Spatial Patterns in Bacterial Community Structure and Function within Shallow Alpine Tarns

---

A thesis

submitted in partial fulfilment

of the requirements for the Degree of

Master of Science

at

Lincoln University

by

Julia Ruth Bellamy

Lincoln University

2013

Spatial Patterns in Bacterial Community Structure and Function within Shallow Alpine Tarns

by

Julia Ruth Bellamy

Small scale spatial variation in bacterial community structure and function in freshwater ecosystems is poorly understood. I investigated the spatial variation of bacterial communities within three tarns located at Tekapo Scientific Reserve using automated ribosomal intergenic spacer analysis (ARISA). I examined the variability in bacterial community structure both within and among tarn locations and explored whether bacterial communities adhere to the same biogeographical patterns commonly reported for communities of larger organisms; these were the taxa-area and distance-decay relationships. I also attempted to identify physicochemical variables that were significantly related to the observed community heterogeneity. To achieve these aims, I collected more than 100 samples in total across the three tarns and measured a range of physicochemical variables (pH, conductivity, total carbon and anion concentrations) for each sample. The ARISA data revealed significant variability in bacterial community structure among the tarns and some variation within the tarns that was related to correlated spatial variability in a range of physicochemical variables such as, pH, total carbon and conductivity. Distance-decay and taxa-area relationships in bacterial community similarity were also observed. There was no correlation between the structural and functional attributes (i.e., carbon substrate utilisation patterns) of the bacterial communities, suggesting that there was some functional redundancy in these bacterial communities in terms of carbon substrate utilisation. This study provides valuable information about how freshwater bacterial biodiversity is maintained and expands our understanding of the link between bacterial community structure and function.

Keywords: Biogeographical patterns, taxa-area, z-value, distance-decay, ARISA, BIOLOG EcoPlates™, contour plots, PRIMER, Tekapo, New Zealand, DistLM, variance partitioning

Acknowledgements

First and foremost, I would like to thank Dr. Gavin Lear at the University of Auckland, NZ, for his patience and guidance throughout the duration of this study. I would also like to thank Dr. Hannah Buckley at Lincoln University, NZ, for her significant contribution towards this study, especially in regards to data analysis and interpretation. I am grateful to Bradley Case, also at Lincoln University, NZ, for his help with the GIS program, ArcMap 10. Finally, the kind assistance with sample collection by Jack Lee and Taiga Yamamura was much appreciated. This work was supported by the RSNZ Marsden Fund (Grant Number LIU0901).

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

An important focus of ecology is to understand how biodiversity is generated and maintained. This can, in part, be achieved through understanding the spatial distribution of organisms (Green and Bohannan, 2006). However, while the distribution of larger organisms is well documented (Rosenzweig, 1995), the biogeography of other groups of organisms, such as, microorganisms remains poorly understood (Fierer and Jackson, 2006; Green and Bohannan, 2006; Martiny *et al.*, 2006). This is partly due to the fact that until recently we relied on culture based techniques to identify microorganisms. It is proposed that less than 0.5 % of microorganisms can be cultured (Torsvik *et al.*, 1990), such that a large proportion of the microbial community is not detectable using such methods. However, with the introduction of molecular techniques, microorganisms can now be rapidly characterised by their DNA, removing the need to grow cells in the laboratory (Malik *et al.*, 2008). Molecular based techniques also allow microorganisms to be detected *in situ* (Gilbride *et al.*, 2006). Consequently, interest in microbial ecology continues to increase.

There is controversy surrounding the distribution of microorganisms. Some argue that all microbial taxa are everywhere (cosmopolitan distribution) (Martiny *et al.*, 2006), since the small size and high abundance of microorganisms means that they are widely distributed in large numbers in global currents of wind and water. Conversely, others claim that microorganisms display biogeographical patterns, such as the species-area and distance-decay relationships (Green and Bohannan, 2006; Horner-Devine *et al.*, 2004; Martiny *et al.*, 2006), which suggests that there are barriers or limitations to microbial immigration or colonisation at certain spatial scales.

1.1 Taxa-area relationship

A common biogeographical pattern reported for communities of macroorganisms is the species, or taxa-area relationship (Crawley and Harral, 2001; Fridley *et al.*, 2005; Rosenzweig, 1995; Ulrich and Buszko, 2003). The taxa-area relationship states that taxa richness increases in accordance with area sampled (Lomolino, 2001). There are a number of formulae used to model the taxa-area relationship (Scheiner, 2003), with the power law, Arrehenius (1921), being one of the most commonly used:

$$S = cA^z$$

where *S* is species or taxa richness (number of species), *A* is area sampled, *c* is a constant that partially determines the slope of the curve with *z* which is a measure of the rate of taxa turnover across space. A larger *z* value represents a greater rate of turnover (Rosenzweig, 1995; Zhou *et al.*, 2008). For example, the taxa-area relationship was modelled using the power law in arithmetic space for South American birds in four biomes types and shows the expected increase in species richness with increasing area sampled (Rosenzweig, 1995) (Figure 1).

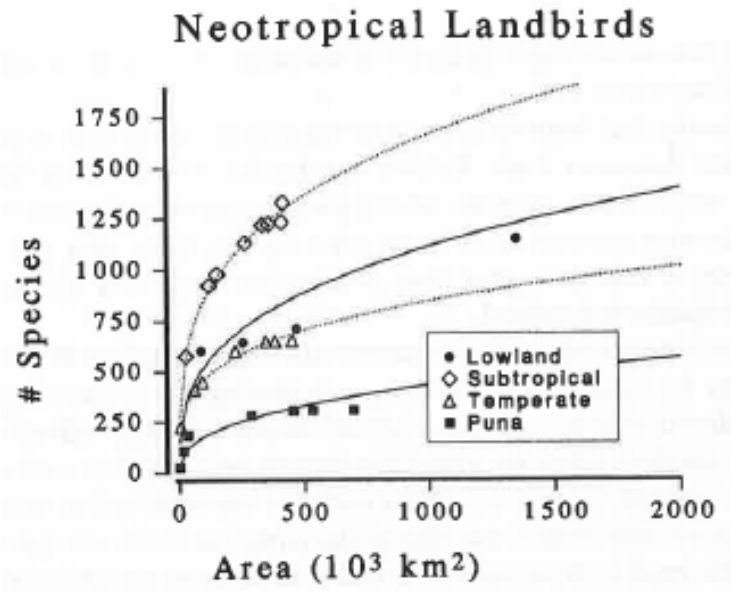

Figure 1. Taxa-area relationship modelled by the power law and plotted in arithmetic space showing an increase in species richness with increasing area. Taken directly from Rosenzweig (1995).

With the taxa-area relationship, it is important to understand that there is a difference in the *z*-values for islands and contiguous habitats, which may be because islands are surrounded by a barrier that hinders colonisation, whereas contiguous habitats can be colonised from adjacent areas (Bell *et al.*, 2005). With islands, *z*-values typically range from 0.25 to 0.35 (Rosenzweig, 1995). However, with contiguous habitats, *z*-values usually range from 0.12 to 0.18 (Rosenzweig, 1995). These *z*-values are for macroorganisms, so it is possible that they may be different for microorganisms. It has been proposed that because the small size and high abundance of microorganisms means that they are widely distributed, this will result in lower *z*-values for microorganisms than for macroorganisms (Bell et *al.*, 2005). Cencini *et al.* (2012) determined that as the size of the local population increases, the species-area curve become shallower (lower *z*-value), which is consistent with the theory that because microorganisms are widely distributed, they will have a higher local population and hence lower *z*-values than macroorganisms. Alternatively, microorganisms may have lower *z*-values than macroorganisms due to decreased local diversification, because large population sizes result in low extinction rates (Fenchel and Finlay, 2004), or low speciation rates due to horizontal gene transfer (Thomas and Nielsen, 2005) and lack of geographical isolation (Finlay, 2002).

The taxa-area relationship can be explained by three possible theories. First, it is likely that more taxa will be encountered when sampling a larger area than a smaller area because there is a higher probability that rare taxa will be encountered (Hill *et al.*, 1994; Kallimanis *et al.*, 2008). Second, a larger area will be more environmentally heterogeneous and will thus support a greater variety of taxa, each of which may be specialised to different environmental conditions (Cam *et al.*, 2002; Kallimanis *et al.*, 2008; Preston, 1962; Rosenzweig, 1995). Finally, the island biogeography theory which states that a dynamic equilibrium exists between immigration and extinction rates on an island (MacArthur and Wilson, 1967), and can be used to qualitatively predict shifts in species richness and turnover rates with changes in area sampled and degree of isolation. One of the first studies to identify a taxa-area relationship for microbial communities was performed by Bell *et al.* (2005). Bell *et al.* (2005) observed a taxa-area relationship for microbial communities in water-filled tree holes (islands). This study was performed over a relatively small spatial scale with the largest island sampled (volume of water) being 20 litres. In contrast to Bell *et al.* (2005), Humbert *et al.* (2009) did not detect a taxa-area relationship in microbial

communities from six reservoirs of different sizes (maximum volumes ranged from 0.6 to 1700 Mm^3) in Burkina Faso or three alpine lakes (maximum volumes ranged from 1124 to 8900 Mm^3) in France. The inconsistency in these findings may be due to differences in the scale of the studies as Bell *et al.* (2005) concluded that with large areas of contiguous habitat the slope of the taxa-area relationship appears reduced. Alternatively, microbial communities in freshwater ecosystems may not display taxa-area relationships because freshwater systems are usually well mixed and environmental heterogeneity is low (Scheiner *et al.*, 2000). These studies highlight the importance of understanding the distribution of microorganisms at various spatial scales.

1.2 Distance-decay relationship

Another common biogeographical pattern described in literature is the distance-decay relationship (Dexter *et al.*, 2012; Finkel *et al.*, 2012; Jones *et al.*, 2012; King *et al.*, 2010; Palmer, 2005; Sommaruga and Casamayor, 2009; Thieltges *et al.*, 2009). The distance-decay relationship states that community similarity declines with increasing geographic distance (Horner-Devine *et al.*, 2004). For example, this relationship was observed by Dexter *et al.* (2012), who identified that the compositional similarity of *Inga* (a genus of nitrogen fixing tree) communities in Madre de Dios, Peru for terra firme and floodplain showed a decline with increasing geographic distance (Figure 2).

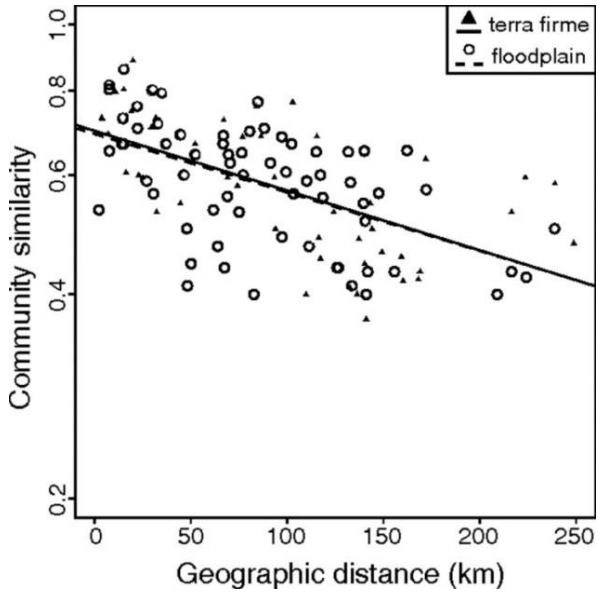

Figure 2. Relationship between the compositional similarity of paired *Inga* communities and geographic distance. Taken directly from Dexter *et al.* (2012).

Two mechanisms may be responsible for the distance-decay relationship. First, because microorganisms may have limited dispersal ability, community similarity will show a decline with increasing geographic distance regardless of environmental features (Soininen *et al.*, 2007). Second, increasing environmental heterogeneity with increasing geographic distance will result in a decline in community similarity with increasing distance. This is because microorganisms are adapted to different environmental conditions which change over space and will therefore be present where favourable conditions exist. For example, the landscape structure and the degree of isolation of the landscape may influence the dispersal rate of microorganisms (Nekola and White, 2004). The similarity of communities within a complex landscape structure that has dispersal barriers will decline more abruptly than a less complex and more open landscape. Also, landscapes that are more isolated will take longer to colonise resulting in a gradient in microbial diversity after a disturbance.

Horner-Devine *et al.* (2004) analysed whether or not microorganisms displayed a taxa-area relationship, using a distance-decay approach, in a New England salt marsh, and determined that environmental heterogeneity affects microbial communities more than geographic distance. It was found that when the effect of geographic distance was removed, habitats with similar environmental conditions had similar microbial communities. However, when the effect of environmental similarity was removed, geographic distance had no effect on the microbial communities. This was supported by Van der Gucht *et al.* (2007), who concluded that spatial distance is insignificant compared to local environmental conditions. Finkel *et al.* (2012) investigated the dispersal limitation in phyllosphere communities on the leaf surfaces of spatially dispersed, salt-excreting *Tamarix* trees. The trees were located over a 500 km east-west transect in the Sonoran Desert of southwestern United States with relatively uniform climate conditions (minimising environmental heterogeneity). They found that some communities of bacterial taxa, such as betaproteobacteria, showed a significant decline in similarity with increasing geographic distance. However, even though the sampling regime was designed to minimise environmental heterogeneity, a weak relationship was observed between geographic distance and environmental heterogeneity. This study indicates that the distance-decay relationship can be observed for microorganisms over a relatively large spatial scale (0 to 500 km). However, will it still be present over finer spatial scales?

In regards to the distance-decay relationship being present over a range of spatial scales, King *et al.* (2010), determined that the dissimilarity in bacterial community composition in soil samples collected from a continuous landscape on the south side of the Green Lakes Valley Watershed, Colorado, USA, only showed an increase between 2 to 240 m. However, in contrast to Finkel *et al.* (2012), between 240 to 2000 m King *et al.* (2010) observed no relationship between dissimilarity in bacterial community composition and geographic distance. Spatial differences in bacterial community composition for samples that were located less than 240 m apart were suggested to be due to the landscape distribution of biogeochemical properties (King *et al.*, 2010). These studies indicate that environmental heterogeneity is a significant driver of the distance-decay relationship for microorganisms, but also suggest that spatial location has some influence on the distance-decay relationship. In addition, it has been identified that the distance-decay relationship might not be present at all spatial scales.

1.3 Relationship between spatial scaling and efficient sampling strategies

The spatial scale of variation in microbial diversity may result in microbial communities that are separated on a small scale of less than a metre being significantly different from each other in terms of their composition (Franklin and Mills, 2003). The difference in microbial communities may be due to variation in local, or microsite, conditions such as pH, organic matter content or interactions with biological neighbours. Unfortunately, many sampling regimes do not take this into consideration and in analysing only a limited number of samples, they underestimate total microbial diversity (Mocali and Benedetti, 2010). The spatial structure of microbes is poorly understood because a large number of samples are typically required to accurately determine spatial variation (Saetre and Bååth, 2000). However, recent advances in the high throughput analysis of microbial DNA now allow the simultaneous analysis of hundreds of samples at a relatively low cost. A previously mentioned study, Humbert *et al.* (2009), involved samples being collected from only one location in reservoirs in individual lakes in Burkina Faso and alpine lakes in France because it was concluded in an earlier study (Dorigo *et al.*, 2006) that one sample provides a good evaluation of the total bacterial diversity in a freshwater ecosystem. However, this is not supported by Jones *et al.* (2012), who performed a study that involved collecting multiple samples from each of two lakes located in Wisconsin, USA. The samples were used to

analyse the variation in bacterial community composition on a small spatial scale (10 m). They identified significant horizontal variation in bacterial community composition across the lakes (Figure 3). These studies suggest that even though microbial diversity may appear to be relatively homogeneous across freshwater ecosystems nevertheless it may display small scale spatial patterns. Therefore it is important to collect multiple samples across freshwater ecosystems in order to provide an accurate representation of the microbial community present.

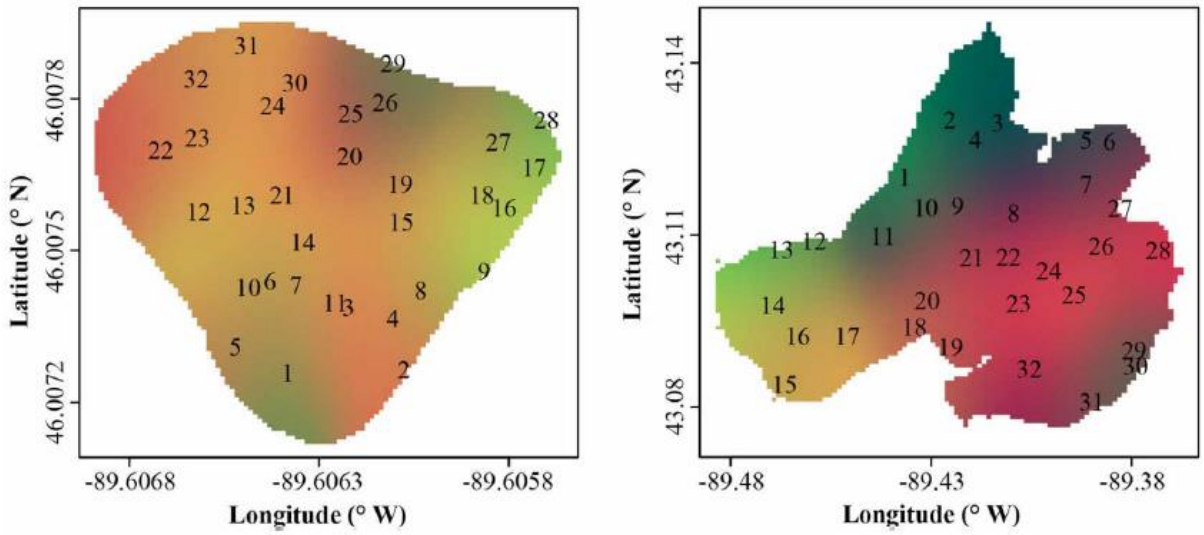

Figure 3. Lake Mendote and Crystal Bog. Principle coordinate analysis (PCoA) was performed on the bacterial community composition in the lakes and the axis of the ordinations were coloured and this information was plotted on the maps. Similar colours denote sites represented by more similar bacterial communities. Taken directly from Jones *et al.* (2012).

1.4 Link between microbial community structure and function

Microorganisms have a significant influence on important ecological processes in the environment, such as trace gas emissions, soil structure and formation, decomposition of organic matter and xenobiotics, and the recycling of essential elements (e.g. carbon, nitrogen, phosphorous, and sulphur) and nutrients (Green and Bohannan, 2006; Horner-Devine *et al.*, 2004; Rastogi and Sani, 2011). The fact that microorganisms are involved in a number of important ecological processes, along with the recent development of molecular techniques, has resulted in an increase in the number of studies (Bell *et al.*, 2005; Cho and

Tiedje, 2000; Horner-Devine *et al.*, 2004; Reche *et al.*, 2005; Van Der Gast *et al.*, 2005) being performed on microorganisms in relation to their distribution, which contributes to our understanding of biodiversity. However, there remains limited information regarding the link between the variation in microbial composition and the functional attributes of the communities in natural ecosystems. It has been proposed that the number of species required to sustain ecosystem functioning is directly correlated with the number of processes considered (Hector and Bagchi, 2007) and there are a number of different scenarios that can be used to model the relationship between functional diversity and species diversity (Figure 4). A number of studies have attempted to identify the link between microbial composition and ecological processes.

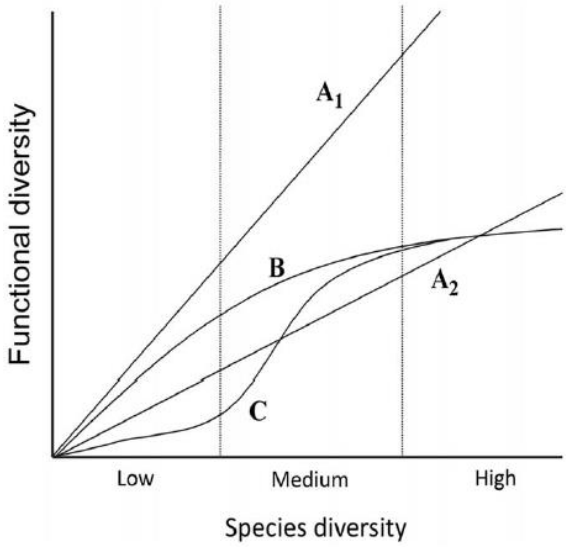

Figure 4. Schematic of possible relationships between functional and species diversity. A1, every species has a unique functional role resulting in a ratio of 1:1 between functional and species diversity; A2, multiple species have similar functional attributes; B, at low species richness, functional diversity rapidly increases and then increases at declining rates with high species diversity, and eventually reaches an asymptote; C, relationship between functional diversity and species diversity is sensitive to changes in environmental variables. Taken directly from Guillemot *et al.* (2011).

