</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Richness Sand %</td>
<td>-0.6966</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Richness Porosity</td>
<td>0.5093</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Richness Distance from E2 (km)</td>
<td>0.0296</td>
<td>0.8049</td>
<td></td>
</tr>
<tr>
<td>Richness Distance from E16 (km)</td>
<td>0.0698</td>
<td>0.5604</td>
<td></td>
</tr>
<tr>
<td>Richness Shannon</td>
<td>0.878</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Shannon Surface Temp (°C)</td>
<td>0.2229</td>
<td>0.0598</td>
<td></td>
</tr>
<tr>
<td>Shannon Surface Conductivity (m/s)</td>
<td>0.2518</td>
<td>0.0329</td>
<td></td>
</tr>
<tr>
<td>Shannon Surface Salinity (ppt)</td>
<td>0.2598</td>
<td>0.0276</td>
<td></td>
</tr>
<tr>
<td>Shannon Bottom Temp (°C)</td>
<td>0.1974</td>
<td>0.0964</td>
<td></td>
</tr>
<tr>
<td>Shannon Bottom Conductivity (m/s)</td>
<td>0.3155</td>
<td>0.0069</td>
<td></td>
</tr>
<tr>
<td>Shannon Bottom Salinity (ppt)</td>
<td>0.295</td>
<td>0.0119</td>
<td></td>
</tr>
<tr>
<td>Shannon Redox (mV)</td>
<td>-0.3086</td>
<td>0.0084</td>
<td></td>
</tr>
<tr>
<td>Shannon Secci (cm)</td>
<td>0.1211</td>
<td>0.3108</td>
<td></td>
</tr>
<tr>
<td>Shannon Depth (m)</td>
<td>0.4247</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Shannon Organic Matter %</td>
<td>0.4016</td>
<td>0.0005</td>
<td></td>
</tr>
<tr>
<td>Shannon TN ug/L</td>
<td>-0.0125</td>
<td>0.917</td>
<td></td>
</tr>
<tr>
<td>Shannon TP ug/L</td>
<td>0.2308</td>
<td>0.0511</td>
<td></td>
</tr>
<tr>
<td>Shannon NO2-/NO3- ug/L</td>
<td>0.0152</td>
<td>0.8992</td>
<td></td>
</tr>
<tr>
<td>Shannon DRP ug/L</td>
<td>-0.088</td>
<td>0.4624</td>
<td></td>
</tr>
<tr>
<td>Shannon NH3 ug/L</td>
<td>0.1537</td>
<td>0.1975</td>
<td></td>
</tr>
<tr>
<td>Shannon Clay %</td>
<td>0.498</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Shannon Silt %</td>
<td>0.611</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Shannon Sand %</td>
<td>-0.5654</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Shannon porosity</td>
<td>0.4325</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Shannon Distance from E2 (km)</td>
<td>0.0249</td>
<td>0.8358</td>
<td></td>
</tr>
<tr>
<td>Shannon Distance from E16 (km)</td>
<td>-0.1212</td>
<td>0.3105</td>
<td></td>
</tr>
</tbody>
</table>

### 3.3 Beta diversity

Multivariate analyses were carried out to explore patterns within the microbial community data based on differences in the relative abundance of taxa within each site.

#### 3.3.1 Principle component analysis (PCA)

Principal component analysis of the relative abundance of lake microbial populations at a genus level clustered sites into a gradient of community types with the first, second and third components explaining 8.89%, 5.84% and
5.49% of the changes in community composition between the sites (Fig. 10). Colour coding of sites by sand percentage demonstrated that sediment texture was associated with clustering of sites over the first component (PC1), and that sandier sites were more variable with respect to component 2 (PC2). The same method showed a correlation between organic matter and PC1 and a potential correlation between redox and PC2 (Fig. 11). This method was unable to link changes in water column chemistry to changes over the most significant principal components.

