The Distributional Dynamics of South Australian Aquatic Bacterio- and Virioplankton

Flinders UNIVERSITY

Lisa M. Dann

BHS, BSc (Honours)

Supervisors: Prof. Jim Mitchell & Dr Rod Oliver (CSIRO)

Thesis submission for Doctorate of Philosophy in Science and Engineering, School of Biological Sciences at Flinders University, South Australia.

October 2015

Lisa M Dann was supported by an Australian Postgraduate Award (APA) and top-up scholarship from the CSIRO (project number 8449).

TABLE OF CONTENTS

SUMMARY

Aquatic microbial communities, such as virio- and bacterioplankton, are important for ecosystem function as they play critical roles in biogeochemical processes, such as the cycling of carbon, phosphorus and nitrogen. These microbial communities are often observed by means of bulk phase sampling whereby relatively large volume samples are extrapolated over larger volumes or areas and considered representative of microbial activity and abundance. However, it is over micrometre to centimetre scales that crucial microbial interactions and processes occur, with microbial activity and abundance changing significantly over such scales. This patchiness indicates bulk phase sampling is not representative of biomass distributions as it implies homogeneity below the sampling scale or volume. This thesis investigates the extent of this patchiness in aquatic environments.

This thesis explores the distribution and composition of virio- and bacterioplankton communities within South Australian water sources, with particular focus given to the Murray River, Australia. Specifically, the microscale (mm) and large scale (km) spatial dynamics of virio- and bacterioplankton populations will be investigated via flow cytometric enumeration and taxonomic identification via primer-based sequencing.

The results of this thesis indicate microscale patchiness is higher than previously reported. This patchiness was observed as heterogeneity between samples from the same site in large scale (km) studies, with increased patchiness and diversity downstream of a small rural town (Chapter 2). This heterogeneity between samples from the same site indicated microscale patchiness may impact taxonomic diversity profiles more so than taxonomic patterns over larger distances, here 3.3 km

(Chapter 2). When testing this patchiness on the microscale, at the sub-centimetre scale of marine interface microenvironments, virio- and bacterioplankton abundance varied 45- and 2500-fold per centimetre, respectively (Chapter 3). This increased patchiness indicates resource competition and the likelihood of viral infection are higher in the small volumes important for individual cell encounters than bulk measurements (Chapter 3). Investigations of this patchiness within a freshwater system showed virio- and bacterioplankton abundance varied up to 107-and 80.5-fold per centimetre (Chapter 4). This indicates significant microbial patchiness extends to scales that may directly impact small populations or individual microbial interactions in freshwater systems.

Taxonomic analysis of these microbial patches revealed hotspots and coldspots, which contained microenvironments distinct from the taxonomy of the background community (Chapter 5). This indicates heterogeneous genus richness and composition exists within a river system and therefore suggests small scale fluid parcels persist together long enough to build up genera numbers. This would enable distinctly different proximate taxonomic microenvironments to form, which will therefore have a significant impact on microbial biogeochemical processes. Furthermore, taxonomic analysis of 1 µl subsamples from hotspot, coldspot and background regions revealed heterogeneity, which was most extreme between hotspot subsamples (Chapter 6). Hotspot subsample heterogeneity revealed two distinct taxonomic patterns: overall increases in genera common to most subsamples or the dominance of specific individual genera. Therefore reiterating microscale microbial hotspots represent discrete microenvironments, which will have important implications for nutrient exchange and cellular interactions and hence overall system function.

The findings of this thesis therefore indicate large scale measurements are not an accurate assessment of the conditions under which microbial dynamics exist in freshwater and marine systems. The high variability reported in this thesis indicates few microbes experience the ‘average’ concentrations frequently measured and that bulk phase sampling underestimates absolute concentration and misses the resources, competition and viral exposure gradients that microbes experience. Therefore bulk phase sampling does not represent microbial processes within aquatic systems, highlighting the need to sample at the microscale to accurately understand globally important microbial biogeochemical processes.

DECLARATION

'I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.'

Lisa Dann

16th October 2015

ACKNOWLEDGEMENTS

Firstly I would like to thank my primary supervisor, Jim Mitchell, for his support and guidance throughout my studies. I would also like to thank my co-supervisor, Rod Oliver from the CSIRO, whom provided insight and support. I am grateful to the CSIRO for the financial support I received during my studies, and for allowing me to work in collaboration with them on an interesting project.

I would like to thank current and past members of my laboratory. A special thanks to Jacqui, Eloise and Jess for all the discussions and fun-filled times we spent together. Thank you to Kelly, James and Tom for assiting me with field trips, methods and for taking the time out of their busy schedules to sit down with me and go through my work.

Thank you to my friends, whom although did not understand half of my science rambling, still offered their ears.

Lastly, I would like to thank my family for their ongoing support. Special thanks to my mum and partner Corey for listening to all my problems and putting up with my tears and tantrums. I can safely say I would not have made it through my studies without your help, encouragement and understanding.

ABSTRACTS ARISING FROM THESIS

Dann LM, Mitchell JG, Speck PG, Newton K, Jeffries T & Paterson J (2014) Virio- and bacterioplankton microscale distributions at the sediment-water interface. *PLoS One* 9(7): 1-14

Dann LM, Paterson JS, Newton K, Oliver R, Mitchell JG (2015) Correlations between virus-like particles and prokaryotes within microenvironments. *PLoS One* 11 (1): e0146984. doi:10.1371/journal.pone.0146984

Chapter 1

Introduction

AQUATIC MICROBIAL COMMUNITIES

Microbial communities, such as virio- and bacterioplankton, dominate the abundance and biomass of aquatic environments, and are crucial to ecosystem function as they are involved in important biogeochemical cycles (Azam 1998; Long & Azam 2001; Zinger *et al.* 2012). Specifically, bacterioplankton communities contribute to organic matter and nutrient cycling, whilst virioplankton affect nutrient and dissolved organic matter release, primary production, genetic exchange between microorganisms and control bacterial community composition and abundance via viral-mediated lysis (Fuhrman 1999; Newton *et al.* 2011; Suttle 2005; Wommack & Colwell 2000; Wilhelm & Matteson 2008; Zinger *et al.* 2012). These important biogeochemical processes often occur at or near cell surfaces with interactions occurring over micrometres to centimetres (Azam 1998; Long & Azam 2001; Mitchell *et al.* 1985; Seymour *et al.* 2004, 2005a, 2005b, 2006; Suttle 2007; Waters *et al.* 2003).

MICROBIAL MICROSCALE DISTRIBUTIONS

Microbial distributions are often measured in large volumes (up to thousands of litres) and then that value is extrapolated over larger areas or volumes in what is sometimes termed a 'mean field' approach (Angly *et al.* 2006; Corinaldesi *et al.* 2003; Hillmer *et al.* 2008; Liu *et al.* 2002; Pan *et al.* 2005, 2007; Seymour *et al.* 2004; Strutton *et al.* 1997). Microbial community distribution has often been considered homogeneous or to follow patterns of random aggregation, leading to the mean-field approach being considered representative of microscale microbial processes (Long & Azam 2001; Seymour *et al.* 2009). However, for marine systems, it is well accepted that microbial abundance and activity differ by orders of magnitude over

micrometre to centimetre scales (Azam 1998; Duarte & Vaque 1992; Long & Azam 2001; Mitchell & Fuhrman 1989; Seymour *et al.* 2004, 2005, 2006).

Associations between microscale microbial distributions are commonly observed as heterogeneous "hotspots" (Azam 1998; Seymour *et al.* 2000, 2004, 2005, 2006). Hotspots are areas of increased microbial abundance, which can form as a result of bacterial accumulation around nutrient patches via chemotaxis, particle aggregation or disintegration, the occurrence of grazing in adjacent areas or small scale water mixing (Barbara & Mitchell 2003; Blackburn *et al.* 1998; Seymour *et al.* 2005a, 2006, 2007b; Stocker & Seymour 2012). These areas of increased bacterial abundance and host density then lead to increased viral production via lysis (Azam 1998; Barbara & Mitchell 2003; Blackburn *et al.* 1998; Mitchell 2002; Seymour *et al.* 2004, 2005, 2006, 2007; Stocker *et al.* 2008).

Lysis events will produce intense local concentrations of viruses, which can spread as local epidemics. This should eventually produce regions of lowered bacterial concentration or bacterial "coldspots", which would be areas of low bacterial abundance. These coldspots can also be caused by grazing or attachment to particles, which could in turn lead to lowered viral abundance due to reduced host density (Waters *et al.* 2003). Grazing and lysis are a significant source of bacterial mortality and each influence bacterial abundance and community composition differently. For instance, grazing is selective as protists choose their prey based on motility, size and surface characteristics, thereby shaping bacterial community composition (Chow *et al.* 2014; Weinbauer *et al.* 2006).

Within the marine environment, protistan grazing can contribute to the mortality of 10 to 250 bacteria per protozoa daily within the water column (Bettarel *et al.* 2003;

Chow *et al.* 2014; Fuhrman & Noble 1995; Weinbauer & Höfle 1998). This grazing leads to reduced bacterial production and the transfer of organic matter to higher trophic levels (Bettarel *et al.* 2003; Weinbauer & Höfle 1998). Viral lysis also contributes to bacterial mortality, with 0 to 30% of bacterial production removed via lysis in marine and freshwater systems (Bettarel *et al.* 2003; Weinbauer & Höfle 1998). Lysis ensures organic matter remains within the microbial loop, with lysis products, such as cell debris, being used primarily by bacteria and therefore stimulating bacterial productivity and carbon uptake (Weinbauer & Höfle 1998). Lysis is often genus-specific and favoured towards high density populations, targeting fast-growing and abundant bacterial populations, whilst slow-growing resistant strains persist in what is termed the kill the winner hypothesis (Weinbauer *et al.* 2006; Thingstad 2000; Thingstad & Lignell 1997). As grazing and lysis control the majority of bacterial productivity, they therefore significantly impact bacterial community composition and system dynamics.

In addition to hotspots and coldspots, microbial abundance is usually high at surfaces, which creates a gradient between the surface and the fluid (Seymour *et al.* 2007b). These surface gradients are often observed in microbial abundance, and are the result of organic matter sinking and its incorporation into the benthos. This is followed by degradation and transformation via microbes, leading to increased nutrient concentrations and hence high microbial concentrations directly above the interface (Baric *et al.* 2002; Seymour *et al.* 2005b, 2006, 2007b).

Studies analysing the microscale distribution of microbial communities at interfaces reveal viral abundances show the highest heterogeneity 1–2 cm above the sediment water interface, with abundances varying by more than 2-fold over a 15 cm sampling distance, and the highest concentrations 1.5 cm from the interface (Seymour *et al.*

2005b, 2006). Similar patterns were observed in Seymour *et al.* (2007b), which showed high microbial abundance closest to the sediment-water interface with increases of 1.4- and 1.5-fold per centimetre in total virus and total bacteria. However, one-dimensional sampling and a lack of replication hindered the determination of whether these abundance changes were single or multiple point maxima or surface gradients. In freshwater environments, microbial distribution studies often still adopt the mean-field approach, therefore excluding direct observations of important ecological associations between microbial populations (Jezbera *et al.* 2012; Lear *et al.* 2014; Liu *et al.* 2011; Palijan 2012; Pollard & Ducklow 2011). These limitations have hindered investigations of the microbial structure that exists at the scale of microbial food webs.

Despite hotspots and coldspots being ubiquitous features of microbial distributions and representing the units of microscale microbial patchiness, they are often vaguely defined in studies. Hotspots are commonly classified as high microbial abundance patches that ‘exceed’ or are ‘elevated above’ background variation across one or two sampling points, sometimes only by a few tens of percentage points (Seymour *et al.* 2004, 2005a, 2008). Chapters within this thesis clarify the extent of patchiness via flow cytometric enumeration of prokaryotic subpopulations, low DNA (LDNA) and high DNA (HDNA), as well as virus-like particle (VLP) subpopulations, whilst also increasing the resolution from past studies and employing two-dimensional rather than one-dimensional sampling.

In addition to microbial abundance, determining taxonomic diversity within microbial communities is important for understanding system dynamics as the persistence, loss and appearance of bacterial taxa in aquatic environments can have a significant impact on system function (Finlay *et al.* 1997; Konopka 2009). Previous

investigations of microbial microscale hotspots and coldspots have been restricted to abundance estimates alone, leaving unanswered whether microbial hotspots represent distinct taxonomic microenvironments rather than temporary abundance increases or decreases of the overall community (Doubell *et al.* 2006; Seuront *et al.* 2007; Seymour *et al.* 2004, 2005ab, 2006, 2007, 2008; Waters *et al.* 2003). Within marine systems, Long and Azam (2001) analysed 16s rRNA gene diversity and richness using microlitre samples of seawater to test for small-scale patchiness in bacterial species. Their results indicated microscale variation in the bacterial community richness of the ocean. However, this study lacked sampling resolution and spatial coherency as 8 sub-samples were collected from a 1 litre seawater sample. Chapter 5 of this thesis will identify the taxonomic composition of microscale microbial hotspots and coldspots using sub-centimetre sampling resolution and two-dimensional sampling.

THE MURRAY RIVER AS A MODEL FRESHWATER SYSTEM

Ecosystems of importance, such as agricultural, industrial and potable water supplies, and their quality are influenced by microbial activity; therefore it is crucial that we understand the microbial community structure within these systems (Karr & Chu 2000). The research within this thesis will use the Murray River as a model system to investigate the flow of microbial communities within freshwater systems, which is particularly important for the Murray River as its water flow has been significantly altered by the addition of man-made obstructions such as locks, weirs, diversions and dams (Gehrke *et al.* 1995; Maheshwari *et al.* 1995). This flow regulation has caused a major shift in the normal flows and water levels of the river, which has impacted on the frequency of flooding events (Maheshwari *et al.* 1995; Walker & Thoms 1993). As many organisms rely on such flooding events for

successful spawning and germination, major declines have occurred in many key indigenous Murray River species (Gehrke *et al.* 1995; Maheshwari *et al.* 1995). With this decline in native species, a surge has been observed in the population numbers of other species better suited to the river's new flow regime, which in some cases are introduced species (Maheshwari *et al.* 1995). This has led to a major negative shift in the Murray River's ecosystem.

In addition, extreme weather patterns have led to the Murray River facing years of drought which has resulted in increased salinity (Bond *et al.* 2008; Mosley *et al.* 2012). Shifts in nutrient levels have also occurred, leading to disequilibrium within the indigenous microbial communities, which often leads to the likelihood of algal bloom formation (Maier *et al.* 2001; Mosley *et al.* 2012). Algal blooms are often toxic, having devastating effects on the fish species within the river as they are asphyxiated from the thick layers of algae that form on the river's surface (Anderson *et al.* 2002).

The Murray River is only one example of the many important freshwater systems found globally. However, it serves as a crucial indicator of the importance of maintaining the natural equilibrium within these systems. In order for an ecosystem to function, factors involved in the health of these systems need to be understood; such as the distribution and composition of indigenous microbial communities.

THESIS AIMS

The main objective of this thesis was to investigate the distributional dynamics of South Australian aquatic virio- and bacterioplankton, predominately using the Murray

River as a model system. Large scale and microscale investigations of the abundance and taxonomy of microbial populations will provide insight into the extent to which microscale microbial patchiness exists. In addition, information will be provided on the indigenous, non-pathogenic microbial communities found in the Murray River system.

Specifically the aims of this thesis are:

- To identify and compare the prokaryotic communities upstream and downstream of an enhanced anthropogenic impact within a river system
- To determine whether microscale patchiness has an effect on mean abundance estimates and therefore determine whether the mean field approach is an accurate estimate of abundance for marine and freshwater systems
- To determine whether microscale patchiness, and specifically heterogeneous hotspots and coldspots, exist within freshwater plankton abundance distributions
- To determine whether virio- and bacterioplankton population abundances are significantly different between the microenvironments of the air-water and sediment-water interfaces
- To identify the species composition of microscale microbial patches, and specifically microbial hotspots and coldspots
- To determine whether heterogeneity exists between 1 µl subsamples from microscale microbial hotspot, coldspot and background regions

THESIS STRUCTURE

This thesis is structured in manuscript format with the results from chapters 2 to 5 either published in peer-reviews journals or submitted to journals for publication, therefore redundancies are present in the introduction and methods sections of the chapters. Chapter 2 describes the river microbial communities upstream and 3.3 km downstream of a small rural town via 16s sequencing, and has been submitted to *CSIRO Marine and Freshwater Research*. Chapter 3 measures the distribution of marine virio- and bacterioplankton with a new two-dimensional technique to determine whether patchiness has an effect on mean abundance estimates and therefore whether the mean field approach is an accurate estimate of abundance, and has been published in the *Public Library of Science (PLOS) One* (9 (7): 1-14, 2014). Chapter 4 follows on from Chapter 3 and investigates the microscale abundance distributions of virio- and bacterioplankton in the lower reaches of the Murray River to determine the extent to which microscale microbial patchiness exists in freshwater systems, and has been published in the *Public Library of Science (PLOS) One* (11 (1): 1-19, 2016). Chapter 5 investigates whether microscale microbial hotspots, coldspots and background regions represent taxonomically distinct microenvironments and hence whether microscale variation in prokaryotic genus richness and composition exists. This Chapter has been submitted to *Aquatic Microbial Ecology.* Chapter 6 investigates the internal structure of microscale hotspot, coldspot and background regions to determine whether heterogeneity exists between 1 µl subsamples, and has been submitted to the *Public Library of Science (PLOS) One*. To reduce redundancies, one collated reference list is provided at the end of this thesis, providing references for all cited literature.

Chapter 2

Freshwater Bacteria Upstream and Downstream of a River System

ABSTRACT

Bacterial species are what shape microbial community composition and influence aquatic ecosystem dynamics. Studies on bacterial community persistence in rivers have primarily focused on microbial source tracking as indicators for faecal source contamination whilst archetypal freshwater species have received minimal attention. This study describes the river microbial communities upstream and 3.3 km downstream of a small rural town. Here, via 16s sequencing, we report three patterns in microbial community composition: persistence, loss and appearance. Persistence was observed via 46% site similarity perhaps due to generalists, which have information lengths that exceed 3.3 km and are capable of adapting to system fluctuations. Loss was observed as 10% site exclusivity upstream perhaps due to removal processes such as predation and lysis during transport downstream. Lastly, appearance was observed as 44% site exclusivity downstream indicating the potential anthropogenic impacts from land run-off on bacterial community composition. High dispersion between downstream samples, as well as sample dissimilarity, present as microscale hotspots of discrete Firmicutes and Cyanobacteria species, indicated higher heterogeneity downstream, and therefore increased patchiness from downstream transport and inputs of bacterial genotypes. These findings suggest three fates for bacterial species of fluvial systems, persistence, loss and appearance, with each having different impacts on system dynamics.

INTRODUCTION

Determining taxonomic diversity within microbial communities is the first step towards understanding system dynamics (Finlay *et al.* 1997; Konopka 2009). Bacterial species are what shape microbial community composition and as bacteria are critical to numerous biogeochemical cycles, such as the cycling of nitrogen, phosphorus and carbon; therefore influence ecosystem dynamics (Finlay *et al.* 1997; Konopka 2009; Meyer 1994). Ecosystems of importance, such as agricultural, industrial and potable water supplies and their quality are influenced by microbial activity; therefore it is crucial that we understand the microbial community structure within these systems (Karr & Chu 2000). However, despite this importance, studies looking at the persistence of bacterial communities in rivers primarily focus on microbial source tracking as indicators for faecal source contamination whilst archetypal freshwater species receive minimal attention (Eleria & Vogel 2005; Fries *et al.* 2006, 2008; Kreader 1998; Traister & Anisfield 2006).

The persistence of bacterial taxa throughout fluvial systems is observed in typical freshwater phylotypes such as Actinobacteria, Proteobacteria (alpha- and beta-subdivisions), the Cytophaga-Flavobacterium-Bacteroides (CFB) group and Verrucomicrobia, which dominate the bacterial divisions present in the water column of lake and river systems (Zwart *et al.* 2002). Due to the continuum of fluvial systems it is suggested that bacterial genotypes are not confined to specific locations, hence allowing persistence throughout the system (McArthur & Tuckfield 1997). The persistence of bacterial genotypes within fluvial systems relates to the nutrient spiralling concept, which is a critical biogeochemical pattern within fluvial networks that explains how nutrients are cycled and hence persist (Newbold *et al.* 2011; Webster 1975). Nutrient spiralling involves regeneration and transportation of

nutrients downstream whilst being cycled continually through several transformations prior to reutilisation. As each cycle is displaced downstream from the previous, this creates a spiral pattern throughout the system (Leff *et al.* 1992; Newbold *et al.* 1983; Webster 1975; Webster & Patten 1979). Newbold (1981) provided a mathematical description of nutrient spiralling, explaining spiralling length, which is the distance a dissolved nutrient atom can travel prior to immobilisation via biotic or abiotic processes. From this, the distance a specific nutrient molecule persists can be quantitatively determined, allowing comparisons of nutrient spiralling between different systems.

As the nutrient spiralling concept is analogous to the flow of genes through fluvial networks it can be used to determine genotype persistence. Leff *et al.* (1992) described the information spiralling concept, which connected the transport of bacteria and their genes in lotic environments to nutrient spiralling, indicating that bacterial genes would be coupled with spatial position in fluvial systems. This relates to the information length, which is the maximum distance a specific gene is effective independent of its host (Leff *et al.* 1992; McArthur & Tuckfield 1997). Previous information length studies, such as Wise *et al.* (1995), which investigated the genetic variability within lotic *Burkholderia (Pseudomonas) cepacia,* found adaptation to heterogeneous microenvironments over 5 km distances (Wise *et al.* 1995). In addition, McArthur *et al.* (1992) showed *B. cepacia* and *Pseudomonas pickettii* underwent genetic shifts over 3.5 km. The information length is coupled with colonisation distance, the maximum distance downstream from the initial formed colony at which a new colony can establish; and together will determine microbial distributions and hence ecosystem function (McArthur & Tuckfield 1997). As the information length and colonisation distance are directly related to the average

temporal energy load of the system, the higher the energy in a system, the shorter the information lengths and therefore the shorter the distance an effective gene will persist in a system, which may lead to the loss of specific bacterial genotypes (McArthur & Tuckfield 1997).