Balser and Firestone (2005) performed a study investigating the link between microbial community composition and soil process rates such as N mineralisation, nitrification, CO_2 and N_2O flux, independent of certain environmental variables (soil temperature and water content). The study involved the transplantation of soil cores from two sites, a grassland and

conifer forest ecosystem, along an elevation gradient with different climates (temperature and precipitation). Not only did they find that there was a significant relationship between microbial community composition and soil processes independent of soil temperature and water content, but soil processes were also related to the environment. Unfortunately, this study only controlled soil temperature and water content and it is possible that there were a number of other environmental variables that would have affected the soil processes. Another study by Eilers *et al.* (2010) analysed shifts in bacterial community phylogenetic structure in soils following the addition of specific carbon compounds (glucose, glycine and citric acid). They found that the magnitude of the change in function (CO_2 respiration) in bacterial communities upon the addition of carbon compounds did not correlate with the change in bacterial community composition. An increase in the abundance of a subset of taxa was observed when the carbon compounds were added to the soil. However, a number of abundant taxa showed no response to the addition of the carbon compounds. This study shows that there is a link between bacterial community composition and ecological processes to some degree but also indicates that microbial communities may exhibit some functional redundancy. Functional redundancy is defined by Wohl *et al.* (2004) as "multiple species, while biologically unique, contributing with similar intensity to the same process within an ecosystem, such as energy flow or nutrient cycling". It is evident from the study by Balser and Firestone (2005) that there is variation in the functional attributes of microbial communities in soil ecosystems. Is this relationship also present in freshwater ecosystems, which can be expected to be less heterogeneous, due to mixing (Scheiner *et al.*, 2000), or will the microbial communities in these ecosystems show functional redundancy?

It is possible that microbial communities contain some functional redundancy. The high species richness that is usually displayed by microorganisms is thought to facilitate functional redundancy because the probability of encountering multiple species that perform the same ecological processes is increased. An important question in microbial ecology is, if a microbial community suffers from a disturbance that results in a loss of biodiversity, will the new community be functionally similar to the original community or will a loss in ecological processes be observed? Functional redundancy can be observed in two different ways: (i) First, the new community may be compositionally different (e.g. contain different taxa). However, it may still be capable of performing the same ecological processes as the original community. (ii) Second, there may have been no change in the composition of

the community but individual taxa may function differently resulting in the same process rate at the community level (Allison and Martiny, 2008). Alternatively, microbial communities may display functional dissimilarity where a loss in the number of microbial taxa corresponds to a loss in ecological processes. Microbial communities showing both functional redundancy and functional dissimilarity have been observed in a range of studies.

Yin *et al.* (2000) performed a study that involved a number of soil samples from a soil reclamation gradient in a tin mine site in Brazil being amended with different carbon compounds (L-serine, L-threonine, sodium citrate, and α-lactose hydrate). A change in the functional attributes of the bacterial communities would be expected upon the addition of the carbon substrates and the corresponding amount of change or no change in the richness and diversity of the bacterial communities would give an indication of the degree of functional redundancy present in the bacterial communities. The richness and diversity of bacterial groups increased along the reclamation gradient in response to the addition of the carbon substrates, indicating some functional dissimilarity in the bacterial communities because the modification in environment conditions supported the growth of new bacterial taxa. However, even though some functional dissimilarity was observed, it was determined that bacterial functional redundancy was positively correlated with changes in soil that supported plant growth. Wohl *et al.* (2004) investigated functional redundancy by analysing, under constant environmental conditions, if an increase in species richness influenced function (cellulose decomposition) and if a decline in species richness was observed over time in functionally redundant communities. It was observed that, in contrast to what was expected, a decline in species richness only occurred for microcosms inoculated with two species (one species was eliminated). Species richness was maintained in the other microcosms that were inoculated with four or more species which was thought to be due to the inherent diversity of multiple species that allowed for different resources to be exploited. This implies that a functionally redundant community can support high species richness through the production of resources by individuals resulting in resource heterogeneity. The study also found that species richness was the main driver of function (cellulose decomposition) with the microcosms that contained four or eight species displaying the highest cellulose decomposition. Finally, Strickland *et al.* (2009) performed a study that involved sterilised litter being inoculated with soil, both of which were collected from three sites within the continental United States. A large proportion of the variation in

litter decomposition (measured by carbon mineralisation) was explained by the inoculum source. This supports the idea of functional dissimilarity, that is, species perform unique functions. These studies are evidence of research being performed to look at the functional attributes of microbial communities in which the link between microbial composition and function remains poorly understood (Guillemot *et al.*, 2011), especially in freshwater ecosystems. Unfortunately, the majority of the aforementioned studies were temporal based and it would have been interesting if the studies had incorporated spatial parameters into the investigation of functional diversity. In addition, previous studies investigated the functional attributes of microbial communities by amending their habitat and then monitoring the response of the microbial communities to this change instead of determining functional redundancy in an undisturbed environment.

1.5 Aims, objectives and hypotheses

My study characterises spatial variability in bacterial community structure and functional characteristics within tarns (shallow mountain lakes) located at Tekapo Scientific Reserve. The influences of geographical distance and sample area on the bacterial communities were investigated to examine concepts related to patterns of distance-decay and taxa-area relationships, respectively. More than 100 bacterial community samples were collected using a custom built sampling apparatus that minimised disturbance and mixing of the tarns. Multiple samples were collected from each tarn allowing an investigation of fine scale spatial variation within the tarns. It was expected that the variation within the tarns would be less than the variation among the tarns. Bacterial community structure was assessed using a DNA finger-printing approach, automated ribosomal intergenic spacer analysis, (ARISA), and differences in bacterial community function were assessed using carbon substrate utilisation analysis (BIOLOG EcoPlates™).

Identification of the most significant factors driving the bacterial community structure of these freshwater ecosystems may improve our ability to adequately sample them for assessment of bacterial community structure and function. I examined the extent to which bacterial communities in alpine tarns adhered to common biogeographical patterns. Specifically: (1) The taxa-area relationship; I characterized the taxa-area relationships for bacterial communities in the tarns. I predicted that the bacterial communities would display

a positive taxa-area relationship, i.e. that taxa richness would increase in relation to the volume of sample. (2)The distance-decay relationship; I characterized distance-decay relationships for the bacterial communities in the tarns. I expected that (i) similarity in bacterial community structure and function would decline with increasing geographic distance between samples, and (ii) there would be more variation in bacterial community structure among the tarns than within the tarns. (3) Spatial and environmental variation; I characterized spatial variation in bacterial community structure and function in the tarns. I then used regression analyses to identify which environmental variables had the most significant relationship with bacterial community structure and function in the tarns. I predicted that (i) spatial location would contribute the most towards variation in bacterial community structure and (ii) the environment would have a stronger relationship with bacterial community function than with bacterial community structure.

Chapter 2

Methods

2.1 Experimental design

2.1.1 Study site

Samples of water were collected from tarns located at the Tekapo Scientific Reserve, New Zealand (Figure 5). Three tarns were selected based on their volume of water, as many of the smaller tarns were dry. Samples were collected consecutively over three days; the 10th, 11th and 12th of January 2012 from Tarns 1, 2 and 3, respectively. Additional physicochemical data (depth, dissolved oxygen and temperature) for all three tarns were collected on the 12th of January 2012. The tarns are located on Department of Conservation (DOC, http://www.doc.govt.nz/) land and therefore permission was obtained from DOC before commencement of sampling.

2.1.2 Sample locations

The tarns had the following surface areas; Tarn 1 = 2,450 m^2, Tarn 2 = 2,800 m^2 and Tarn 3 = 3,900 m^2. From each tarn, samples were collected using a grid format, of approximately 7 m x 7 m (Figure 6). For the largest tarn, Tarn 3, additional samples were collected from three random locations using a smaller grid format, of approximately 3.5 m x 3.5 m. The coordinates from where the samples were collected were recorded using two GPS systems (Rino 650, Garmin and Trimble ProXT Differential GPS). A total of 36, 33 and 53 samples were collected from Tarns 1, 2 and 3, respectively.

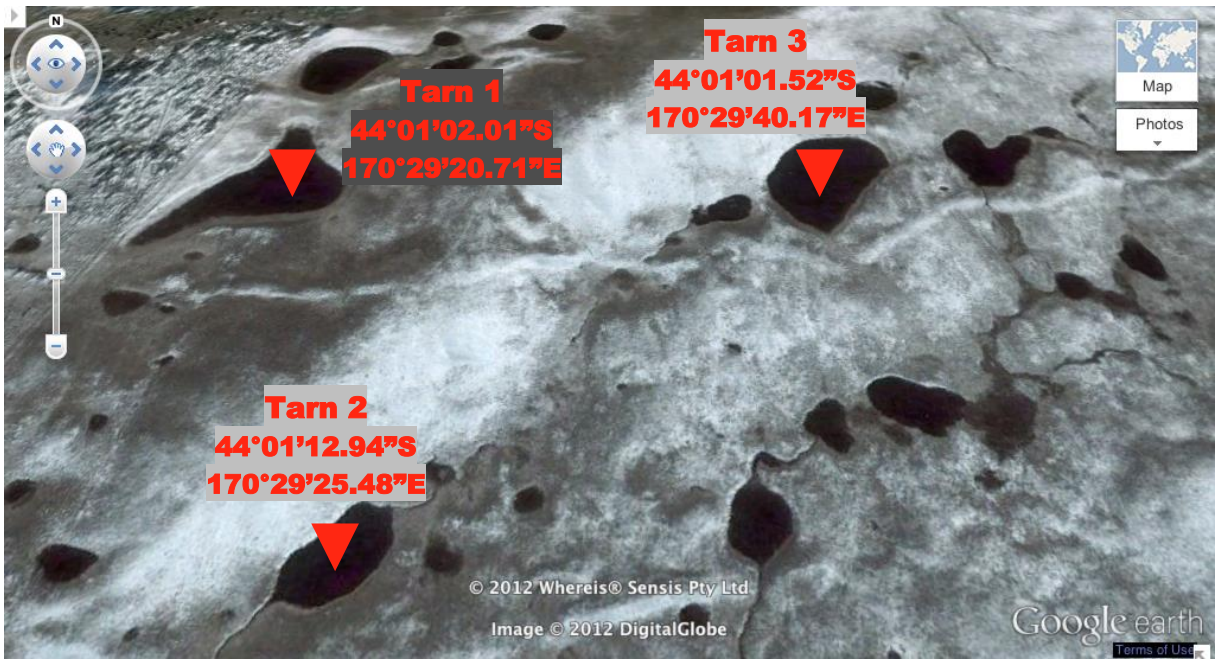

Figure 5. Location of tarns near Lake Tekapo, New Zealand.

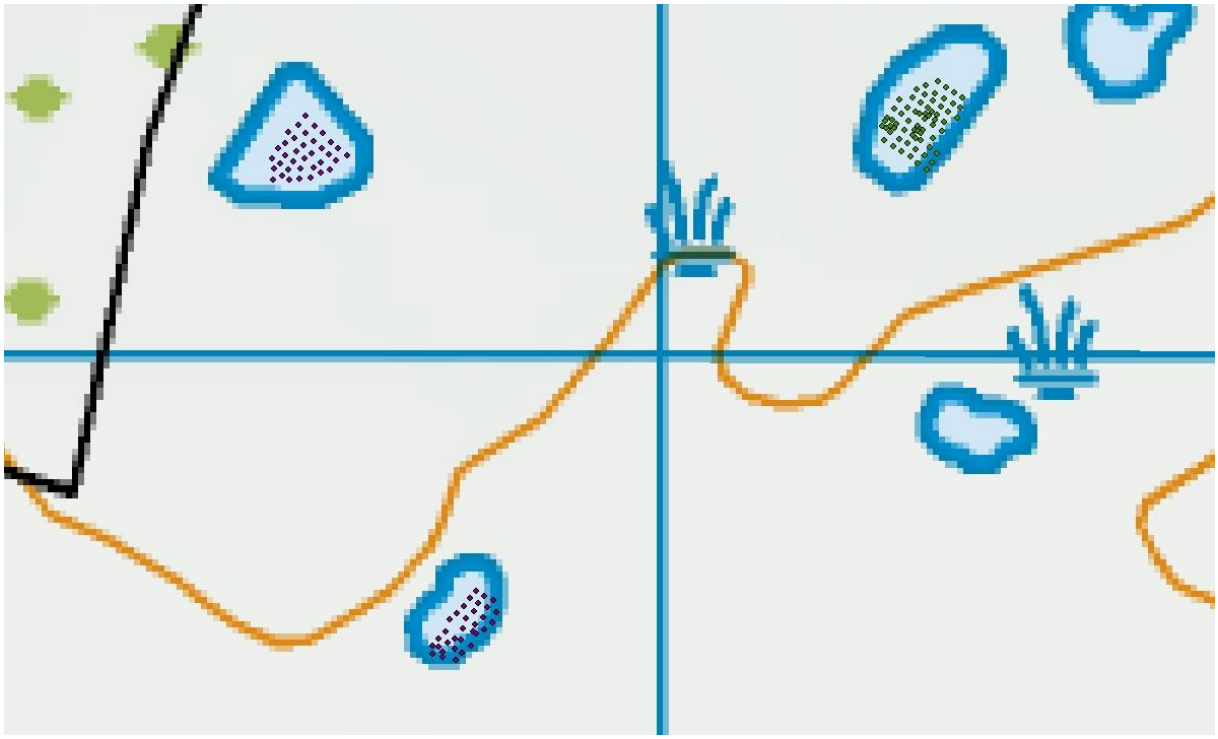

Figure 6. Sample locations in the three tarns from which samples were collected. Each dot shows where water samples, and other data were collected.

2.1.3 Sample collection

A custom designed sampling apparatus was constructed to allow the collection of samples with minimal disturbance of the tarns while also being portable and easy to set up. The sampling apparatus (Figure 7 & Figure 8) consisted of; (i) two adjustable support poles that were erected either side of the tarns, (ii) a length of rope with tubing attached, connecting the two poles together, (iii) a floatation device (Figure 9) that allowed the end of the tubing to be positioned at a set depth of 5 cm below the water's surface and finally (iv) a petrol-engine pump which was used to transfer the water through the tubing. To ensure that each sample was not contaminated with water from the previous sample, after the collection of each sample, both ends of the tubing were briefly removed from the water to allow the formation of an air bubble and then the tubing was flushed until approximately 10 seconds after the air bubble had been expelled.

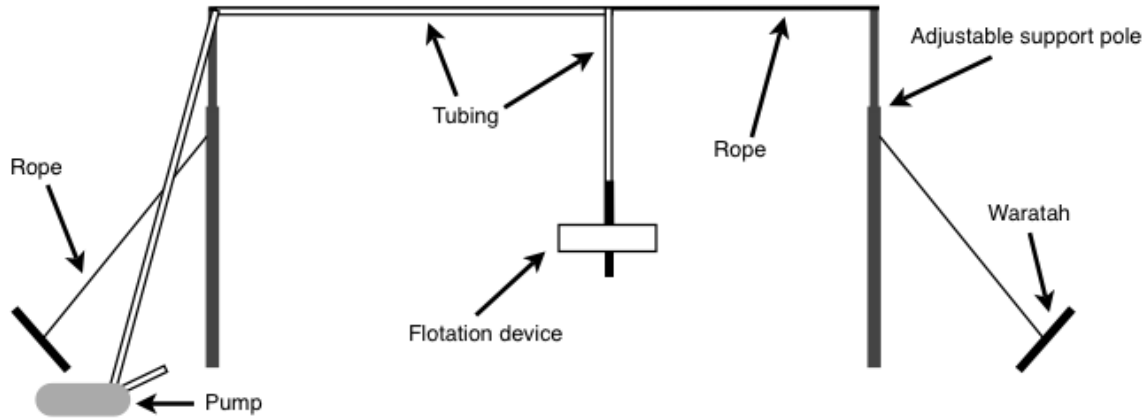

Figure 7. Schematic showing the components of the sampling apparatus. The adjustable support poles extended to a height of 2.5 m and the rope was 100 m long giving us the ability to sample across a distance of 50 m (not to scale).

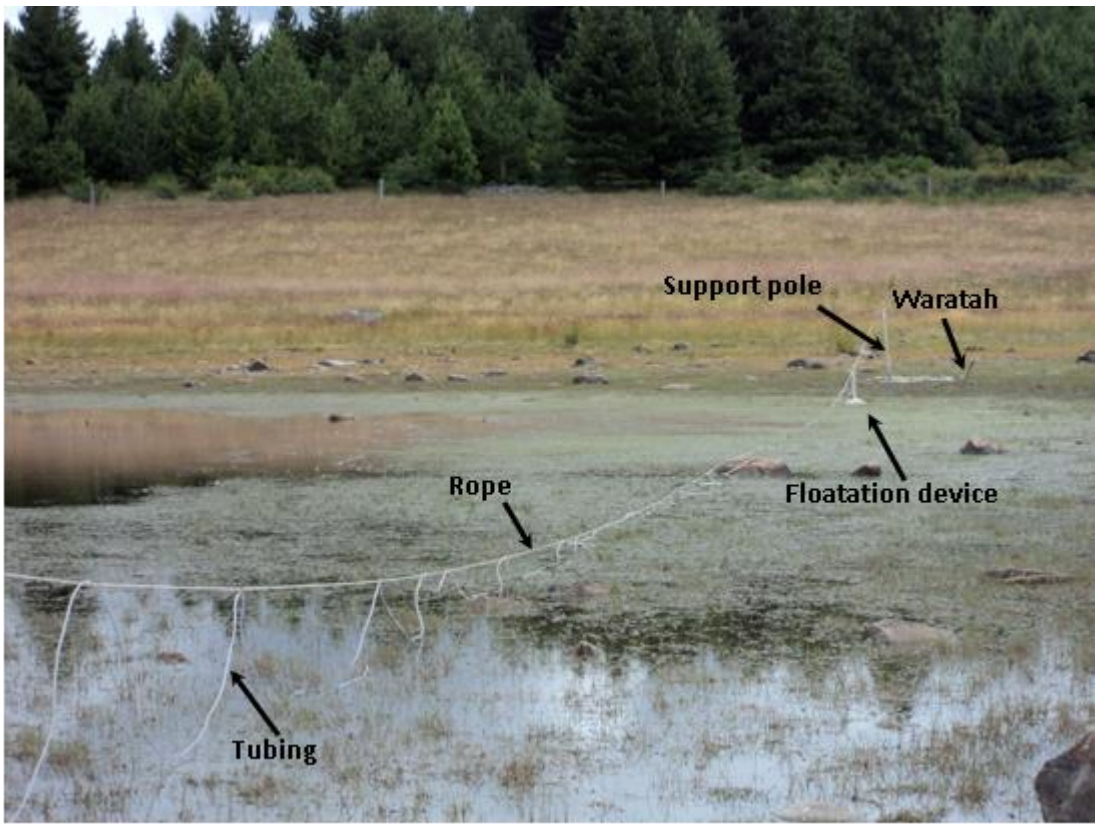

Figure 8. One of the support poles with attached rope and tubing of the sampling apparatus.

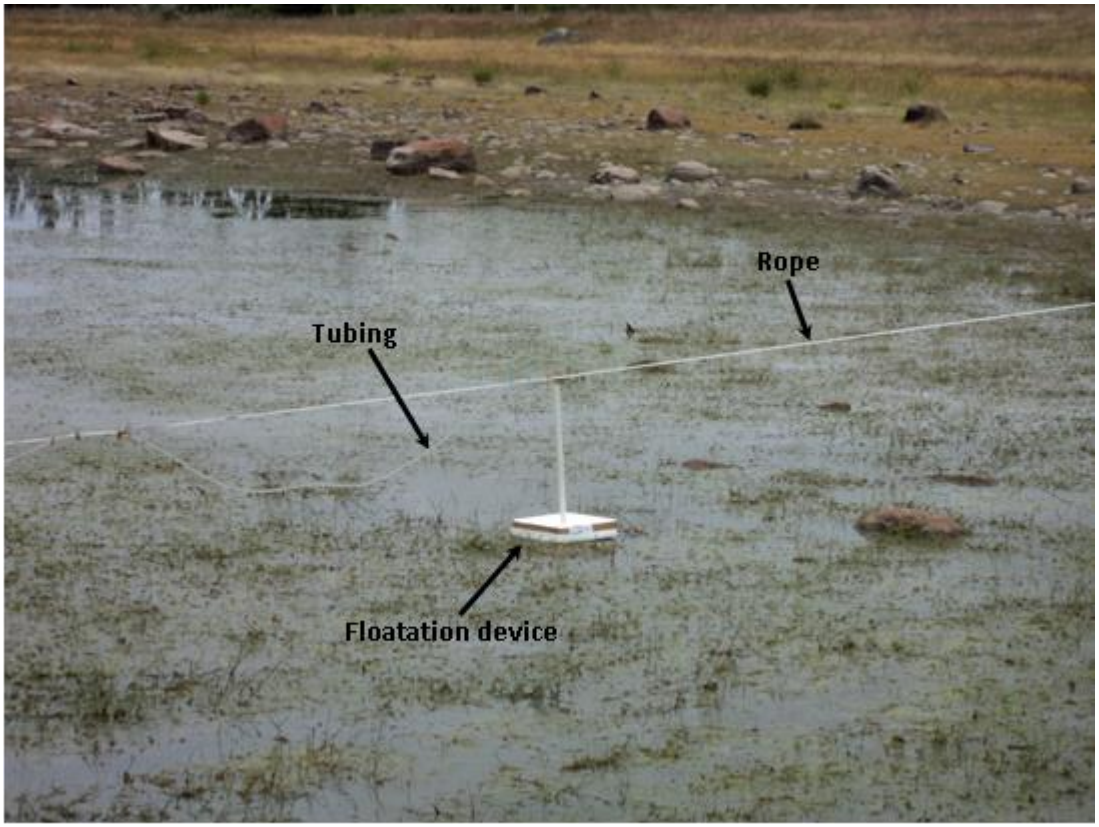

Figure 9. Floatation device that was used to collect water from a depth of 5 cm.