Phyla level PCA and cluster analysis demonstrated that the texture-associated changes seen at the genera level were also active at the phyla level (Fig. 12, 13). High levels of sand, or low levels of silt, were associated with enrichment of Actinobacteria, Planctomycetes and unclassified bacteria. The phyla Bacteroidetes, Chlorobi and unclassified organisms followed the opposite trend being enriched in areas of high silt and low sand. Changes in sediment organic matter content were similarly associated with these community/texture changes. Though not obvious through our cluster analysis, linear regression was able to show tentative links between lake water chemistry and phyla level community changes over PC3. TN and TP showed a significant (p<0.0001, one way anova) but weak positive correlation to PC3 (Fig. 14).
Figure 10: Clustering of sites based on PCA of genus level relative abundance data and color coding based on sand percentage at each site. Genus level relative abundance calculated using Qiime and 16S rRNA gene sequencing data at a depth of 17400 sequences per sample over four replicates per site. Percentages on axes show the variation explained by each component. Blue (Low sand%), grey (intermediate sand%), red (high sand%) points represent each of 4 individual site sampling replicates for 18 sites (72 points).

Figure 11: Clustering of sites based on PCA of genus level relative abundance data and colour coding based on organic matter content (percent of total dry sediment composition) (Left) and redox (Right). Genus level relative abundance calculated using the Qiime. Percentages on axes show the variation explained by each component. Blue (low), grey (intermediate), red (high) points represent each of 4 individual site sampling replicates for 18 sites (72 points). A single site (E15) was excluded as an outlier for organic matter analysis.
**Figure 12:** Clustering of sites based on PCA of phylum level relative abundance data and color coding based on sand percentage at each site. Phylum level relative abundance calculated using the Qiime. Percentages on axes show the variation explained by each component. Blue (low sand%), grey (intermediate sand%), red (high sand%) points represent each of 4 individual site sampling replicates for 18 sites (72 points).

**Figure 13:** Two-way cluster analysis of phyla by sample (based on 16S rRNA gene OTU\textsubscript{97} relative abundance and classification at the phylum level). Relative abundance data analyzed using Qiime 1.8 from each of 4 sampling replicates from 18 lake sites. Correlated drivers of community change identified by comparing to physical (right panel) and chemical (left panel) data. Colour...
gradient indicates relative abundance with intensity of red representing high abundance relative to the sample mean and blue relatively low abundance.

**Figure 14:** Water nutrients show correlation to phyla level PCA components. (Left) Linear regression of TN and Phyla PCA component 3. (Right) Linear regression of TP and Phyla PCA component 3. Shaded bars indicate confidence of line fit. Points indicate the comparison of 72 principal component co-ordinates and chemical data extrapolated from a one-off site measurement at 18 sites to 72 replicates.

### 3.4 UniFrac

Clustering of sites was also analysed using an alternative metric of community phylogenetic similarity (UniFrac). UniFrac principal coordinate analysis (PCoA) clustered sites into a gradient of community types with the first, second and third components explaining 42.2, 17.98 and 7.87% of the variability in the data (Fig. 15). Again overlaying sand percentage data demonstrated that sediment texture quality was associated with changes over component 1. Non-parametric correlation of genus level, phyla level and uniFrac principal components showed that these analyses were separating the lake microbial community in a similar fashion though genus level PC2 and 3 appear to be switched in relation to phyla and uniFrac PC2 and 3 with genus level PC2 correlating more to phyla and uniFrac PC3 (Table 3)
Figure 15: Principal coordinate analysis (PCoA) at the genus level separates lake site communities based on sand percentage. Genus level relative abundance calculated and analysed by PCoA using Qiime. Percentages on axes show the variation explained by each component. Coloured points represent each of 4 individual site sampling replicates for 18 lake sites (72 points). Colours indicate sediment sand content at the 18 lake sites.

Table 3: Non parametric analysis of PCA plot principle components

| Variable | by Variable | Spearman \( \rho \) | Prob>|\( \rho \)| |
|----------|------------|-----------------|-------------|
| Phyla PC1 | uniFrac PC1 | 0.9048 | <.0001 |
| Phyla PC2 | uniFrac PC2 | 0.7452 | <.0001 |
| Phyla PC3 | uniFrac PC3 | 0.5775 | <.0001 |
| 16S PC1  | uniFrac PC1 | -0.916 | <.0001 |
| 16S PC1  | Phyla PC1  | -0.9099 | <.0001 |
| 16S PC2  | uniFrac PC3 | -0.4583 | <.0001 |
| 16S PC3  | Phyla PC3  | -0.5816 | <.0001 |
| 16S PC3  | uniFrac PC2 | -0.7399 | <.0001 |
| 16S PC3  | Phyla PC2  | -0.6713 | <.0001 |
3.5 Community structure of organisms involved in nitrogen cycling