In addition to short information lengths, the loss of bacterial taxa throughout river systems can also be a result of removal processes, or top-down controls, such as viral lysis or protistan grazing (Chow *et al.* 2014). Grazing and lysis are a significant source of bacterial mortality and each influence bacterial abundance and community composition differently. For instance, grazing is selective, with protists choosing prey based on motility, size and surface characteristics, therefore shaping bacterial community composition. Protistan grazing can contribute to the mortality of 10 to 250 bacteria per protozoa daily within the water column and this grazing leads to a reduction in bacterial production and its transfer to higher trophic levels (Bettarel *et al.* 2003; Chow *et al.* 2014; Fuhrman & Noble 1995; Weinbauer & Höfle 1998). This Viral lysis equally contributes to bacterial mortality, with recent studies indicating 0 to 30% of bacterial production is removed via lysis in marine and freshwater systems (Bettarel *et al.* 2003; Weinbauer & Höfle 1998). Lysis leads to bacterial production remaining within the microbial loop, with lysis products, such as cell contents, being used primarily by bacteria, stimulating bacterial productivity and carbon uptake (Weinbauer & Höfle 1998). Lysis is typically genus-specific and favoured at high population densities, targeting abundant and fast-growing bacteria whilst slow-growing resistant strains persist in what is referred to as the kill the winner hypothesis (Weinbauer *et al.* 2006; Thingstad 2000; Thingstad & Lignell 1997). As lysis and grazing control most bacterial productivity, they therefore significantly impact bacterial community composition and system dynamics.

Additional to persistence and loss is the appearance of bacterial taxa in aquatic environments, which can be attributed to reintegration from the benthos, detachment from biofilms via erosion, abrasion, grazing, sloughing or dislodgement via higher organisms, inputs from land run-off, morphological or hydrophobic changes, or atmospheric deposition (Jones *et al.* 2008; Leff *et al.* 1992; Moore *et al.* 2000). These introduced cells, whether planktonic or particle-associated, can then colonise their new environment (Leff *et al.* 1992). Therefore, the appearance of bacterial genotypes in aquatic system can have a significant impact on system function.

This study aims to describe the bacterial community composition of the Murray River upstream and downstream of a small rural town to assess the persistence and fluctuations in microbial spatial dynamics over 3.3 km. We hypothesise three species composition patterns: persistence from genotype conservation in fluvial systems; loss due to sampling distances that should exceed genotype information lengths; and appearance downstream, coupled with higher abundance and diversity via the perturbations from anthropogenic processes between sampling sites.

MATERIALS AND METHODS

Sample Sites

River water samples were collected upstream of Murray Bridge at Mobilong (-35.105, 139.275) and 3.3 km downstream at Long Island reserve, South Australia (-35.128, 139.295) on February 11th, 2014. Triplicate samples were taken at each site to determine silica, nitrate, nitrite, ammonium, phosphate, iron and total chlorine (free and combined) concentrations. Nutrient concentrations were determined via a LF

2400 photometer using Aquaspex water testing products. A two-tailed student t-test with assumed unequal variance was used to test whether nutrient concentrations were significantly different between sample sites. The Long Island sampling site had an island dividing the river channel, whilst the upstream site at Mobilong was an open water channel site with minimal vegetation. A communal ferry under daily use, houseboats and agricultural lands primarily raising cattle, occur between the two sampling sites. The Murray River is characterised by highly regulated flows and turbid waters with monitoring stations showing a water level of 0.55 metres AHD at the time of sampling (DEWNR 2014).

Heterotrophic bacterial enumeration via flow cytometry

At each sampling site, triplicate samples were collected for heterotrophic bacterial enumeration via flow cytometry. Samples were collected according to Dann *et al.* (2014 - Chapter 3) where 1 mL samples were transferred into 2 mL cryovials containing 20 µl of glutaraldehyde (0.5% final concentration) and stored at 4 ˚C in the dark for 15 minutes (Dann *et al.* 2014 - Chapter 3). Samples were then snap frozen in liquid nitrogen and stored at -80 ˚C (Brussaard 2004; Dann *et al.* 2014 - Chapter 3). Flow cytometric analysis was performed within three weeks to avoid deterioration (Brussaard 2004).

Flow cytometric enumeration followed the same protocol as described in previous studies (Roudnew *et al.* 2012, 2014; Seymour *et al.* 2000, 2004, 2007 and 2008). Samples were thawed and diluted 1:100 with Tris-EDTA buffer (pH 7.4, 0.2 µm filtered, 10 mM Tris, 1 mM EDTA) and stained with SYBR Green I at 1:500 dilution (Molecular Probes). The addition of 1 µm fluorescent beads (Molecular Probes) at 10^5 beads ml^{-1} per sample served as internal size and concentration standards.

Measured flow cytometry parameters were normalised to bead concentration and fluorescence (Brussaard *et al.* 2000; Brussaard 2004; Dann *et al.* 2014 – Chapter 3; Gasol & Del Giorgio 2000; Marie *et al.* 1997, 1999; Schapira *et al.* 2009).

Samples were run on a FACSCanto II flow cytometer (BD) equipped with a red (633 nm, 17 mW), blue (488 nm, 20 mW, air-cooled) and violet (405 nm, 30 mW) laser using a phosphate-buffered saline (PBS) sheath fluid. For each sample, green fluorescence (SYBR I), right-angle light scatter (SSC) and forward-angle light scatter (FSC) were acquired. Each sample was run for two minutes at a medium flow rate setting to achieve approximately < 1000 events per second. Triplicate negative control samples were analysed in each flow cytometry session to eliminate potential background noise created in sample preparation or from flow cytometer artifacts. Each negative control contained Tris-EDTA buffer (0.2 µl filtered) stained with SYBR Green I (Dann *et al.* 2014 – Chapter 3). Resulting histograms and cytograms were exported as FSC 3.0 files and used to enumerate heterotrophic bacterial populations via FlowJo (Tree Star, Inc.) (Smith *et al.* 2015). Discrimination of heterotrophic bacterial populations was based on side-scatter (SSC), indicative of cell size; and SYBR Green fluorescence, indicative of nucleic acid content (Brussaard 2004; Dann *et al.* 2014 – Chapter 3; Marie *et al.* 1997, 1999).

Bacterial 16s sequencing

- Bacterial concentration, purification and sequencing

Triplicate 5 L samples were collected at each sampling site using sterile carboys for bacterial 16s sequencing. During sampling, care was taken not to disrupt the river bed as water column samples were desired. Samples were processed immediately by prefiltering through 5 µm filters (Whatman) and a series of Nalgene filtering units

and vacuum pumps to remove large suspended particulate matter. Each sample was then filtered onto MoBio waterclean 0.22 µm filters and stored at -80°C until extraction. Bacterial DNA was extracted using a MoBio PowerWater DNA isolation kit (GeneWorks). DNA quality and concentration were determined via a Quibit fluorometer (Quant-iT dsDNA high-sensitivity Assay Kit; Invitrogen). Barcoded amplicon sequencing, referred to as bTEFAP, was performed on the extracted bacterial DNA. Briefly, 16S universal Eubacterial primers 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 518R (5'-ATTACCGCGGCTGCTGG-3') were added to the samples. A single-step 30 cycle PCR was then performed at 94^0C for 3 minutes using HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA), followed by 28 cycles at 94^0C for 30 seconds; 53^0C for 40 seconds and 72^0C for 1 minute. An elongation step was then carried out at 72^0C for 5 minutes. The amplicon products from all the samples were then combined in equal concentrations and purified via Agencourt Ampure beads (Agencourt Bioscience Corporation, MA, USA). Roche 454 FLX titanium instruments and reagents were used to sequence the bacterial DNA according to the manufacturer's guidelines yielding an average of 10 k sequences per sample (Molecular Research LP, Shallowater, TX, USA).

- Taxonomic and statistical analyses

For taxonomic analysis, sequenced bacterial DNA in the format of FASTA and quality mapping files were converted to FASTQ files and barcodes were stripped. Sequences were quality filtered and length truncated with reads discarded if shorter than 250 bp or containing > 0.5% expected errors for all bases. Sequences void of a recognisable barcode or forward PCR primer were also discarded. Full length dereplication was performed via the USEARCH pipeline followed by abundance

sorting with a minimum size of 2 to ensure singletons were removed (Edgar 2010). Singletons were removed to ensure the removal of spurious OTUs resulting from PCR errors and/or sequencing artifacts (Edgar 2013; Liu *et al.* 2015). OTUs were clustered at 97% identity via the cluster_otus command using UPARSE in USEARCH v8 (Edgar 2013). Reference-based chimeric filtering was performed via UCHIME using the gold database (Edgar *et al.* 2011; Pylro *et al.* 2014). Reads were then globally mapped to the OTUs using an identity threshold of 97%. Taxonomic assignment was performed using RDP Classifier via the utax command in USEARCH v8 (Edgar 2013; Wang *et al.* 2007; Zhang *et al.* 2014). OTU tables were constructed using python scripts and the resulting OTU tabbed text file was used to determine average abundances of OTUs.

Alpha rarefaction curves were constructed via the alpha_rarefaction.py command within QIIME using median sequence count as the upper limit of rarefaction depths (Caporaso *et al.* 2010). These rarefaction curves were used to determine whether sequencing efforts were adequate to enumerate species richness within the samples (Simberloff 1978). Rank abundance graphs were constructed of the lowest classification sequence reads to assess whether bacterial diversity displays power law behaviour (Edwards & Rohwer 2005). For rigor, maximum likelihood estimation was used to fit a power law and test significance (Clauset *et al.* 2009). Specifically, an optimal power law ($y=Ax^{-\alpha}$) was found for *n* species, counting from the most to least abundant, for each site (Clauset *et al.* 2009). The single most abundant species was removed as an outlier, and the model fit to species of equal or higher abundance than xmin. The resulting model was tested for significance via a Kolmogorov-Smirnov (KS) goodness of fit test. If the data and model fit the same distribution, the KS test statistic D_n is less than the KS critical value (95%

confidence). Bacterial OTU tables were run through multivariate statistical software package PRIMER (version 7) where data was square root overall transformed and similarity percentage (SIMPER) analysis was performed. SIMPER analysis was used to determine what is driving Bray-Curtis dissimilarity within and between sites. As the sample size in this dataset was smaller than what is required for conventional non-parametric tests (< 1000), PERMANOVA tests performing permutations were used (Anderson *et al.* 2008). Data was overall transformed and Bray-Curtis resemblance was performed to produce dissimilarity matrices which were analysed via principal coordinates (PCO) analysis and PERMdisp to determine the percentage of total variation and level of dispersion between site samples (Anderson *et al.* 2008). Metric multidimensional scaling (MDS) using Bootstrap Average analysis was performed on each sample to determine the level of spread between the upstream and downstream samples and produce smoothed and small 95% bootstrap regions for each factor, i.e. sample site. Metric MDS ordination used 500 bootstrap averages of the centroid of each group to show where 95% of the averages of the true centroid lies within multivariate space. Presence vs. absence overall transformed data with Bray-Curtis dissimilarity was used to determine site exclusivity of bacterial OTUs, where site exclusivity refers to bacterial OTUs present at one sampling site but not at the other. Two-sample t-tests assuming unequal variances were used to determine significant differences between the flow cytometrically defined heterotrophic bacterial subpopulations of both sampling sites.

RESULTS

Nutrient analysis

At the upstream site, the mean nutrient concentration for silica was 0.97 mg/L (95%CI = 0.07, n = 3), nitrate was 0.93 mg/L (95%CI = 0.13, n = 3), nitrite was 0.02 mg/L (95%CI = 0.01, n = 3), ammonium (total NH_3/total NH^{4+}) was 0.04 mg/L (95%CI = 0, n = 3), iron was 0.20 mg/L (95%CI = 0.11, n = 3), total chlorine was 0 mg/L (95%Ci = 0, n = 3), and phosphate was 0.83 mg/L (95%CI = 0.24, n = 3).

At the downstream site, the mean nutrient concentration for silica was 0.87 mg/L (95%CI = 0.26, n = 3), nitrate was 0.37 mg/L (95%CI = 0.17, n = 3), nitrite was 0.01 mg/L (95%CI = 0, n = 3), ammonium (total NH_3/total NH^{4+}) was 0.07 mg/L (95%CI = 0.04, n = 3), iron was 0.27 mg/L (95%CI = 0.17, n = 3), total chlorine was 0.03 mg/L (95%CI = 0.07, n = 3), and phosphate was 0.33 mg/L (95%CI = 0.24, n = 3). A two-tailed student t-test with assumed unequal variance identified nitrate and phosphate concentrations were significantly higher at the upstream site ($p < 0.009$, $p < 0.045$).

Heterotrophic bacterial abundance

Two prokaryotic subpopulations were identified via monoparametric histograms of SYBR Green I fluorescence and biparametric cytograms of side-scatter (SSC) and SYBR Green I fluorescence. The level of fluorescence was used to differentiate prokaryotes into high-DNA (HDNA) and low-DNA (LDNA) subpopulations. Upstream the mean LDNA abundance was 3.3×10^7 cells ml^{-1} (95%CI = 1.2×10^7, n = 3), the mean HDNA abundance was 1.6×10^7 cells ml^{-1} (95%CI = 6.2×10^5, n = 3) and the mean total heterotrophic bacterial abundance was 5.0×10^7 cells ml^{-1} (95%CI = 1.2×10^7, n = 3). Downstream the mean LDNA abundance was 2.3×10^7 cells ml^{-1} (95%CI = 5.8×10^6, n = 3), whilst the mean HDNA abundance was 1.1×10^7 cells ml^{-1} (95%CI = 4.7×10^6, n = 3) and the mean total heterotrophic bacterial abundance was 3.4×10^7 cells ml^{-1} (95%CI = 2.9×10^7, n = 3) (Table. S1). Two-sample t-tests

revealed the mean LDNA and HDNA abundances were not significantly different between the upstream and downstream sampling sites ($p > 0.2$).

Upstream and downstream bacterial taxonomic composition

A total of 108,178 matched primer sequences were quality filtered resulting in 86,882 (80.3%) reads passed, 3,633 short reads discarded (< 250 bp) and 17,663 low quality records discarded (expected errors > 0.5). Dereplication identified 16,285 unique sequences and 11,341 singletons. OTU clustering at 97% minimum identity revealed 655 OTUs and 458 chimeras (9.3%). Of these 655 OTUs, chimera removal using a reference database revealed 30 OTUs were chimeric. Taxonomy assignment via the RDP Classifier assigned the sequences to a total of 613 OTUs. Of these 613 OTUs, there were 78,511 total reads with upstream samples containing between 12,027 and 15,500 reads (mean = 14030.3, std. dev. = 1796.9) and downstream samples containing between 10,473 and 13,849 reads (mean = 12140, std. dev. = 1688.4). Rarefaction plots revealed the upstream site samples may each contain > 269, > 290 and > 273 OTUs and the downstream site samples may each contain > 232 and > 269 OTUs. Each sample did not reach asymptote. The outlier sample downstream, Downstream 1, had a steeper slope compared to the other samples (Fig. 1).

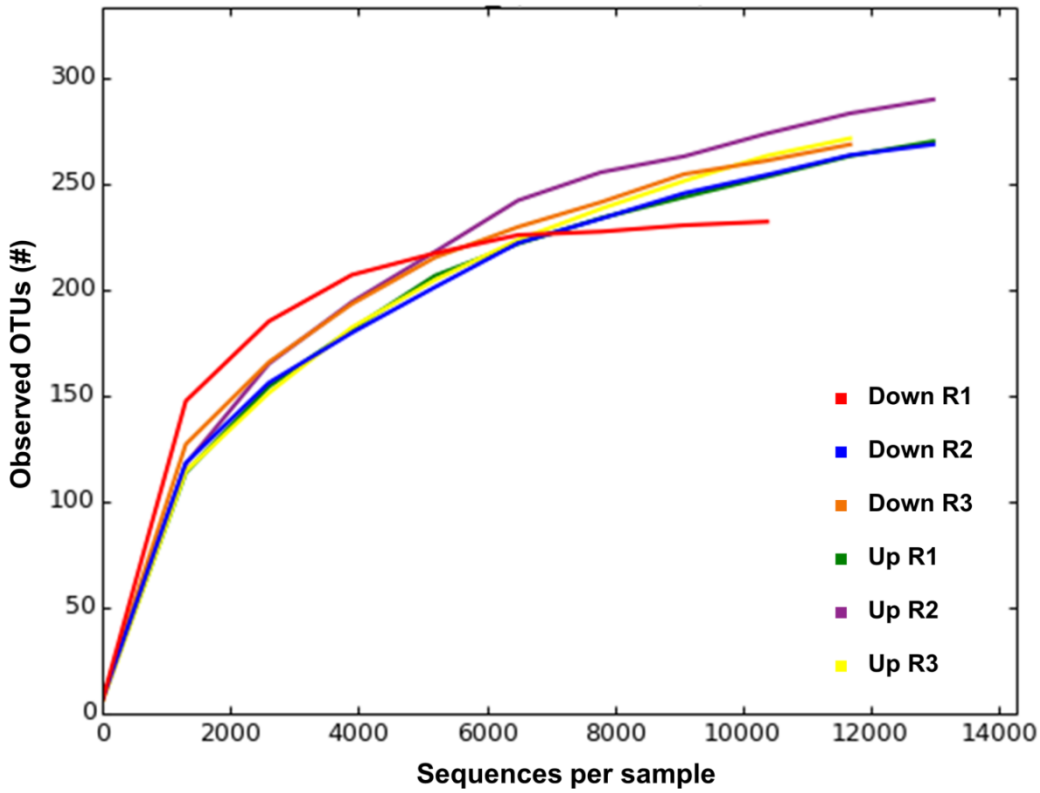


Figure 1. Rarefaction curve for observed OTUs in each site sample. Down = downstream, Up = upstream and R# = sample number. OTUs determined via the RDP Classifier reference database within the UPARSE pipeline (Edgar 2013; Wang *et al.* 2007).

Rank abundance graphs of the lowest sequence classification showed significant power law relationships at all sites and therefore best described the community structure at the upstream and downstream sites (Fig. 2 & Table S2). The slopes for the upstream samples were 1.71, 1.75 and 1.77 whilst the downstream sample slopes were 1.80, 1.61 and 1.63 (Table S2).

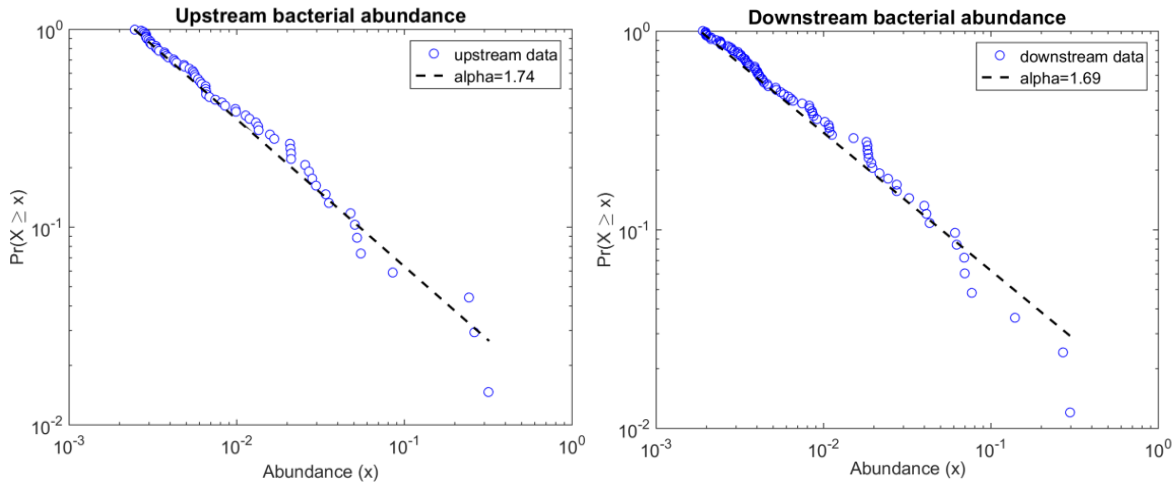


Figure 2. Rank abundance graphs of the bacterial species **A** upstream and **B** downstream.

Taxonomic analysis revealed Actinobacteria (total 45%, n = 6) and Proteobacteria (total 34.2%, n = 6) dominated the sequences at both sites. Of the Proteobacteria, the Alphaproteobacteria (total 71.8%, n = 6) dominated the sequence reads, followed by Betaproteobacteria (total 17.7%, n = 6). Acidobacteria (total 5.8%, n = 6), Chloroflexi (total 4.5%, n = 6) and Bacteroidetes (total 2.6%, n = 6) were also abundant at both sampling sites. Verrucomicrobia (0.9%, n = 6), Chlorobi (0.15%, n = 6) and TM7 (0.1%, n = 6) were also present at both sites (Fig. 3).