For Tarn 1, two 250 mL opaque sample bottles (1(A), 1(B)) were filled to capacity with water for each sample (Table 1). Two 250 mL bottles (3(A), 3(B)) were also collected for Tarn 3. Due to resource limitations, for Tarn 2, only one 250 mL bottle (2(A)) was collected per sample. However, an additional 15 mL of water was collected in a centrifuge tube (2(B)) for each sampling location in Tarn 2.

Table 1. Samples required for analyses and storage conditions.

Analysis	Source of sample	Storage
DNA analysis	1(A), 2(A), 3(A)	Stored on ice during sample collection then at 4 ℃
HPLC	1(B), 2(A), 3(B)	Stored on ice during sample collection then at -20 ℃
BIOLOG	1(A), 2(B), 3(A)	Stored on ice during sample collection then at 4 ℃

2.1.4 Physicochemical properties of the tarns

A range of physicochemical properties were measured onsite. These included; pH, dissolved oxygen, temperature and depth. To measure pH, 250 mL of water from each sample location was collected in a measuring cylinder and the pH was recorded using a portable pH meter (handylab pH 11, SCHOTT Instruments) within five minutes of sample collection. The dissolved oxygen content and the temperature of the tarns were measured using a dissolved oxygen probe (550A YSI, Yellow Springs, OH). The depths of the tarns were measured using a pole marked at 5 cm intervals. A minimum of 12 random sites in each tarn were analysed for the suite of physicochemical properties. Samples were also collected for chemical analysis. These samples were analysed by High Performance Liquid Chromatography (HPLC), by Joy Jiao, Lincoln University. HPLC allows the concentrations of various anions; chloride, bromide, nitrate-N, phosphate-P and sulphate-S to be determined.

2.2 Molecular methods

2.2.1 Filtration procedure

In preparation for DNA extraction and subsequent analysis, water samples were filtered using a Buchner filter (Figure 10) to collect bacterial cells on filters.

Figure 10. Buchner filter showing a Buchner funnel (Nalgene®, Thermo SCIENTIFIC, Rochester, NY, U.S.A.) and the waste collection container (KIMAX®, KIMBLE, Vineland, NJ, U.S.A.).

Using a Buchner filter, 100 mL of each sample was filtered through a new filter (Millipore Express® PLUS Membrane, polyethersulfone, hydrophilic, 0.22 µm, 47 mm, MERCK Millipore, Billerica, MA, U.S.A.). Using sterile forceps the filter was placed in a bead beater tube (Micro tube, 2 mL, Global Science and Technology, Auckland, New Zealand), containing 0.5 g of 0.1 mm zirconia/silica beads and 0.5 g of 2.3 mm zirconia/silica beads (BioSpec, Inc., Bartlesville, OK, U.S.A.). The bead beater tubes were then stored in a freezer at -20 °C. After each sample had been filtered, the filtering apparatus was rinsed with tap water and then distilled water. The forceps were sterilised between each sample to avoid the transfer of contaminants by dipping them in 100 % ethanol and then flaming them.

2.2.2 DNA extraction

In order to investigate the microbial diversity in the tarns the DNA had to first be extracted from the filters prepared in the previous procedure (see section 2.2.1). The DNA extraction procedure utilised in this study was adapted from Miller *et al.* (1999). To extract the DNA, 270 µL of phosphate buffer (100 mM, pH 8.0) and 300 µL of Sodium Dodecyl Sulphate (SDS) lysis buffer (100 mM, Tris pH 8.0, 10% SDS) were added to each bead beater tube containing a filter. The samples were then gently mixed by hand and then 300 µL of chloroform:isoamyl alcohol (24:1) was added to each tube. Tubes were shaken in a TissueLyser II (QIAGEN®, bio-strategy, Auckland, New Zealand) at 3 m s^{-1} for 1 min. The tubes were then transferred to the other side of the TissueLyser II and shaken for another minute. The tubes were transferred to a bench top centrifuge (Spectrafuge 24D; Labnet, Woodbridge, NJ, U.S.A.) and were centrifuged at full speed (13,300 rpm) for 5 min to pellet debris. The supernatant (approximately 650 µL) was transferred to a clean 1.5 mL microtube (AXYGEN, Auckland, New Zealand) and then 360 µL of 7 M ammonium acetate (NH_4OAc) added to the tubes to achieve a final concentration of 2.5 M. The tubes were gently mixed by hand and then centrifuged at full speed (13,300 rpm) for 5 min. After centrifugation, two distinct phases were observed, with the SDS forming a gel-like interphase. The upper phase (approximately 580 µL) was transferred to a clean 1.5 mL microtube and 315 µL of ice-cold isopropanol (Analar) was added to each tube. The tubes were then incubated at room temperature for 15 min before centrifugation at full speed (13,300 rpm) for 5 min to pellet the DNA. The supernatant was then discarded and the pellets were washed using 1 mL of 70% ethanol. The tubes were again centrifuged at full speed (13,300 rpm) for 5 min and the supernatant discarded. The pellets were then dried in a desiccator jar under vacuum for 15 - 45 min. Once the pellets were dry, they were resuspended in 50 µL of Nuclease-Free water (Promega, Sydney, Australia).

2.2.3 Polymerase chain reation (PCR)

PCR was performed on the DNA extracts to amplify bacterial DNA that was of interest to us in the present study. The PCR that was performed utilised primers designed for automated ribosomal intergenic spacer analysis (ARISA). ARISA primers (Invitrogen™ Ltd., Victoria, Australia) target a specific intergenic spacer (or 'inter-gene') region between the bacterial

16S rRNA and the 23S rRNA genes. This region is highly variable among different microbial species and thus a PCR utilising ARISA primers will result in the production of many fragments of different lengths. The resulting fragment pattern can therefore be used to characterize differences in microbial community composition among samples. For the PCR, a master mix containing all the different components required for a PCR was made (Table 2). The GoTaq® reagent (Hot Start Green Master Mix, 2X, Promega, Sydney, Australia) used in the master mix contained Taq DNA polymerase, dNTPs, $MgCl_2$ and reaction buffers at optimal concentrations.

Table 2. Reagents used to create a mastermix suitable for the PCR amplification of DNA for ARISA for one sample.

Reagent	Volume
Nuclease-Free water	18 µL
*Forward primer, SD Bact (10 µM)	2 µL
**Reverse primer, LD Bact (10µM)	2 µL
Bovine serum albumin BSA, 10% (Invitrogen™ Ltd., Victoria, Australia)	1 µL
GoTaq® (Hot Start Green Master Mix, 2X, Promega, Sydney, Australia)	25 µL
DNA	(2 µL)
Total volume (master mix)	48 µL
Total PCR volume	50 µL

*SD Bact (5'-TGCGGCTGGATCCCCTCCTT-3') (Invitrogen™ Ltd., Victoria, Australia)

** LD Bact (5'-CCGGGTTTCCCCATTCGG-3') (Invitrogen™ Ltd., Victoria, Australia)

For each sample, 48 µL of master mix was added to each well of a 96 well plate (Scientific Specialties Inc., Lodi, CA, U.S.A.) and then 2 µL of DNA extracts added to each well, except for two DNA-free controls which were included to ensure that PCR mastermixes were not contaminated with bacterial DNA. PCR was performed in a thermal cycler (Veriti, 96 Well Thermal Cycler, Applied Biosystems, Victoria, Australia) under the following conditions listed in Table 3.

Table 3. Conditions under which the PCR was performed.

Temperature (°C)	Time (minutes)	Cycles
95.0	5.00	1
95.0	0.30	30
61.5	0.30	
72.0	1.30	
72.0	7.00	1
Hold at 15.0		∞

2.2.4 Agarose gel electrophoresis

Agarose gel electrophoresis was performed on the PCR samples to confirm that the DNA extraction and PCR both worked. An agarose gel was prepared by combining 1.5 g of agarose (AppliChem low EEO powder, Auckland, New Zealand) in 100 mL of 1X Tris/Borate/EDTA (TBE) buffer in a 500 mL Schott bottle. The agarose was dissolved by heating in a microwave for two minutes, or until fully dissolved. The liquid agarose was allowed to cool to the touch before 10 µL of SYBR® safe DNA stain (invitrogen, Victoria, Australia) was added and then the solution was poured into a rectangular gel mould and left to set at room temperature (20 min). The solid agarose gel was placed into the electrophoresis apparatus (ENDURO, Labnet, Woodbridge, NJ, U.S.A.) and 1X TBE was added until the gel was completely submerged. The gel was then loaded with 6 µL of 1 Kb Plus DNA ladder (E-Gel®, invitrogen, Victoria, Australia) into the first well. With the samples, 6 µL of each sample (including the negative controls) was loaded directly into individual wells of the gel because the GoTaq® master mix already contains loading dye. Electrophoresis was performed at 110 V for 30 to 40 min. The gel was then visualised under UV light using a gel doc (UVItec Limited, Cambridge, UK).

2.2.5 DNA purification

The agarose gel electrophoresis confirmed that DNA was present in the samples and no contamination was observed in the negative controls, therefore DNA purification was performed. DNA was purified (using a DNA Clean & Concentrator^{TM}-5 Kit; ZYMO RESEARCH, Irvine, CA, U.S.A.) to remove the primers and other unwanted DNA sequences in the samples.

2.2.6 Automated ribosomal intergenic spacer analysis (ARISA)

After DNA purification had been performed 1 µl of each sample was combined with 10 µl HiDi formamide and an internal LIZ1200 standard (Applied Biosystems Ltd., New Zealand). The samples were then heat treated (95 °C, 5 min) and cooled on ice. To generate ARISA profiles, the samples were run on a 3130XL Capillary Genetic Analyser (Applied Biosystems Ltd.) using a 50 cm capillary and standard Genemapper protocol, but with an increased run time (15 kV, 65 000 s).

2.3 Functional analysis (carbon substrate utilisation)

To provide a measure of the functional capability of the bacterial community in each sample, BIOLOG EcoPlates™ (BIOLOG Inc, Hayward, CA, U.S.A., Figure 11) were inoculated by aliquoting 120 µL of each sample into individual wells which contain a range of carbon substrates (Appendix A). The plates were incubated in the dark, at room temperature, for approximately 48 hours. The absorbance of the colour produced in the individual wells was measured at a wavelength of 590 nm using a plate reader (BMG LABTECH FLUOstar Omega model). The absorbance recorded for the well containing water was used as a blank and the value was subtracted from the values for the other wells.

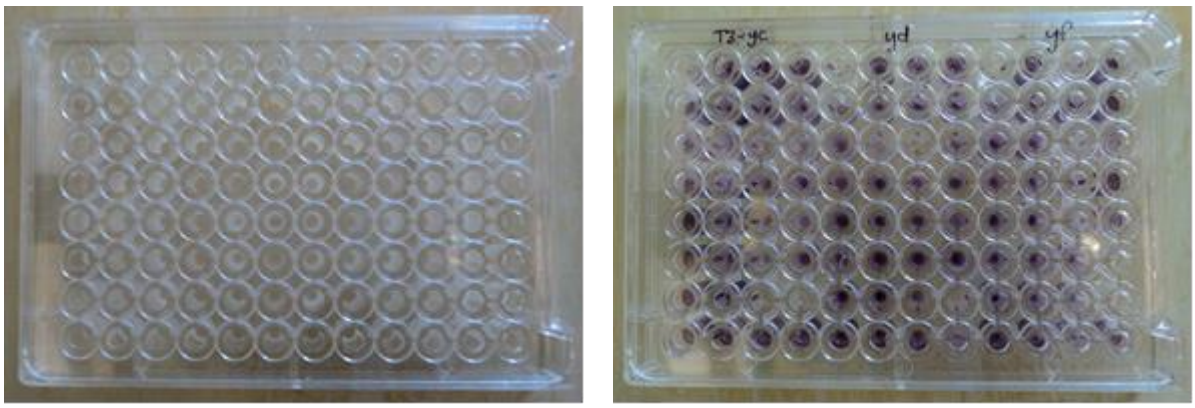

Figure 11. Unused BIOLOG EcoPlate™ (left) and BIOLOG EcoPlate™ (right) showing colour development.

2.4 Statistical analysis

GENEMAPPER software (v 3.7; Applied Biosystems Ltd.) was used to produce electropherograms from the fluorescence data obtained during the ARISA procedure. The electropherograms provide a comparison of the proportional quantities of different-sized DNA fragments in each sampled community (Lear *et al.*, 2008). The protocol of Ramette *et al.* (2009) was then used to identify 'true peaks' (i.e., removing background 'noise' generated during automated analysis) and bin fragments of similar size. Each number greater than zero, in the tabulated ARISA data (Appendix B), represented a detected taxon. The ARISA and BIOLOG data were standardised, in which relative percentages, disregarding the effect of total abundance (ARISA) or intensity (BIOLOG) were calculated. A Bray-Curtis similarity matrix (Legendre and Legendre, 1998) was then produced for both bacterial community structure (ARISA data) and function (BIOLOG data) in PRIMER (version 6.1.12; Primer-E Ltd., Plymouth, UK) to allow a quantitative description of the relative differences among the samples. The following Bray-Curtis equation was used to produce the matrices;

$$Bray\ Curtis\ similarity = 100\left(1 - \frac{\sum_i |y_{i1} - y_{i2}|}{\sum_i y_{i1} + \sum_i y_{i2}}\right), \qquad \text{eq. 1}$$

where y_{i1} is the standardised peak size of taxon *i* from sample 1 and $y_{i2}$is the standardised peak size of taxon *i* from sample 2 (Clarke and Gorley, 2006). This was calculated for all possible pairs of samples.

2.4.1 Evaluation of the taxa-area hypothesis

To determine whether positive taxa-area relationships existed in the tarns, a species accumulation curve was plotted (using the tabulated ARISA data) for each tarn using the 'exact' argument (Kindt *et al.*, 2006) in the 'specaccum' function in the 'vegan' package in R (version 2.15.2) (R_Core_Team, 2012). A taxa-area curve usually involves species richness being plotted against area sampled. However, with this study, the method that was used to determine bacterial richness, ARISA, only has resolution to taxa (ARISA peak) level, not species level, and therefore taxa richness was plotted instead of species richness. The cumulative number of samples was used to represent sample 'area' because each analysed sample was approximately the same volume of water (100 mL) (samples were taken approximately 7 m apart from each other in a grid format). After the curves were plotted, a power law function was fitted to the data using the non-linear regression function 'nls' in R (version 2.15.2) (R_Core_Team, 2012) (see Appendix C for R script). For Tarn 3, the samples collected from the smaller grid format of 3.5 x 3.5 m were excluded to allow a direct comparison of these data with data collected from Tarns 1 and 2, from which such samples were not collected.

2.4.2 Evaluation of the distance-decay hypothesis

Distance-decay plots were produced for both the ARISA and BIOLOG data. The 'vegan' package in R (version 2.15.2) (R_Core_Team, 2012) was used to calculate a Bray-Curtis similarity matrix (see eq. 1) for both bacterial community structure (ARISA data) and function (BIOLOG data). These similarity matrices were then used to compare similarity in bacterial community data with their geographic (straight-line) distance apart. For Tarn 3, the samples that were collected from the smaller grid format of 3.5 x 3.5 m were included in the analysis to increase the sensitivity of the approach. For bacterial community structure an exponential curve was fitted to the data using the non-linear regression function 'nls' in R (version 2.15.2) (R_Core_Team, 2012). The function took the form, $S = e^{a\ x\ distance}$, where S is the pairwise Bray-Curtis similarity, a is a constant and distance is pairwise distance. A linear function was fitted to the data for bacterial community function (see Appendix D for R script).

2.4.3 Evaluation of the space vs environment hypothesis

Variation among the tarns

Using PRIMER (version 6.1.12; Primer-E Ltd., PERMANOVA+ add on, Plymouth, UK), the overall variation in the tarns was visualised by producing non-metric multi-dimensional scaling (MDS) plots. Two-hundred and fifty re-starts were conducted for both bacterial community structure and function using the respective Bray-Curtis matrix. MDS plots are 2-d or 3-d ordination plots in which the relative distances apart of all points are in the same rank order as the relative Bray-Curtis similarities of the samples. This results in samples that have more similar bacterial community structure or function being located closer together. Permutational MANOVA (PERMANOVA) was then performed to determine if there were significant differences in bacterial community structure and function among the tarns, and to identify the tarns with the most similar/dissimilar communities. Type III (partial) analyses with 9999 unrestricted permutations of the raw data were used (Lear *et al.*, 2008). Like the more traditional, multivariate analysis of variance (MANOVA), PERMANOVA is used to determine if the means of several groups are significantly different. The advantages of using PERMANOVA over MANOVA are that PERMANOVA can be used for unequal group sizes and also, because PERMANOVA is a permutation method, it is unaffected by the statistical distribution of the samples.

To visualise spatial differences in bacterial community structure and function among the tarns, contour plots of the data were constructed for each study site. The values plotted on the contour plots were the 1-d 'configuration scores' sourced from the MDS plots of the ARISA and BIOLOG Bray-Curtis for all tarns. The data from the contour plots were interpolated using the inverse distance weighting (IDW) interpolation function in the Geostatistical Wizard, ArcMap10 (Esri, Wellington, New Zealand). The IDW function predicts values for unmeasured locations by using measured values surrounding the prediction locations. More weight is given to values close to the prediction locations than values that are further away. After the predicted values had been calculated, the range of values were classified into ten even classes and each class was assigned a colour from a colour gradient (lowest value = dark blue to highest value = light green) with more similar colours representing more similar values. Similar plots were produced by Jones at al. (2012) (see section 1.3). The result of this analysis was that for each tarn, a 'heat map' was produced, in

which sampling locations that contained similar bacterial communities were represented by a similar colour.

Differences in environmental variables among the three tarns were determined by calculating the mean concentration and standard error of each environmental variable (pH, conductivity, total carbon, chloride, nitrite-N, bromide, nitrate-N, phosphate-P, sulphate-S) and performing Tukey's test in MiniTab 16 to confirm significant differences ($P \leq 0.05$). Tukey's test identifies means that are significantly different by comparing all possible pairs of means.

Variation within the tarns

The following analyses were all performed in PRIMER (version 6.1.12; Primer-E Ltd., PERMANOVA+ add on, Plymouth, UK). An MDS plot was produced for each tarn, based on a resemblance matrix constructed from Bray-Curtis similarity data. For the same sample data a cluster diagram, or dendogram was plotted using the 'group average' approach. Bacterial samples that were determined to be less than or equal to 50 % similar using this approach were then marked as separate clusters on the MDS plots. Clusters were not overlaid on the MDS plots for bacterial community function because all the samples for each tarn were constrained in one cluster at the 50 % similarity level (Appendix E). The significant differences in bacterial community structure among the clusters were investigated using PERMANOVA (type III (partial), unrestricted permutation of raw data, 9999 permutations). The total variation in bacterial community structure among the samples within each tarn was analysed using the multivariate dispersion (MVDISP) function which calculates a factor dispersion value in which larger values are indicative of greater variation in the samples. The clusters were plotted on maps of each tarn (where different data clusters were represented by different colours) allowing the visualisation of the position of the clusters within the tarns. The same approach that was used to produce the contour plots showing the variation among the tarns (see section 2.4.3; 'Variation among the tarns') was used to produce contour plots for both bacterial community structure and function within the individual tarns, using the ARISA and BIOLOG Bray-Curtis data for the individual tarns to identify any spatial patterns present.

To analyse the environmental variation in the tarns, regression analyses were performed using distance-based linear models (DistLM) as used by Lear *et al.* (2008) to identify any environmental variables that had a significant relationship with bacterial community structure or function after any correlated environmental variables (Draftsmans plot, $r < -0.05$; $r > 0.05$) had been removed. DistLM was carried out using the following parameters in the Primer-E program: selection procedure; 'Forward selection', which adds one variable, the variable that improves the selection criterion the most, at each step until no improvement in the selection criterion is possible; selection criteria, ' R^2', which is the proportion of explained variation and finally 9999 permutations were performed. DistLM is a regression analysis that models the relationship between a resemblance matrix (e.g. Bray-Curtis similarity matrix) and a set of predictor variables, which in this study were a range of environmental variables. For any environmental variable that was identified as having a significant relationship with bacterial community structure or function, the IDW interpolation function in the Geostatistical Wizard, ArcMap10, was used to produce a contour plot from the raw data for the environmental variable. This allowed the visualisation of any spatial patterns present in these environmental variables in the tarns.

Finally, variance partitioning was performed using DistLM to determine the proportion of total variation explained by either spatial location, environment variables or a combination of both for each tarn separately. The spatial factors and environmental variables were divided into two categories using the 'Indicator' function in PRIMER (version 6.1.12; Primer-E Ltd., PERMANOVA+ add on, Plymouth, UK). One of the categories, Space (S) consisted of the complete trend surface regression (Eastings (E), Northings (N), E^2, N^2, EN, E^2N, N^2E, E^3, N^3) (Legendre and Legendre, 1998) and the other category, Environment (E) included any uncorrelated (see section 2.4.3 'Variation within tarns') environmental variable data. DistLM analysis was carried out using the following parameters in the Primer-E program: Group variables (S and E); selection procedure; 'All specified', which fits the predictor variables (S and E) in the order given under the 'Force inclusion' column in the 'Selection' dialog, 'R^2' and finally 9999 permutations were performed. Dividing the data into two categories, S and E, allowed the proportion of total variation explained by one of these categories to be determined directly from the DistLM output. This value was used to manually calculate the proportion of total variation explained by the other category and space and environment combined. Any variation that was not accounted for by space, environment or a combination

of the two was identified as unexplained variation (Anderson *et al.*, 2008). More details of this approach are provided in Appendix F. For each tarn, the proportion of total variation attributable to pure spatial variation that is independent of any environmental variables, pure environmental variation that is independent of any spatial factors, spatially structured environmental variables, and unexplained variation was visually represented on a column graph (Figure 12) produced in Excel (2010).