To understand the structure and distribution of organisms involved in the cycling of nitrogen within Lake Ellesmere genes involved in denitrification (i.e. loss of nitrogen) or nitrogen fixation (i.e. biologically linked inflow of nitrogen) were assessed via quantitative PCR (qPCR) of the nirS, nosZ I, nosZ II and nifH genes. Mean nifH, nirS, nosZI , nosZII copy numbers changed significantly over the lake floor (one way ANOVA, nifH-nirS-nosZI p<0.0001, nosZII p=0.0259) ranging from 4609, 20766, 4615 and 128437 to 122311, 173838, 11278 and 350219 copies respectively (Fig. 16). All sites showed within site variation but this effect was proportionally larger at higher copy number sites (Fig. 16).

The distribution of nitrogen cycling genes was not clearly interrelated. Abundance of nitrogen fixers, as determined by nifH abundance, tended to be lower around the perimeter of the lake with high values at sites 9, 10 and 11(Fig. 16A,E). Mapping extrapolations (Fig. 16G,H) suggested nosZI and II, both of which catalyze the conversion of N$_2$O to N$_2$, were negatively correlated to one another but this could not be confirmed by linear regression (data not shown). The distribution of nirS carrying denitrifiers did not relate to either nosZI or nosZII carrying denitrifiers with lowest abundances at mid-lake sites 11 and 12 where nosZ abundances were found to be relatively high (Fig. 16B-F).

A denitrification enzyme assay was carried out for sediments from all 18 Lake sites by David Hamilton, University of Waikato. DEA quantifies the denitrification activity of a site when given unlimited carbon and/or nitrogen
sources. Activity was highest for sediments amended both carbon and nitrogen (DEA +C+N)(Fig. 17) and showed lowest values for sites E2, E10 and E18 at 18.8, 10.8 and 19.2ppmv of N₂O respectively. Nitrate only (DEA+N) amended sediments showed a similar distribution but with much lower N₂O production overall and much higher relative activity compared to DEA +C+N for site E5 but much lower relative activity for site E9 (Fig17).
Figure 16: Distribution of nitrogen fixation and denitrifying genes throughout Lake Ellesmere. Mean *nifH* (A,E), *nirS* (B,F), *nosZI* (C,G), *nosZII* (D,H) gene copy number per 5ng total sediment extracted DNA as measured by quantitative PCR (qPCR). Four biological replicates, as well as four technical replicates, were analysed for each site within the lake. Mapping extrapolations were carried out using the Kriging gridding method on Surfer 12 (Golden software). Five sites were eliminated in analysis of *nosZII* due to the measurement of an unwanted amplification product. Error bars indicate one std error from the mean. *nifH-nirS-nosZI* p<0.0001, *nosZII* p=0.0259 by one way ANOVA.
Figure 17: Denitrification enzyme assay over 18 lake sites. Sediment samples were amended with nitrate + carbon (glucose) (A,C) or just nitrate (B,D) and N₂O emissions measured over 6 hours. Map figures C and D represent 18 data points extrapolated by Kriging gridding.

3.5.1 Potential nitrogen cycling community drivers

To determine the relationship between the nitrogen cycling community and other lake variables, nitrogen cycling gene copy numbers and DEA results were added to the Chemical/Physical/Alpha diversity PCA. With the exception of the \textit{nifH} copy number, the newly added variables tended to segregate away from physical and chemical factors into less significant components suggesting that their distribution was not related to these factors (Fig. 18).
Figure 18: PCA separates physical and chemical gradients and biological within Lake Ellesmere. (Bottom panel) Separation of 18 lake sites by PCA of lake water chemistry (TN, TP, NH₃, NO₂⁻/NO₃⁻, DRP, bottom salinity), sediment chemistry (organic matter content, redox), physical factors (Depth, sand%, silt%, clay%, secci depth, porosity, bottom temperature, distance from site E2, distance from site E16), alpha diversity measures (richness, shannon diversity) and denitrification parameters (blue)(Denitrification enzyme assay, nirS copy number, nifH copy number, nosZI copy number, nosZII copy number). (Top panel) Factor loadings for inputted variables: alpha diversity measures (green), denitrification parameters (blue). Axis percentage values indicate the percentage of variability within the data explained by that component.
Non-parametric correlation (Table 4) and linear regression (Fig. 19) was used in lieu of a conclusive PCA to look for individual factors associated with the nitrogen fixing and denitrifying community.