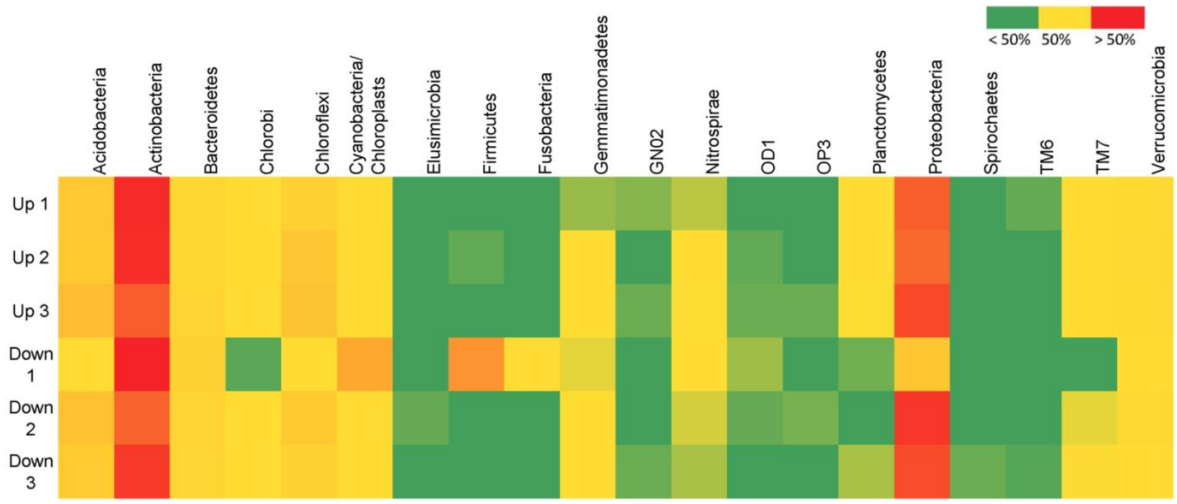


Figure 3. Taxonomic heatmap of the average abundance percentages of bacterial phyla upstream and downstream. Colour scale bar indicates percent average abundance. Numbers 1, 2 and 3 indicate the sample number. Sequences were matched to the RDP Classifier reference database via the UPARSE pipeline (Edgar 2013).

However, one downstream sample, Downstream 1, contained a different taxonomic profile to the other samples (Fig. 3 & 4). Within this sample there was an abundance of Actinobacteria 52.6%, Firmicutes 7% and Cyanobacteria 5.2% (Fig. 3). This abundance of Actinobacteria was due to *Propionibacterium acnes* which accounted for 43.5% average abundance and was absent in all other samples. The other samples were dominated by Actinomycetales (13.6%, std. dev. 0.01, n = 5) and *Actinomycetales ACK-M1* (26.7%, std. dev. 0.06, n = 5), whilst their average abundances in the atypical downstream sample were 0.9% and 3.2% respectively. The abundance of Firmicutes in this single sample was due to Staphylococcus (7.0%), Finegoldia (2.2%), Anaerococcus (1.9%) and Ruminococcaceae (1.1%) dominating the Clostridia sequence hits, and Streptococcus (2.0%) dominating the Bacilli hits. The abundance of Cyanobacteria was due to Streptophyta (14.8%).

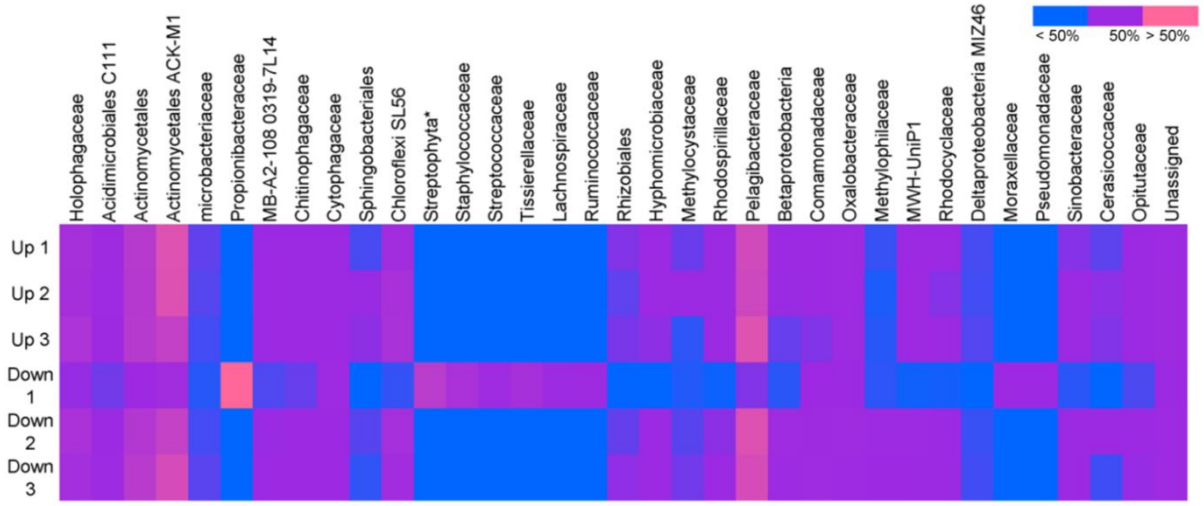


Figure 4. Taxonomic heatmap showing the bacterial families upstream and downstream. Colour scale bar indicates percent average abundance. When family level description was unavailable, the lowest classification level was given. * indicates sequences chloroplast in origin. Numbers 1, 2 and 3 indicate the sample number. Sequences were matched to the RDP Classifier reference database via the UPARSE pipeline (Edgar 2013).

A total 216 bacterial taxa were identified with 119 taxa found upstream and 194 found downstream. Presence vs absence data revealed 117 site exclusive bacteria, with the upstream site containing 22 exclusive bacteria and downstream containing 95 exclusive bacteria (Table 1 & 2). Kiloniellales and Ellin6067 were the only bacteria present in all three samples of the upstream sampling site and absent from the downstream site. Within the 95 exclusive taxa downstream, 81 were in a single sample, Downstream 1 (Table 2). The majority of families found only in this sample were human- or animal-associated, whilst the majority of bacteria in the other samples were typical of freshwater aquatic environments (Table 1 & 2).

Table 1. Bacterial taxa exclusive upstream.

Upstream	
Solibacteraceae Candidatus Solibacter	OP3 PBS-25
Rhodococcus	Gemmata
Micrococcaceae	DH61
Chloroplast - Stramenopiles	Kiloniellales
Oscillatoriophycideae	Beijerinckiaceae
Phormidium	Erythrobacteraceae
GN02 3BR-5F	NB1-j MND4
Alcaligenaceae	Aeromonadaceae
Ellin6067	Shewanella
Dechloromonas	Crenothrix
Pedosphaerales auto67_4W	Chthoniobacteraceae

Table 2. Bacterial taxa exclusive downstream. *Indicates taxa found only in downstream outlier sample

Downstream		
Chloracidobacteria RB41*	Moraxellaceae*	Finegoldia*
Acidimicrobiales*	Acinetobacter*	Gallicola*
Varibaculum*	Acinetobacter rhizosphaerae*	Peptoniphilus*
Corynebacterium*	Staphylococcus*	Tissierellaceae WAL_1855D*
*Microbacterium maritypicum**	Enhydrobacter*	Anaerostipes*
Nocardioidaceae*	Moraxella*	Blautia*
*Propionibacterium acnes**	Pseudomonadaceae*	Bacillus*
Propionibacterium granulosum*	Pseudomonas*	*Veillonella dispar**
Pseudonocardia*	*Pseudomonas nitroreducens**	Erysipelotrichaceae*
Coriobacteriaceae*	*Thermomonas fusca**	Fusobacterium*
Coriobacteriaceae Atopobium*	Herbaspirillum	Leptotrichia*
*Collinsella aerofaciens**	Nitrosomonadaceae	Clostridiaceae*
Rubrobacter*	Tremblayales	Lachnospiraceae*
Gaiellaceae*	Bacteriovoracaceae	Lactococcus*
Chitinophagaceae Niabella *	Bdellovibrio	Streptococcus*
Prevotella*	Haliangiaceae	*Streptococcus infantis**
Paraprevotella*	Coxiellaceae	Clostridiales *
Bacteroides*	*Hydrocarboniphaga effusa*	Tissierellaceae*
*Bacteroides uniformis**	Spirochaetaceae Treponema	Anaerococcus*
Porphyromonas*	Gemm-3	*Faecalibacterium prausnitzii**
*Porphyromonas endodontalis**	OD1 SM2F11	Oscillospira*
Prevotella;*	OP3 koll11GIF10kpj58rc	Ruminococcus*
*Prevotella copri**	Ochrobactrum*	*Staphylococcus aureus**
Rikenellaceae*	Agrobacterium*	Facklamia*
Cytophagaceae Adhaeribacter*	Rhodospirillales	Granulicatella*
Weeksellaceae*	*Phaeospirillum fulvum*	Planococcaceae*
Chryseobacterium*	Sphingobium*	Jeotgalicoccus*
Cloacibacterium*	Comamonas*	Haemophilus parainfluenzae*
*Capnocytophaga ochracea**	Herbaspirillum*	Coprococcus*
Gitt-GS-136*	Ralstonia*	Ruminococcaceae*

Chloroplast - Streptophyta*	Neisseriaceae*	Enterobacteriaceae*
Elusimicrobia MVP-88;	Campylobacter*	

SIMPER analysis at the lowest classification level revealed an average similarity of 86.60 (n = 3) for the upstream samples and an average similarity of 44.61 (n = 3) for the downstream samples. The main contributors to similarity for the upstream site were Pelagibacteraceae, ACK-M1, Actinomycetales, Holophagaceae and Chloroflexi SL56, which contributed to 12.7%, 11.3%, 8.7% and 5.9% similarity respectively. Downstream the main contributors were also ACK-M1, Pelagibacteraceae and Actinomycetales, which contributed to 11.5%, 9.3% and 7.8% similarity respectively. There was an average dissimilarity of 36.21 (n = 6) between upstream and downstream with ACK-M1 and *Propionibacterium acnes* contributing to > 5% dissimilarity.

Hierarchical agglomerative clustering of samples via CLUSTER analysis revealed the highest similarity between upstream sample 1 and 2 which fused at 87% similarity, followed by downstream sample 3, which fused with these upstream samples at 85% and upstream sample 3 and downstream sample 2 which also fused at 85% similarity (Fig. 5). All these samples fused at 83% similarity. However, downstream sample 1 fused at a similarity of only 28% (Fig. 5). PERMdisp analysis revealed the mean dispersion was 8.4 (SE 0.5) for the upstream samples and 32.4 (SE 7.5) for the downstream samples.

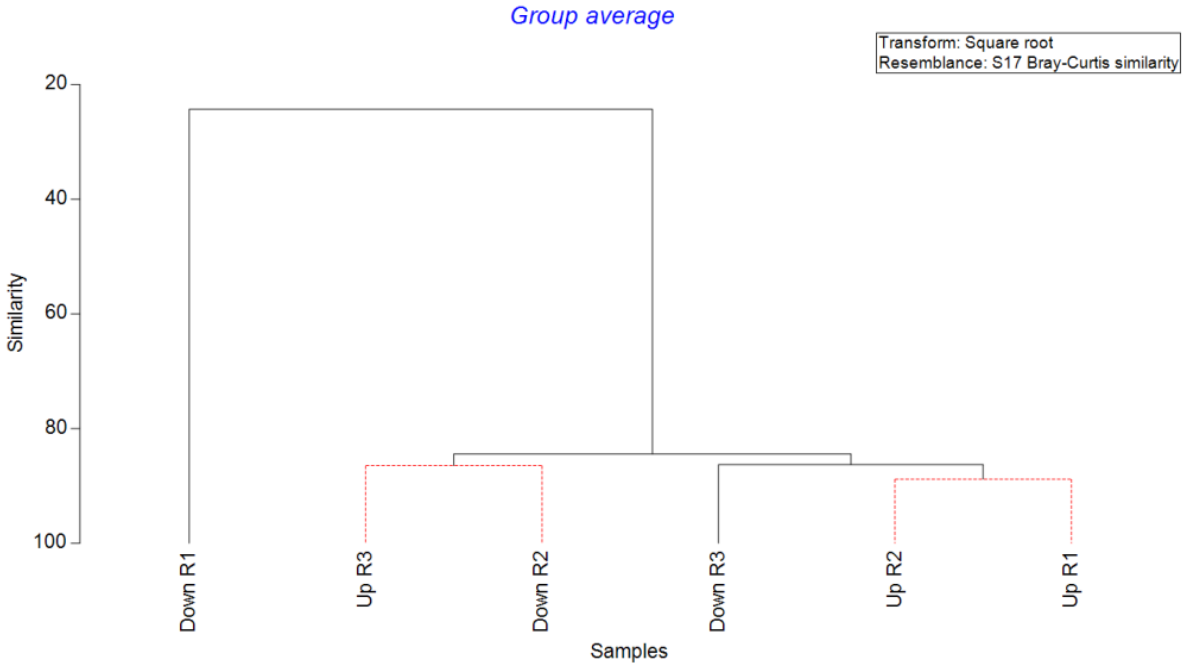


Figure 5. CLUSTER dendrogram showing similarity of the upstream (UP R1, R2 & R3) and downstream (DOWN R1, R2 & R3) samples.

PCO analysis showed 95% of the total variation between samples could be explained by two axes, with 91.6% variation explained by the PCO1 axis (Fig. 6). Metric MDS using Bootstrap Averages analysis showed no overlap between the site samples indicating very different behaviour between the factors, i.e. sample sites (Fig. 7).

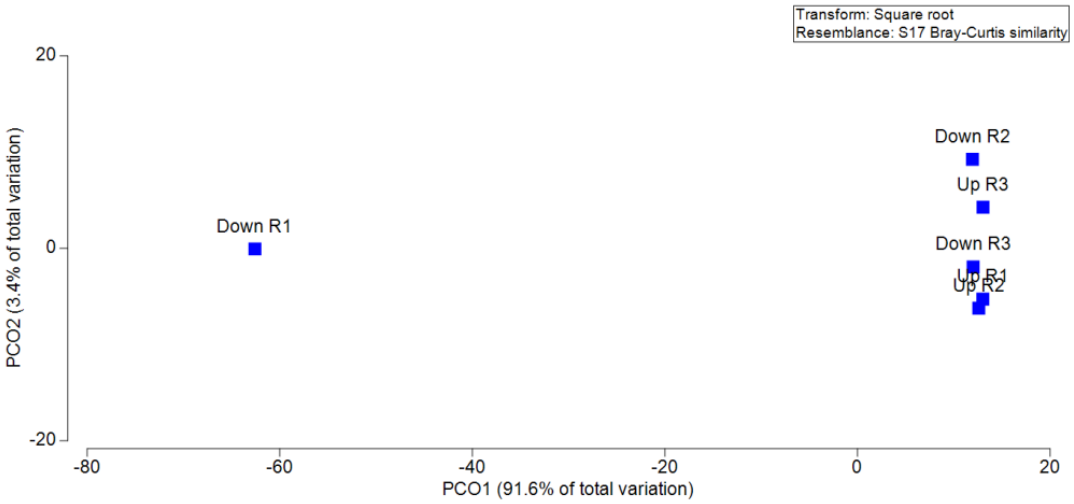


Figure 6. PCO analysis plot of upstream and downstream samples. R1, R2 and R3 indicate sample number.

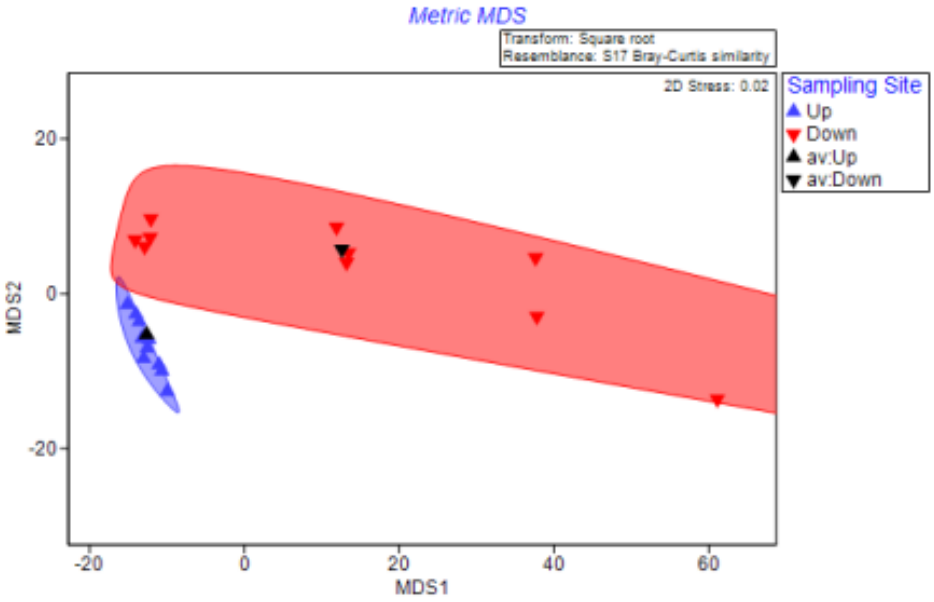


Figure 7. Metric MDS plot of the upstream and downstream samples using Bootstrap Averages analysis.

DISCUSSION

Site similarities reveal persistence

As fluvial systems are continua it is unlikely bacterial genotypes will be conserved to specific locations (McArthur & Tuckfield 1997). Similar to what has been observed in microbial communities of the human gut, microbes must pass from an upstream to a downstream location where they compete with residential microorganisms in order to establish (Wise *et al.* 1996). The transport of these microbial genotypes and the distance at which they will persist is reflected in the information length and colonisation distance. Here the ACK-M1 and Pelagibacteraceae were persistent within the river system implying information lengths and colonisation distances equal to or greater than 3.3 km (Fig. 4). Previous studies have shown Actinomycetales

ACK-M1, now known as the acI lineage, is one of the most common bacterial lineages in freshwater systems (Newton *et al.* 2007). Acl consists of small cells capable of avoiding some bacterivorous protist attack therefore providing this lineage an advantage over other bacteria as it can actively thrive and distribute itself widely throughout many diverse freshwater systems (Dennis *et al.* 2013; Jezbera *et al.* 2009; Newton *et al.* 2006, 2007; Warnecke *et al.* 2004). In addition, Newton *et al.* (2006) found this bacterial lineage was an exception to the rule of continual temporal variation within the bacterial community composition of a humic lake. It was suggested that the acl lineage is an exclusive pelagic freshwater lineage capable of sustainable growth within freshwater lakes and is not present due to transient soil re-suspension (Newton *et al.* 2006, 2007; Warnecke *et al.* 2004; Wu *et al.* 2007; Zwart *et al.* 2002).

The Pelagibacteraceae family, also known as the SAR11 clade, is ubiquitous in marine environments and has been shown to dominate heterotrophic bacterial community composition (Brown *et al.* 2012). Although the SAR11 clade is restricted to aquatic environments, it shows a bifurcation into a freshwater lineage, referred to as LD12, therefore indicating its ability to adapt to varying salinity levels (Brown *et al.* 2012; Zwart *et al.* 2002*)*. Interestingly, the Actinomycetales ACK-M1 lineage clusters with a marine organism, CRO-FL14, and it was suggested by Zwart *et al.* (2002) that this may indicate frequent transitions and/or higher salinity range tolerances. As the Murray River feeds into the Indian Ocean via the Murray Mouth and several lakes with high and varied salinity ranges this would indicate that these Proteobacteria and Actinobacteria lineages, given they persist throughout the system via long information lengths and colonisation distances, may have an advantage in transitioning from their current freshwater environment into a marine habitat (Kiem &

Verdon-Kidd 2011; Zwart *et al.* 2002). In addition these lineages may reflect the high salinity levels encountered in the Murray River due to saline groundwater influences (Goss 2003; Jolly *et al.* 2001; van Dijk *et al.* 2007).

Persistence may also relate to the kill the winner hypothesis where there is a trade-off between growth rate and defence (Thingstad 2000; Thingstad & Lignell 1997). Here, the persistent bacteria would represent slow-growing defensive strains which are abundant in systems in contrast to fast-growing competitive strains which are subject to viral lysis (Rodriguez-Brito *et al.* 2010; Rodriguez-Valera *et al.* 2009; Thingstad 2000; Winter *et al.* 2010).

In addition, the nutrient concentrations reported here show persistence in silica and iron, with no significant concentration difference between sites indicating either nutrient uptake lengths at or above 3.3 km or replenishment sources (McArthur & Tuckfield 1997). In addition, phosphate and nitrate were significantly higher upstream, yet still present downstream, indicating shorter nutrient uptake lengths, and hence greater retention efficiency for these specific nutrients, perhaps due to their removal via benthic biotic and abiotic processes (Haggard *et al.* 2005; Newbold *et al.* 1981). Elwood *et al.* (1983) estimated the average distance a phosphate ion travelled prior to its uptake by stream microbes as 6 metres during peak autumn leaf fall and 165 metres during summer indicating a temporal difference in nutrient uptake (Mulholland *et al.* 1985). As sampling occurred during summer, this would suggest the presence of long nutrient spiralling lengths.

Acidobacteria, Chloroflexi and Bacteroidetes were also abundant at both sampling sites indicating their persistence (Fig. 3). Previous freshwater studies have found an abundance of Bacteroidetes*,* which are filamentous bacteria known to be associated

with the human gut microbiota (Andersson *et al.* 2008; Ballesté & Blanch 2010; Newton *et al.* 2006; Zwart *et al.* 2002). However, Acidobacteria and Chloroflexi are less commonly reported in freshwater studies and this may be due to seasonal patterns as this study sampled in the last month of summer, whilst previous studies sampled at the start/during winter (Zwart *et al.* 2002), at the end of spring/start of summer (Crump *et al.* 1999) or did not state the sampling period (Newton *et al.* 2006). A study conducted seasonally that looked at the bacterial community composition of a lake in China found the highest diversity was during Summer, where Acidobacteria and Chloroflexi were observed, and it was proposed that this related to the occurrence of cyanobacterial blooms, which correlates with the findings of this study (Crump *et al.* 2003; Wu *et al.* 2007). The presence of Cyanobacteria at both sites supports this statement.