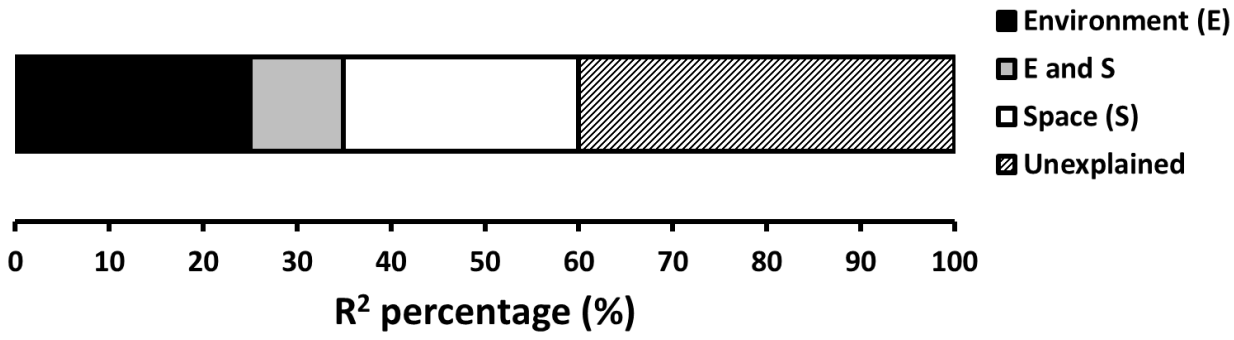

Figure 12. Example of variance partitioning displayed on a column plot. Four different components are shown: (hatched) unexplained variation; (black) pure environmental variation that is independent of any spatial factors; (white) pure spatial variation that is independent of any environmental variables; (grey) spatially structured environmental variables.

Chapter 3

Results

3.1 Overview

Three tarns located in the Tekapo Scientific Reserve, New Zealand were sampled. The samples in the three tarns were collected over a total range of 2 to 61 metres. The mean number of taxa in each tarn was 54, 71 and 68 for Tarns 1, 2 and 3, respectively. In addition, the maximum number of taxa in each tarn was 83, 101 and 93, and the minimum number of taxa was 12, 17 and 29 in Tarn 1, Tarn 2 and Tarn 3 respectively. The mean % Bray-Curtis similarity values for bacterial community structure were 32, 44 and 42, and for bacterial community function were 78, 72 and 74 comparing samples within Tarn 1, Tarn 2 and Tarn 3 respectively.

3.2 Taxa-area relationship

A positive taxa-area relationship was observed in each tarn with the following equations, for Tarn 1 (± standard error): $S = 81 \pm 2.4 * A^{0.32 \pm 0.010}$, for Tarn 2 (± standard error): $S = 98 \pm 2.9 * A^{0.29 \pm 0.010}$ and for Tarn 3 (± standard error): $S = 101 \pm 2.4 * A^{0.26 \pm 0.007}$ (Figure 13).

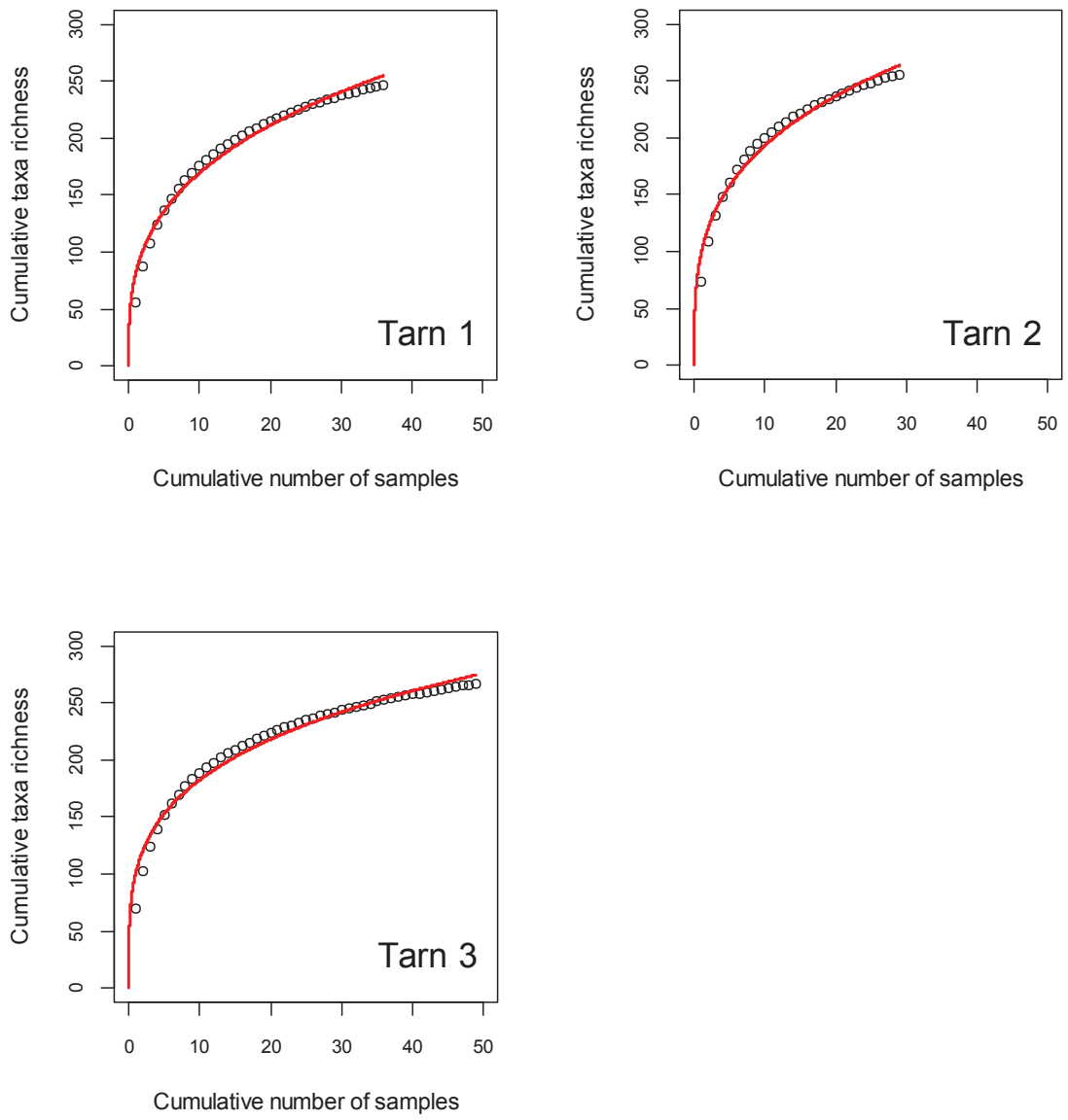

Figure 13. Taxa-area relationship for each tarn. The cumulative number of samples (mean taxa richness was calculated for each combination of samples) was used to represent area. The relationship in each tarn was modelled by the power law ($S = cA^z$, red line = fitted values).

3.3 Distance-decay relationship

A decline in the similarity in bacterial community structure between paired samples with increasing geographic distance was observed (Figure 14). The coefficient for the exponential curve for bacterial community structure in each tarn was -0.06 (± 0.002 S.E.) for Tarn 1, -0.04 (± 0.001 S.E.) for Tarn 2 and -0.04 (± 0.001 S.E.) for Tarn 3. In contrast, no significant declines in similarity were observed with increasing geographic distance for bacterial community function for Tarns 1 and 3 ($P > 0.05$). However, the slope of the relationship for Tarn 2 was significant ($P = 0.004$). The slope coefficients of the linear functions for bacterial community function in each tarn were 0.00008 (± 0.0002 S.E.) for Tarn 1, -0.00113 (± 0.0004 S.E.) for Tarn 2 and -0.00007 (± 0.0001 S.E.) for Tarn 3. There was more variation in bacterial community structure than function, as measured by the range in Bray-Curtis similarity among tarns. The overall mean ranges in Bray-Curtis similarity across the three tarns were 78 % in compositional similarity compared to 44 % in functional similarity.

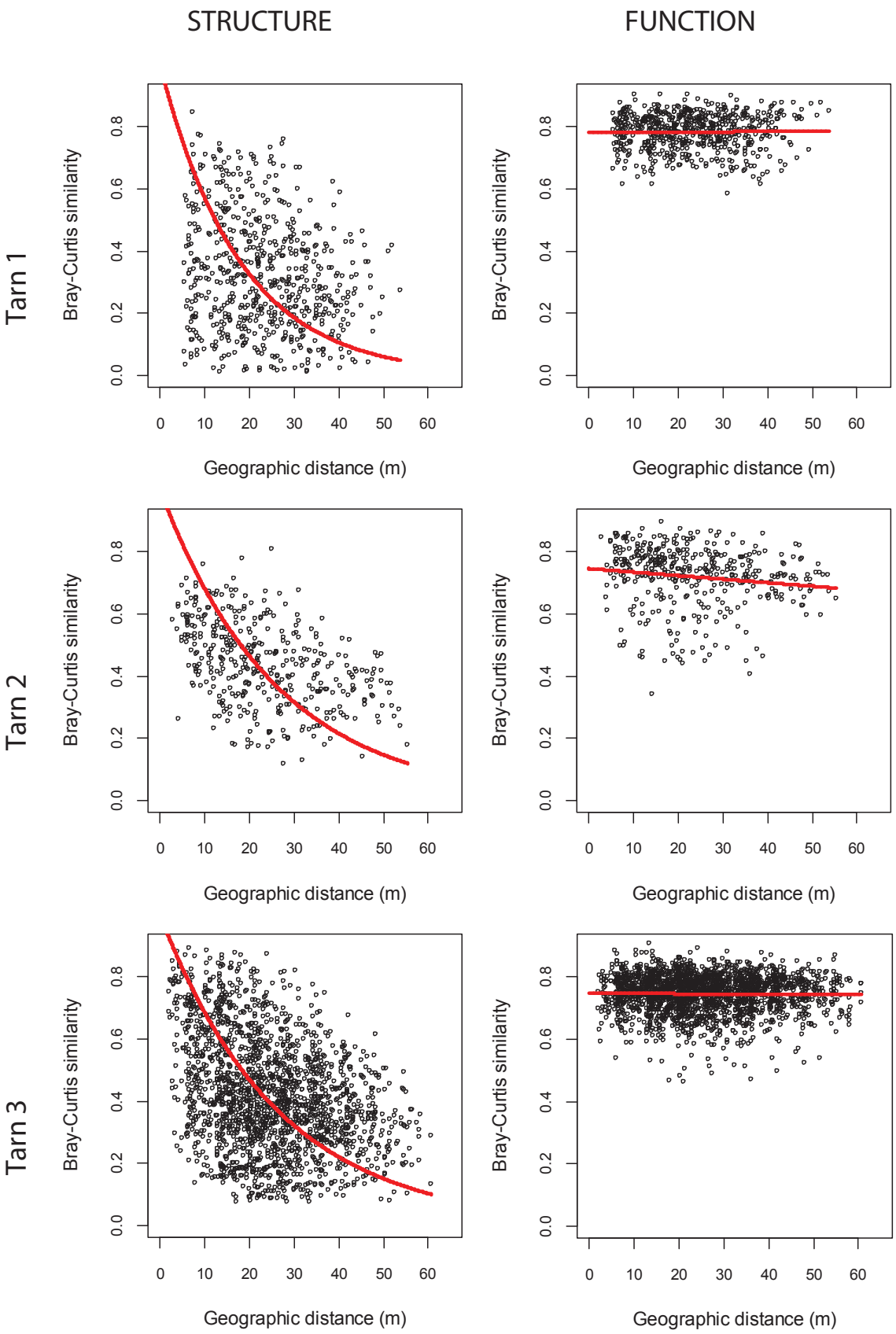

Figure 14. Distance-decay curve for bacterial community structure and function in each tarn. Each data point represents the Bray -Curtis similarity score for paired samples (y-axis) and their respective geographic distance (x-axis). For bacterial community structure, an exponential curve was fitted to the data in each tarn . For bacterial community function a linear regression line was fitted.

3.4 Spatial and environmental variation

3.4.1 Among tarn variation

The three tarns differed significantly in bacterial community structure (PERMANOVA, *P* = 0.0001). This is visually represented on the MDS plot for bacterial community structure in which three separate distinct clusters can be observed (Figure 15). In addition, Tarns 2 and 3 were identified as having the most similar bacterial community structure with a mean Bray-Curtis similarity of 22.6 % and Tarns 1 and 2 were identified as having the most dissimilar bacterial community structure with a mean Bray-Curtis similarity of 13.4 %. Tarns 1 and 3 had a mean Bray-Curtis similarity of 16.9 % for bacterial community structure.

The clusters representing bacterial community function in each tarn overlap each other in the MDS plot for bacterial community function indicating that there is some similarity in bacterial community function in the three tarns (Figure 15). However, bacterial community function in the three tarns was still identified as being significantly different (PERMANOVA, *P* = 0.0001). Tarns 1 and 3 were identified as having the most similar bacterial community function with a mean Bray-Curtis similarity of 74.8 % and Tarns 1 and 2 were identified as having the most dissimilar bacterial community function with a mean Bray-Curtis similarity of 66.1 %. Tarns 2 and 3 had a mean Bray-Curtis similarity of 67.5 % for bacterial community function.

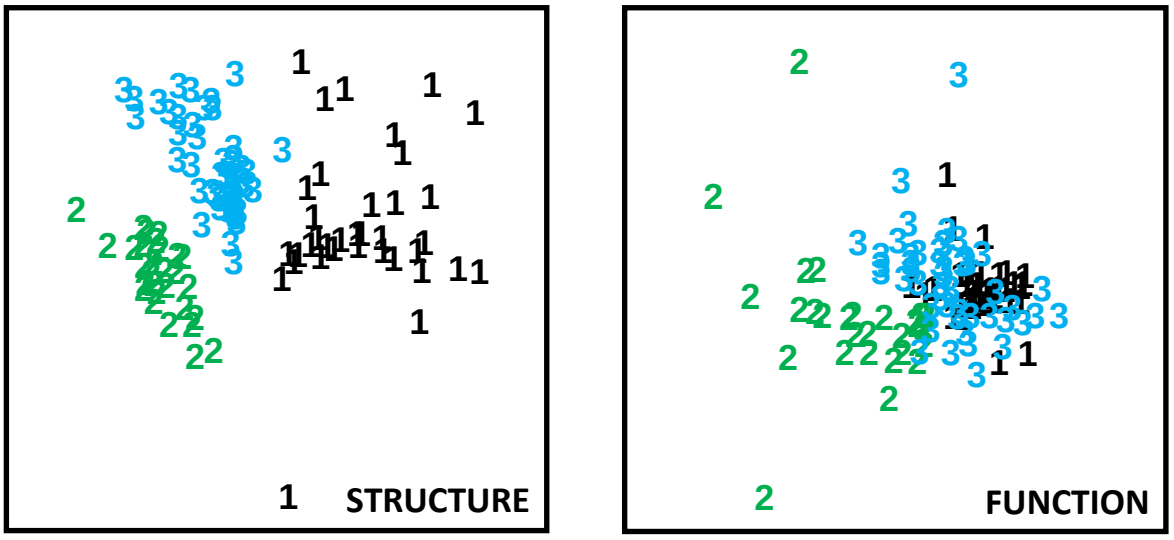

Figure 15. Differences in bacterial community structure (left) and bacterial community function (right). Plots are non-metric multidimensional scaling of ARISA data and carbon substrate utilisation data, using Bray-Curtis similarity matrices of samples. Data points

relate to samples abstracted from Tarn 1, 2 or 3. Two-dimensional stress values are 0.20 and 0.18 for plots of bacterial community structure and function, respectively.

The contour plots indicate that both bacterial community structure and function show significant spatial variation among the different tarns (Figure 16). In agreement with the Bray-Curtis similarity values calculated during the PERMANOVA analysis (see section 3.4.1), Tarns 1 and 2 have the most dissimilar bacterial community structure and function, i.e. the most different colours along the colour gradient. No strong spatial patterns can be observed with any of the tarns. This may be due to a lack of sensitivity because the MDS values used to produce the contour plots were calculated for all tarns, and thus colour differences on the maps largely represent differences in the data occurring among, rather than within, the tarns.

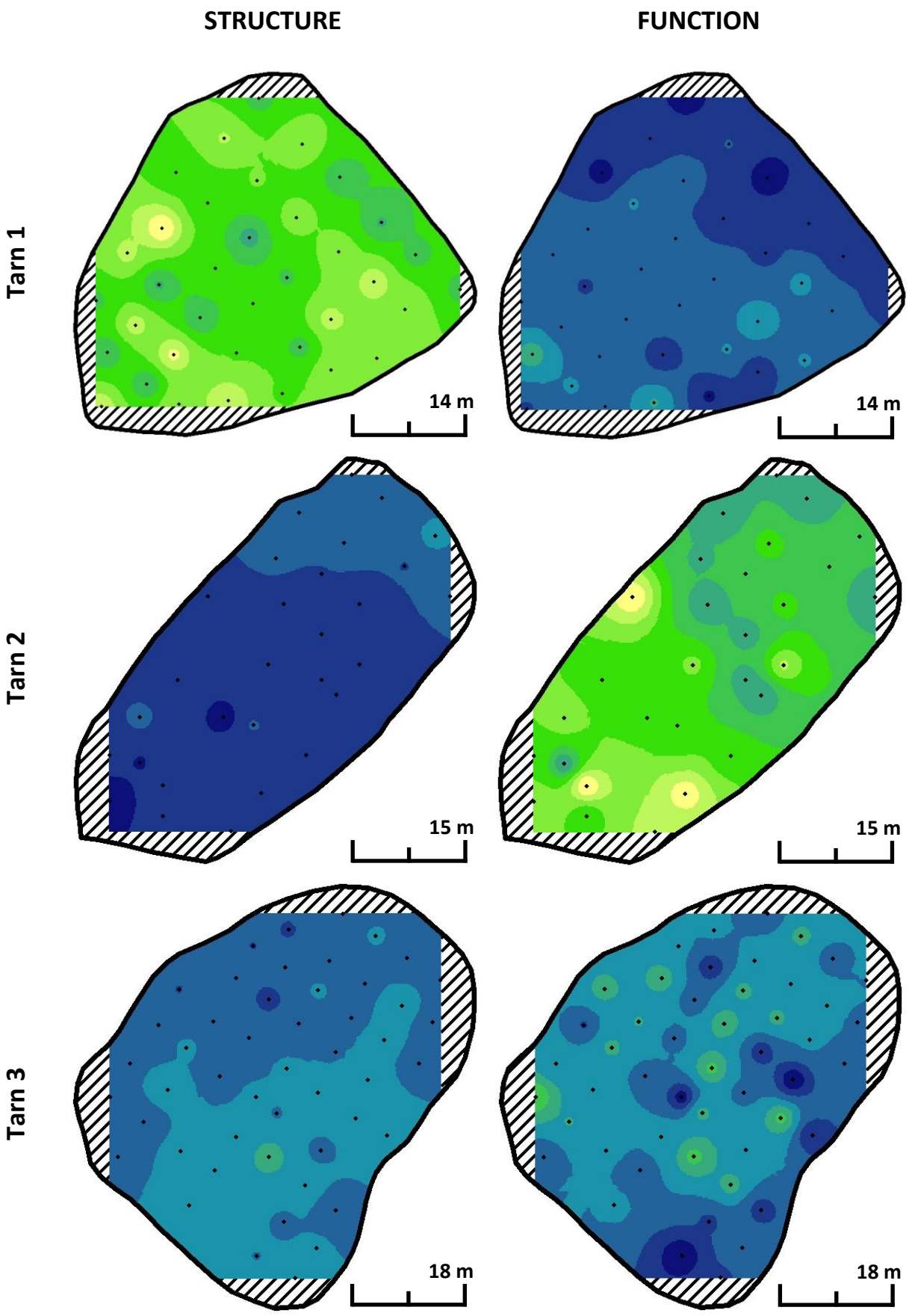

Figure 16. Similarity in bacterial community structure and function among tarns. Bacterial community data for each tarn were subjected to a data reduction procedure by non-metric multidimensional scaling of Bray-Curtis similarity data. The differences between the

highest and lowest 1-d configuration scores for each plot were then used to classify the configuration into ten equally sized classes. The bacterial community data falling within each class was assigned a different colour, across a gradient from dark blue (lowest 1-d configuration score) to light green (highest 1-d configuration score). The outcome of this approach is that samples hosting more similar bacterial community data are represented by more similar colours on each map.