*nifH* levels were correlated to the sediment texture/depth gradient across the lake and associated factors such as richness and salinity. As sediments became more coarse *nifH* levels fell (Fig. 19A). Similarly *nosZII* was found to be at higher in deeper finer sediments but this association was much weaker (Table 4, Fig. 19C). *nosZI* followed the opposite trend, being enriched at shallow, showing significant but even weaker association (Table 4, Fig. 19D). Variation in levels of *nirS* did not correlate strongly with any physical or chemical gradients across the lake or to biological measures. Significantly, there was no positive association between *nirS* copy number and DEA which should both act as proxies for the conversion of NO$_2^−$ to N$_2$O (Table 4, Fig. 19B).

**Table 4:** Non parametric analysis of lake nitrogen cycling genes

<p>| Variable       | by Variable           | Spearman ρ | Prob&gt;|p| |
|----------------|-----------------------|------------|------|
| nifH copy #    | Sand %                | -0.8169    | &lt;.0001|
| nifH copy #    | Shannon               | 0.5365     | &lt;.0001|
| nifH copy #    | Surface Conductivity (m/s) | 0.6213    | &lt;.0001|
| nifH copy #    | Organic Matter %      | 0.6281     | &lt;.0001|
| nifH copy #    | Surface Salinity (ppt) | 0.6343    | &lt;.0001|
| nifH copy #    | porosity              | 0.6851     | &lt;.0001|
| nifH copy #    | Bottom Conductivity (m/s) | 0.7317    | &lt;.0001|
| nifH copy #    | Bottom Salinity (ppt) | 0.7338     | &lt;.0001|
| nifH copy #    | Depth (m)             | 0.7584     | &lt;.0001|
| nifH copy #    | Richness              | 0.7615     | &lt;.0001|
| nifH copy #    | Clay %                | 0.7742     | &lt;.0001|
| nifH copy #    | Silt %                | 0.8066     | &lt;.0001|
| nirS copy #    | Richness              | 0.3212     | 0.0059|
| nirS copy #    | nifH copy #           | 0.3143     | 0.0072|
| nirS copy #    | Shannon               | 0.2922     | 0.0128|</p>
<table>
<thead>
<tr>
<th>nosZI copy #</th>
<th>Depth (m)</th>
<th>-0.4003</th>
<th>0.0005</th>
</tr>
</thead>
<tbody>
<tr>
<td>nosZI copy #</td>
<td>Clay %</td>
<td>-0.3447</td>
<td>0.003</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>DEA +C+N</td>
<td>0.337</td>
<td>0.0038</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>TP ug/L</td>
<td>-0.3271</td>
<td>0.005</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>Distance from E2 (km)</td>
<td>0.3162</td>
<td>0.0068</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>DEA +N</td>
<td>0.2986</td>
<td>0.0108</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>porosity</td>
<td>-0.2974</td>
<td>0.0112</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>Organic Matter %</td>
<td>-0.2934</td>
<td>0.0124</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>Sand %</td>
<td>0.2921</td>
<td>0.0128</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>Shannon</td>
<td>0.2354</td>
<td>0.0465</td>
</tr>
<tr>
<td>nosZI copy #</td>
<td>NH3 ug/L</td>
<td>-0.2324</td>
<td>0.0495</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>Sand %</td>
<td>-0.4423</td>
<td>0.0016</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>nifH copy #</td>
<td>0.3536</td>
<td>0.0137</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>NH3 ug/L</td>
<td>0.3804</td>
<td>0.0077</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>Depth (m)</td>
<td>0.407</td>
<td>0.0041</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>Silt %</td>
<td>0.4283</td>
<td>0.0024</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>Organic Matter %</td>
<td>0.4502</td>
<td>0.0013</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>Clay %</td>
<td>0.4571</td>
<td>0.0011</td>
</tr>
<tr>
<td>nosZII copy #</td>
<td>porosity</td>
<td>0.4834</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

*significance cut-off at p>0.05

*Spearmans p cut-off at <0.4 for nifH
Figure 19: Sediment denitrification genes correlate poorly with physical and chemical lake gradients. Linear regression of \( nifH \) (A), \( nirS \) (B), \( nosZII \) (C), \( nosZI \) (D) to a variety of lake physical, chemical and biological factors. Shaded bars indicate confidence of line fit.
4 Discussion

The overarching goal of this project was to increase our understanding of the fate of anthropogenic nitrogen from farmland upon entering Lake Ellesmere and the impact of eutrophication on the microbial community in the lake. Specifically, the focus was on the role of the benthic microbial community and the process of denitrification. The aims of this research were to:

1) Determine the overall microbial community structure using 16S rRNA gene sequencing.