Site dissimilarities and exclusivity reveal loss and appearance

Taxonomic analysis revealed shifts within community structure between upstream and downstream with the loss of some species and gain of others (Table 1 & 2). The loss of species from upstream can be attributed to genotypes with information lengths and colonisation distances shorter than the 3.3 km sampling interval, perhaps due to removal processes such as grazing and lysis, which could have led to the loss of specific genotypes (Bettarel *et al.* 2003; Chow *et al.* 2014; Leff *et al.* 1992; McArthur & Tuckfield 1997; Weinbauer & Höfle 1998). This loss and appearance led to site exclusivity, which was evident in 10% of the species upstream and 44% of the species downstream. The taxa exclusive upstream were predominantly environmental bacteria, such as Erythrobacteraceae, Kiloniellales and Phormidium, indicating a microbial community dedicated to system productivity (Teneva *et al.* 2005; Tonon *et al.* 2014; Wiese *et al.* 2009) (Table 1). In contrast,

downstream the majority of exclusive taxa were human- and/or animal-associated, such as Lachnospiraceae*,* Prevotellaceae and Tissierellaceae (Alauzet *et al.* 2014; De Menezes *et al.* 2011; Meehan & Beiko 2014), indicating domestic and/or agricultural inputs, perhaps due to houseboats and run-off from agricultural lands, leading to perturbations within the microbial community composition between the two sampling sites (Table 2).

The presence of Propionibacteriaceae downstream and its absence upstream was the main contributor to dissimilarity between sampling sites (Fig. 4). Propionibacteriaceae are primarily known for their role in the cheese industry, being involved in flavour development and the formation of “eyes” in cheeses (Stackebrandt *et al.* 2006). In addition, *Propionibacterium acnes* has been isolated from chronically infected sinuses, post-operative infections, healthy nasal cavities, female genital tracts, prostate tissue, faeces and skin sebaceous glands (Bassis *et al.* 2014; Goldschmidt *et al.* 2008; Stackebrandt *et al.* 2006). Therefore Propionibacteriaceae may have been present downstream due to dairy production sites and/or agricultural farms or the dense human population throughout the town of Murray Bridge where waste may have entered the river water column via effluent or land run-off.

These changes in bacterial community composition due to species loss upstream to downstream, could relate to succession mechanisms produced by *in situ* changes in new cell production within different bacterial populations. These changes occur via a multifaceted mixture of environmental factors such as the abundance of and selection by grazing protists, phytoplankton species composition and viral lysis (Crump et al. 2003; Hahn & Höfle 1999; Van Hannen et al. 1999). These bacterial community composition shifts could also relate to the introduction of allochthonous

bacterial species via advection from land run-off or inflowing streams/water inputs (Crump et al. 2003).

For protistan grazing and viral lysis, all bacterial strains able to grow faster than the grazing rate can establish. However, growth that is faster than the grazing rate is compensated by viral lysis; therefore even if bacteria are able to overcome the effect of grazing they must also simultaneously avoid viral lysis to survive (Rodriguez-Brito *et al.* 2010; Winter *et al.* 2010). This indicates the species lost from upstream may have either (i) slow growth rates that were overcome by the grazing rate or (ii) fast growth rates that were susceptible to viral attack. The latter is observed via selective sweeps caused by viruses which kill all non-resistant bacteria (Rodriguez-Brito *et al.* 2010; Rodriguez-Valera *et al.* 2009).

Sample dissimilarity reveals small-scale variability

Sample dissimilarity due to heightened Firmicutes and Cyanobacteria abundance in a single sample downstream indicated microscale patchiness (Fig. 3). Microbial patchiness is well established in marine systems (Azam 1998; Azam & Malfatti 2007; Blackburn *et al.* 1998; Dann *et al.* 2014 – Chapter 3; Seymour *et al.* 2000, 2004, 2005ab, 2007, 2009; Stocker 2012; Stocker *et al.* 2008), with taxonomic differences observed between 1 µl seawater samples (Long & Azam 2001). Within freshwater systems, heterogeneous microscale hotspots and coldspots in abundance were reported recently in this environment (Dann *et al.* submitted – Chapter 4), with bacteria exhibiting an 80.5-fold variation over 0.9 cm (Dann *et al.* submitted – Chapter 4). Here we report hotspots and coldspots in discrete taxa, which led to large differences between samples from the same site (Fig. 5 & 6). The Firmicutes abundance in the dissimilar downstream sample may indicate hotspots of faecal matter, which would act as fertilizer. This would explain the heightened

Cyanobacteria abundance, which could form the beginning of a future cyanobacterial blooms (Oliver & Ganf 2002). As the primary taxonomic differences between the sampling sites were due to sample dissimilarity, local small-scale variability may have a greater impact on taxonomic diversity profiles compared to taxonomic patterns over the 3.3 km compared here. These results agree with previous findings that showed the bias of using the "mean field approach" to estimate microbial abundance and here it is also shown for taxonomic profiles in freshwater systems (Dann *et al.* 2014 – Chapter 3).

Large dispersion between the downstream samples indicated higher heterogeneity downstream (Fig. 7). Dispersal and competition determine microbial community structure and occur via local species sorting or mass effects, such as emigration from water flow rate out of a habitat (Adams *et al.* 2014). Mass effects dominate as the mechanism for community structure when dispersal is high, whilst local species sorting dominates when dispersal is low (Adams *et al.* 2014). As Murray Bridge is characterised by low flow rates, local species sorting would dominate in this environment. Local species sorting results from selective pressures within a habitat impacting on the bacterial species present, thus determining which species thrive (Adams *et al.* 2014; Sax *et al.* 2007; Vrba & Gould 1986). Therefore, the influx and loss of species, followed by local species sorting would result in a disturbance in microbial biogeographic patterns within the river (Adams *et al.* 2014; Van der Gucht *et al.* 2007). If this sorting occurs at the scale relevant to bacterial cells, micrometres to centimetres, this would lead to the formation of heterogeneous microenvironments which would explain the larger dispersion, as well as dissimilarity between the downstream samples.

Community diversity structure

Rarefaction curves did not reach asymptote for either sampling site indicating sequencing effort was insufficient for the microbial diversity within this environment, corresponding with previous freshwater studies (Liu *et al.* 2015; Wang *et al.* 2012) (Fig. 1). The downstream outlier sample, despite having lower sequence reads, had a steeper slope than the other samples indicating lower diversity within this sample. As there was an abundance of Cyanobacteria and Firmicutes in this outlier sample, the lowered diversity could imply these phyla hindered the establishment of other taxa therefore creating lower diversity. The rarefaction curves for the outlier sample indicated > 232 OTUs, whilst the remaining samples indicated > 269 to > 290 OTUs per sample, implying a sampling effort of > 11,000 to > 13,000 sequences per sample, rather than sample site, is required to reasonably determine microbial diversity within this river system (Fig. 1). However, as singletons were removed during quality filtering, horizontal asymptote is hard to achieve (Edgar 2013). If singletons were included asymptote may have been reached, however, here we chose to discard singletons to ensure conservative results as some singletons are associated with sequencing errors and/or PCR artifacts (Edgar 2013). In addition, the lack of horizontal asymptote suggests the use of Illumina sequencing technology, rather than 454 pyrosequencing, as higher throughput can be achieved. However, with an increase in sequencing throughput comes a decrease in read length and therefore makes OTU identification and taxonomic assignment less resolved, which would be undesirable for this river system as it has not been taxonomically defined previously (Liu *et al.* 2015).

Power law trends in bacterioplankton

Here, power law trends best described bacterial taxonomic community structure in all samples (Fig. 2). Power laws have been observed in a variety of systems, such as

bacterial abundances in mangrove estuary and marine environments (Dann *et al.* 2014 – Chapter 3), marine phage communities (Edwards & Rohwer 2005), gene expression levels (Ueda *et al.* 2004) and the intermolecular interactions of different proteins (Park *et al.* 2001), and are typically representative of rich-grow-richer mechanisms, or the kill the winner hypothesis in the case of marine phage communities (Hoffmann *et al.* 2007; Thingstad 2000). The slope value for the downstream outlier sample was substantially different to the other sample slopes perhaps indicating a shift in the mechanisms driving the power law. We suggest here a difference in slopes between samples could potentially be used as an indicator for shifts in system dynamics, such as heightened patchiness.

CONCLUSION

Here we report persistence, loss and appearance in bacterial genotypes upstream and 3.3 km downstream of a river system. Persistence via site similarity of dominant archetypal freshwater taxa suggests information lengths and colonisation distances equal to or greater than 3.3 km for some bacterial taxa. The loss of microbial species observed as site exclusivity upstream and downstream was attributed to mechanisms involved in the kill the winner hypothesis where bacteria with high growth rates were selected against via viral sweeps and protistan grazing. The appearance of microbial species was suggestive of immigration from domestic and agricultural land, resulting in the appearance of human- and/or animal-associated microbial communities.

Sample dissimilarity due to microscale heterogeneous hotspots of discrete Firmicutes and Cyanobacteria species, as well as high dispersion at the downstream

site, indicated microbial patchiness exists in freshwater systems. Local species sorting was suggested as the dominant mechanism for these discrete microenvironments in the microbial community structure. As the primary taxonomic differences between upstream and downstream were due to dissimilarity in samples from the same site, this indicates microscale patchiness may impact taxonomic diversity profiles more so than taxonomic patterns over larger distances, here 3.3 km. These results further demonstrate the bias of using the "mean field approach" to estimate microbial community dynamics in freshwater systems.

SUPPLEMENTARY INFORMATION

Table S1. Heterotrophic bacterial abundance determined via flow cytometry. 1, 2 and 3 indicates sample number.

Sample	**Subpopulation abundance** x 10^7 cells ml^{-1} (95%CI, n)		
	LDNA	HDNA	Total
Upstream 1	2.91	1.57	4.48
Upstream 2	4.54	1.67	6.22
Upstream 3	2.54	1.66	4.20
Upstream Average	3.33 (1.2 x 10^7, 3)	1.64 (6.2 x 10^5, 3)	4.97 (1.2 x 10^7, 3)
Downstream 1	2.06	0.79	2.85
Downstream 2	2.85	1.58	4.43
Downstream 3	1.89	0.97	2.86
Downstream Average	2.27 (5.8 x 10^6, 3)	1.11 (4.7 x 10^6, 3)	3.38 (2.9 x 10^7, 3)

Table S2. Resulting values for maximum likelihood testing of upstream (Up) and downstream (Down) sites and their corresponding samples (R1, R2 and R3). N = number of species.

Site	***n***	**Minimum abundance (*xmin*)**	**Alpha (α)**	**KS statistic (D_n)**	**KS critical value (95% confidence)**

Down R1	34	1.97E-03	1.803	0.0904	0.2372
Down R2	33	1.18E-03	1.609	0.0612	0.2399
Down R3	35	1.06E-03	1.629	0.0747	0.2331
Up R1	29	1.62E-03	1.707	0.1094	0.2555
Up R2	24	2.47E-03	1.748	0.0846	0.2803
Up R3	23	2.69E-03	1.773	0.1402	0.2862
Down Combined	83	1.84E-03	1.695	0.0774	0.1542
Up Combined	68	2.45E-03	1.744	0.0615	0.1607

Chapter 3

Virio- and Bacterioplankton Microscale Distributions at the Sediment-Water Interface

Published as: Dann LM, Mitchell JG, Speck PG, Newton K, Jeffries T & Paterson J (2014) Virio- and bacterioplankton microscale distributions at the sediment-water interface. *PLoS One* 9 (7): 1-14

ABSTRACT

The marine sediment-water interface is an important location for microbially controlled nutrient and gas exchange processes. While microbial distributions on the sediment side of the interface are well established in many locations, the distributions of microbes on the water side of the interface are less well known. Here, we measured that distribution for marine virio- and bacterioplankton with a new two-dimensional technique. Our results revealed higher heterogeneity in sediment-water interface biomass distributions than previously reported with a greater than 45– and 2500-fold change cm^{-1} found within bacterial and viral subpopulations compared to previous maxima of 1.5- and 1.4-fold cm^{-1} in bacteria and viruses in the same environments. The 45-fold and 2500-fold changes were due to patches of elevated and patches of reduced viral and bacterial abundance. The bacterial and viral hotspots were found over single and multiple sample points and the two groups often coincided whilst the coldspots only occurred over single sample points and the bacterial and viral abundances showed no correlation. The total mean abundances of viruses strongly correlated with bacteria ($r = 0.90$, $p < 0.0001$, $n = 12$) for all three microplates ($n = 1350$). Spatial autocorrelation analysis via Moran's I and Geary's C revealed non-random distributions in bacterial subpopulations and random distributions in viral subpopulations. The variable distributions of viral and bacterial abundance over centimetre-scale distances suggest that competition and the likelihood of viral infection are higher in the small volumes important for individual cell encounters than bulk measurements indicate. We conclude that large scale measurements are not an accurate measurement of the conditions under which microbial dynamics exist. The high variability we report indicates that few microbes experience the 'average' concentrations that are frequently measured.

INTRODUCTION

Marine ecosystems are dependent on the microbial loop to carry out critical marine biogeochemical processes such as carbon and nitrogen fixation (Azam & Malfatti 2007; Cotner & Biddanda 2002; Elser *et al.* 1995; Gasol & Del Giorgio 2000; Pace 1997; Pontes *et al.* 2007; Rodriguez-Brito *et al.* 2010; Venter *et al.* 2004). These marine biogeochemical processes often occur at or near cell surfaces with interactions occurring over micrometres to centimetres (Azam 1998; Long & Azam 2001; Mitchell *et al.* 1985; Seymour *et al.* 2004, 2005a, 2005b, 2006; Suttle 2007; Waters *et al.* 2003). However, marine microbe distributions are commonly measured in large volumes and then that value is extrapolated over larger areas or volumes in what is sometimes termed a 'mean field' approach (Angly *et al.* 2006; Corinaldesi *et al.* 2003; Hillmer *et al.* 2008; Liu *et al.* 2002; Pan *et al.* 2005, 2007; Seymour *et al.* 2004; Strutton *et al.* 1997).

There is growing awareness of deficiencies in this mean field approach for productivity, biomass and process estimates (Curtis *et al.* 2006; Grunbaum 2002; Karl & Dore 2001; Medlin *et al.* 2000; Seymour *et al.* 2004; Waters & Mitchell 2002). Studies on patch dynamics in terrestrial and aquatic ecosystems show most organisms have heterogeneous spatial distributions. This patchiness would be missed if the sampling resolution used was not appropriate for the size of the organism studied (Genin *et al.* 1994; Grünbaum 2012; Levin & Whitfield 1994; Pickett & Cadenasso 1995). For instance, Brentnall *et al.* (2003) found that the mean field production of planktonic communities greatly underestimated actual production due to small scale patchiness. The little work on picoplankton distributions indicates that there is microscale spatial heterogeneity of bacteria and viruses, and that 'hotspots'

and ‘coldspots’ are common, with bacterial and viral abundance changing 5-fold and 1.4-fold cm^{-1} (Seymour *et al.* 2000, Seymour *et al.* 2006).

Hotspots are areas of high microbial abundance which are believed to result from interactions between microbes and organic matter (Azam 1998; Seymour *et al.* 2004, 2005a, 2006, 2007b). These viral and bacterial hotspots were often found to be due to bacterial response to nutrient patches, as marine bacteria have been found to accumulate around areas of high nutrient concentration via chemotaxis (Azam 1998; Mitchell 2002; Barbara & Mitchell 2003; Blackburn *et al.* 1998; Seymour *et al.* 2010; Stocker *et al.* 2008), This aggregation of bacteria then results in bacterial hotspots that are often followed by increased viral production via lysis due to close host proximity (Seymour *et al.* 2006). In contrast, coldspots are areas of low microbial abundance that are believed to be a result of lysis or grazing events in the case of lowered bacterial abundance and grazing or attachment in the case of lowered viral abundance (Water *et al.* 2003).

The presence of these hotspots and coldspots are the units of microscale patchiness. The lower size limit of patches is controlled by the Batchelor scale, which is the smallest scale that nutrient gradients can occur before being dispersed by diffusion (Stocker 2012; Taylor & Stocker 2012). Chemotactic bacteria can exploit ephemeral nutrient gradients above the Batchelor scale. This exploitation leads to hotspots in bacterial abundance (Barbara & Mitchell 2003; Blackburn *et al.* 1998). Patchiness in bacterial and viral distributions are also dependent on the season and time of sampling (Seymour *et al.* 2005a, 2010). For instance, previous studies have shown bacterial abundance and activity were higher in the late afternoon (Seymour *et al.* 2005a). This study ensured sampling was taken at the same time of day and within the same season to avoid confounding from these low frequency signals.

Although hotspots are a ubiquitous feature of microbial distributions they are often vaguely defined in studies, primarily being classified as patches of high microbial abundance that 'exceed' or are 'elevated above' background variation across one or two sampling points (Seymour *et al.* 2004, 2005a, 2008). The methods for determining background are not given and the amount of elevation is rarely stated and differs considerably between studies making comparative analysis difficult. For instance, in the case of Seymour *et al.* (2004) the data shown for hotspots suggests that elevation above background is less than one order of magnitude whereas in Seymour et al. (2008) a conservative value of ≥ 4-fold higher is considered as a hotspot. In addition to this, coldspots have also lacked a definitive characterisation despite being observed in the microscale distributions of microbial communities previously (Seymour *et al.* 2004, 2006; Waters *et al.* 2003).

As well as hotspots and coldspots, surface gradients are a common pattern seen in microbial distributions, and are seen when bacterial and viral abundance is highest at an interface and then dissipates into the water column as the distance from the interface increases. These surface gradients arise from the sinking of organic matter and its incorporation into the benthos followed by degradation and transformation by microbes leading to high nutrient concentrations and hence high bacterial and viral concentrations directly above the sediment-water interface (Seymour *et al.* 2007b).

The microscale distributions of microbial communities on the sediment side of the interface are well established, whilst the distributions of microbes on the water side of the interface are less well known and where this has been investigated the focus has been largely on bacteria (Seymour *et al.* 2000, 2004; Waters *et al.* 2003). However, of the studies analysing the water side, surface gradients have been seen in the distributions of bacterial and viral subpopulations above coral colonies (Seymour *et*

al. 2005), in the water column above an anchialine sinkhole (Seymour *et al.* 2007a) and in coastal (Seymour *et al.* 2004) and estuary environments (Seymour *et al.* 2006, 2007b). Viral studies show the greatest level of heterogeneity is found 1–2 cm above the sediment water interface with abundance at St Kilda mangroves in South Australia varying by more than 2-fold, over a 15 cm sampling distance, about 1.13-fold cm^{-1}, with the highest concentrations found 1.5 cm from the sediment-water interface (Seymour *et al.* 2005b, 2006, 2007b). Similar patterns were seen in the vertical profiles in Seymour *et al.* (2007b), which showed high viral and bacterial abundance closest to the sediment-water interface with approximate increases of 1.4- and 1.5-fold cm^{-1} in total virus and total bacteria. However, due to one-dimensional sampling and lack of replication, determining whether these changes in abundance were single or multiple point maxima or surface gradients was unknown.

The null hypothesis tested in this study was that patchiness has no effect on mean biomass estimates and therefore the mean field approach is an accurate estimate of biomass. To test this hypothesis, two-dimensional sampling with extensive replication was used which builds upon previous work by Seuront *et al.* (2007), Seymour *et al.* (2005a, 2005b, 2006, 2007b, 2008) and Waters *et al.* (2003). Here, a new sampler with 9 mm resolution provided replication via the collection of 8 vertical profiles per sampler or 24 profiles per environment, which is much larger than the previous maxima of 5 single vertical profiles per environment (Seymour *et al.* 2007b). This replication allowed a detailed picture of microbial distributions immediately above the sediment water interface. The value of this design is that it shows how much the ‘mean field’ approach underestimates absolute concentration and misses the resources, competition and viral exposure gradients that microbes experience.

MATERIALS AND METHODS

Study site

Samples were collected from Saint Kilda mangroves and Port Noarlunga in South Australia. These environments were chosen to build on previous work using these sites (Seymour *et al.* 2000, 2004, 2005a, 2006, 2007b). The site at Port Noarlunga is located near the mouth of the Onkaparinga River and is bordered by sand dunes covered in remnant vegetation. Previous studies have described the water at Port Noarlunga as oligotrophic (Seymour *et al.* 2004). Samples were collected from Port Noarlunga (35°16'S, 138°47'E) on the 12th of July 2011 at 11:20 am; the temperature of the water was 12.3°C with a pH of 9.4, a dissolved O_2 level of 14.3 ppm and a conductivity value of 40.7 ppt.

The St Kilda site was characterised by hypersaline lagoons less than 1 m deep in dense mangrove forests. Previous studies have shown St Kilda mangroves host a microbial community that is characterised by highly productive microbial mats and sulphur oxidising bacteria (Barbara & Mitchell 1996; Seymour *et al.* 2004, 2005a, 2007b). Samples were collected from St Kilda (34°74'S, 138°54'E) on the 6th of July 2011 at 10:00 am; the temperature of the water was 11.5°C with a pH of 8.5, a dissolved O_2 level of 14.3 ppm and a conductivity value of 41.5 ppt. Environmental conditions data was collected using a Hydrolab Datasonde 4a sensor. Permission to access the sampling site from the St Kilda mangrove trail site was provided by the National Parks and Wildlife Rangers of the City of Salisbury Council. Specific permission to access the Noarlunga sampling site was not required. The field studies did not involve endangered or protected species.