A number of environmental variables (pH, conductivity, total carbon, chloride) were determined to have significantly different mean concentrations for the three tarns (Tukey's test $P < 0.05$, Table 4). In general, the nutrient levels, disregarding chloride, were relatively consistent across the tarns. Significant differences were detected for a few of the nutrient levels, such as nitrite-N for Tarn 3, nitrate-N for Tarn 1 and sulphate-S for Tarn 1, among the tarns (Tukey's test $P < 0.05$). However, because the concentrations were so low the differences are comparatively minor.

Table 4. Average water chemistry obtained for samples abstracted from each tarn. Data are means ± standard error.

	Tarn 1	Tarn 2	Tarn 3
pH	7.66 ± 0.74^{a}	6.68 ± 0.19^{b}	8.48 ± 0.77^{c}
Conductivity (µs cm^{-1})	82 ± 17^{a}	49 ± 4^{b}	106 ± 11^{c}
Total carbon (ppm)	59 ± 7^{a}	41 ± 5^{b}	34 ± 4^{c}
Chloride (mg l^{-1})	6.9 ± 1.8^{a}	2.1 ± 0.4^{b}	10.4 ± 1.1^{c}
Nitrite-N (mg l-1)	0.013 ± 0.004^{a}	0.012 ± 0.002^{a}	0.010 ± 0.004^{b}
Bromide (mg l^{-1})	0.065 ± 0.016^{a}	*N.D.*	0.068 ± 0.021^{a}
Nitrate-N (mg l^{-1})	0.26 ± 0.06^{a}	0.22 ± 0.03^{b}	0.24 ± 0.04^{b}
Phosphate-P (mg l^{-1})	0.11 ± 0.04^{a}	*N.D.*	0.10 ± 0.05^{a}
Sulphate-S (mg l^{-1})	1.09 ± 0.21^{a}	0.90 ± 0.17^{b}	0.95 ± 0.19^{b}

N.D. are values that were below detection limits in each row.
Data in columns not linked by the same letter (a, b or c in superscript) are significantly different (Tukey's test; $P < 0.05$).

3.4.2 Within tarn variation

Cluster plots were produced to analyse the variation in bacterial community structure and function within each tarn (Figure 17). For variation in bacterial community function, MDS plots were not superimposed with group average clusters because at the chosen similarity level of 50 %, for each tarn, all the samples were constrained within one cluster (Appendix E). The MDS plots with superimposed group average clusters for bacterial community structure indicate that six, five and five main clusters were observed for Tarns 1, 2 and 3 respectively (Figure 18). Tarn 1 had a significant number of samples that did not fall into the main clusters because of significant differences in Bray-Curtis similarities. This indicates significant variation (multivariate dispersion index of 1.2) among the samples in Tarn 1. In contrast, most of the samples for Tarns 2 and 3 fell within the main clusters, which indicates less variation (multivariate dispersion index of 0.80 and 0.95 respectively) within these tarns than for Tarn 1.

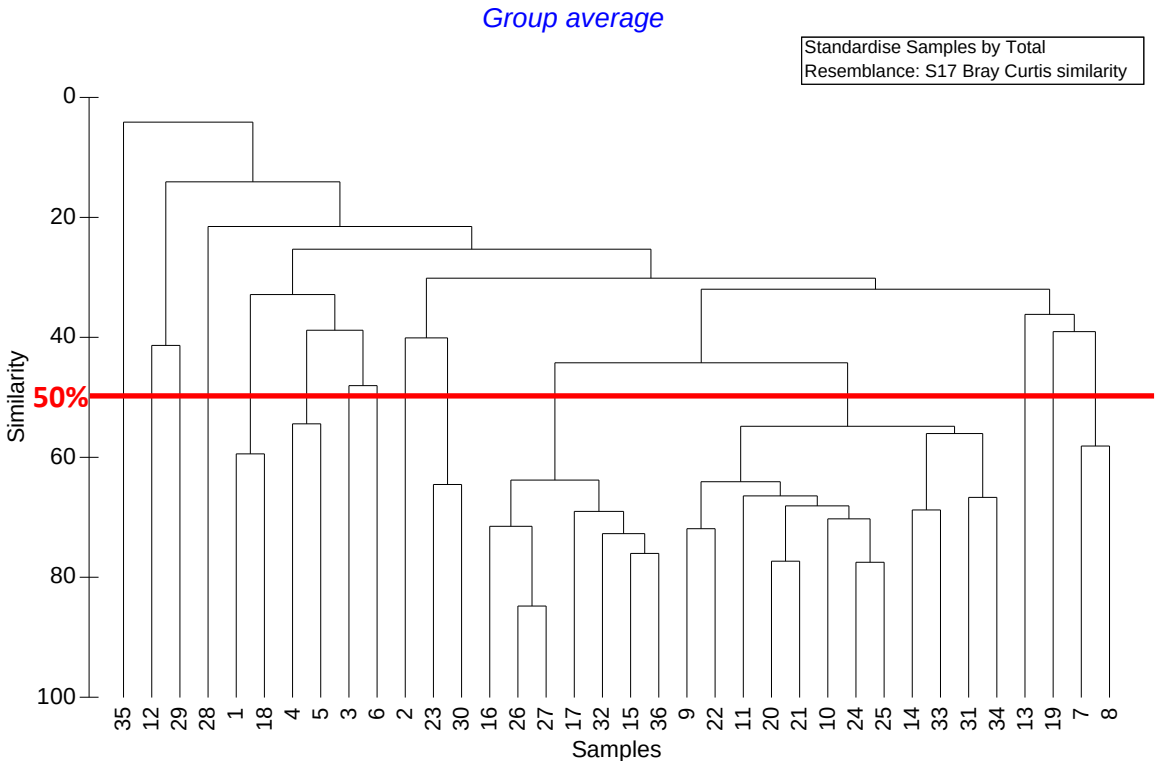

Figure 17. Example of a dendrogram for bacterial community structure in Tarn 1 showing the cut off at 50 % (red line) Bray-Curtis similarity. Each branch that the red line crosses is a different cluster in the tarn at the 50 % Bray-Curtis similarity level.

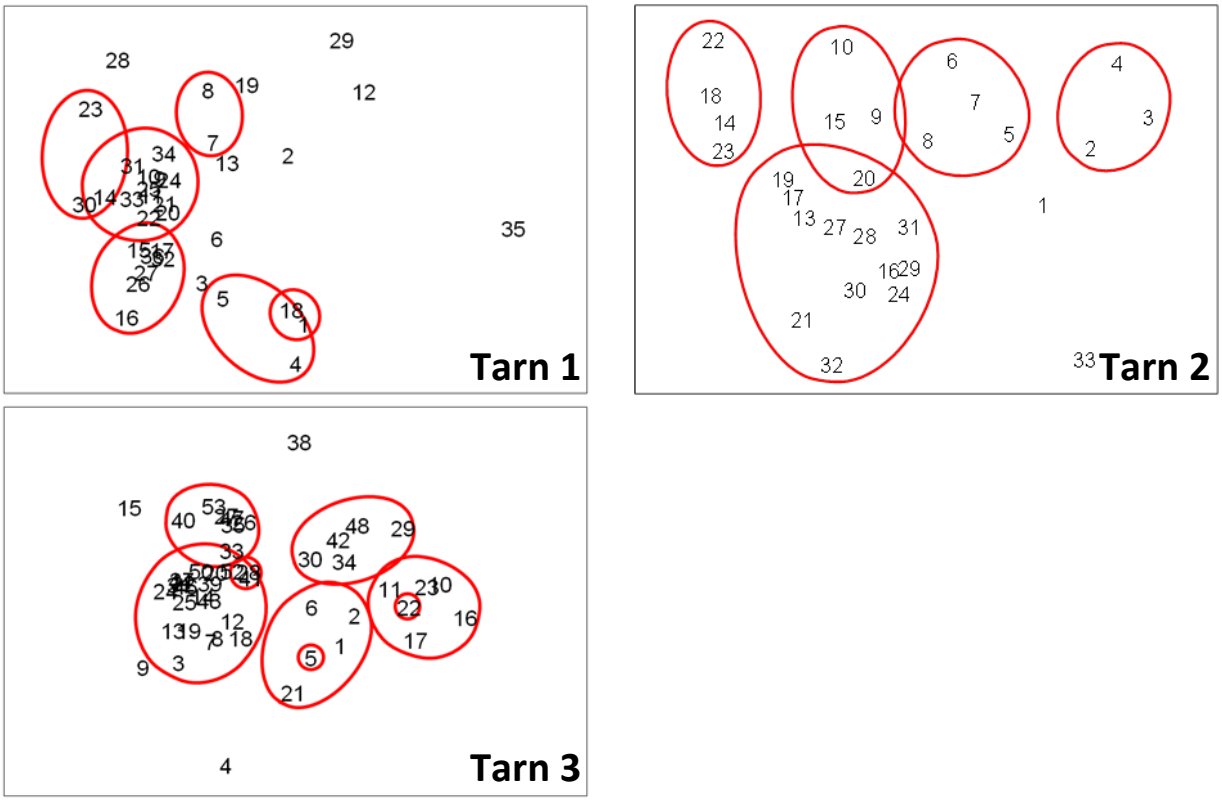

Figure 18. Variation in bacterial community structure in each tarn. Plots are derived from non-metric multidimensional scaling of ARISA data using a Bray-Curtis similarity measure. Group average clusters are superimposed on each plot (red circles) at the level of 50 % Bray-Curtis similarity. Each sample that is not constrained was in its own cluster. Two-dimensional stress values are 0.16, 0.13 and 0.13, respectively.

A side-by-side comparison of MDS contour plots and maps containing the position of the group average clusters was used to visualise patterns in bacterial community structure in the tarns (Figure 19). Dividing the samples into clusters allows any spatial variation within the tarns to be visualised. Tarn 1 shows the most variation in bacterial community structure among samples, with no distinct group average clusters. Samples represented by more similar colours, i.e. more similar bacterial communities, on the contour plots in Tarns 2 and 3, are grouped closely together and are located in distinct areas in the tarns. For Tarn 1, the clusters that had the most similar bacterial community structure were the light green and yellow clusters (average similarity between groups, 44.5 %) closely followed by the green and yellow clusters (average similarity between groups, 44.4 %). The green and red clusters had the most dissimilar (average similarity between groups, 14.5 %) bacterial community structure. In Tarn 2, the clusters that had the most similar bacterial community structure were the orange and light orange clusters (average similarity between groups, 49.2 %). These two clusters are located next to each other in the tarn. The clusters with the most

dissimilar bacterial community structure were the light green and red clusters (average similarity between groups, 19.9 %). In Tarn 3, the green and light orange clusters had the most similar (average similarity between groups, 49.8 %) bacterial community structure and the clusters with the most dissimilar bacterial community structure, were clusters, orange and light orange (average similarity between groups, 18.4 %).

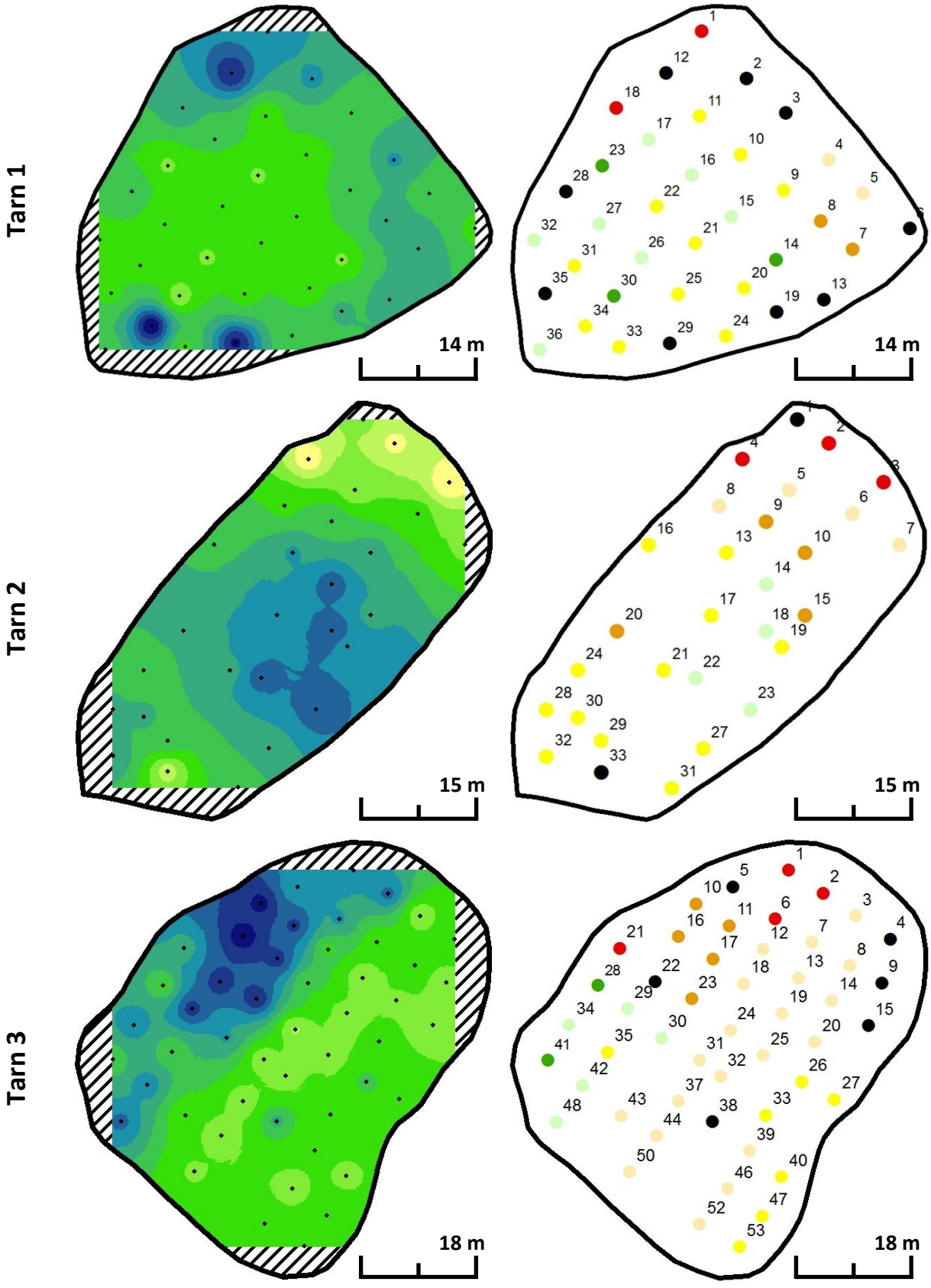

Figure 19. The plots on the left show similarity in bacterial community structure within the tarns. Bacterial community data for each tarn were subjected to a data reduction procedure by non-metric multidimensional scaling of Bray-Curtis similarity data. The differences between the highest and lowest 1-d configuration scores for each plot were then used to classify the configuration into ten equally sized classes. The bacterial

community data falling within each class was assigned a different colour, across a gradient from dark blue (lowest 1-d configuration score) to light green (highest 1-d configuration score). The outcome of this approach is that samples hosting more similar bacterial community data are represented by more similar colours on each map. The plots on the right show the physical location of the different clusters within the tarns. Each colour represents a different cluster with black representing individual samples that did not fall within the main clusters.

Another set of MDS contour plots (using MDS 1-d configuration scores from plots of the individual tarns) was produced for both bacterial community structure and function to allow any spatial patterns to be identified within the tarns (Figure 20). The MDS contour plots for the tarns show that with bacterial community structure, there appears to be a strong spatial pattern present in each tarn. First, in Tarn 1, the bacterial community structure in the centre of the tarn is most dissimilar to the bacterial community structure around the edge of the tarn. Second, for Tarn 2, there appears to be a gradient in bacterial community structure with the bacterial community structure becoming more dissimilar with increasing distance. Finally, in Tarn 3, the bacterial community structure appears to be divided into two, significantly different, adjacent groups. In contrast to bacterial community structure, with bacterial community function, there appear to be no strong spatial patterns present in any of the tarns.

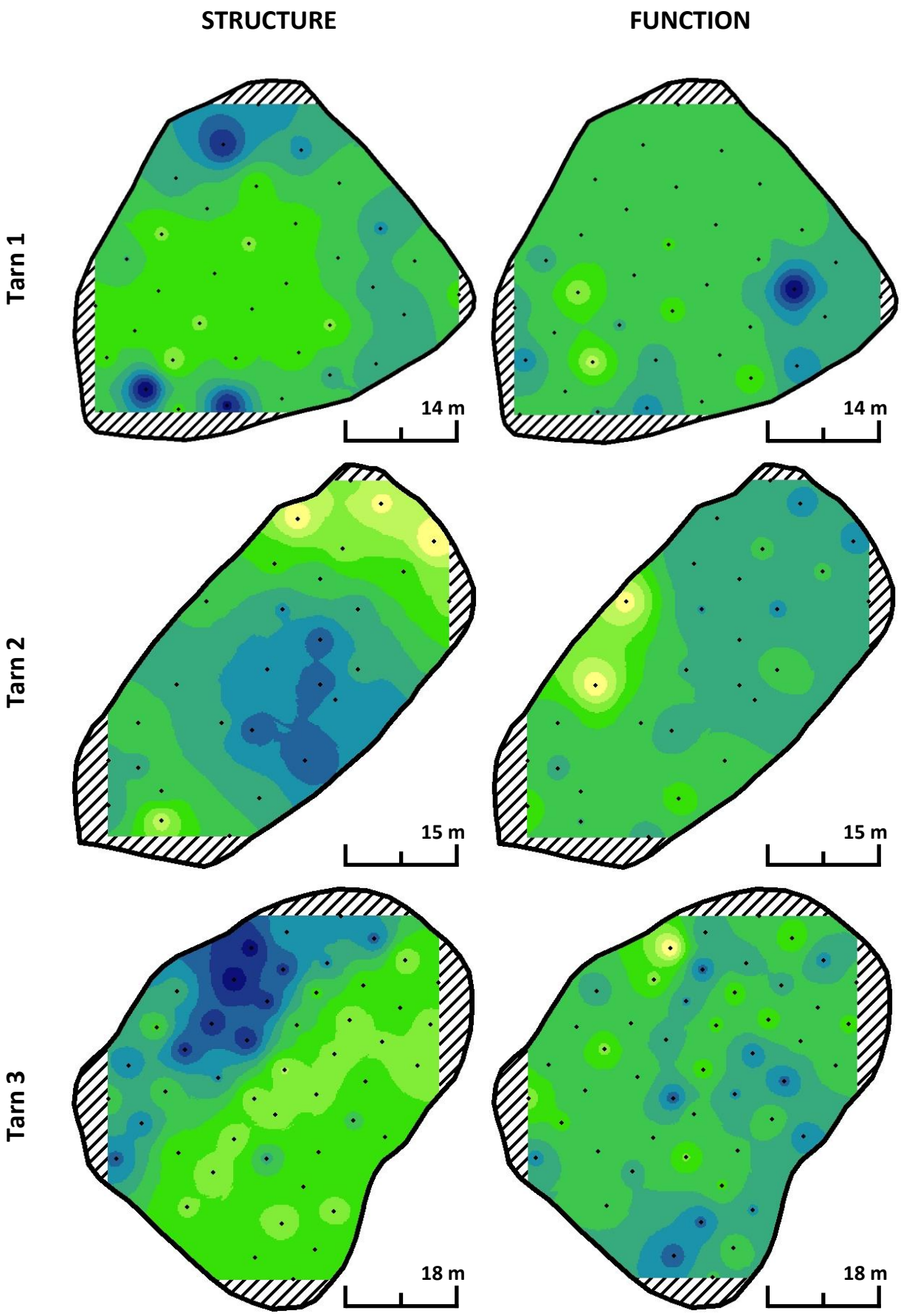

Figure 20. Similarity in bacterial community structure and function within the tarns. Bacterial community data for each tarn were subjected to a data reduction procedure by

non-metric multidimensional scaling of Bray-Curtis similarity data. The differences between the highest and lowest 1-d configuration scores for each plot were then used to classify the configuration into ten equally sized classes. The bacterial community data falling within each class was assigned a different colour, across a gradient from dark blue (lowest 1-d configuration score) to light green (highest 1-d configuration score). The outcome of this approach is that samples hosting more similar bacterial community data are represented by more similar colours on each map.

A number of physicochemical environmental variables were used to determine the relationship between the environment and bacterial community structure and function. The variables that were included in the analysis were; pH, conductivity, total carbon and concentrations of a range of nutrients. pH, total carbon and nitrite-N were identified as having a significant relationship with bacterial community structure across all tarns (DistLM test, $P < 0.05$, Table 5). In regards to the individual tarns, different environmental variables for each tarn were identified as having a significant relationship with bacterial community structure (DistLM test, $P < 0.05$, Table 5). In contrast to bacterial community structure, for bacterial community function, no environmental variables were identified as having a significant relationship with bacterial community function in the individual tarns (DistLM test, $P > 0.05$). However, pH and total carbon were identified as having a significant relationship with bacterial community function across all tarns (DistLM test, $P < 0.05$, Table 6).

Table 5. Relationship of different environmental variables with bacterial community structure. Environmental variables in bold were identified as being significant (*P* < 0.05). Subscript numbers indicate the order of significance of the environmental variables (forward selection) and the % represents the proportion that each environmental variable contributed towards total variation. Depth was only measured for samples in Tarn 3.