2) Determine the structure and distribution of benthic microbial denitrification and nitrogen fixing communities within Lake Ellesmere sediments using qPCR of three denitrification genes (\textit{nirS}, \textit{nosZI} and \textit{nosZII}) and one nitrogen fixation gene (\textit{nifH}).

3) Relate the community structures determined in aims one and two to the chemical and physical gradients observed within the lake.

It was hypothesized that as agriculturally derived nitrogen flowed into the lake it would produce a concentration gradient that would determine the state of overall and denitrifying communities within the lake sediments.

4.1 State of the lake-chemical factors

It is clear that the distribution of nutrients in Lake Ellesmere is complicated. TN, TP and DRP concentrations were reduced at more distant sites from the Selwyn River mouth (Ellesmere’s main tributary), supporting the hypothesis that agriculturally derived nitrogen inflow drives a nutrient gradient within the lake (Fig. 5). However, this gradient was weak and not seen for all water
nutrients. $\text{NO}_2^-/\text{NO}_3^-$ levels, which are the inputs for denitrification, did not respond in the same manner, suggesting that the production of $\text{NO}_2^-/\text{NO}_3^-$ could be occurring within the lake at the time of sampling. Prior work suggested that on average approximately 80% of total nitrogen entering through Lake Ellesmere’s tributaries was DIN (dissolved inorganic nitrogen), 95-99% of which was in the form of $\text{NO}_2^-/\text{NO}_3^-$ (Larnerd and Schallenberg, 2006). This statement is not necessarily at odds with our findings given the noted variability of tributary and lake chemistry (Hamill and Schallenberg, 2013). Rather, low $\text{NO}_3^-$ levels observed may represent normal seasonal changes and a general reduction of nitrogen levels within the lake over recent years. Similarly $\text{NH}_3$ within the lake did not appear to correlate to inflow from the Selwyn suggesting that production may occur within the lake.

Lake salinity, an important factor in a lake intermittently open to the ocean, was relatively high which suggests significant influence from the marine environment. Lake water nearer the top (northern end) of the lake was less saline, likely the result of freshwater inflow from tributaries (Fig. 2M,N). Although the lake was not open to the ocean at the time of sampling a significant period of opening had recently occurred from which the lake still appeared to be recovering. Halophilic microbial genera (Haladaptatus, Haloarcula, Halococcus, Haloferax, Halomarina) were identified within the lake on the basis of 16S taxonomy but their abundance did not appear to respond to the salinity gradient across the lake.
4.2 State of the lake physical factors

A clear gradient of sediment texture size within the lake was observed which is likely a result of the effects of particle size on sedimentation and wind induced movement in the water column over the lake (Fig. 4). Similar depth effects were observed by Kelderman et al., (2011) in a shallow windswept lake in Sweden. It is likely that at deeper lake sites, finer sediments accumulate due to reduced movement in the water column by external factors such as wind, thus allowing settling that might not be possible at more shallow sites. Among chemical parameters, organic matter was most strongly linked to the physical gradient of texture changes. Such an association is a common observation within sediments (Legg, Zheng, et al., 2012; Pelletier, Campbell, et al., 2011). The relationship is believed to be a result of adsorption of organic molecules to sediment particle surfaces. Given the larger surface area of smaller particles such as clays and silts they have more area to react with organic molecules (Bergamaschi, Tsamakis, et al., 1997).

4.3 Community structure and drivers

One of the primary aims of this project was to determine the overall microbial community structure in Lake Ellesmere’s sediments using 16S rRNA gene sequencing and to relate this to the chemical and physical gradients within the lake. Cluster and principal component analyses indicated that changes in sediment texture, depth and organic matter content within the lake were the most important determinant for overall community structure (Fig. 10, 11, 12, 13). Changes in these factors correlated with changes in phyla (Fig.12, 13) and genus level relative abundances (Fig. 10, 11) and measures of alpha diversity.
such as richness and the Shannon diversity index (Table 2). Due to the collinearity of these parameters it is difficult to understand which of the associated factors is most important for the microbial community, though Shannon diversity and richness show stronger non-parametric correlation with silt and sand percentage (Table 2).