Sampling device (and supplementary methods)

Microplate triplicates were used to measure the microscale spatial distribution of marine viral and bacterial populations. Corning Costar TC-treated cell culture cluster microplates were used. They were 8 cm x 12.3 cm, nonpyrogenic, polystyrene and sterile, consisting of 96 flat bottom 8 mm x 8 mm wells which held 360 µl per well. The centre to centre distance between wells was 9 mm and the well perimeters were 7.2 cm at the ends and 10.8 cm on the sides resulting in a collection area of about 77.8 cm^2. This allowed the collection of 8 vertical profiles per microplate, which each consisted of 12 sampling points (Figure S1). Microplate triplicates gave 24 vertical profiles. The sampler was placed vertically against the respective surface and removed in a vertical motion with a 16 cm x 16 cm glass plate used as a cover to minimise mixing and turbulence when collecting samples.

Sample collection and preservation

All samples were transferred into 2 ml cryovials containing 4 µl of glutaraldehyde (0.5% final concentration) using a pipette and stored at 4°C in the dark for 15 minutes. Samples were then snap frozen in liquid nitrogen and stored in a -80°C laboratory freezer (Brussaard 2004). Flow cytometric analysis was performed within three weeks to avoid deterioration (Brussaard 2004).

Flow cytometry sample preparation and analysis

Many previous studies have used flow cytometry for the enumeration of bacteria and viruses due to its rapidity and accuracy (Brussaard *et al.* 2000; Marie *et al.* 1999). Here we use the same flow cytometry technique as previous studies (Seymour *et al.* 2000, 2004, 2005a, 2005b, 2006, 2007b, 2008) and built on these studies by adding

extensive replication. Briefly, samples were defrosted and prepared for flow cytometric viral and bacterial enumeration by diluting each sample 1:10 with Tris-EDTA buffer (pH 8.0, 0.2 mm filtered, 10 mM Tris, 1 mM EDTA) and staining with a nucleic acid-specific dye, SYBR Green I (1:500 dilution commercial stock; Molecular Probes). Samples were then incubated at 80°C in the dark for 10 minutes to optimise viral counts (Brussaard 2004; Marie *et al.* 1999; Schapira *et al.* 2009; Seymour *et al.* 2007). Reference beads (1 mm diameter, Molecular Probes) were added as an internal concentration and size standard with a final concentration of approximately 10^5 beads ml^{-1} in each sample. Measured flow cytometry parameters were normalised to the fluorescence and concentration of these beads (Gasol & Del Giorgio 2000; Schapira *et al.* 2009).

Flow cytometry was conducted on a FACSCanto II flow cytometer (BD) equipped with a blue (488 nm, 20 mW, aircooled), red (633 nm, 17 mW) and violet (405 nm, 30 mW) laser. Forward-angle light scatter (FSC), right-angle light scatter (SSC) and green fluorescence (SYBR I) were acquired for each sample and a phosphate-buffered saline (PBS) solution was used as a sheath fluid. The flow cytometer settings were normalised to fluorescence and bead concentration (Seymour *et al.* 2007b; Schapira *et al.* 2009). Each sample was run for two minutes at a medium flow rate setting for Noarlunga samples and a low flow rate setting for St Kilda samples due to more suspended particulate matter being present in St Kilda samples. In each flow cytometry session, triplicate blank control samples were analysed which consisted of 500 ml of 0.2 µl filtered Tris-EDTA buffer stained with 12.5 µl of SYBR Green I to eliminate any background noise that may have been created during the sample preparation or from flow cytometer artifacts.

Sample processing

The resulting cytograms, density plots and histograms acquired from flow cytometry were exported as listmode files and analysed using Win Midi 2.9 (Joseph Trotter) to enumerate the bacterial and viral populations present (Seymour *et al.* 2007). Viral and bacterial populations were discriminated via side-scatter (SSC), which indicates cell size; and SYBR Green fluorescence, which is indicative of nucleic acid content (Brussaard 2004; Marie *et al.* 1997, 1999).

Data analysis and representation

For this study, rank abundance graphs were used to discriminate between hotspots, coldspots and background values. The background was determined according to Weibe (1970) where the median value of the dataset is used as the background due to inclusion or exclusion of hotspots or coldspots not overly affecting the values obtained.

Surfer 10 (Golden Software, Inc.) was used to create two dimensional contour plots of the spatial distribution of viral and bacterial populations. A minimum contour interval value of at least 1000 was chosen when constructing plots due to this being larger than the maximum FACSCanto II flow cytometer error observed. This error refers to the background noise seen within triplicate blank control samples during flow cytometric analysis. This value is conservative as the maximum machine error was less than 36 events μl^{-1}.

Correlations were determined via Pearson's coefficient with the α of 0.05 being reduced by sequential Bonferroni (Holm 1979). Spatial autocorrelation statistics were used to determine whether correlations were present two-dimensionally and multi-directionally between sample points that were proximate and had similar values.

From this, values were derived that illustrated the spatial autocorrelation present in the dataset. The two most common geospatial statistical tests are Moran's I and Geary's C statistics and both of these were employed as they allow statistical testing on smaller datasets (Waters *et al.* 2003).

- Moran's I Statistic

Moran's I spatial autocorrelation statistics test within CrimeStat 3.3 (courtesy of Ned, Levine software) was used to determine the level of randomness within each population. The equation is as follows:

$$I = \frac{N \sum i \sum j W_{ij} (Xi - \overline{X})(X_j - \overline{X})}{(\sum i \sum j W_{ij}) \sum i (X_i - \overline{X})^2}$$

Where N is the sample number, X_i and X_j represent the variable values at a specific locations, i and j (where $i \neq j$), X is the mean of the variable and W_{ij} is the weight applied to the i and j comparison. A weighted Moran's I test was used which gave a weight value of 1 to sample points adjacent and a weight value of 0 to sample points not adjacent.

The Moran's I statistical test (Moran 1950) is a global statistical test used to determine spatial autocorrelation within a set of values. It is multi-directional, being able to use vertical, horizontal and diagonal directional analysis to consider all position correlations. The Moran's I value ranges from +1 to 21, with +1 indicating perfect clustering where high values are proximate and low values are proximate, 21 indicating perfect dispersion where high values are found far apart and low values are found far apart and zero being indicative of a random distribution. Critical cut-off values for most statistical tests are dependent on the scale and values used and are

often well established through previous literature using the same or similar types of studies. However, in the instance of Moran's I values for small-scale spatial studies of bacteria and viruses there is only 1 study to compare which I values are indicative of significance. Waters et al. (2003) looked at phytoplankton distributions on a 2 cm and 4 cm scale. From looking at the I values obtained in the Waters *et al.* (2003) study, it appears I values between 0.08 and 0.18 indicate a considerable level of clustering, as opposed to a random distribution, which is far less than the general clustering I value of +1.

- Moran Correlograms

Moran correlograms were constructed using the Moran's I statistic, which is applied to pairs of sample values separated by a lag distance, which in this case was the distance between each sampling well (0.9 cm). Geostatistical analysis requires each lag distance to have ≥ 30 pairs of sample points to enable statistical reliability; therefore sample intervals with less than this were not included in the correlograms (Rossi *et al.* 1992; Waters *et al.* 2003). The standard error obtained via Moran's I analysis was used to determine if correlograms and lag distance values were significant. Output values were used to determine whether spatial autocorrelation or independence was present at each distance interval from the sediment-water interface.

- Geary's C Statistic

Geary's C statistics (Geary 1954) test within CrimeStat 3.3 (courtesy of the Ned, Levine software) is a global statistic that can be used in conjunction with Moran's I as it is more sensitive to local clustering and is able to identify deterministic patterning of

extreme values and non-spatially related chance phenomena within the dataset, which cannot be identified using Moran's I alone. Geary's C is similar to Moran's I, however it calculates the spatial autocorrelation of a dataset by the deviation in intensity of each sample value's location compared to one another, whereas in Moran's I spatial autocorrelation is calculated by the cross-product of the deviations from the mean within the dataset. The equation is as follows:

$$C = \frac{(N-1)\left[\sum i \sum j W_{ij} (X_i - X_j)^2\right]}{2\left(\sum i \sum j W_{ij}\right) \sum i (X_i - \overline{X})^2}$$

All terms are the same as Moran's I. For Geary's C, the values range from 0 to approximately 2, however there is no definitive upper limit (Sokal 1978). A value of 1 indicates spatial independence, whilst values < 1 indicate positive spatial autocorrelation and values > 1 indicate negative spatial autocorrelation. Therefore this statistical test is inversely related to the Moran's I test (Griffith 1987). The Moran's I and Geary's C can be used in conjunction to identify deterministic patterning of extreme values (Sokal 1978).

RESULTS

Flow cytometry analysis

Flow cytometric analysis revealed three distinct bacterial subpopulations and two distinct viral subpopulations at both sites which were comparable to those identified previously in marine systems (Marie *et al.* 1999; Seymour *et al.* 2004, 2006). Subpopulation discrimination was based on the presence of aggregated regions on biparametric cytograms of side scatter and SYBR green fluorescence and discrete peaks in monoparametric histograms of SYBR green fluorescence (Figure. 1A–D).

Bacterial subpopulations were differentiated and separated into high-DNA (HDNA) and low-DNA (LDNA) groups (Figure. 1A–D). The HDNA 1 bacterial population is the intermediate bacterial population that contains bacteria with a size and nucleic acid content that is intermediate to the LDNA and HDNA 2 populations. The two viral subpopulations were identified and separated into two virus-like particle subpopulations (VLP 1 and VLP 2).

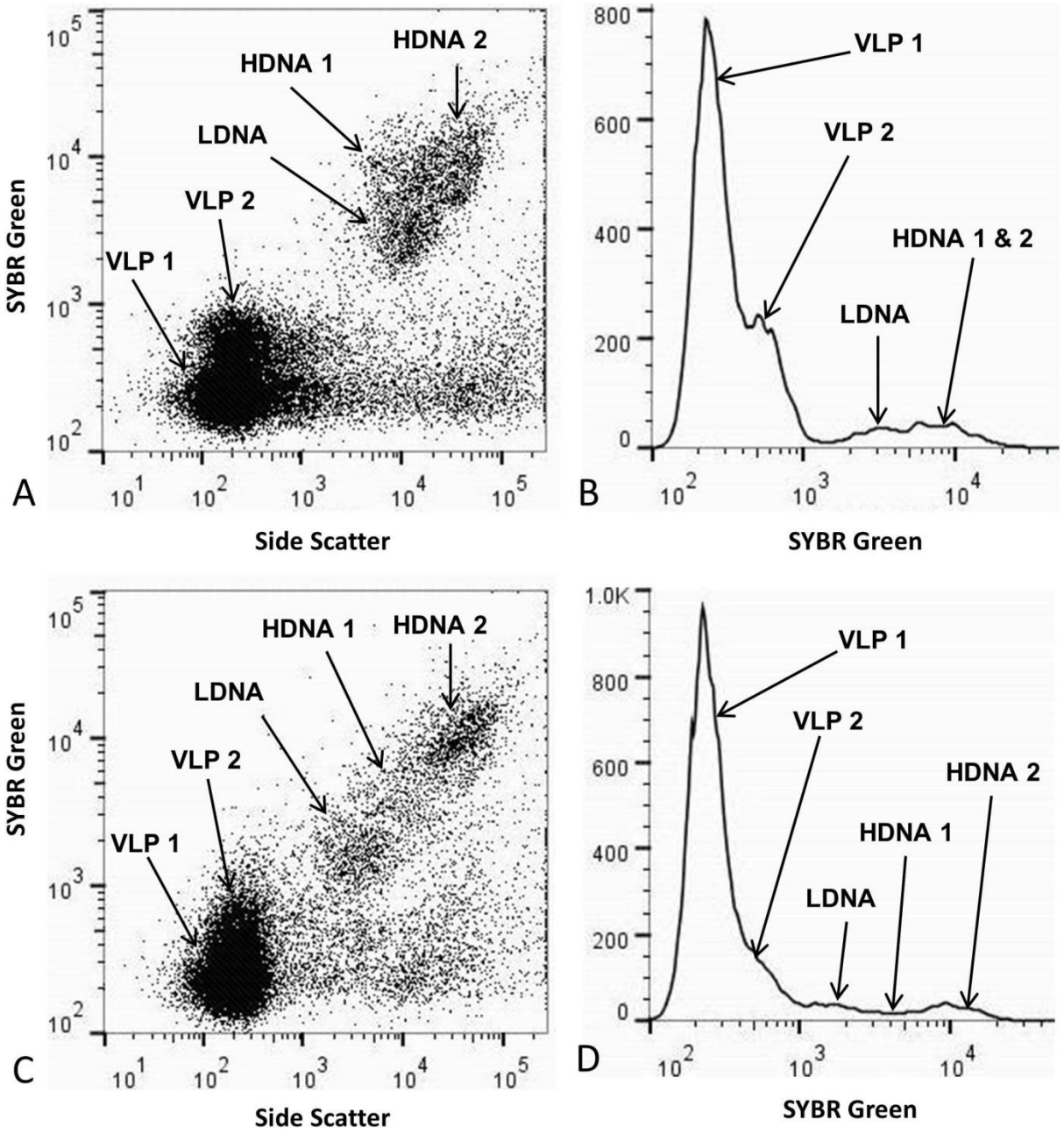


Figure 1. Identification of bacterial and viral subpopulations via flow cytometry. Flow cytometric cytograms of side-scatter light versus green fluorescence (SYBR Green) and histograms of green fluorescence (SYBR Green). **Noarlunga: A** cytogram **B** histogram; **St Kilda**: **C** cytogram **D** histogram showing two distinct viral populations (VLP 1 and VLP 2) and three distinct bacterial populations (LDNA, HDNA 1 and HDNA 2).

At Noarlunga, the total mean abundance of the VLP 1 subpopulation was 3.7 x 10^6 cells ml^{-1} (95%CI= 0.3 x 10^6 cells ml^{-1}, n =269), whereas it was 1.3 x 10^6 cells ml^{-1} for the VLP 2 subpopulation (95%CI= 0.1 x 10^6 cells ml^{-1}, n= 269) (Figure. 1A and 1B). Hence, the total mean abundance of VLP 2 was 2.8-fold less than the total mean abundance of VLP 1. At St Kilda, the total mean abundance of VLP 2 was also less than VLP 1, but to a greater extent as the total mean abundance of VLP 1 was 75 x 10^5 cells ml^{-1} (95%CI = 5.1 x 10^5 cells ml^{-1}, n= 209) compared to 6.1 x 10^5 cells ml^{-1} (95%CI= 0.45 x 10^5 cells ml^{-1}, n =209) for VLP 2 (Figure. 1C and 1D). Hence, the total mean abundance of VLP 2 was 12.3-fold less than the total mean abundance of VLP 1.

For the bacteria, the LDNA subpopulation at Noarlunga was less abundant than the HDNA 1 and the HDNA 2 subpopulations. The total mean abundance for LDNA was 0.4 x 10^5 cells ml^{-1} (95%CI = 0.03 x 10^5 cells ml^{-1}, n = 269) compared to total mean abundances of 3.1 x 10^5 cells ml^{-1} (95%CI= 0.2 x 10^5 cells ml^{-1}, n = 269) for HDNA 1 and 2.7 x 10^5 cells ml^{-1} (95%CI =0.2 x 10^5 cells ml^{-1}, n= 269) for HDNA 2 (Figure. 1A and 1B). Hence, the total mean abundance of LDNA was up to 7- fold less than the total mean abundance of HDNA 1 and HDNA 2. Conversely at St Kilda, HDNA 1 was the least abundant bacterial subpopulation (Figure. 1C and 1D) with total mean abundances of 3.0 x 10^5 cells ml^{-1} (95%CI= 0.32 x 10^5 cells ml^{-1}, n =209) for LDNA,

2.1 x 10^5 cells ml^{-1} (95%CI= 0.47 x 10^5 cells ml^{-1}, n =209) for HDNA 1 and 5.9 x 10^5 cells ml^{-1} (95%CI =1.7 x 10^5 cells ml^{-1}, n= 209) for HDNA 2. Hence, the total mean abundance of HDNA 1 was 2-fold less than the total mean abundances of LDNA and HDNA 2 (Table S1 and S2).

Vertical profiles

At Noarlunga, the mean of 24 vertical profiles had an r value of 0.9 ($p < 0.0001$, n= 12) and showed almost identical patterns between the total mean bacterial and total mean viral population (Figure. 2A), as well as the total mean LDNA, HDNA 1 and HDNA 2 subpopulations (Figure. S2A), the total mean VLP 1 and VLP 2 subpopulations (Figure. S2B) and the mean vertical profile of VLP 1 and LDNA from microplate 1, which all had an r value of 0.99 ($p < 0.0001$, n = 12) (Figure. S3).

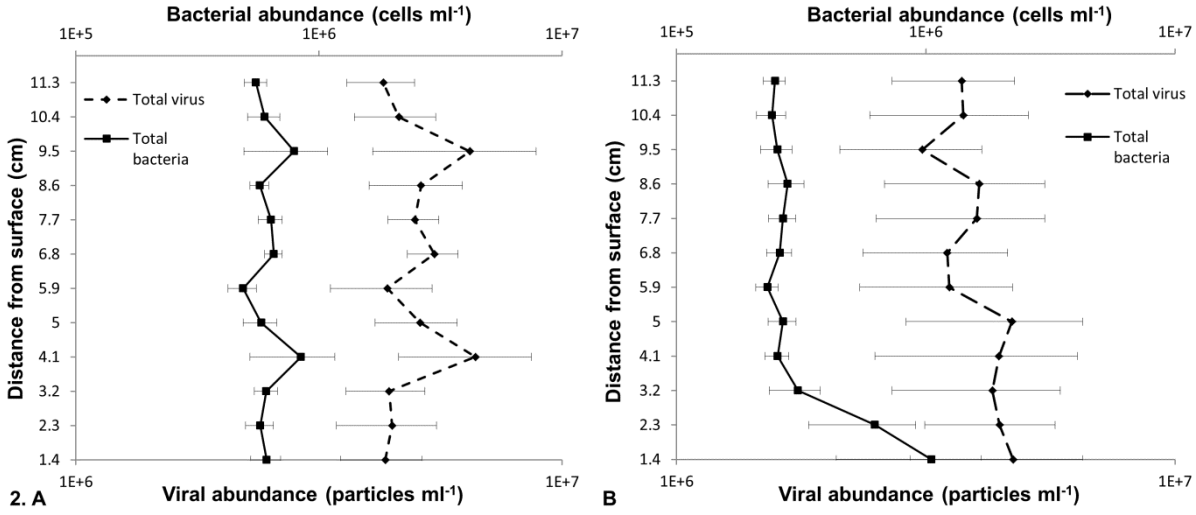


Figure 2. Comparison of total mean bacterial and viral populations at Noarlunga via vertical depth profiles. Total mean bacterial and total mean viral population abundance within all three microplates. **A** Noarlunga, (n = 1350). **B** St Kilda, (n = 1045). Error bars represent the 95% confidence intervals obtained from all three replicates (n = 12).

In addition, 145 of the 240 possible bacterial and viral single vertical profile subpopulation correlations were significantly correlated with a p-value generally less than 0.0001, such as the single vertical profile of VLP 1 and LDNA from microplate 1 which had an r value of 0.98 ($p < 0.0001$, $n = 12$) (Figure. 3B). However, 95 of the single vertical profiles were not significantly correlated, such as the single vertical profile of VLP 1 and LDNA from microplate 3 which had an r value of 2 0.33 ($p = 0.12$, $n = 12$) (Figure. 3A). At St Kilda, the mean of 24 vertical profiles was not significantly correlated (Figure. 2B). However, the total mean LDNA, HDNA 1 and HDNA 2 populations were correlated ($r = 0.98$, $p < 0.0001$, $n = 12$), as well as the total VLP 1 and total VLP 2 populations ($r = 0.64$, $p = 0.02$, $n = 12$). Of the 240 possible bacterial and viral single vertical profile subpopulation correlations, only 64 were significantly correlated with a p-value generally less than 0.0001, whilst 176 were not correlated.

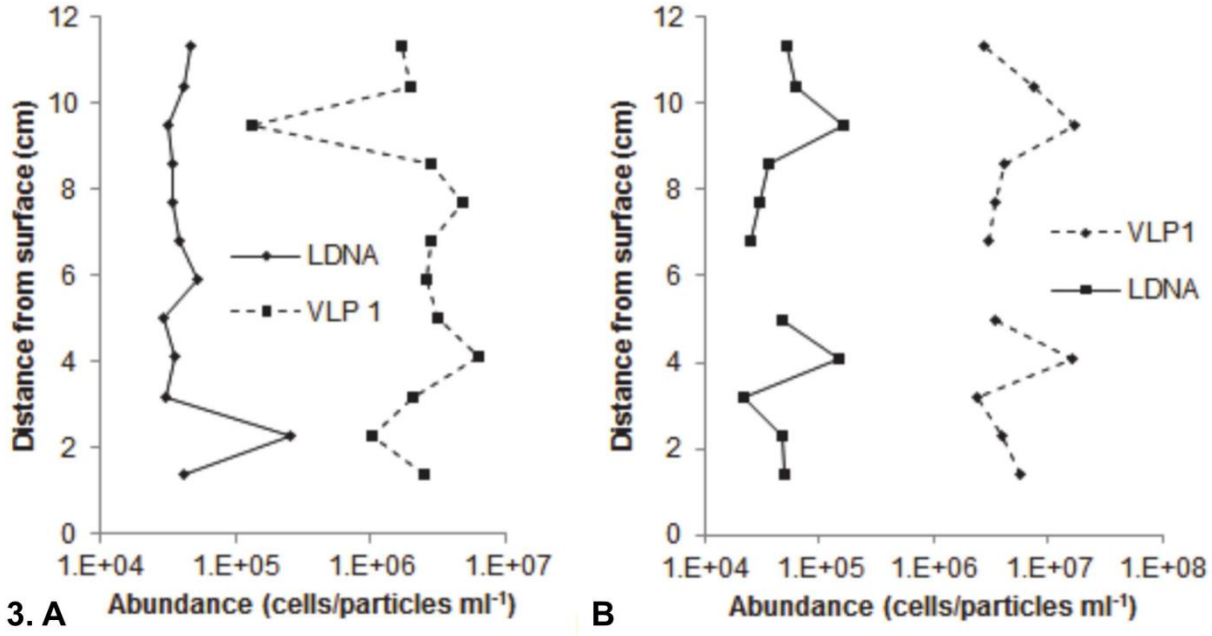


Figure 3. Single vertical profile of VLP 1 and LDNA populations at Noarlunga. **A** Microplate three showing little to no association. **B** Microplate one showing association. Gap in profile indicates a missing data point.