Variable	All Tarns		Tarn 1		Tarn 2		Tarn 3	
	P value	*%*	*P value*	*%*	*P value*	*%*	*P value*	*%*
pH	**^{2}0.0001**	**8.0**	^{6}0.627	2.2	^{4}0.673	2.2	**^{1}0.0005**	**9.3**
Total carbon (ppm)	**^{1}0.0001**	**15.0**	^{5}0.508	2.5	**^{1}0.0001**	**17.3**	**^{2}0.014**	**5.5**
Chloride (mg L^{-1})	—	—	^{4}0.280	3.1	—	—	—	—
Nitrite-N (mg L^{-1})	**^{3}0.005**	**1.5**	**^{1}0.003**	**7.2**	^{5}0.844	1.7	^{5}0.555	1.4
Nitrate-N (mg L^{-1})	^{5}0.502	0.6	^{2}0.070	4.5	^{3}0.103	4.6	^{6}0.616	1.4
Phosphate-P (mg L^{-1})	—	—	—	—	—	—	^{7}0.946	0.8
Sulphate-S (mg L^{-1})	^{4}0.054	1.1	^{3}0.312	3.1	**^{2}0.040**	**5.9**	^{4}0.064	3.4
Depth (cm)	—	—	—	—	—	—	**^{3}0.050**	**3.7**

Table 6. Relationship of different environmental variables with bacterial community function. Environmental variables in bold were identified as being significant (*P* < 0.05). Subscript numbers indicate the order of significance of the environmental variables (forward selection) and the % represents the proportion that each environmental variable contributed towards total variation. Depth was only measured for samples in Tarn 3.

Variable	All Tarns		Tarn 1		Tarn 2		Tarn 3	
	P value	*%*	*P value*	*%*	*P value*	*%*	*P value*	*%*
pH	**^{1}0.0001**	**6.8**	^{1}0.125	4.1	^{1}0.144	5.3	^{1}0.135	3.0
Total carbon (ppm)	**^{2}0.0001**	**7.3**	^{5}0.709	2.1	^{4}0.879	1.9	^{4}0.515	1.8
Chloride (mg L^{-1})	—	—	^{4}0.449	2.7	—	—	—	—
Nitrite-N (mg L^{-1})	^{4}0.768	0.5	^{6}0.936	1.3	^{5}0.986	1.0	^{3}0.369	2.2
Nitrate-N (mg L^{-1})	^{3}0.576	0.7	^{2}0.307	3.2	^{2}0.308	4.1	^{7}0.799	1.3
Phosphate-P (mg L^{-1})	—	—	—	—	—	—	^{5}0.439	2.0
Sulphate-S (mg L^{-1})	^{5}0.922	0.4	^{3}0.064	4.8	^{3}0.804	2.2	^{6}0.540	1.8
Depth (cm)	—	—	—	—	—	—	^{2}0.258	2.5

For the environmental variables that had a significant relationship with bacterial community structure in the individual tarns, contour plots were produced to identify any gradients or patterns present for these environmental variables in the tarns (Figure 21). The patterns observed for nitrite-N in Tarn 1, total carbon in Tarn 2 and finally pH and total carbon in Tarn 3 mirrored the patterns displayed for the MDS contour plots for bacterial community structure in the respective tarns (Figure 21).

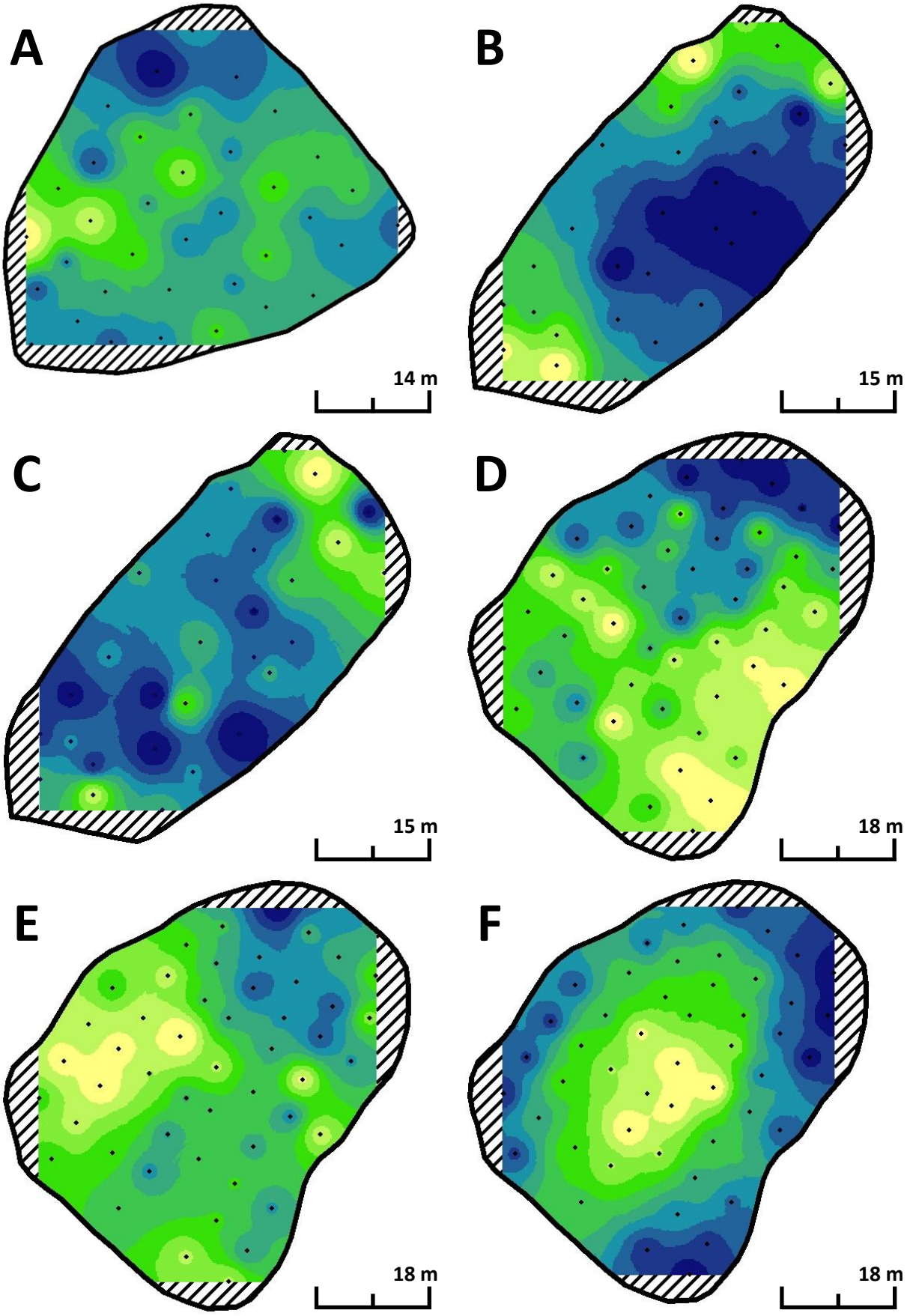

Figure 21. Contour plots of significant environmental variables. A = nitrite-N, Tarn 1; B = total carbon, Tarn 2; C = sulphate-S, Tarn 2; D = pH, Tarn 3; E = total carbon, Tarn 3; F = depth, Tarn 3. The difference between the highest and lowest value for each

environmental variable was used to classify the data for each variable into ten equally sized classes. The environmental data falling within each class was assigned a different colour, across a gradient from dark blue (lowest 1-d configuration score) to light green (highest 1-d configuration score). The outcome of this approach is that samples hosting more similar environmental values (for the selected environmental variable) are represented by more similar colours on each map.

Variance partitioning was performed on the data to determine whether spatial location or environmental variables contributed the largest proportion towards explaining the total variation in bacterial community structure and function. Pure spatial variation contributed the largest proportion towards both bacterial community structure (range = 24 - 38 %, Figure 22) and function (range = 18 – 37 %, Figure 23). A similar proportion of variation was explained by environmental variation independent of spatial factors for both bacterial community structure (range = 4 – 19 %) and function (range = 1 – 18 %). However, a larger proportion of variation in bacterial community structure was explained by pure environmental variation and spatially structured environmental variables (range = 22-32 %) than it was for bacterial community function (range = 14-18 %). Overall, more total variation is explained for bacterial community structure (range = 50-70 %) than it is for bacterial community function (range = 33-52 %).

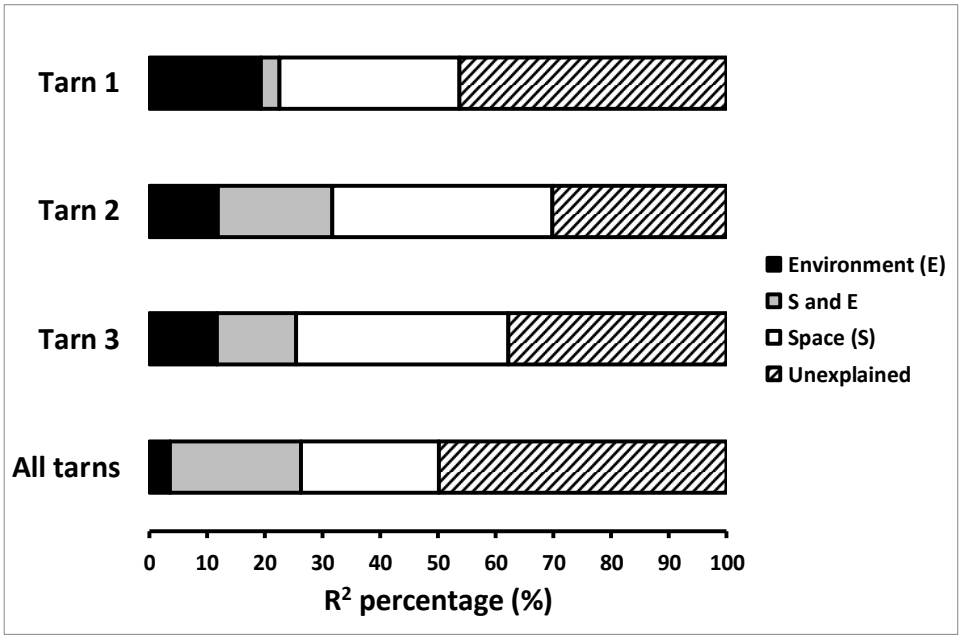

Figure 22. Results of partial regression analysis, partitioning the variation in bacterial community structure within each of the three tarns (data for each tarn analysed individually), or among all tarns (data for all three tarns analysed together). Four different components are shown: (hatched) unexplained variation; (black) pure environmental variation that is independent of any spatial factors; (white) pure spatial variation that is independent of any environmental variables; (grey) spatially structured environmental variables.

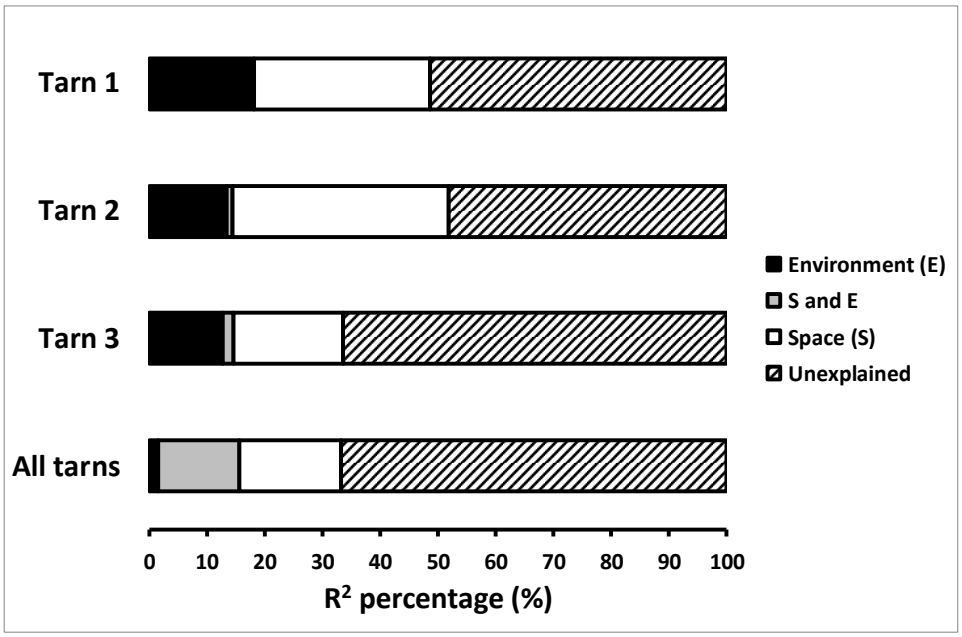

Figure 23. Results of partial regression analysis, partitioning the variation in bacterial community function within each of the three tarns (data for each tarn analysed individually), or among all tarns (data for all three tarns analysed together). Four different components are shown: (hatched) unexplained variation; (black) pure environmental variation that is independent of any spatial factors; (white) pure spatial variation that is independent of any environmental variables; (grey) spatially structured environmental variables.

Chapter 4

Discussion

4.1 Overview

The main aims of this study were (1) to determine if bacteria in freshwater ecosystems display similar biogeographical patterns to macroorganisms, such as, the taxa-area and distance-decay relationships, (2) to describe variation in community structure and function among and within the tarns analysed, and (3) to relate the variation in bacterial community structure and function to spatial location and environmental variables. My results show that bacterial communities in freshwater ecosystems (tarns) do indeed display biogeographical patterns. A positive taxa-area relationship was observed among samples within each tarn. A distance-decay relationship was observed for bacterial community structure in that there was a negative relationship between the paired compositional similarity of samples and increasing geographic distance. However, this distance-decay relationship was only mirrored by bacterial community function, measured as potential carbon use, in one of the tarns (Tarn 2). In each of the tarns, significant spatial variation in bacterial community structure was observed that correlated with measured environmental variables. However, spatial patterns in bacterial community function were weaker, and apparently unrelated to environmental variation.

4.2 Taxa-area relationship

For the first objective of this study, I characterised the taxa-area relationships for bacterial communities in tarns. I predicted that there would be a positive taxa-area relationship for the bacterial communities, i.e., taxa richness would increase in relation to the volume of sample. In agreement with my prediction, a positive taxa-area relationship was observed in all three tarns. The *z*-values for the taxa-area relationships ranged from 0.26 to 0.32. These values are relatively high in comparison to the *z*-values calculated for microorganisms in some studies (Fierer and Jackson, 2006; Green *et al.*, 2004), but are comparable to others (Bell *et al.*, 2005; Noguez *et al.*, 2005; Van Der Gast *et al.*, 2005). Green *et al.* (2004) determined that the taxa-area relationship for terrestrial microbial eukaryotes in soil,

sampled in a nested manner (contiguous habitat), was quite flat (*z*-value = 0.074). Fierer and Jackson (2006) also calculated a low *z*-value of 0.03 for unique taxa in 98 soil samples from North and South America (islands). In contrast, Bell (2005), determined a *z*-value of 0.26 for microbial communities in water-filled treeholes and van der Gast *et al.* (2005), calculated *z*-values ranging from 0.246 to 0.294 for bacterial communities colonising metal-cutting fluids in machines with different sump tank sizes. Both of these studies analysed habitats that would be regarded as islands. Another study performed by Noguez *et al.* (2005) that involved the nested sampling (contiguous habitat) of free-living prokaryotes in two sites in the tropical forest of Chamela, Jaliscon, on the western coast of Mexico found that the prokaryotes displayed OTU-area relationships with *z*-values of 0.47 and 0.42 for the two sites. A high *z*-value is indicative of a significant turnover in bacterial community structure among sampling sites and because the *z*-values calculated in this study for bacterial communities in a contiguous habitat are relatively high, this suggests that the bacterial community structure varied significantly across the tarns (Woodcock *et al.*, 2006).

The taxa-area relationship has been characterised for microorganisms, but how do the *z*-values calculated in this study compare to *z*-values for macroorganisms? With macroorganisms, for areas of contiguous habitat, *z*-values usually range from 0.12 to 0.18, whereas *z*-values for islands typically range from 0.25 to 0.35 (Rosenzweig, 1995). The *z*-values calculated in this study for bacterial communities sampled in a nested manner (contiguous habitat) are higher than the expected range for macroorganisms in areas of contiguous habitat (0.12 to 0.18) (Rosenzweig, 1995). However, other studies (Crawley and Harral, 2001; Fridley *et al.*, 2005; Ulrich and Buszko, 2003) indicate that the *z*-values for macroorganisms may be higher than the ranges stated in Rosenzweig (1995). Fridley *et al.* (2005) investigated the species-area relationship for vascular plant richness located across southeastern United States and calculated a mean *z*-value of 0.372 for the species-area relationship for vascular plants as area increased from 0.01 to 1000 m^2. Likewise, another study by Crawley and Harral (2001), performed on vascular plants in Berkshire in southeast England at different scales found that they also displayed a positive species-area relationship with *z*-values of 0.1 to 0.2 at small scales (less than 100 square meters), 0.4 to 0.5 at intermediate scales (1 hectare to 10 square kilometres) and 0.1 to 0.2 for large scale transitions (more than 10 square kilometres). Both of these studies involved collecting samples in a nested manner (contiguous habitat). Ulrich and Buszko (2003) analysed the

species-area relationship for European butterfly species and calculated a *z*-value of 0.10 for northern and eastern European countries (contiguous habitat) and a *z*-value of 0.49 for Mediterranean countries (islands). These studies indicate that *z*-values for macroorganisms in areas of contiguous habitat can be as low as 0.1 (Crawley and Harral, 2001; Rosenzweig, 1995; Ulrich and Buszko, 2003); however, they can also be as high as 0.35 - 0.5 (Crawley and Harral, 2001; Fridley *et al.*, 2005). The *z*-values calculated in this study fit within this range, which suggests that bacteria display similar taxa/species-area relationships to macroorganisms. In addition, the study by Crawley and Harral (2001) shows that *z*-values can vary depending on the scale at which data is collected. In this study, in terms of the body size and richness of bacteria, samples were collected on a relatively large scale (7 m inter-sample distances). This scale is probably comparable to sampling at a much larger scale for macroorganisms. With the study by Crawley and Harral (2001) that determined species-area relationships for vascular plants over a range of spatial scales, *z*-values of 0.4 - 0.5 were calculated for the species-area relationship at 1 hectare to 10 square kilometres. It is probably more appropriate to compare the *z*-values obtained in this study for bacterial communities to the *z*-values calculated at this larger scale for macroorgansisms than at the same scale.

4.3 Distance-decay relationship in bacterial community structure

The second objective of this study was to characterise the distance-decay relationship in each tarn for both bacterial community structure and function. I predicted that similarity in bacterial community structure and function would decline with increasing geographic distance. As expected, with bacterial community structure, a decline in similarity between paired samples was observed with increasing geographic distance. Because a distance-decay relationship was observed for bacterial community structure this indicates that there is geographical differentiation which may be due to dispersal limitation, environmental heterogeneity and/or habitat structure (Nekola and White, 2004; Soininen *et al.*, 2007). It also suggests that freshwater ecosystems may not be as well mixed as previously thought (Scheiner *et al.*, 2000), because a flat distance-decay relationship would be expected for a well-mixed freshwater ecosystem.

A distance-decay relationship was also observed for microorganisms in a number of other studies (Finkel *et al.*, 2012; Jones *et al.*, 2012; King *et al.*, 2010; Sommaruga and Casamayor, 2009) and for macroorganisms (Palmer, 2005; Thieltges *et al.*, 2009). Sommaruga and Casamayor (2009) investigated if the bacterial community composition in six remote, high-altitude lakes located in the Mount Everest region was related to distance on a relatively small scale of less than six kilometres. They determined that similarity in bacterial community composition was related to distance however, this pattern was driven by environmental variables. Jones *et al.* (2012) collected 32 samples from each of two lakes in Wisconsin, USA (Lake Mendota and Crystal Bog Lake), to investigate the variation in bacterial community composition over space and time. They identified a significant decline in bacterial community similarity between paired samples with increasing geographic distance. Similar findings have been reported for communities of macroorganisms (Dexter *et al.*, 2012; Palmer, 2005; Thieltges *et al.*, 2009). For example, Palmer (2005) analysed trees in an old-growth neotropical forest at the La Selva Biological Station in the Atlantic Lowlands of Costa Rica and identified a significant but weak distance-decay relationship. The relationship appeared to be driven by environmental variables over the majority of the distance, apart from between 300 and 500 m. Thieltges *et al.* (2009) also found that the similarity in trematode (fluke worm) communities from three prominent coastal molluscs, the gastropods *Littorina littorea* and *Hydrobia ulvae* and the bivalve *Cerastoderma edule*, in the north-east Atlantic showed a decline with increasing geographic distance. They determined that the decline in similarity at local scales was due to limited dispersal ability whereas the decline at larger scales (regions) was because of the low occurrence of environmental variables that supported the various hosts. The distance-decay relationship has been observed for both microorganisms and macroorganisms suggesting that these two groups display similar biogeographical patterns.