The relationship between texture change and the structure of microbial communities has been widely demonstrated in both soils and sediments (Chau, Bagtzoglou, et al., 2011; Lauber, Strickland, et al., 2008; Perryman, Rees, et al., 2011; Wang, Liu, et al., 2013; Zheng, Wang, et al., 2014). A number of mechanisms behind these associations have been identified. Sessistch et al., (2001) suggest that larger particles provide less protection from grazing microbial predators; organisms large enough to resist this predation might be selected for in sands where silts and clays would select for a diverse range of organisms. Differences in community composition may arise due to the separation of different forms of organic matter preferentially into different sized fractions (Yeager and Sinsabaugh, 1998). Perhaps the most common explanation is that finer grained soils and sediments provide more surface area for microbial colonization (Santmire and Leff, 2007; Yeager and Sinsabaugh, 1998).

At the phylum level, organisms belonging to the Planctomycetes and Actinobacteria were enriched in coarser grained sediments while those associated with the Bacteroidetes and Chlorobi were depleted (Fig. 13). This is somewhat inconsistent with previous studies. Jackson and Weeks (2008) found Bacteroidetes, Verrucomicrobia and Planctomycetes were depleted in coarser grained sediments (1000µM-2000µM). Inconsistency between the distribution
of the Planctomycete phyla between this study and ours is likely a result of grain size classification. I note that Jackson and Weeks’ study identified higher levels of Planctomycete 16S sequences in the 250-500µm fraction than their 63-125µm (The same is true for Verrucomicrobia). In our project coarse sediments (sand) were identified as particles of greater than 63µm. It is possible that Planctomycetes prefer larger grained sediments up to a cutoff size where larger and larger grained sediments are associated with their depletion, though it is unclear why.

A central hypothesis of this project was that changes associated with anthropogenic nutrient inflow in the lake would determine, at least in part, the structure of the overall microbial community. Strong correlation between water column chemical parameters and measures of alpha and beta diversity were not observed. There could be any number of explanations for this observation, among the most simple and promising, water column nutrient concentrations may not reflect sediment concentrations. Gradients in chemical factors may not be strong enough to drive microbial community changes or the levels of available nutrients to microorganisms may not be particularly limiting, leaving only the effect of physical factors detectable. Physical factors have repeatedly been demonstrated as the major drivers of microbial community structure while chemical factors act as modulators. Sessitsch et al. (2001) showed that microbial community structure in organic matter amended soils was primarily determined by soil texture and not organic matter treatments. More recently it was demonstrated that denitrification community structure in Australian streams was more related to changes in sediment texture and associated sediment
organic matter content than water column nitrogen concentrations (Perryman, Rees, et al., 2011).

4.4 Determinants of nitrogen cycling community structure

Although 16S rRNA studies allow us to explore entire communities, targeting specific functional populations gives us information on key groups catalyzing reactions. I hypothesized that the concentration of nitrogenous compounds (particularly NO$_2^-$/NO$_3^-$) within the lake waters would drive the distribution of nitrogen cycling organisms (measured by quantification of nitrogen cycling genes) within the lake as denitrification rates and denitrification gene abundances have been shown to respond strongly to these inputs (Dong, Smith, et al., 2009b; Piña-Ochoa and Álvarez-Cobelas, 2006). Interestingly, principal component analysis of lake parameters resulted in the segregation of denitrification gene copy numbers away from chemical, physical and biological measures into separate, less significant components suggesting that the distribution of denitrifying organisms is not strongly influenced by these factors. The same plot suggested that the distribution of nitrogen fixation gene *nifH* was strongly influenced by physical gradients within the lake (Fig. 18). With the exception of *nifH*, non parametric correlations (Table 4) and linear regression (Fig. 19) showed only minor correlations between the measured nitrogen cycling genes and chemical, physical and biological measures within the lake, confirming the implications of the PCA plot.

The abundance of separate denitrifying genes was not strongly interrelated (Table 4). This may be explained by the varying responsiveness of the organisms carrying these genes to environmental parameters. Though it is hard
in the absence powerful correlations in this system to speculate on what those parameters are and how they work to determine gene abundances. The absence of powerful correlations between denitrification gene abundances and environmental parameters likely indicates that no singular parameter is important for determining the structure of the denitrifying community in Lake Ellesmere, rather many parameters are at work in concert. In a terrestrial system Zhang et al. (2013) used a systematic approach to test the effects of various parameters individually and in concert on the abundances of nitrogen cycling genes and nutrients. They demonstrated clearly that organisms carrying different nitrogen cycling genes responded with differing sensitivities to different combinations of parameters. The utilization of a similar systematic approach may be necessary to understand what drives nitrogen cycling in Lake Ellesmere.