Determination of hotspots and coldspots

To determine the presence of hotspots and coldspots in the viral and bacterial distributions, rank abundance graphs were created (Figure. 4A–D). In some instances only one linear trend was seen within the background values and in others there were breaks in the fits that indicated two or three separate linear trends. St Kilda showed primarily one linear trend, for instance in the VLP 1 and HDNA 1 subpopulations (Figure. 4A–B) whilst Noarlunga showed primarily two linear trends within their background values, for instance in the VLP 1 and HDNA 2 subpopulations (Figure. 4C– D). The hotspots were identified as the sample points that exceeded this linear fit (Figure. 4A–D). Exponent values for these hotspots ranged from -0.20 to -1.1 for the bacteria and -0.14 to -0.63 for the viruses at Noarlunga, whilst exponent values at St Kilda ranged from -0.27 to -1.4 for the bacteria and -0.23 to -0.66 for the viruses (Figure. 4A–D). Lastly, the coldspot sample values were identified as sample points that fit a linear trend but exhibited steeper slopes than what were seen in the background values due to large differences between adjacent sample points (Figure. 4A–D).

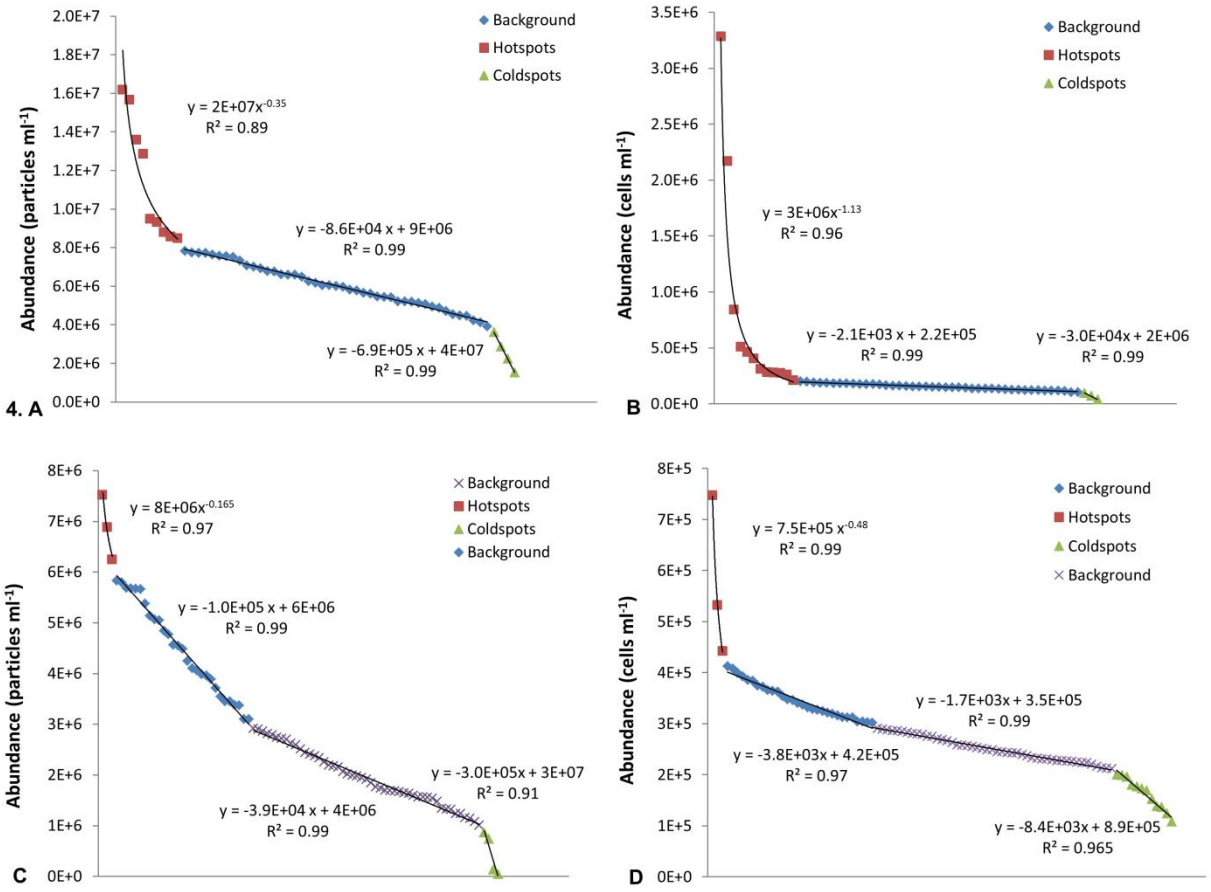


Figure 4. Rank abundance graphs used to differentiate hotspot and coldspot values from background values. One linear trend within the background values were characteristic of St Kilda whilst two linear trends were primarily seen at Noarlunga. **St Kilda: A** VLP 1, microplate 2. **B** HDNA 1, microplate 2. **Noarlunga: C** VLP 1, microplate 3. **D** HDNA 2, microplate 2.

Two-dimensional distributions

Two-dimensional contour plots of both sampling sites showed heterogeneous distributions with patches or hotspots of high abundance that occurred over single or multiple sample points as well as gaps or coldspots of low abundance that occurred only over single sample points (Figure. 5A–D).

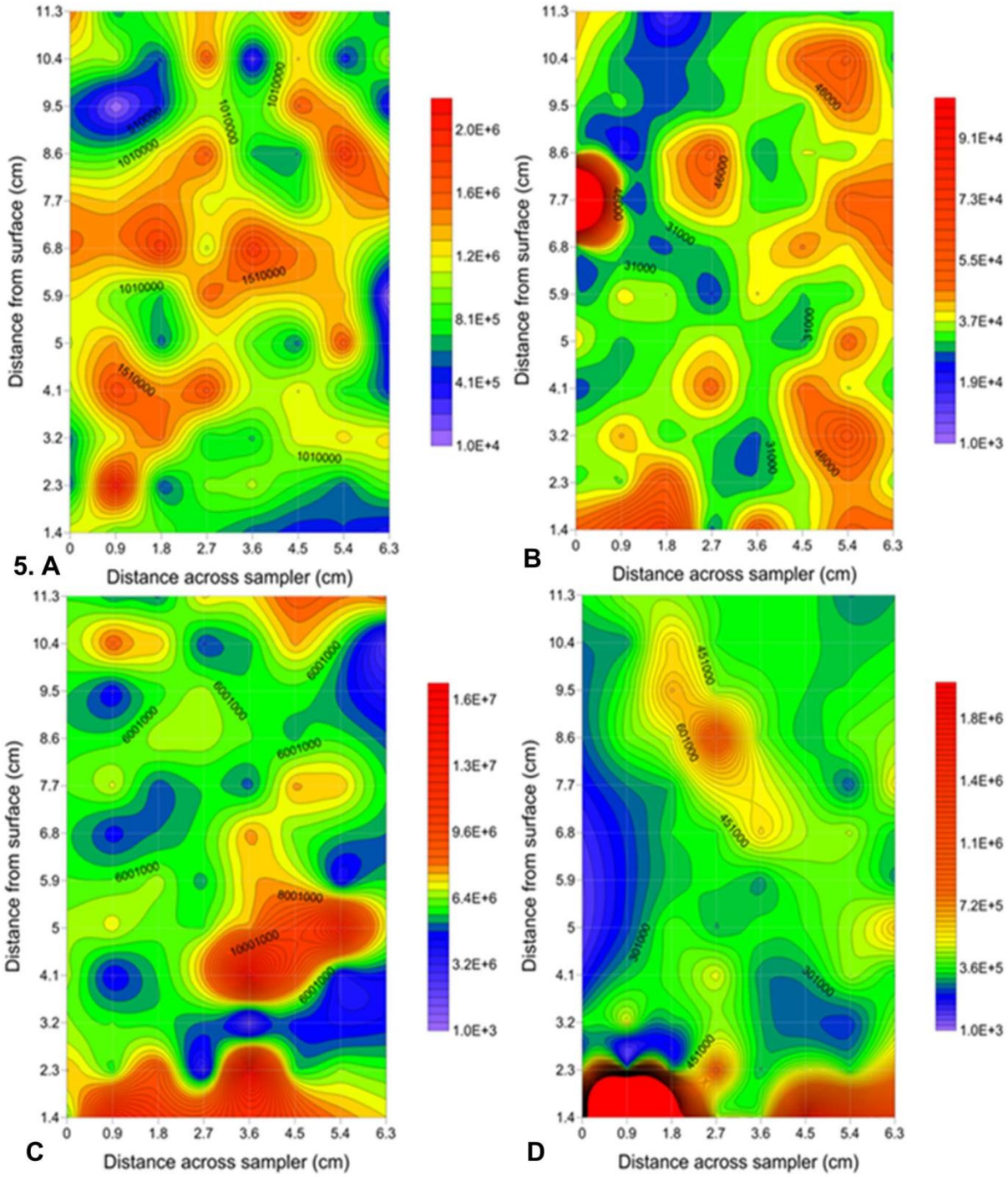


Figure 5. Two-dimensional contour plots showing the highest change in heterogeneity due to the presence of hotspots and coldspots within bacterial and viral subpopulations. Hotspots and coldspots seen across a distance of 6.3 cm x 11.3 cm using Surfer 10 (Golden Software, Inc.). **Noarlunga: A** VLP 2 showing a 2585-fold change in abundance over 0.9 cm. **B** LDNA showing a 12.9 fold change in abundance over 0.9 cm. **St Kilda: C** VLP 1 showing a maximum 10.52-fold change

in heterogeneity seen over 0.9 cm. **D** HDNA 2 showing a maximum 45.2-fold change in heterogeneity seen over 0.9 cm. There were a range of heterogeneities over 0.9 cm (Fig. S4) indicating a variety of intensities for hotspots and coldspots. Abundance levels are indicated by a colour intensity scale in units of cells/particles ml^{-1}. Solid red circles indicate areas of abundance higher than the maximum contour level selected. A minimum contour interval value of at least 1000 was chosen based on maximum machine error. The faint gridlines show sample interval.

At Noarlunga, high levels of heterogeneity were seen, with the VLP 2 subpopulation showing the highest level of heterogeneity of the two viral subpopulations. It was due to the presence of coldspots in the VLP 2 subpopulation that resulted in a 2585-fold change in abundance over 0.9 cm, with a coldspot of 6.64×10^2 particles ml^{-1} below a background of 1.07×10^6 particles ml^{-1} (Figure. 5A). For the bacterial subpopulations, the highest level of heterogeneity was seen in the LDNA subpopulation with a maximum hotspot of 2.59×10^5 cells ml^{-1} above a background of 3.47×10^4 cells ml^{-1} resulting in a 12.9-fold change in abundance over 0.9 cm (Figure. 5B).

At St Kilda, the highest viral heterogeneity was present in the VLP 1 subpopulation which had a maximum hotspot of 1.62×10^7 particles ml^{-1} above a background of 6.12×10^6 cells ml^{-1} which resulted in a 10.52-fold change in heterogeneity seen over 0.9 cm (Figure. 5C). For the bacterial subpopulations at St Kilda a 45.2- fold change in heterogeneity was seen in the HDNA 2 subpopulation due a maximum hotspot of 1.45×10^7 cells ml^{-1} above a background of 3.7×10^5 cells ml^{-1} (Figure. 5D). There were a range of heterogeneities over 0.9 cm (Figure. S4) indicating a variety of intensities for hotspots and coldspots. Representative adjacent orthogonal gradient

distributions as rank abundances show the distribution of abundance values for selected bacterial and viral populations at St Kilda and Noarlunga (Figure. S4).

Spatial autocorrelation

- Moran's I

At Noarlunga, the Moran's I values for all of the bacterial subpopulations were non-significant. In contrast, at St Kilda significant Moran's I values were found in 33% of the bacterial subpopulations; the LDNA subpopulation of microplate 2 and 3 and the HDNA 2 subpopulation and total bacteria of microplate 2. These significant Moran's I values ranged from 0.039–0.070, with all the non-significant Moran's I values being below these values (≤ 0.036). For these bacterial subpopulations z scores under the assumption of randomisation were much larger than the cut-off point of +1.96, with z scores ranging from 2.5 to 3.8. All the viral subpopulations at Noarlunga and St Kilda had non-significant Moran's I values. All of the significant Moran's I values for the bacterial subpopulations were positive (Table S3 and S4). This trend is evident in the 2 dimensional contour plots of HDNA 2 and LDNA at St Kilda which show the presence of sediment-water interface surface gradients (Figure 6A–B). Statistical significance was not found for negative Moran's I values (Table S3 and S4). In geographic data, statistically significant positive values are more common than statistically significant negative values.

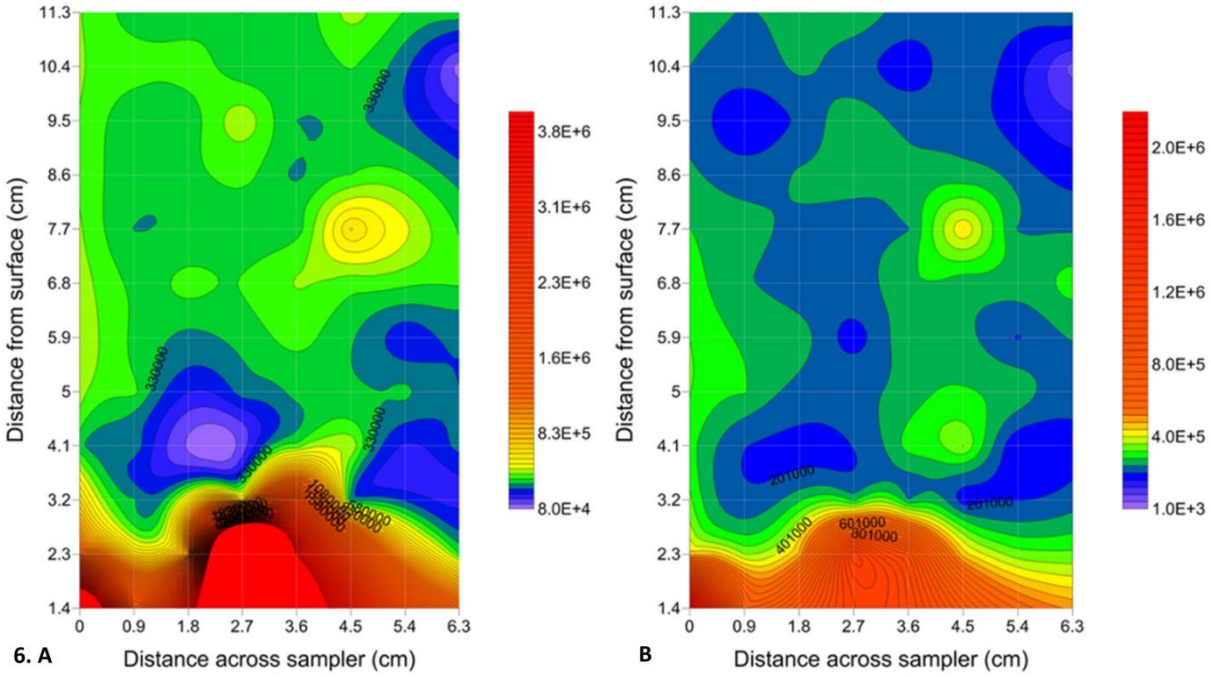


Figure 6. Two-dimensional contour plots showing surface gradients within bacterial subpopulations at St Kilda. Surface gradients seen across a distance of 6.3 cm x 11.3 cm using Surfer 10 (Golden Software, Inc.). Microbial abundance levels are indicated by a colour intensity scale in units of cells ml^{-1}. **A** HDNA 2, microplate 2. **B** LDNA, microplate 2. Solid red circles indicate areas of abundance higher than the maximum contour level selected. A minimum contour interval value of at least 1000 was chosen based on maximum machine error. The faint gridlines show sample interval.

- Moran correlograms

Moran correlograms for the bacterial subpopulations that had significant Moran's I values showed a general trend of positive to negative spatial autocorrelation as the distance from the sediment-water interface increased. On closer inspection, the correlograms for these bacterial subpopulations showed positive spatial correlation in the sample points located within 5.9 cm of the sediment-water interface and spatial independence as this distance from the sediment-water interface increased at > 5.9

cm (Figure. 7A–D). This is indicative of surface gradients, which are shown in figures 6A–B. Moran correlograms for the LDNA, HDNA 1 and total bacteria in microplate 1, and HDNA 1 in microplate 2 at St Kilda also followed this trend despite having non-significant Moran's I values. Geary's C statistical testing was used to further analyse these subpopulations. At Noarlunga and St Kilda, all of the viral subpopulation correlograms were non-significant and showed no general trend, with their distributions alternating between positive and negative spatial autocorrelation throughout the sampling distance indicating the presence of hotspots and coldspots.

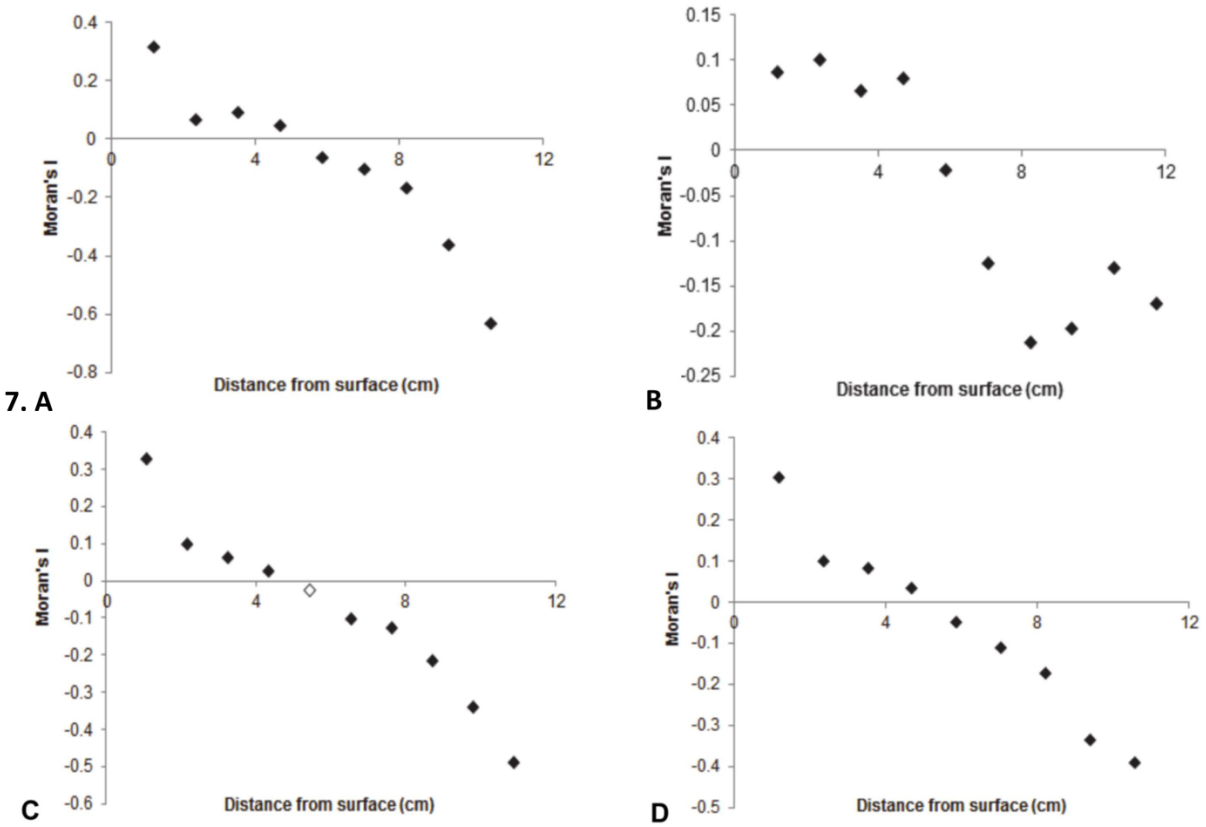


Figure 7. Significant Moran correlograms of non-randomly distributed bacterial subpopulations at St Kilda. **A** LDNA, microplate 2. **B** LDNA, microplate 3. **C** HDNA 2, microplate 2. **D** Total bacteria, microplate 2. Filled and unfilled data points indicate significant and non-significant Moran's I values ($p \leq 0.01$). Only sample points with ≥ 30 pairs of values were included.

- Geary's C

At Noarlunga, significant Geary's C values were obtained from all of the bacterial subpopulations in microplate 1, whilst all of the bacterial subpopulations in microplate 2 and 3 were non-significant. Geary's C values were between 0.87 and 1.02; however, significance was only seen in C values of 0.87 (Table S3 and S4).

In contrast, at St Kilda significant Geary's C values were seen in 75% of the bacterial subpopulations; all of the subpopulations in microplate 1 and 2, as well as LDNA within microplate 3. Significant C values were between 0.73 and 0.92 with C values between 0.96 and 0.99 showing non-significance (Table S3 and S4).