4.4 Distance-decay relationship in bacterial community function

In contrast to bacterial community structure, bacterial community function showed a decline in similarity between paired samples with increasing geographic distance only for Tarn 2. Because bacterial community function did not display a similar distance-decay relationship to bacterial community structure in all the tarns, this indicates that bacterial communities in freshwater ecosystems may exhibit functional redundancy in carbon use/digestion, that is,

multiple species are capable of performing the same function (Wohl *et al.*, 2004). A recent study by Burke *et al.* (2011) determined that the bacterial species on different *Ulva australis*, a green alga, had high phylogenetic variability with only 15 % similarity between samples. However, similarity in functional composition was higher at around 70 %. This shows that bacterial community composition varied to a greater degree than functional composition. This would not be observed if every bacterial taxon had unique functional attributes and this indicates that at less some of the bacteria in this study were likely to be functionally redundant. Another study by Teske *et al.* (2011), who investigated the relationship between bacterial community composition and function in seawater and sediment samples collected at different depths (surface-water (2 m), bottom-water (195 m), sediments) at an arctic fjord of Svalbard, used enzyme activity as a measure of function. The extracellular enzymatic hydrolysis rates of ten substrates were measured. It was found that bacterial community composition varied between seawater and sediments, surface-water and bottom-water, and surface and subsurface sediments. These differences were reflected in functional differences between seawater and sediments. However, the enzyme activity in surface and bottom waters was very similar. These studies (Burke *et al.*, 2011; Teske *et al.*, 2011) reinforce what was found in the present study, that even though there might be significant variation in bacterial community structure this may not be reflected in bacterial community function due to functional redundancy.

Alternatively, the observation of functional redundancy in bacterial communities in this study may be because the technique (BIOLOG EcoPlates™) that was used to measure bacterial community function was not appropriate, perhaps bacteria are not selective when it comes to the carbon source that they metabolise. In regards to non-selectivity of carbon sources by bacteria, Cottrell and Kirchman (2000) determined that a diverse bacterial community was required to metabolise complex dissolved organic matter in the ocean. Complex dissolved organic matter contains a variety of carbon substrates, and therefore its complete degradation will indicate the capacity of a bacterial community to utilise different carbon substrates. Cottrell and Kirchman (2000) divided the bacterial community into phylogenetic groups, and investigated the consumption of a range of organic compounds by these groups. They found that the consumption of organic compounds varied between the different groups. It was concluded that comparing dissolved organic matter consumption with bacterial community structure will contribute to the understanding of the structure-

function relationship in aquatic bacterial communities. This suggests that BIOLOG EcoPlates™ are in fact an appropriate technique to use to identify variation in bacterial community function (carbon substrate utilisation) and gives weight to the evidence for functional redundancy within these aquatic tarn communities.

4.5 Spatial and environmental variability among the tarns

For the last objective, I characterised spatial and environmental variation in bacterial community structure and function among and within the tarns. The three tarns had significantly different bacterial community structure and function. Spatial location and/or the environment variables of the three tarns may be responsible for driving the differences in bacterial community structure and function. Schöttner *et al.* (2012) found that there were significant differences in the bacterial community structure among sites, which in this study were four different cold-water coral reef systems along the Norwegian continental margin. In addition, the bacteria appeared to display habitat specificity indicating that environmental variables are important in driving bacterial community structure. In a separate study, Lear *et al.* (2008) identified that the bacterial community structure for biofilms in two stream locations, Cascade Stream and Opanuku Stream, both located in the Waitakere District of Auckland, was significantly different between the sites. In addition, the variation in bacterial community structure between the two streams was determined to be greater than the variation in bacterial community structure within the individual streams. From the range of environmental variables that were measured, water temperature was identified as having the largest effect on bacterial community structure.

In terms of variation in environmental variables driving bacterial community structure, Berdjeb *et al.* (2011) identified that the temporal variation in bacterial community structure in the mesotrophic Lake Bourget, located in France, was related to changes in a range of environmental variables, such as, water temperature, conductivity, dissolved oxygen concentration, total organic carbon, and nutrient concentrations (total nitrogen, dissolved ammonium (NH_4-N), dissolved nitrates (NO^3-N), total phosphorus (TP) and orthophosphates (PO_4-P)) and also related to chlorophyll α fluorescence. The changes in environmental and biological variables explained a significant proportion, around 60 %, of temporal variation in bacterial community structure at different depths (2 and 50 m). Even though this study was

not analysed in a spatially explicit manner, it nonetheless illustrates how bacterial community structure can vary with changes in environmental variables. In this study, it was observed that a range of environmental variables displayed significant variation across the tarns, for example, total carbon. These variables may, in part, be responsible for driving the differences in bacterial community structure in the same way that Berdjeb *et al.* (2011) observed.

4.6 Spatial and environmental variability within the tarns

With the overall variation in bacterial community structure and function, it was predicted that (i) spatial location would explain the most variation in bacterial community structure, whereas (ii) the environment would have a stronger relationship with bacterial community function than with bacterial community structure. It was expected that bacterial community structure would be closely related to spatial factors due to the dispersal abilities of bacteria. Even though it has been assumed in the past that microorganisms display a cosmopolitan distribution (Martiny *et al.*, 2006), because their small size and high abundance facilitates dispersal, evidence now suggests that microorganisms might in fact display biogeographical patterns (Green and Bohannan, 2006; Horner-Devine *et al.*, 2004; Martiny *et al.*, 2006). This indicates that microorganisms are dispersal limited to some degree. Microorganisms have a role in important ecological processes in the environment such as trace gas emissions, soil structure and formation, decomposition of organic matter and xenobiotics, and the recycling of essential elements (e.g. carbon, nitrogen, phosphorous, and sulphur) and nutrients (Green and Bohannan, 2006; Horner-Devine *et al.*, 2004; Rastogi and Sani, 2011). It is likely that the capacity of microorganisms to perform these ecological processes is related to certain environmental variables. Studies indicate a relationship between environmental heterogeneity and microbial richness (Cam *et al.*, 2002; Kallimanis *et al.*, 2008; Preston, 1962; Rosenzweig, 1995), which is probably because specific microbial taxa are adapted to certain environmental conditions due to their functional capabilities.

4.6.1 Spatial variability within the tarns

With the spatial variation in the tarns, Tarn 1 displayed the most variation, with no distinct clusters of sample data whereas with Tarns 2 and 3, there was some spatial variation, but the sample data formed distinct clusters. Because there were clusters of sample data with unique bacterial community structure in Tarns 2 and 3, this indicates that either the bacteria were dispersal limited or the tarns displayed environmental heterogeneity, potentially in the form of geographical isolation by barriers. With the previously mentioned study by Jones *et al.* (2012), significant variation in bacterial community composition between samples within the individual lakes was identified with horizontal variation in bacterial community composition being observed in both lakes. In addition, the bacterial community composition for samples that were located close together (triplicate samples collected within a 1 m^2 site) were significantly more similar than sites located further away. Yannarell and Triplett (2004) also investigated spatial variation, ranging from 10 to over 100 metres, in bacterial community composition among and within 13 lakes in northern and southern Wisconsin. They identified significant variation in bacterial community composition among the different lakes, especially those that were isolated from each other. They also identified horizontal variation in bacterial community composition within some of the lakes, but it was less than the variation displayed among the lakes. It was concluded that the horizontal variation in bacterial community composition in some of the lakes was due to restricted water flow. These studies (Jones *et al.*, 2012; Yannarell and Triplett, 2004) reinforce the findings in the present study by supporting the theory that bacteria display biogeographical patterns, rather than having a cosmopolitan distribution.

4.6.2 Environmental variability within the tarns

With the contour plots that were produced for the environmental variables that were identified as being significantly related to bacterial community structure, some of the spatial patterns for the environmental variables, such as nitrite-N for Tarn 1, total carbon for Tarn 2 and pH and total carbon for Tarn 3, showed similar spatial variation to the corresponding MDS contour plots. This indicates that these environmental variables might be partly responsible for driving the spatial variation in bacterial community structure in these tarns. Alternatively, the bacterial community structure could have affected the physicochemical

properties of the tarns through the use and synthesis of environmental variables (Reed and Martiny, 'in press'). A number of past studies have identified various environmental variables that are partly responsible for driving bacterial community structure in a variety of ecosystems (Berdjeb *et al.*, 2011; Lear *et al.*, 2008; Schöttner *et al.*, 2012; Yannarell and Triplett, 2005; Zeng *et al.*, 2009). For example, Zeng *et al.* (2009) determined that total nitrogen, ammonia and pH were significant drivers of bacterioplankton community structure in the eutrophic Lake Xuanwu, China. Distinct clusters of samples were formed based on the environmental characteristics of eight samples collected from the lake, which indicates that there was horizontal variation in the environmental variables across the lake. In addition, the bacterioplankton community structure in the clusters was observed to be significantly different. Another study by Yannarell and Triplett (2005) found that pH and Secchi depth (measure of water clarity) were significant drivers of variation in bacterial community composition in 30 lakes located in the Northern Highlands Lake District (Vilas and Oneida countries) in Wisconsin and in several counties in southern Wisconsin. These studies indicate that environmental variables, especially pH, may be responsible for either indirectly or directly influencing the spatial variation in bacterial community structure in freshwater ecosystems.

4.7 Variance partitioning of spatial factors and environmental variables

Even though it appears that environmental variables were important in driving bacterial community structure in the tarns, spatial factors were identified as explaining a greater proportion of the total variation in bacterial community structure than environmental variables. Spatial factors also explained a larger proportion of the total variation for bacterial community function. It was originally hypothesised, that spatial location would contribute the most towards total variation in bacterial community structure and that the environment would have a stronger relationship with bacterial community function than with bacterial community structure. However, it appears as though the environmental variables that were analysed in this study have a stronger relationship with bacterial community structure than with bacterial community function, and spatial factors contribute the most towards the variation in both bacterial community structure and function. In support of the findings in this study, Jones *et al.* (2012) performed a highly resolved intra-lake survey and observed spatial heterogeneity in bacterial community composition, but did not observe habitat-

specific patterns of bacterial community composition. This indicates the importance of spatial location in driving bacterial community structure. Langenheder and Ragnarsson (2007) analysed the variation in bacterial community composition due to environmental variables and spatial location among 35 rock pools along the Baltic Sea Coast in central Sweden. The study calculated that of the approximately 25 % explained variation; spatial location contributed approximately 9 % towards total variation, whereas environmental variables contributed approximately 14 % towards total variation. Therefore, even though less variation in bacterial community composition is explained by spatial location than it is by the environment Langenheder and Ragnarsson (2007) confirm that bacterial composition is likely influenced by spatial location. Van der Gucht *et al.* (2007) performed a study that determined the importance of spatial location and environmental variables on bacterial community composition in 98 shallow lakes across Europe (Denmark, 32 lakes; Belgium and The Netherlands, 34 lakes; southern Spain, 32 lakes). When the three regions were analysed together, only 3 % of the total variation was explained by spatial location while 20 % was explained by significant environmental variables and 6 % was explained by spatial location and significant environmental variables combined. This left a remaining 71 % of unexplained variation. Both of these studies (Langenheder and Ragnarsson, 2007; Van der Gucht *et al.*, 2007) indicate that environmental variables should contribute a greater proportion towards total variation in bacterial community structure than spatial location; however, in the present study the opposite was observed. It is possible that environmental variables that were not measured in this study were confounded by spatial factors and if these environmental variables had been analysed, they may have increased the proportion of total variation explained by the environment and decreased the proportion attributable to spatial location. Here, insufficient data for a number of important environmental variables, such as temperature, dissolved oxygen and depth for Tarns 1 and 2, were collected. Due to sampling restrictions, data for these environmental variables were collected from only 12 random sites within each tarn with the expectation that the interpolation of these data would allow values to be calculated for each sampling site. However, not enough sites were sampled to allow interpolation to be performed.

4.8 Future research

This study analysed spatial variation in bacterial community structure and function but did not account for temporal variation. If additional samples were collected in subsequent seasons or years, then the temporal variation in bacterial community structure and function could be investigated. It is possible that significant temporal variation would be observed due to the fact that it appears as though the tarns dry up on an annual basis (see section 2.1.1; smaller tarns were dry). This would result in mass extinctions, and therefore the composition of the resulting bacterial communities would likely be related to stochastic recruitment from the surrounding environment (Burke *et al.*, 2011).

An extension on this study could be to analyse the variation in bacterial community structure and function among a larger number of tarns. It would be interesting to determine if the bacterial community structure and function remains as distinct across a greater number of tarns, or whether some of the tarns would contain similar data clusters in bacterial community structure and function. In addition, if a larger number of tarns of varying size were analysed then the 'island' taxa-area relationship (number of taxa in each tarn plotted against the area of the tarn) could be investigated.

To determine bacterial community structure in this study, a DNA-based 'finger-printing' approach, ARISA, was used. Unfortunately, DNA-based fingerprinting techniques, ARISA included, are not sensitive enough to discriminate among bacteria at the species level and only report the abundance of the most common taxa, such that rare organisms often go undetected (Woodcock *et al.*, 2006). However, as DNA sequencing continues to advance and becomes less expensive high throughput sequence analysis could be used instead of ARISA. This would not only enable the detection of a greater proportion of the bacterial population but would also allow the contribution of individual species to the taxa-area and distance-decay relationships to be determined. This is of particular relevance since Horner-Devine *et al.* (2004) found that the turnover of bacteria differs between taxonomic groups.

I used BIOLOG EcoPlates™ to measure bacterial community function, which unfortunately only analyse carbon substrate utilisation, one of many possible ecological processes. Future research could include the analysis of multiple functions simultaneously through the use of functional gene arrays. These allow a number of genes whose products are responsible for certain functions, such as nitrogen, carbon, sulphur and phosphorus transformations and

cycling, metal reduction and resistance and organic xenobiotic degradation to be detected (McGrath *et al.*, 2010; Reeve *et al.*, 2010).

4.9 Conclusions

This study has demonstrated the existence of biogeographical patterns that are commonly observed for macroorganisms within aquatic bacterial communities sampled at a fine scale within three alpine tarns. The tarns differed significantly in both bacterial community structure and function. Strong spatial patterns in bacterial community structure were present within the tarns but little spatial pattern in bacterial community function was observed. In addition, the relationship between bacterial community structure and the environment was stronger than the relationship between the environment and bacterial community function. Finally, spatial location contributed more towards total variation for both bacterial community structure and function than the environment did. Not only has this study provided valuable information about how freshwater bacterial biodiversity is maintained but it has also expanded our understanding of the link between bacterial community structure and function and has highlighted the potential of functional redundancy in bacterial communities.

Appendix A

BIOLOG EcoPlate™

Table A. 1. Carbon sources in BIOLOG EcoPlate™ (BIOLOG Inc, Hayward, CA, U.S.A.). Each plate contains triplicate replicates of the carbon sources.

A1 Water	A2 β-Methyl-D-Glucoside	A3 D-Galactonic Acid γ-Lactone	A4 L-Arginine	A1 Water	A2 β-Methyl-D-Glucoside	A3 D-Galactonic Acid γ-Lactone	A4 L-Arginine	A1 Water	A2 β-Methyl-D-Glucoside	A3 D-Galactonic Acid γ-Lactone	A4 L-Arginine
B1 Pyruvic Acid Methyl Ester	B2 D-Xylose	B3 D-Galacturonic Acid	B4 L-Asparagine	B1 Pyruvic Acid Methyl Ester	B2 D-Xylose	B3 D-Galacturonic Acid	B4 L-Asparagine	B1 Pyruvic Acid Methyl Ester	B2 D-Xylose	B3 D-Galacturonic Acid	B4 L-Asparagine
C1 Tween 40	C2 i-Erythritol	C3 2-Hydroxy Benzoic Acid	C4 L-Phenylalanine	C1 Tween 40	C2 i-Erythritol	C3 2-Hydroxy Benzoic Acid	C4 L-Phenylalanine	C1 Tween 40	C2 i-Erythritol	C3 2-Hydroxy Benzoic Acid	C4 L-Phenylalanine
D1 Tween 80	D2 D-Mannitol	D3 4-Hydroxy Benzoic Acid	D4 L-Serine	D1 Tween 80	D2 D-Mannitol	D3 4-Hydroxy Benzoic Acid	D4 L-Serine	D1 Tween 80	D2 D-Mannitol	D3 4-Hydroxy Benzoic Acid	D4 L-Serine
E1 α-Cyclodextrin	E2 N-Acetyl-D-Glucosamine	E3 γ-Hydroxybutyric Acid	E4 L-Threonine	E1 α-Cyclodextrin	E2 N-Acetyl-D-Glucosamine	E3 γ-Hydroxybutyric Acid	E4 L-Threonine	E1 α-Cyclodextrin	E2 N-Acetyl-D-Glucosamine	E3 γ-Hydroxybutyric Acid	E4 L-Threonine
F1 Glycogen	F2 D-Glucosaminic Acid	F3 Itaconic Acid	F4 Glycyl-L-Glutamic Acid	F1 Glycogen	F2 D-Glucosaminic Acid	F3 Itaconic Acid	F4 Glycyl-L-Glutamic Acid	F1 Glycogen	F2 D-Glucosaminic Acid	F3 Itaconic Acid	F4 Glycyl-L-Glutamic Acid
G1 D-Cellobiose	G2 Glucose-1-Phosphate	G3 α-Ketobutyric Acid	G4 Phenylethyl-amine	G1 D-Cellobiose	G2 Glucose-1-Phosphate	G3 α-Ketobutyric Acid	G4 Phenylethyl-amine	G1 D-Cellobiose	G2 Glucose-1-Phosphate	G3 α-Ketobutyric Acid	G4 Phenylethyl-amine
H1 α-D-Lactose	H2 D,L-α-Glycerol Phosphate	H3 D-Malic Acid	H4 Putrescine	H1 α-D-Lactose	H2 D,L-α-Glycerol Phosphate	H3 D-Malic Acid	H4 Putrescine	H1 α-D-Lactose	H2 D,L-α-Glycerol Phosphate	H3 D-Malic Acid	H4 Putrescine

Appendix B

Automated ribosomal intergenic spacer analysis (ARISA) data

Table A. 2. Sample of tabulated ARISA data. Columns are samples and the data in each row provides the relative abundance of each bacterial taxon (ARISA peak). Each number greater than zero represents a taxon that was detected in the respective sample.

	1-1.	1-2.	1-3.	1-4.	1-5.	1-6.	1-7.	1-8.
158.2	0	0	0	0	0	0	0	0
160.2	0	0	0	0	0	0	0	0
166.2	0	0	0	0	0	0	0	0
168.2	0	0	0	0	0	0	0	0
170.2	0	0	0	0	0	0	0	0
172.2	0	0	0	0	0	0	0	0
192.2	0	0	0	0	0	0	0	0
258.2	0.382371	0	0	0	0.167645	0	0	0
260.2	0	0	0	0	0	0	0	0
262.2	0	0	0	0	0	0	0	0
264.2	0.201248	0	0	0	0	0	0	0
266.2	0	0	0	0	0	0	0	0
270.2	0	0	0	0	0	0	0	0
282.2	0	0	0	0	0	0	0	0
284.2	0	0	0	0	0	0	0	0
286.2	0.482995	26.58	0.250601	0.267352	0	2.255254	2.091795	5.49945
288.2	0	0	0	0	0	0	0	0
290.2	0	0	0	0	0	0	0	0
292.2	0	0	0	0	0	0	0	0
298.2	0	0	0	0	0	0	0	0
300.2	0	0	0	0	0	0	0	0
302.2	0	0	0	0	0	0	0	0
304.2	0	0	0	0	0	0	0	0
306.2	0.100624	0	0	0	0	0	0	0
308.2	0	0	0	0	0	0	0	0
310.2	0	0	0	0	0	0	0	0
312.2	0	0	0	0	0	0	0	0.559944
314.2	0	0	0	0	0	0	0	0
318.2	0	0	0	0	0	0	0	0
322.2	1.247736	0	0.37089	0.102828	0.167645	0	0	0
324.2	1.026363	2.28	0.400962	0	0.136211	0.133265	0	0.569943
326.2	0	0	0	0	0	0	0	0
328.2	0.181123	0	0	0	0	0	0	0
330.2	0.986114	0	0	0.236504	0	0	0	0

Appendix C

R script for taxa-area relationship

```
setwd("H:\\research_project\\FINAL RESULTS\\R scripts\\SAR - Text files")

library(vegan)


ARISA_data <- read.table("ARISA.txt",header=T,sep="\t")  # read in ARISA data

ARISAm <- ARISA_data[,-1] # remove site names column and convert data to a matrix


tarn <- specaccum(ARISAm,method="exact")

rich <- as.numeric(tarn$richness)

sites <- as.numeric(tarn$sites)

summary(pow.tarn <- nls(rich~c*(sites^z),start=list(c=1,z=1)))


#Plot graph

x11(4,4.5)

plot(rich~sites,xlim=c(0,50),ylim=c(0,300), cex.axis=0.8, ylab="Cumulative taxa
richness",xlab="Cumulative number of samples",cex.lab=1, cex=0.7)

new.x <- seq(0,max(sites1),0.001)

new.y <- coef(pow.tarn)[1]*(new.x^coef(pow.tarn)[2])

lines(new.y~new.x,lwd=2,col="red")

text(42,30,"Tarn",cex=1.5)
```