It is not clear which factors drive denitrification gene distribution within Lake Ellesmere however the individual and relative abundances of such genes are still of interest. A recent study identified an additional clade (clade II) of nitrous oxide reductases and demonstrated that the relative proportions of nitrous oxide reductase genes attributable to clade I and II enzymes change significantly over different environments. In the littoral and limnetic zones of Lake Erken, Sweden it was shown that nosZII levels were two fold higher than nosZI. In French arable soils more significant ratios were seen, showing 10 fold differences between the abundances of these two genes (Jones, Graf, et al., 2013b). In this study, I found that nosZII was the dominant nitrous oxide reductase gene with on average 30 fold higher levels than nosZI. Moreover, my results show that the levels of nitrous oxide reductase genes outnumber
measured nitrite reductase genes (nirS) by approximately 2.5 to 1. It is a very important consideration that the contribution of the nitrite reductase nirK has not yet been quantified in this study, however nirS has been shown to be the more dominant nitrite reductase in a number of estuarine systems (Abell, Revill, et al., 2010; Nogales, Timmis, et al., 2002). Nevertheless, the high relative abundance of nitrous oxide reductase genes to nitrite reductase genes in Lake Ellesmere might indicate a propensity toward N₂ emission over N₂O which would constitute a considerable ecosystem service carried out by this lake. Unfortunately, initially planned N15 isotopic tracer experiments to identify the contribution of different nitrogen cycling processes to N₂ and N₂O emission in Lake Ellesmere were unable to be carried out. Therefore I am unable to comment on how denitrification gene copy numbers relate directly to N₂ and N₂O emissions within Lake Ellesmere.

The nitrogen fixation gene nifH was quantified in this study to assess the potential for alternative non anthropogenic nitrogen inputs into the lake. A large population of nitrogen fixers was observed within the sediments of this lake (Fig. 16A,E) (though I cannot exclude the possibility that nitrogen fixing bacteria from the water column have contributed to these measured levels) suggesting that benthic nitrogen fixation could be active, even in the presence of high levels of anthropogenic nitrogen. Such an observation is not entirely unexpected; benthic nitrogen fixation appears to continue in a number of eutrophied systems (Beversdorf, Miller, et al., 2013; Knapp, 2012; Rao and Charette, 2012) even though increased levels of nitrogenous compounds such as NH₃ have been shown to inhibit nitrogen fixation (Knapp, 2012).
Levels of *nifH* appeared to be associated with the sediment texture and organic matter gradient (Table 4). Study of the role of sediment organic matter levels on nitrogen fixation is still ongoing but it has been shown that reduced organic matter levels promote net nitrogen fixation over net denitrification in coastal sediments (Fulweiler, Nixon, et al., 2007), though a decrease in the ratio of these two processes does not necessitate a decrease in nitrogen fixation. Here, I observed high levels of the *nifH* gene associated with high sediment organic matter content. This result does not address the question of the relative rates of nitrogen fixation and denitrification in sediments but it does indicate the potential for benthic nitrogen fixation within Lake Ellesmere, a result that will need to be considered in future nitrogen budgets of this lake.

It is surprising that the levels of NH$_3$ in the water column was not identified above the Spearman's cut-off for associations with *nifH* (Table 4). Considering the hypothesis that NH$_3$ is produced within the lake it would be expected that the abundance of nitrogen fixing bacteria capable of producing this NH$_3$ would follow its distribution. It seems likely that mixing or transcription effects act to reduce this association.

Interestingly the distribution of the *nirS* gene which catalyses the reduction of NO$_2^-$ to Nitric oxide (NO) was not found to correlate with N$_2$O emissions as measured by DEA, which should act as a proxy for all but the final step of denitrification. The implication is that functional gene abundance in this lake is unrelated to emission potential. Nevertheless, DEA measures a very specific response to a specific carbon and nitrogen source. Considering that the activity of other unmeasured genes (eg *nirK*), the distinction between gene abundance
and active transcription and the alternative responses of different denitrifiers to different carbon sources, this result is not surprising.