At Noarlunga, significant Geary's C values were found in the total virus and VLP 2 subpopulation in microplate 1 whilst all other viral subpopulations were non-significant. Significance was seen in C values of 0.87, whilst C values of 0.98–1.02 showed nonsignificance. At St Kilda, all of the viral subpopulations had nonsignificant Geary's C values. C values were higher than what was seen in the bacteria at St Kilda, being between 0.96 and 1.02, which was similar to the C values obtained for the viruses at Noarlunga (Table S3 and S4). Using Geary's C to look in more detail at the randomly distributed subpopulations that showed positive to negative spatial autocorrelation within the Moran correlograms revealed they had significant Geary's C values (Table S3 and S4). These subpopulations were the bacterial subpopulations in microplate 1, as well as the HDNA 1 subpopulation in microplate 2 at St Kilda and the bacterial subpopulations and VLP 2 and total virus populations in microplate 1 at Noarlunga.

DISCUSSION

Bacterioplankton and virioplankton distributions

- Surface gradients.

Strong surface gradients were found at St Kilda, which are well established along with characteristic low dissolved oxygen levels (Seymour *et al.* 2007b). These strong gradients rising from the sediment surface showed up to a 15-fold change in abundance over the 11.3 cm vertical sampling distance collected above the sediment-water interface, The presence of these gradients is perhaps due to the strong input of organic matter from the mangroves (Barbara & Mitchell 1996). In support of this, Seymour *et al.* (2007b) found microscale microbial distributions were highly correlated to nutrient concentration gradients. Seymour *et al.* (2006) found the shear velocities at St Kilda at the time of sampling were not high enough for the resuspension of microbes. Given the protected nature of the environment, resuspension may be minimal compared to exposed sites and would not play a role in the surface gradients observed.

Bacteria and viruses immediately above the sediment are exposed to a variety of biological and non-biological parameters that will affect their distribution (Seymour *et al.* 2007b). At both sampling sites clear abundance gradients were often seen in relation to distance from the sediment-water interface. The remineralisation of organic nutrients is known to cause gradients with dissolved and particulate matter, which usually show decreasing concentrations as the distance from the sediment-water surface increases (Baric *et al.* 2002; Seymour *et al.* 2007b).

- Hotspots and coldspots.

Previous studies could not determine whether a high or low point in a one-dimensional profile was a gradient, a multiple point hotspot/coldspot or a single point maximum/minimum (Seymour *et al*. 2007b). In this study, two-dimensional sampling revealed the presence of single and multiple point hotspots and single point

coldspots in bacterial and viral populations at Noarlunga and St Kilda. In previous studies, local increases in abundance of up to 2 times above background over a distance of 1 cm have been considered hotspots (Seymour *et al*. 2005a, 2005b, 2007b, 2008), whereas this study showed hotspots with increases of up to 2585 and 45.2 times cm^{-1} in viral and bacterial subpopulations above the sediment-water interface. These hotspots were high in bacteria and viruses, consistent with bacteriophage infection being dependent on host density (Seymour *et al.* 2006). The correlation of these bacterial and viral hotspots suggests a repeated viral infection and lysis cycle for at least some of the bacterial species present (Seymour *et al.* 2004, 2005a, 2006), which, in turn, suggests that groups of picoplankton, at least close to the sediment interface remain close together for extended periods. Through entrainment in the overlying water these near-sediment hotspots may act to inoculate the overlying water column. With viral hotspots showing increases of up to 2585 times cm^{-1} this would contribute substantially to the abundance of viral particles hence playing a role in viral production and succession. In the case of coldspots, correlations were seen with viral subpopulations but not with bacterial subpopulations. There are many possible dynamics at play here. For instance, this may indicate these coldspots are capturing the stable bacterial community numbers, which have not yet been exposed to viral communities or perhaps areas in which compatible viral particles were not present hence leading to viral decay or extensive grazing as the viruses would not have been able to protect themselves from external sources by infecting the bacteria present. Another possible explanation for these coldspots in viruses could be that they are microenvironments that favour lysogeny over lysis, such as areas of low nutrient concentration (Weinbauer & Suttle 1999). The wide range of hotspot values (Figure. S4) suggests that all of these and possibly

more mechanisms are generating hotspots. Of course, the actual processes occurring in these hotspots and coldspots cannot be determined with biomass values alone and call for microscale nutrient and species analysis. At St Kilda, patchy viral and bacterial hotspots and coldspots were identified and, in the case of the hotspots, often coincided. These hotspots and coldspots resulted in changes in spatial heterogeneity of 10.5-fold cm^{-1} for viruses and 45-fold cm^{-1} for bacteria. These results are much higher than the previously found 1.4- and 1.5-fold cm^{-1} increases found in total virus and total bacteria by Seymour *et al.* (2007b). As these hotspots and coldspots often occurred within single sampling points of 0.9 cm distance this highlights the importance of high sampling resolution and two-dimensional arrays when studying the distributions of microbial communities.

The null hypothesis tested in this study was that patchiness has no effect on mean biomass estimates and therefore the mean field approach is an accurate estimate of biomass. This null hypothesis was rejected with the presence of patchy hotspots and coldspots in this study showing differences of up to four orders of magnitude in biomass. In the most extreme case, if the mean field approach had been employed without sufficient replication, estimates in microbial biomass would be off by a factor of 1,560. Previous studies have also lacked a definition for hotspots and coldspots making comparative analysis impossible. In this study rank abundance graphs have been used as a method for determining what sample values are hotspots, coldspots or background. Background values fitted a linear trend as these values are random and are an indicator of equilibrated additive and reductive processes. Such additive processes in bacterial and viral communities could be reproduction or aggregation, whilst reductive processes could relate to grazing, decay or lysis events. Hotspots were identified as the sample points that exceeded the linear fit and exhibited a steep

power law trend, indicating their non-random nature whilst the coldspots were characterised by a linear trend with slope values much higher than those of the background values due to large differences seen between adjacent value points on the graph (Figure. 4A–D). The hotspots indicate a favouring towards additive processes and coldspots indicate a favouring towards reductive processes resulting in a non-equilibrated state when compared to the background values. Exponent values for the hotspots were similar at Noarlunga and St Kilda and ranged from -0.1 to -1.4 suggesting this is the level of structure seen within bacterial and viral communities. Exponent values at and below -1 indicated considerable self-organisation and strong patches (Mitchell 2004). The background values fit single or multiple linear trends with St Kilda being characterised primarily by single linear trends whilst Noarlunga was characterised primarily by two linear trends within its background values for bacterial and viral communities. This could be an indication of different randomisation processes present. For instance, when considering bacteria, the presence of two linear trends seen at Noarlunga could indicate equilibrium between grazing and aggregation in one linear trend whilst the second may indicate equilibrium between reproduction and lysis. By using these rank abundance graphs, not only are hotspots and coldspots identified but patterns within the background values are able to be observed. Experimental work will be required to distinguish between randomisation processes.

Bacterial and viral abundances

The LDNA and HDNA bacterial subpopulations observed have been described previously as different phylogenetic groups present within the environment (Wang *et al.* 2009) or as bacteria with differing activity levels (Gasol *et al.* 1999; Lebaron *et al.* 2001, 2002). In the former context, the total mean abundance of HDNA 2 at St Kilda

and HDNA 1 at Noarlunga being higher than the total mean abundance of the other LDNA and HDNA subpopulations, may be explained by the species within these subpopulations being more efficient at utilising resources or avoiding predatory attack (Wang *et al.* 2009). The total mean abundances also indicated HDNA 1 at St Kilda and LDNA at Noarlunga had the lowest abundances of the bacterial subpopulations which could be due to the species within HDNA 1 at St Kilda and LDNA at Noarlunga being more susceptible to infection or grazing (Bouvy *et al.* 2004; Seymour *et al.* 2007b; Wang *et al.* 2009). In addition, higher concentrations of predators are found within the first few centimetres above the sediment-water interface, such as heterotrophic nanoflagellates (Seymour *et al.* 2007b). However, in the context of differing bacterial activity, where the LDNA subpopulation represents dormant bacterial cells and the HDNA subpopulations represent active cells, the higher abundances of the HDNA subpopulations seen at both sites when compared to the LDNA subpopulation would be explained by the ability of the HDNA subpopulations to replicate and grow, hence leading to their higher abundances (Gasol *et al.* 1999; Lebaron *et al.* 2001, 2002; Seymour *et al.* 2007b). In the context of different species, the hotspots would represent accumulation of chemotactic bacteria and the coldspots areas they had left. Another factor controlling the abundances of the bacterial subpopulations is the viral populations present. The total mean abundance of the VLP 1 subpopulation was larger than the total mean abundance of the VLP 2 subpopulation at both sites. This has been found in previous studies, with aquatic samples being dominated by the viral population that contains small viruses between 30 and 60 nm, which corresponds to the VLP 1 subpopulation within the cytograms (Figure.1A and C) (Bergh *et al.* 1989; Marie *et al.* 1999).

Bacterial and viral coupling

Seymour *et al*. (2007b) did not find a correlation between total bacteria and viruses at this scale. In this study, a significant relation existed between total bacteria and total viruses at Noarlunga, whilst at St Kilda, a relation did not exist (Figure. 2A and B). This may be due to the presence of different microbial communities and the difference in environmental parameters between the two sampling sites. The dynamic nature of the environments may have led to short-lived events, such as nutrient patches, causing the aggregation and dispersion of bacteria over the course of minutes to hours. This initial increase in bacterial abundance would lead to increased viral production via lysis hence decoupling the relationship between viral and bacterial populations which may be why no relationship was seen between bacteria and viruses at St Kilda (Blackburn *et al.* 1998; Seymour *et al.* 2006). However, stages in between the bacterial aggregation and bacterial death via viral lysis may have been captured at Noarlunga and would explain the correlation between the bacteria and viruses in this environment.

While Seymour *et al.* (2007b) found no correlation between total bacteria and viruses; correlations were seen between the VLP 1, VLP 2 and LDNA subpopulations and the VLP 2 and HDNA 2 subpopulations. Therefore, showing subpopulations have independent and potentially more complex associations than what is seen in total bacteria and virus populations. These results imply that total populations reflect system biomass and the processes that may impact on this, such as grazing; whilst single subpopulations reflect the phylogenetic changes in the community from such processes as phage infection and lysis. As the individual subpopulations can be decoupled, with phylogenetic fluctuations being independent of the total system biomass, this explains why correlations were seen in the totals but not in the

individual subpopulations. While the correlations are not mechanisms, they show that possible relationships exist that could be further investigated. These statements are based on biomass only. Taxonomic identification of the community will be part of a manuscript in preparation.

Spatial autocorrelation

- Non-random vs random distribution

At St Kilda, non-random distributions were seen in the LDNA subpopulation of microplate 2 and 3 and the HDNA 2 subpopulation and total bacteria of microplate 2 due to these subpopulations having significant Moran's I values and z scores larger than the +1.96 cutoff value. Conversely all viral subpopulations were randomly distributed due to having non-significant Moran's I values. This may be due to motile bacterial subpopulations congregating around nutrient gradients and patches or resuspension of bacteria from the sediment leading to the observed surface gradients (Thar & Kühl 2003). These non-random distributions may not have been seen in the viral subpopulations as viruses are not motile and are unable to respond to chemical signals.

The significant Moran's I values seen in the bacterial subpopulations at St Kilda were positive, indicating clustering and spatial dependence. This clustering indicates that high values are located close together and low values are located close together, indicating the presence of 'hotspots' and 'coldspots' or gradients of high or low values. Moran's I values ranged from 0.04 to 0.07 for clustered distributions. These values are low when considering a Moran's I value of +1 indicates perfect clustering. These values are also low compared to Waters *et al.* (2003) which had significant Moran's I values between 0.082–0.180. These low values may indicate a low level of clustering present within these distributions.

At Noarlunga, all the bacterial and viral subpopulations were randomly distributed due to having non-significant Moran's I values. This indicates that no spatial autocorrelation was present amongst the bacterial and viral subpopulation distributions and that there is no spatial dependency within the distributions of the bacterial or viral subpopulations locally or globally. On a large scale this implies that environmental processes, such as turbulence, randomize the spatial distribution of bacterial and viral subpopulations at Noarlunga.

The non-random and random bacterial distributions found at St Kilda and Noarlunga coincide with the findings of Waters *et al.* (2003), which found higher spatial complexity within low-energy, high-chlorophyll environments such as St Kilda mangroves than in high-energy, low-chlorophyll environments such as Noarlunga. As St Kilda mangroves is a low-energy system with high productivity, due to the presence of dense microbial mats and mangrove forests causing stagnation in the lagoons, this may allow the development and maintenance of nutrient surface gradients (Barbara & Mitchell 1996; Seymour *et al.* 2004, 2005a, 2007b).

- Correlogram autocorrelation patterns

At St Kilda, the positive spatial autocorrelation seen within 5.9 cm of the sediment-water interface in most bacterial subpopulations indicates the presence of potential nutrient gradients that lead to bacterial clustering and the formation of surface gradients. The negative spatial autocorrelation seen beyond 5.9 cm from the sediment-water interface shows the area at which such gradients dissipate, leading to dispersed bacterial distributions. This trend may be due to eddies, Nyquist frequency limitation or a nepheloid layer. The presence of eddies could cause resuspension of particles via eddy penetration leading to a short-time mixing effect and turbulence (Boegman & Ivey 2009; Braaten *et al.* 1990; Nicholson 1988).

Nyquist frequency, whereby sampling frequency should be at least twice the highest frequency present in the sample data, may have contributed to this trend as if the sampling frequency was insufficient this would lead to the presence of aliases and distortions in the dataset (Kester 2009). The presence of a 5.9 cm nepheloid layer would result in high levels of suspended particulate matter being present at sediment-water interface and the presence of such layer may largely depend on the time of sampling as nepheloid layers are subject to change depths depending on the season (Rutgers van der Loeff *et al*. 2002; Sandilands & Mudroch 1983).

Taking into account the study by Sokal (1979), which showed that high Moran's I values not accompanied by high Geary's C values are representative of deterministic patterning of extreme values and the reverse implies that extreme values are randomly distributed and are hence spatially independent. Then for the bacterial subpopulations which showed positive spatial autocorrelation 5.9 cm from the sediment-water interface and then negative spatial autocorrelation beyond 5.9 cm, with relatively low corresponding Geary's C values between 0.73 and 0.83 suggests there was deterministic patterning of extreme values. In contrast, the positive to negative spatial autocorrelation trend within the Moran correlograms of the bacterial subpopulations that were found to be randomly distributed and had low Moran's I values which were less than the estimated cut-off value for clustered distributions but also had relatively low Geary's C values between 0.85 and 0.89 may indicate that globally these populations are spatially independent and hence randomly distributed, which is why the Moran's I value is low, but locally they show non-random distributions. In addition, the LDNA bacterial subpopulation at St Kilda in microplate 3 showed non-random distributions in the Moran's I testing however it had a higher Geary's C value than the other non-randomly distributed populations (0.92 compared

to 0.73–0.83) which was more similar to the Geary's C values of the randomly distributed subpopulations that showed local association of 0.85–0.89. According to Sokal (1979) this suggests that this subpopulation may be randomly distributed as the non-random result from the Moran's I was skewed by extreme values within the data set. This again highlights the importance of using Moran's I in conjunction with Geary's C as globally bacterial subpopulations may be spatially independent however locally they may be spatially dependent. As Geary's C is able to account for extreme outliers that may skew the overall result it is important to use in microbial distribution studies which have shown increases of up to 45 and 2500 times cm^{-1} in bacteria and viruses.

'Mean field' approach limitations

The microscale spatial heterogeneity of bacterial and viral populations in this study would be underestimated, by more than 3 orders of magnitude in some cases, had the mean field approach been employed (Alonso *et al.* 2001; Drake *et al.* 1998; Thomson *et al.* 2010). This is due to the presence of single point hotspots and coldspots containing more than 30% and less than 0.0008% of the total biomass within each microplate. As the distributions were random at Noarlunga but non-random at St Kilda the mean would not be representative of microbial biomass.

CONCLUSION

Spatial heterogeneity in microbial abundance has been found within the 11.3 cm directly above the sediment-water interface at Noarlunga estuary and St Kilda mangroves. This heterogeneity was present in the form of single and multiple point hotspots at Noarlunga and hotspots and surface gradients at St Kilda which resulted in high levels of viral and bacterial abundance being found near the sediment-water

interface. These patterns in microbial distributions have been seen in previous studies but to a lesser extent. At St Kilda, gradients in bacterial and viral abundance were much higher than previously reported with a 45- and 10.5-fold change cm^{-1} found compared to the previously reported 1.5-fold change in abundance cm^{-1} in bacteria (Seymour *et al.* 2007b) and 1.4-fold change cm^{-1} in viruses (Seymour *et al.* 2006, 2007b). At Noarlunga, bacterial and viral abundance showed a 12.9– and 2584-fold change cm^{-1}. This large fold-change found cm^{-1} would have been missed if lower resolution sampling was used. This high level of microscale heterogeneity highlights the importance of analysing microbial spatial distribution on the level of what individual cells experience rather than using a 'mean field' approach. In addition, this high level of microscale heterogeneity also illustrates the necessity and applicability of the novel two-dimensional high resolution sampler used in this study which was capable of collecting eight vertical profiles. The use of this sampler allowed the discrimination of single and multiple point hotspots and/or surface gradients, which was not possible in past studies as one-dimensional samplers were used.

Bacterial subpopulations were found to be randomly distributed at Noarlunga and non-randomly distributed at St Kilda with reasons for this finding relating to the low energy-nature of St Kilda allowing the formation of sediment-water interface nutrient and microbial surface gradients. Viral subpopulations only showed random distributions at both sites. Moran's I and Geary's C reveal that spatial autocorrelation amongst bacterial subpopulations at St Kilda show a positive to negative trend 5.9 cm from the sediment-water interface, indicating a mechanism, whether this be sampling alias or particulate resuspension, is present leading to this clustering and dispersal of bacterial populations.

Despite hotspots being ubiquitous features of microbial distributions, past studies lacked a universal definition. They qualitatively classified patches as regions of high microbial abundance that ‘exceed’ or are ‘elevated above’ background values (Seymour *et al.* 2004; 2005a, 2008). This paper provides a new quantitative method for determining hotspots, coldspots and background values within microscale microbial distributions. It is hoped that this method will be used in future microbial microscale distributions to allow comparative analyses among studies.

SUPPLEMENTARY INFORMATION

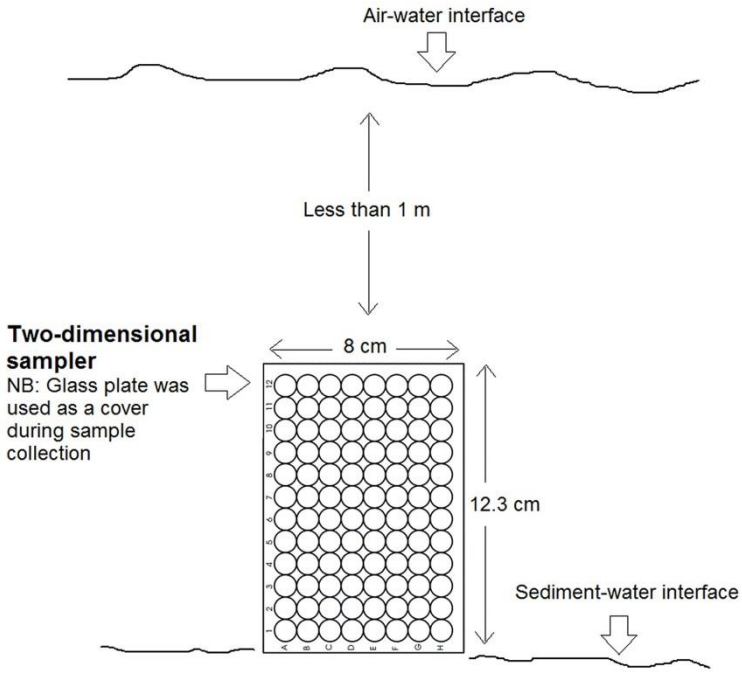


Figure S1. Collection of Vertical Profiles. The use of an 8×12 96-well microplate allowed the two-dimensional array of 8 vertical profiles (row A, B, C, D, E, F, G and H) per microplate which consisted of 12 sampling wells per vertical profile.

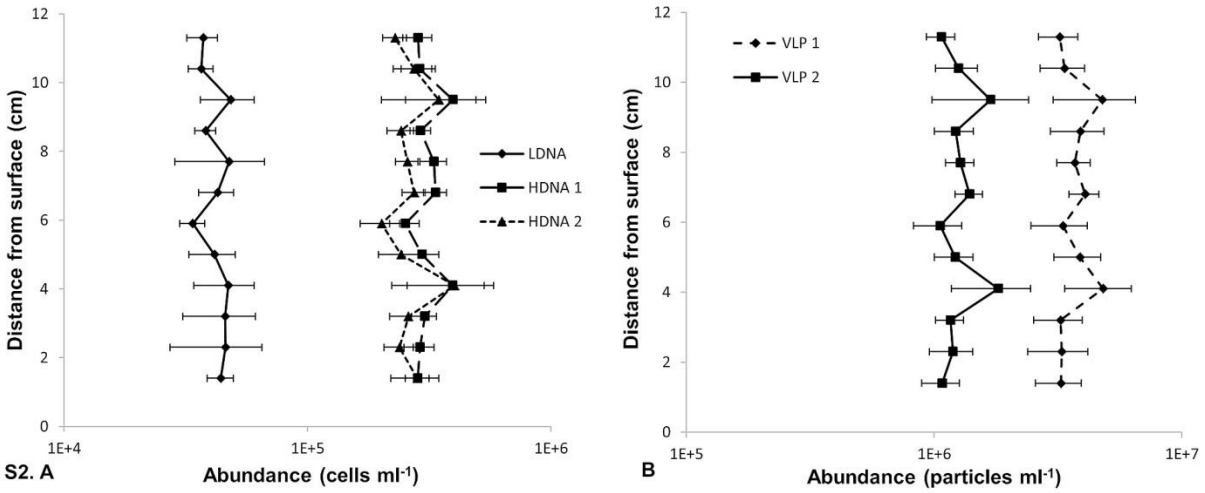


Figure S2. Comparisons of bacterial and viral subpopulations and viral subpopulations at Noarlunga via vertical depth profiles. Bacterial and viral subpopulations of all three microplates (n = 270). **A** Total mean LDNA, HDNA 1 and HDNA 2; **B** Total mean VLP 1 and VLP 2. Error bars represent the 95% confidence intervals obtained from each subpopulation of all three replicates (n = 12).