Appendix D

R script for distance-decay relationship

```
setwd("H:\\research_project\\FINAL RESULTS\\R scripts\\Structure - Text files")

library(vegan)

library(plotrix)


ARISAm <- read.table("ARISA_data.txt",header=T,sep="\t")  # read in ARISA data

spati <- read.table("spatial_data.txt",header=T,sep="\t") # read in spatial data


ARISA <- ARISAm[,-1] # remove site names column and convert data to a matrix

spat <- spati[,2:3] # extract just the eastings and northings and convert data to a matrix


ARISA.bc <- as.numeric(vegdist(ARISA, method="bray")) # calculate bray-curtis distance matrix for ARISA data and convert to a vector (uses the 'vegan' package)

similarity <- 1-ARISA.bc # convert the dissimilarity to similarity (similarity is now a number between 0 and 1, where 1 is completely similar)

distance <- as.numeric(vegdist(spat, method="euclidean")) # calculate euclidean distance matrix for spatial data and convert to kilometres and put into a vector (uses the 'vegan' package)

hist(distance) # examine the distribution of distances among pairs of sites

dd <- as.data.frame(cbind(distance,similarity)) # put distances in a dataframe with the site labels
```

```
#Plot graph

x11(4,4.5)

plot(dd$similarity~dd$distance, xlim=c(0, 65) ,ylim=c(0, 0.9), cex.axis=0.8, xlab="Geographic
distance (m)",ylab="Bray-Curtis similarity", cex.lab=1, cex=0.6)


#Fit a linear function (FUNCTION)

#summary (m1 <- lm(dd$similarity~dd$distance))

#new.x <- seq(0,max(dd$distance),0.001)

#new.y <- coef(m1)[1] + coef(m1)[2]*new.x

#lines(new.y~new.x, lwd=5)


#Fit a non-linear curve (STRUCTURE)

summary(m2 <- nls(similarity~exp(a*distance),start=list(a=0.1),data=dd))

new.x <- seq(0,max(dd$distance),0.001)

new.y <- exp(coef(m2)*new.x)

lines(new.y~new.x, lwd=3, col='red')
```

Appendix E

Cluster plots for bacterial community function

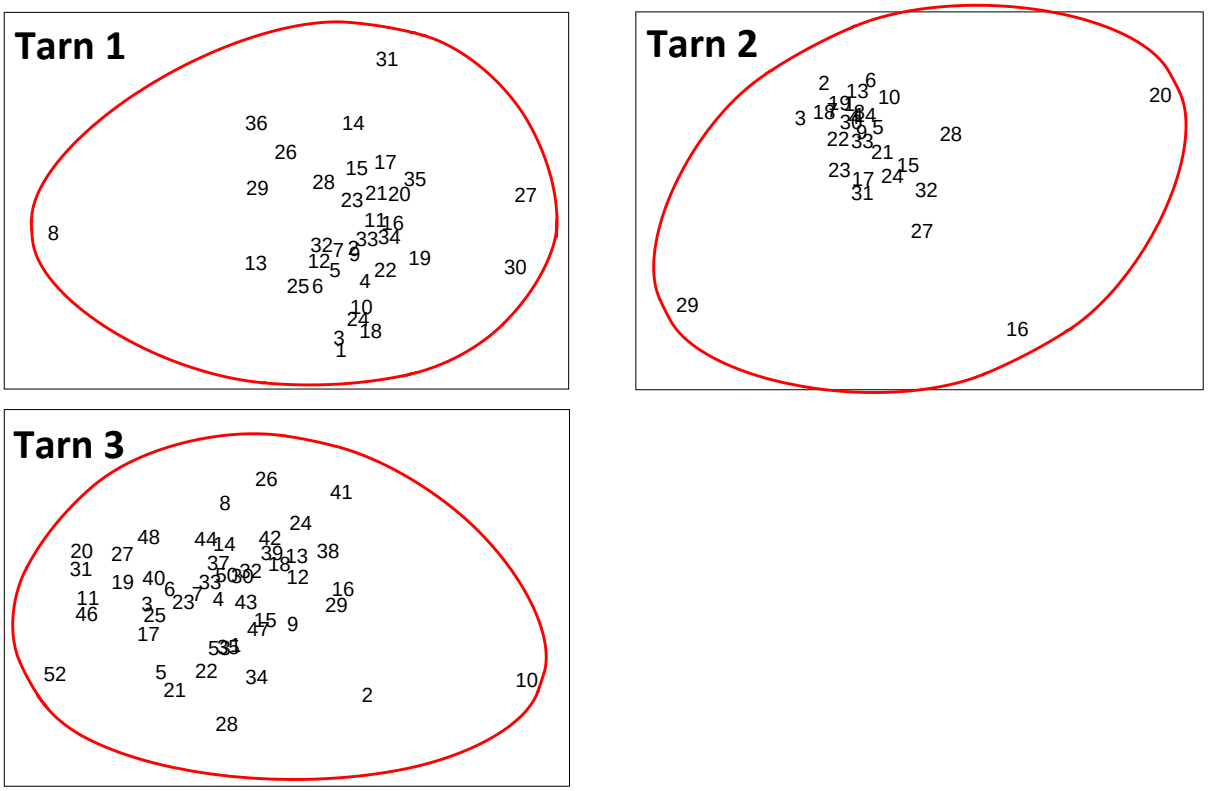

Figure A. 1. Variation in bacterial community function in each tarn. Plots are derived from non-metric multidimensional scaling of ARISA trace data using a Bray-Curtis similarity measure. Group average clusters are superimposed on each plot at the level of 50 % Bray-Curtis similarity. Two-dimensional stress values are 0.18, 0.14 and 0.19, for tarns 1, 2, and 3, respectively.

Appendix F

Variance Partitioning

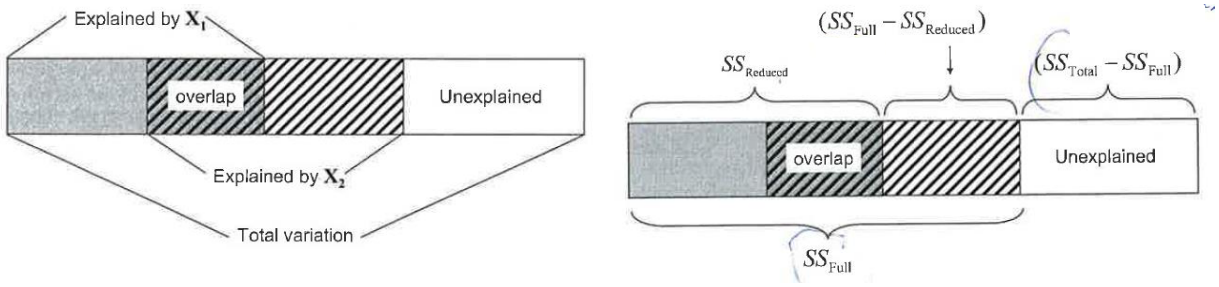

Figure A. 2. Diagram of how the variation attributable to different components is calculated. Taken directly from Clarke and Gorley (2006)

FINAL

DistLM
Distance based linear models

Resemblance worksheet
Name: Resem21
Data type: Similarity
Selection: All
Standardise Samples by Total
Resemblance: S17 Bray Curtis similarity

Predictor variables worksheet
Name: T1 enviro intra
Data type: Other
Sample selection: All
Variable selection: 1,3-5,7,9-18

Group variables, Indicator: ES
Selection criterion: R^2
Selection procedure: Specified

GROUPS
1 E Trial
2 S Trial

Total SS(trace): 85921

MARGINAL TESTS

Group	SS(trace)	Pseudo-F	P	Prop.	res.df	regr.df
E	19357	1.4055	0.0373	0.22529	29	7
S	29676	1.5243	0.0048	0.34539	26	10

NO STARTING TERMS

SEQUENTIAL TESTS

Group	R^2	SS(trace)	Pseudo-F	P	Prop.	Cumul.	res.df	regr.df
+E	0.22529	19357	1.4055	0.0348	0.22529	0.22529	29	7
+S	0.53826	26891	1.5062	0.0066	0.31297	0.53826	20	16

Specified solution

R^2	RSS	No.Groups	Selections
0.53826	39673	2	All

Figure A. 3. DistLM output in PRIMER-E showing the proportion of variation calculated for the different components. The numbers in the coloured bubbles are the values used in the calculations for variance partitioning.

Manual calculations for variance partitioning using the DistLM output and Tarn 1 as an example

Proportion of total variation explained by space = 31.297 %

Space and environment, 34.539 – 31.297 = 3.242 %

Environment, 22.529 – 3.242 = 19.287 %

Unexplained variation, 100 – (19.287 + 3.242 + 31.297) = 46.174%

References

Allison, S. D., and Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences, 105*(Supplement 1), 11512-11519.

Anderson, M. J., Gorley, R. N., and Clarke, K. R. (2008). PERMANOVA+ for PRIMER: guide to software and statistical methods. Plymouth, UK: PRIMER-E Ltd

Balser, T. C., and Firestone, M. K. (2005). Linking microbial community composition and soil processes in a California annual grassland and mixed-conifer forest. *Biogeochemistry, 73*(2), 395-415.

Bell, T., Ager, D., Song, J.-I., Newman, J. A., Thompson, I. P., Lilley, A. K., and van der Gast, C. J. (2005). Larger islands house more bacterial taxa. *Science, 308*(5730), 1884.

Berdjeb, L., Ghiglione, J., Domaizon, I., and Jacquet, S. (2011). A 2-year assessment of the main environmental factors driving the free-living bacterial community structure in Lake Bourget (France). *Microbial Ecology, 61*(4), 941-954.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences, 108*(34), 14288-14293.

Cam, E., Nichols, J. D., Hines, J. E., Sauer, J. R., Alpizar-Jara, R., and Flather, C. H. (2002). Disentangling sampling and ecological explanations underlying species-area relationships. *Ecology, 83*(4), 1118-1130.

Cencini, M., Pigolotti, S., and Muñoz, M. A. (2012). What ecological factors shape species-area curves in neutral models? *PLoS ONE, 7*(6), e38232.

Cho, J.-C., and Tiedje, J. M. (2000). Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology, 66*(12), 5448-5456.

Clarke, K. R., and Gorley, R. N. (2006). PRIMER v6: user manual/tutorial. Plymouth, UK: PRIMER-E Ltd

Cottrell, M. T., and Kirchman, D. L. (2000). Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Applied and Environmental Microbiology, 66*(4), 1692-1697.

Crawley, M. J., and Harral, J. E. (2001). Scale dependence in plant biodiversity. *Science, 291*(5505), 864-868.

Dexter, K. G., Terborgh, J. W., and Cunningham, C. W. (2012). Historical effects on beta diversity and community assembly in Amazonian trees. *Proceedings of the National Academy of Sciences, 109*(20), 7787-7792.

Dorigo, U., Fontvieille, D., and Humbert, J. F. (2006). Spatial variability in the abundance and composition of the free-living bacterioplankton community in the pelagic zone of Lake Bourget (France). *FEMS Microbiology Ecology, 58*(1), 109-119.

Eilers, K. G., Lauber, C. L., Knight, R., and Fierer, N. (2010). Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. *Soil Biology and Biochemistry, 42*(6), 896-903.

Fenchel, T. O. M., and Finlay, B. J. (2004). The ubiquity of small species: patterns of local and global diversity. *BioScience, 54*(8), 777-784.

Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America, 103*(3), 626-631.

Finkel, O. M., Burch, A. Y., Elad, T., Huse, S. M., Lindow, S. E., Post, A. F., and Belkin, S. (2012). Distance-decay relationships partially determine diversity patterns of phyllosphere bacteria on Tamrix trees across the Sonoran Desert. *Applied and Environmental Microbiology, 78*(17), 6187-6193.

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science, 296*(5570), 1061-1063.

Franklin, R. B., and Mills, A. L. (2003). Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology, 44*(3), 335-346.

Fridley, J. D., Peet, R. K., Wentworth, T. R., and White, P. S. (2005). Connecting fine- and broad-scale species-area relationships of southeastern U.S. flora. *Ecology, 86*(5), 1172-1177.

Gilbride, K. A., Lee, D. Y., and Beaudette, L. A. (2006). Molecular techniques in wastewater: understanding microbial communities, detecting pathogens, and real-time process control. *Journal of Microbiological Methods, 66*(1), 1-20.

Green, J., and Bohannan, B. J. M. (2006). Spatial scaling of microbial biodiversity. *Trends in Ecology & Evolution, 21*(9), 501-507.

Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., Gillings, M., and Beattie, A. J. (2004). Spatial scaling of microbial eukaryote diversity. *Nature, 432*(7018), 747-750.

Guillemot, N., Kulbicki, M., Chabanet, P., and Vigliola, L. (2011). Functional redundancy patterns reveal non-random assembly rules in a species-rich marine assemblage. *PLoS ONE, 6*(10), e26735.

Hector, A., and Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature, 448*(7150), 188-190.

Hill, J. L., Curran, P. J., and Foody, G. M. (1994). The effect of sampling on the species-area curve. *Global Ecology and Biogeography Letters, 4*(4), 97-106.

Horner-Devine, M. C., Lage, M., Hughes, J. B., and Bohannan, B. J. M. (2004). A taxa-area relationship for bacteria. *Nature, 432*(7018), 750-753.

Humbert, J.-F., Dorigo, U., Cecchi, P., Le Berre, B., Debroas, D., and Bouvy, M. (2009). Comparison of the structure and composition of bacterial communities from temperate and tropical freshwater ecosystems. *Environmental Microbiology, 11*(9), 2339-2350.

Jones, S. E., Cadkin, T. A., Newton, R. J., and McMahon, K. D. (2012). Spatial and temporal scales of aquatic bacterial beta diversity. *Frontiers in microbiology, 3*, 318.

Kallimanis, A. S., Mazaris, A. D., Tzanopoulos, J., Halley, J. M., Pantis, J. D., and Sgardelis, S. P. (2008). How does habitat diversity affect the species–area relationship? *Global Ecology and Biogeography, 17*(4), 532-538.

Kindt, R., Damme, P., and Simons, A. J. (2006). Patterns of species richness at varying scales in western Kenya: planning for agroecosystem diversification. *Biodiversity & Conservation, 15*(10), 3235-3249.

King, A. J., Freeman, K. R., McCormick, K. F., Lynch, R. C., Lozupone, C., Knight, R., and Schmidt, S. K. (2010). Biogeography and habitat modelling of high-alpine bacteria. *Nature Communications, 1*, 53.

Langenheder, S., and Ragnarsson, H. (2007). The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology, 88*(9), 2154-2161.

Lear, G., Anderson, M. J., Smith, J. P., Boxen, K., and Lewis, G. D. (2008). Spatial and temporal heterogeneity of the bacterial communities in stream epilithic biofilms. *FEMS Microbiology Ecology, 65*(3), 463-473.

Legendre, P., and Legendre, L. (1998). Numerical ecology, *Developments in environmental modelling 20* (2nd English ed.). Amsterdam, The Netherlands: Elsevier Science

Lomolino, M. V. (2001). The species-area relationship: new challenges for an old pattern. *Progress in Physical Geography, 25*(1), 1-21.

MacArthur, R. H., and Wilson, E. O. (1967). The theory of island biogeography. Princeton, NJ: Princeton University Press

Malik, S., Beer, M., Megharaj, M., and Naidu, R. (2008). The use of molecular techniques to characterize the microbial communities in contaminated soil and water. *Environment International, 34*(2), 265-276.

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Ovreas, L., Reysenbach, A.-L., Smith, V. H., and Staley, J. T. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology, 4*(2), 102-112.

McGrath, K. C., Mondav, R., Sintrajaya, R., Slattery, B., Schmidt, S., and Schenk, P. M. (2010). Development of an environmental functional gene microarray for soil microbial communities. *Applied and Environmental Microbiology, 76*(21), 7161-7170.

Miller, D. N., Bryant, J. E., Madsen, E. L., and Ghiorse, W. C. (1999). Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Applied and Environmental Microbiology, 65*(11), 4715-4724.

Mocali, S., and Benedetti, A. (2010). Exploring research frontiers in microbiology: the challenge of metagenomics in soil microbiology. *Research in Microbiology, 161*(6), 497-505.

Nekola, J. C., and White, P. S. (2004). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography, 26*(4), 867-878.

Noguez, A. M., Arita, H. T., Escalante, A. E., Forney, L. J., García-Oliva, F., and Souza, V. (2005). Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecology and Biogeography, 14*(3), 241-248.

Olof, A. (1921). Species and area. *Journal of Ecology, 9*(1), 95-99.

Palmer, M. W. (2005). Distance decay in an old-growth neotropical forest. *Journal of Vegetation Science, 16*(2), 161-166.

Preston, F. W. (1962). The canonical distribution of commonness and rarity: part I. *Ecology, 43*(2), 185-215.

R_Core_Team. (2012). R: a language and environment for statistical computing. Vienna, Austria. Available from http://www.R-project.org/

Ramette, A. (2009). Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. *Applied and Environmental Microbiology, 75*(8), 2495-2505.

Rastogi, G., and Sani, R. K. (2011). Molecular techniques to assess microbial community structure, function, and dynamics in the environment. *Microbes and Microbial Technology*, 29-57.

Reche, I., Pulido-Villena, E., Morales-Baquero, R., and Casamayor, E. O. (2005). Does ecosystem size determine bacteria richness? *Ecology, 86*(7), 1715-1722.

Reed, H. E., and Martiny, J. B. H. ('in press'). Microbial composition affects the functioning of estuarine sediments. *The ISME Journal*. doi:10.1038/ismej.2012.154

Reeve, J. R., Schadt, C. W., Carpenter-Boggs, L., Kang, S., Zhou, J., and Reganold, J. P. (2010). Effects of soil type and farm management on soil ecological functional genes and microbial activities. *The ISME Journal, 4*(9), 1099-1107.

Rosenzweig, M. L. (1995). Species diversity in space and time. Cambridge: Cambridge University Press

Saetre, P., and Bååth, E. (2000). Spatial variation and patterns of soil microbial community structure in a mixed spruce–birch stand. *Soil Biology and Biochemistry, 32*(7), 909-917.

Scheiner, S. M. (2003). Six types of species-area curves. *Global Ecology and Biogeography, 12*(6), 441-447.

Scheiner, S. M., Cox, S. B., Willig, M., Mittelbach, G. G., Osenberg, C., and Kaspari, M. (2000). Species richness, species-area curves and Simpson's paradox. *Evolutionary Ecology Research, 2*(6), 791-802.

Schöttner, S., Wild, C., Hoffmann, F., Boetius, A., and Ramette, A. (2012). Spatial scales of bacterial diversity in cold-water coral reef ecosystems. *PLoS ONE, 7*(3), e32093.

Soininen, J., McDonald, R., and Hillebrand, H. (2007). The distance decay of similarity in ecological communities. *Ecography, 30*(1), 3-12.

Sommaruga, R., and Casamayor, E. O. (2009). Bacterial 'cosmopolitanism' and importance of local environmental factors for community composition in remote high-altitude lakes. *Freshwater Biology, 54*(5), 994-1005.

Strickland, M. S., Lauber, C., Fierer, N., and Bradford, M. A. (2009). Testing the functional significance of microbial community composition. *Ecology, 90*(2), 441-451.

Teske, A., Durbin, A., Ziervogel, K., Cox, C., and Arnosti, C. (2011). Microbial community composition and function in permanently cold seawater and sediments from an arctic fjord of Svalbard. *Applied and Environmental Microbiology, 77*(6), 2008-2018.

Thieltges, D., Ferguson, M. D., Jones, C., Krakau, M., Montaudouin, X., Noble, L., Reise, K., and Poulin, R. (2009). Distance decay of similarity among parasite communities of three marine invertebrate hosts. *Oecologia, 160*(1), 163-173.

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology, 3*(9), 711-721.

Torsvik, V., Goksøyr, J., and Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology, 56*(3), 782-787.

Ulrich, W., and Buszko, J. (2003). Species–area relationships of butterflies in Europe and species richness forecasting. *Ecography, 26*(3), 365-373.

Van Der Gast, C. J., Lilley, A. K., Ager, D., and Thompson, I. P. (2005). Island size and bacterial diversity in an archipelago of engineering machines. *Environmental Microbiology, 7*(8), 1220-1226.

Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., Jeppesen, E., Conde-Porcuna, J.-M., Schwenk, K., Zwart, G., Degans, H., Vyverman, W., and De Meester, L. (2007). The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences, 104*(51), 20404-20409.

Wohl, D. L., Arora, S., and Gladstone, J. R. (2004). Functional redundancy supports biodiversity and ecosystem function in a closed and constant environment. *Ecology, 85*(6), 1534-1540.

Woodcock, S., Curtis, T. P., Head, I. M., Lunn, M., and Sloan, W. T. (2006). Taxa–area relationships for microbes: the unsampled and the unseen. *Ecology Letters, 9*(7), 805-812.

Yannarell, A. C., and Triplett, E. W. (2004). Within- and between-lake variability in the composition of bacterioplankton communities: investigations using multiple spatial scales. *Applied and Environmental Microbiology, 70*(1), 214-223.

Yannarell, A. C., and Triplett, E. W. (2005). Geographic and environmental sources of variation in lake bacterial community composition. *Applied and Environmental Microbiology, 71*(1), 227-239.

Yin, B., Crowley, D., Sparovek, G., De Melo, W. J., and Borneman, J. (2000). Bacterial functional redundancy along a soil reclamation gradient. *Applied and Environmental Microbiology, 66*(10), 4361-4365.

Zeng, J., Yang, L., Du, H., Xiao, L., Jiang, L., Wu, J., and Wang, X. (2009). Bacterioplankton community structure in a eutrophic lake in relation to water chemistry. *World Journal of Microbiology and Biotechnology, 25*(5), 763-772.

Zhou, J., Kang, S., Schadt, C. W., and Garten, C. T. (2008). Spatial scaling of functional gene diversity across various microbial taxa. *Proceedings of the National Academy of Sciences, 105*(22), 7768-7773.