4.5 Limitations of the experimental design

It is important to note that the exploratory nature of this study limited our ability to determine causation, a goal which was outside the scope of this study. Indeed, the approach used in this study is not conducive to understanding whether microbial communities are responding too or creating the gradients seen in lake physical and chemical data. Even in the presence of correlations, collinearity of multiple variables can blind true association. For example, a texture gradient across the lake was, collinear with depth and organic matter content, and was strongly correlated to \( \text{nifH} \) gene abundance in sediments. It is unclear whether the \( \text{nifH} \) containing microbial community is driven by a preference for finer sediments or organic matter content or perhaps that sediment texture drives organic matter distribution which is important for the maintenance of a nitrogen fixing community.

Many of the correlations to chemical gradients measured in this study rely on the assumption that chemistry within the water column is indicative of underlying sediment chemistry. Although the levels of a chemical substrate in the water column must affect sediment levels based on diffusion, the flux of nutrients between the sediment water interface in aquatic environments has been demonstrated to rely on a host of factors including bioturbation effects from macrofauna (Widdicombe, Beesley, et al., 2013), temperature changes (Duan and Kaushal, 2013) and water flows (Santos, Eyre, et al., 2012). Modelling sediment water interface flux is an ongoing process (Fennel, Brady,
et al., 2008) though it has already been demonstrated in a number of aquatic systems that, sediments can act as sources (Cheng, Zeng, et al., 2014) or sinks (Ogilvie, Nedwell, et al., 1997) for water column nitrogen. Indeed it is also important to consider that the distribution of a chemical substrate within the sediment is not continuous and that this distribution cannot simply be predicted, based on water column measures (Trolle, Hamilton, et al., 2010). Significant correlations between the water nutrient levels and the benthic denitrifying community were not made in this study. It is likely that sediment nutrient levels are markedly related to the distribution of the denitrification community or at least the their transcription.

4.6 Future directions

First and foremost, qPCR of the nitrite reductase gene nirK needs to be carried out to complete our knowledge of the nitrite reductase pool in Lake Ellesmere. This could allow more direct comment on the potential of this system to emit N₂O. Similarly the potential for alternative nitrogen removal processes to denitrification including anammox needs to be assessed by functional gene quantification as their contribution to nitrogen removal remains unknown.

In this project distance measures from the Selwyn River mouth were used to determine the influence of nutrient inflow on water column nutrients within the lake. It would be of great interest to concurrently measure the concentration of key nutrients in Ellesmere’s major tributaries during further lake sampling trips so that better linkages could be developed between anthropogenic inputs and the state of the lake.
This study made no attempt to measure the effect of temporal changes on this lake system, but they are likely of considerable importance to the microbial community. Seasonal, yearly, and climatic changes and lake opening are known to impact the state of the lake chemistry and inflowing nutrients (Hamill and Schallenberg, 2013; Schallenberg, Larned, et al., 2010). At times of extreme nutrient inflow it is possible that chemical gradients become a much more significant driver of benthic microbial community structure. Given that inflowing NO$_3^-$ levels peak during winter (Hamill and Schallenberg, 2013) we might see more response from the denitrifying community at that time. Further sampling trips will address this issue.

However it is clear that more than just further sampling trips are needed to describe this complex system. Laboratory work using controlled micro ecosystems (microcosms) will be carried out on sediments from Tomahawk Lagoon in Dunedin to characterize the effects of individual parameters on denitrification activity in New Zealand’s coastal lagoons. The hope is that such experiments will allow prediction of lake community structure based on measured lake parameters and may explain some of the trends or lack there of that we have observed in this study.

### 4.7 Conclusions

Physical factors are the most significant drivers of the benthic microbial community structure within Lake Ellesmere. Data supported the role of sediment texture and organic matter content as drivers of microbial diversity, richness and community composition. Importantly, the data suggests that at the time of sampling neither the total microbial community (determined based on
16S rRNA gene sequencing) or the nitrogen cycling community responded directly to inflowing anthropogenic nutrients from the Selwyn River. The fate of anthropogenic nitrogen entering Lake Ellesmere remains unknown but the quantification of a considerable number of denitrification genes within the lake sediments shows that the potential for denitrification is certainly present within this lake.
References


nitrite reductase and homology of this gene to DNA of other denitrifiers. *Applied and Environmental Microbiology*, 59(1), 250–4.