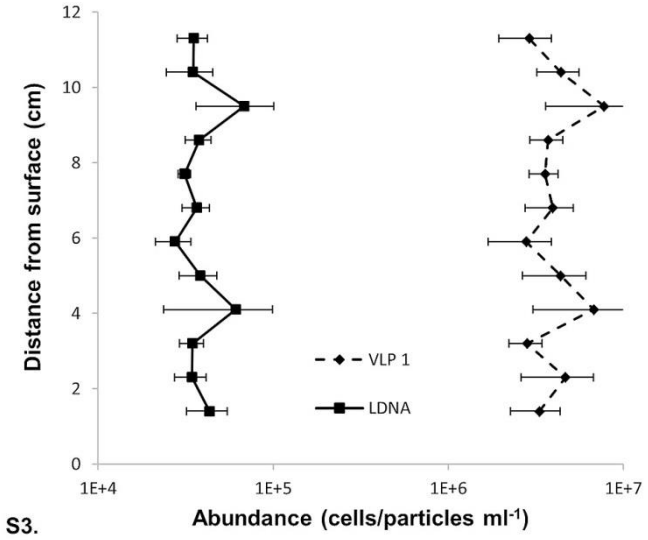


Figure S3. Mean vertical profile of VLP 1 and LDNA, microplate one at Noarlunga (n=90). Error bars represent the 95% confidence intervals obtained from one replicate (n = 12).

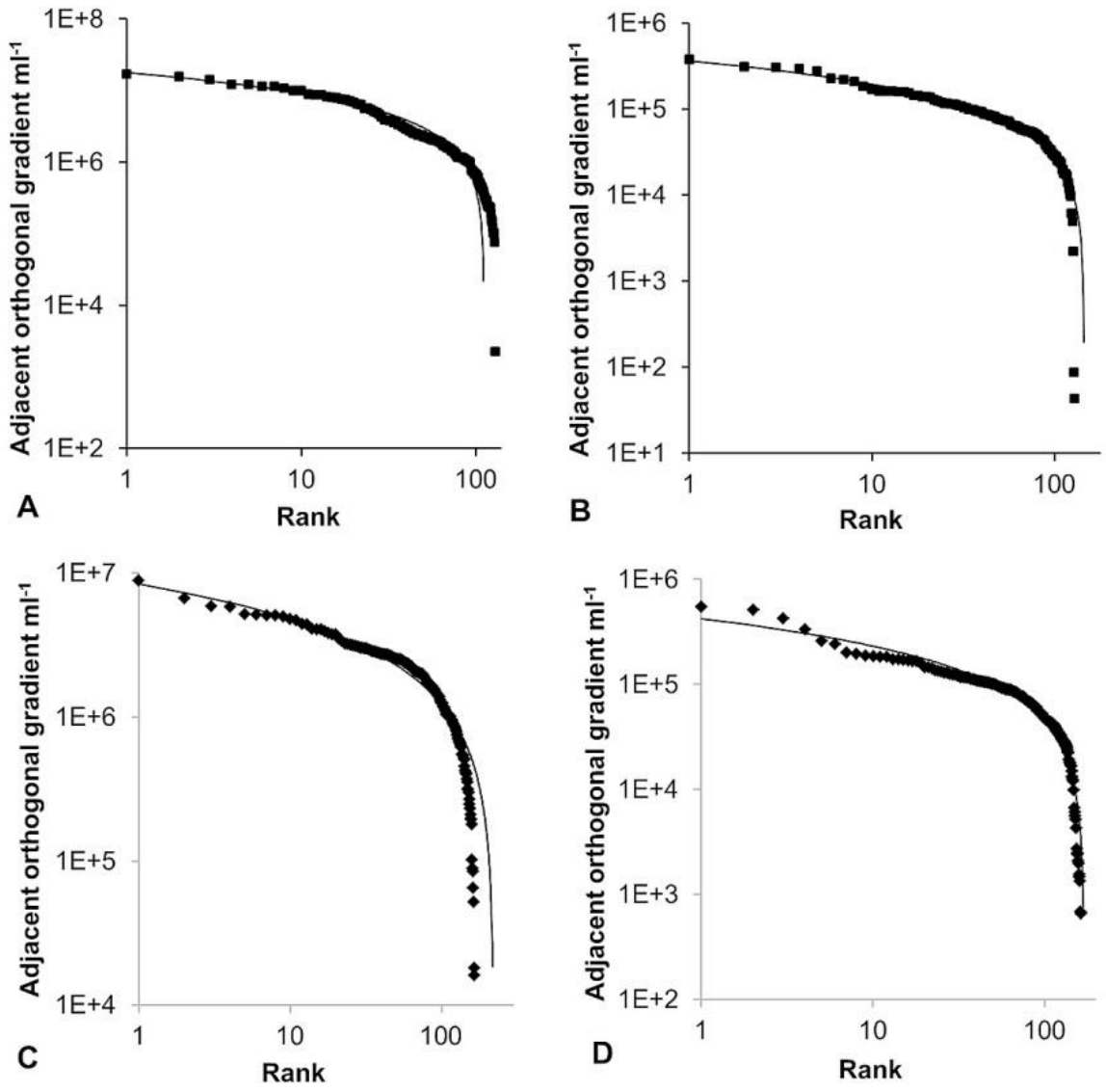


Figure S4. Adjacent orthogonal gradient distributions of bacteria and viruses as rank abundance. **St Kilda: A** VLP 1, **B** HDNA 2. **Noarlunga: C** VLP 2, **D** HDNA 2. Each distribution is ordered as a rank abundance. In each case the first rank was used for the 2 dimensional plots. A logarithmic trend-line was the best fit for each distribution, with the equations and R^2 for **A** $y = -4\times10^6\ln(x)+2\times10^7$ (R^2=0.97), **B** $y = -7\times10^4\ln(x)+4\times10^6$ (R^2=0.99), **C** $y = -2\times10^6\ln(x)+8\times10^6$ (R^2=0.98), **D** $y = -8\times10^4\ln(x)+4\times10^5$ (R^2=0.92). The p values are ≤ 0.05 in all cases.

Table S1. Mean viral abundances per microplate at Noarlunga and St Kilda.

Population	Microplate Number	VLP Abundance x 10^6 particles ml^{-1} (95%CI, n)	
		Noarlunga	**St Kilda**
VLP 1	1	4.3 (0.6, 89)	8.4 (1.0, 69)
	2	4.0 (0.37, 94)	6.6 (0.7, 58)
	3	2.9 (0.35, 86)	7.2 (0.8, 82)
VLP 2	1	1.5 (0.3, 89)	0.7 (0.1, 69)
	2	1.3 (0.1, 94)	0.6 (0.07, 58)
	3	1.0 (0.1, 86)	0.5 (0.05, 82)
Total VLP	1	5.75 (0.9, 89)	9.15 (1.1, 69)
	2	5.3 (0.4, 94)	7.25 (0.75, 58)
	3	3.9 (0.4, 86)	7.8 (0.8, 82)

Table S2. Mean bacterial abundances per microplate at Noarlunga and St Kilda.

Population	Microplate Number	Prokaryotic Abundance x 10^5 cells ml^{-1} (95%CI, n)	
		Noarlunga	**St Kilda**
LDNA	1	0.4 (0.05, 89)	3.4 (0.65, 69)
	2	0.4 (0.05, 94)	3.3 (0.8, 58)
	3	0.5 (0.1, 86)	2.4 (0.1, 82)
HDNA 1	1	3.6 (0.5, 89)	2.4 (0.9, 69)
	2	2.9 (0.2, 94)	2.8 (1.3, 58)
	3	2.9 (0.2, 86)	1.4 (0.1, 82)
HDNA 2	1	3.2 (0.6, 89)	6.8 (4.1, 69)
	2	2.8 (0.2, 94)	8.1 (3.7, 58)
	3	2.1 (0.2, 86)	3.5 (0.2, 82)
Total Prokaryotes	1	7.25 (1.15, 89)	9.8 (1.4, 68)
	2	6.05 (0.3, 94)	14.1 (5.6, 58)
	3	5.5 (0.35, 86)	7.4 (0.3, 82)

Table S3. Comparison of the Moran's I values and Geary's C values obtained for each bacterial subpopulation at Noarlunga and St Kilda.

Microplate Number	Population	Noarlunga		St Kilda	
		Moran's *I* (p-value)	**Geary's C (p-value)**	**Moran's *I* (p-value)**	**Geary's C (p-value)**

1	LDNA	-0.008 (n.s)	0.87 (0.0001)	0.021(n.s)	0.85 (0.0001)
	HDNA 1	-0.008 (n.s)	0.87 (0.0001)	0.004 (n.s)	0.86 (0.0001)
	HDNA 2	-0.008 (n.s)	0.87 (0.0001)	0.006 (n.s)	0.85 (0.0001)
	Total Prokaryotes	-0.008 (n.s)	0.87 (0.0001)	0.019 (n.s)	0.89 (0.001)
2	LDNA	-0.017 (n.s)	0.96 (n.s)	0.062 (0.001)	0.73 (0.0001)
	HDNA 1	-0.004 (n.s)	0.98 (n.s)	0.036 (n.s)	0.75 (0.0001)
	HDNA 2	-0.004 (n.s)	0.98 (n.s)	0.070 (0.001)	0.83 (0.0001)
	Total Prokaryotes	0.007 (n.s)	0.94 (n.s)	0.070 (0.001)	0.77 (0.0001)
3	LDNA	-0.017 (n.s)	0.98 (n.s)	0.039 (0.01)	0.92 (0.001)
	HDNA 1	-0.011 (n.s)	0.99 (n.s)	0.0002 (n.s)	0.99 (n.s)
	HDNA 2	-0.009 (n.s)	1.01 (n.s)	0.002 (n.s)	0.97 (n.s)
	Total Prokaryotes	-0.012 (n.s)	0.98 (n.s)	0.007 (n.s)	0.96 (n.s)

Table S4. Comparison of the Moran's I values and Geary's C values obtained for each viral subpopulation at Noarlunga and St Kilda.

Microplate Number	**Population**	**Noarlunga**		**St Kilda**	
		Moran's *I* (p-value)	**Geary's C (p-value)**	**Moran's *I* (p-value)**	**Geary's C (p-value)**
1	VLP 1	-0.001 (n.s)	1.02 (n.s)	-0.026 (n.s)	1.01 (n.s)
	VLP 2	-0.008 (n.s)	0.87 (0.0001)	-0.023 (n.s)	0.96 (n.s)
	Total VLP	-0.008 (n.s)	0.87 (0.0001)	-0.021(n.s)	1.00 (n.s)
2	VLP 1	0.007 (n.s)	1.01 (n.s)	-0.044 (n.s)	1.02 (n.s)
	VLP 2	0.004 (n.s)	0.99 (n.s)	-0.029 (n.s)	0.97 (n.s)
	Total VLP	0.009 (n.s)	1.01 (n.s)	-0.044 (n.s)	1.00 (n.s)
3	VLP 1	0.014 (n.s)	1.00 (n.s)	0.015 (n.s)	0.96 (n.s)
	VLP 2	0.003 (n.s)	0.98 (n.s)	0.0007 (n.s)	1.00 (n.s)
	Total VLP	0.013 (n.s)	0.99 (n.s)	0.018 (n.s)	0.96 (n.s)

Chapter 4

Correlations Between Virus-Like Particles and Prokaryotes Within Microenvironments

Published as:

Dann LM, Paterson JS, Newton K, Oliver R, Mitchell JG (2016) Correlations between virus-like particles and prokaryotes within microenvironments. *PLoS One* 11 (1): e0146984. doi:10.1371/journal.pone.0146984

ABSTRACT

Microbial interactions are important for ecosystem function, but occur at the microscale and so are difficult to observe. Previous studies in marine systems have shown significant shifts in microbial community abundance and composition over scales of micrometres to centimetres. This study investigates the microscale abundance distributions of virus-like particles (VLPs) and prokaryotes in the lower reaches of a river to determine the extent to which microscale microbial patchiness exists in freshwater systems. Here we report local hotspots surrounded by gradients that reach a maximum 80 and 107 fold change in abundance over 0.9 cm for prokaryotic and VLP subpopulations. Changes in prokaryotic and VLP hotspots were tightly coupled. There were no gradients at tens of centimetres across the boundary layers, which is consistent with strong mixing and turbulence-driven aggregation found in river systems. Quantification of the patchiness shows a marked asymmetry with patches 10 times greater than background common, but depletions being rare or absent in most samples. This consistent asymmetry suggests that coldspots depleted by grazing and lysis are rapidly mixed to background concentrations, while the prevalence of hotspots indicates persistence against disruption. The hotspot to coldspot relative abundance may be useful for understanding microbial river dynamics. The patchiness indicates that the mean- field approach of bulk phase sampling misses the microbially relevant community variation and may underestimate the concentrations of these important microbial groups.

INTRODUCTION

Heterotrophic prokaryotes are crucial to freshwater systems as they cycle important nutrients, such as nitrogen, carbon and phosphorus (Azam 1998; Long & Azam 2001). Similarly, viruses are likely key players in freshwater ecosystems as they affect primary production, nutrient and dissolved organic matter release, genetic exchange between microorganisms and control bacterial community composition and abundance via viral-mediated lysis (Fuhrman 1999; Newton *et al.* 2011; Suttle 2005; Wilhelm & Matteson 2008; Wommack & Colwell 2000). The distributions of these microbial communities have often been considered homogeneous, leading to mean-field approach sampling whereby bulk phase millilitre to litre samples were considered representative of microbial microscale processes (Long & Azam 2001). However, previous studies on patch dynamics within aquatic and terrestrial ecosystems confirm heterogeneous spatial distributions are common in most organisms. For instance, biodiversity 'hotspots' are observed in flora and terrestrial vertebrate species (Brooks *et al.* 2002; Margules & Pressey 2000), with 20% of all plant diversity found in 18 hotspots that accounted for only 0.5% of the Earth's land area (Mittermeier *et al.* 1998; Myers *et al.* 2000). In addition, for marine environments, it is well accepted that microbial abundance and activity differ by orders of magnitude over micrometres to centimetres (Azam 1998; Long & Azam 2001; Duarte & Vaqué 1992; Mitchell & Fuhrman 1989; Seymour *et al.* 2004, 2005, 2006), with previous studies showing 45 to 2584 fold cm^{-1} variation in bacterial and viral abundance (Dann *et al.* 2014), and therefore indicating deficiencies within the mean-field sampling approach.

As microbial communities interact at micrometre to millimetres scales, to understand these microscale processes and dynamics, it is important to analyse their

distributions at such scales. Microscale sampling is an important supplement when precise abundance estimates are necessary, the local interactions may reveal processes not observed by bulk sampling or for localizing particular populations. Determining the microscale abundance distribution is important as it indicates the microbial structure present at the scale of microbial food webs. Determining the microscale abundance distribution is important as it indicates the microbial structure present at the scale of microbial food webs. However, in freshwater systems, studies of microbial distributions often still adopt the mean-field approach, therefore excluding direct examination of important ecological associations between microbial communities (Jezbera *et al.* 2012; Lear *et al.* 2014; Liu *et al.* 2011; Palijan 2012; Pollard & Docklow 2011).

Associations between microbial populations are often observed within microscale distributions in the form of heterogeneous “hotspots” (Azam 1998; Duarte & Vaqué 1992; Lear *et al.* 2014; Mitchell & Fuhrman 1989; Mittermeier *et al.* 1998; Myers *et al.* 2000). Hotspots are areas of elevated bacterial or viral abundance, which can result from bacterial accumulation around high nutrient concentration areas via chemotaxis, aggregation or disintegration of particles, small scale water mixing and the occurrence of grazing in adjacent areas or interactions with suspended particles (Barbara & Mitchell 2003; Blackburn *et al.* 1998; Seymour *et al.* 2005; Stocker & Seymour 2012).

Within fluvial systems suspended particulate inorganic and organic matter are often found in high concentrations (Peduzzi 2015). This particulate organic matter (POM) can comprise a range of different materials, such as biofilms, soils, riparian vegetation or autochthonous algae, and is an important source of energy as bacterial communities contribute to the decomposition and remineralisation of this POM

(Kirchman & Mitchell 1982; Luef *et al.* 2009; Simon *et al.* 2002; Zimmermann-Timm 2002). Previous studies have found higher heterogeneity, and often abundances, within particle-associated, rather than free living communities, with particle-associated bacteria contributing to approximately 30.34% of total bacterial abundance (Besemer *et al.* 2005; Luef *et al.* 2007; Peduzzi & Luef 2008). Particle-associated bacterial communities, such as those found on river snow particles, are important to system function as they contribute to a significant amount of production and activity (Griffith *et al.* 1990; Grossart & Simon 1993, 1998; Iriberri *et al.* 1987; Luef *et al.* 2007).

The presence of these particles also impacts viral communities, with viral abundance on suspended particulate matter ranging from 10^5 to 10^{11} particles ml^{-1} or between 0.4% and 35% of total viral abundance (Luef *et al.* 2007; Peduzzi & Luef 2008). Viral attachment to particles and/or the presence of particulate matter can result in different ecological consequences; for instance, a loss of infectivity as a result of viral adsorption to solid particles, which causes a reduction in viral infection and lysis and a subsequent increase in free-living prokaryotic growth (Luef *et al.* 2009). In addition, prolonged survival or increased phage production and transduction can result from viral particle attachment as high viral abundances on riverine particles may represent microenvironments of heightened viral infection and lysis of bacterial communities (Kapuscinski & Mitchell 1980; Kokjohn *et al.* 1991; Luef *et al.* 2007, 2009; Ripp & Miller 1995; Suttle & Chen 1992).

Lysis events will produce intense local concentrations of viruses, which may spread as local epidemics. This should produce areas of depleted bacterial concentration. However, this is not observed in turbulent environments, presumably because shear caused by turbulence easily disperses the immotile, non-aggregating microbes.

Hotspot formation and maintenance, whether viral or bacterial, is a balance of aggregation and dispersion (Abraham 1998; Siegel 1998; Young *et al.* 2001). Dispersion at the microscale in turbulent systems such as rivers is driven by Kolmogorov eddies. These are the smallest possible eddies for a given fluid viscosity (Kolmogorov 1991). For rivers, where the viscosity is close to that of pure water, the Kolmogorov eddy length is approximately 1 – 10 mm (Hondzo & Lyn 1999; Taylor & Stocker 2012). When eddies and heterogeneities are the same size, mixing is the most efficient homogenizing signals. This is relevant for this paper because our sampling interval is right at the scale where there should be the most homogeneity (Siegel 1998).

While the high shear of Kolmogorov eddies efficiently erase gradients of microbes and nutrient signals (Durham & Stocker 2012; Durham *et al.* 2009; Mitchell *et al.* 1985), many eddies are much larger, which can create conditions for clustering (Mitchell *et al.* 1985). This is because low shear environments are characterised by eddies with long lifetimes, which allow microscale nutrient patch formation and consequently chemotactic swarming of bacteria. The lifetimes of these Kolmogorov eddies differ depending on turbulence, with freshwater environments having eddy lifetimes of approximately 1,000 seconds and lengths of greater than 3 cm (Hondzo & Lyn 1999; Mitchell *et al.* 1985). Shear then helps determine where hotspots and coldspots can form, their size and their lifetime (Hondzo & Lyn 1999; Mitchell *et al.* 1985, 2008; Pollard & Ducklow 2011).

Previously, hotspot discrimination has remained qualitative, identifying abundance regions ‘elevated’ above background. However, a quantitative method developed by Dann *et al.* (2014) discriminated hotspot, coldspot and background via rank abundance graphs, separating sample values based on their slope and line of best

fit. Hotspots were shown to have steep slopes and follow a power law best fit, coldspots and background values follow linear best fits, while coldspots had steeper slopes than background values (Dann *et al.* 2014). The aim of this study was to test the hypothesis thatmicroscale VLP and prokaryote abundance variations occur in rivers. To test this, the microscale distributions of prokaryote and VLP subpopulations within the Murray River were analysed at millimetre resolution.

MATERIALS AND METHODS

Sample collection

Freshwater samples were collected from the Murray River at Murray Bridge, South Australia (-35˚12, 139˚28). Sampling occurred on June 14th, 2012 at 11 am. Daily summary data showed a water level of 0.45 metres, and electrical conductivity (EC) of 325 uS/cm (Long Island Site ID: A4261162) (DEWNR 2012). Water flow rates ranged from 0.01 m/s to 0.09 m/s. The flow rates were determined via velocity profiles using a Flo-Mate (Model 2000) current and flowmeter. Velocity profiles were taken via fixed time averaging with 60 second intervals. Specific permission to access the sampling site was not required. The field study did not involve endangered or protected species.

Triplicate samples were collected at the sediment-water and air-water interface from three locations each separated by 10 metres. Samples were collected using a two-dimensional sampler trialled previously (Dann *et al.* 2014). The sampler, a 3 x 12-well microplate with a glass cover collected three vertical profiles, each containing 12

sample points, ranging from 1.4 cm to 11.3 cm from the sediment- and air-water interface with a sampling resolution of 0.9 cm.

Once collected, samples were transferred into 2 ml cryovials containing 4 µl of glutaraldehyde (0.5% final concentration) and stored in the dark at 4 ˚C for 15 minutes. Samples were then quick frozen in liquid nitrogen and stored at -80 ˚C until analysis (Brussaard 2004; Paterson *et al.* 2012). Flow cytometric analysis was performed within three weeks to avoid sample deterioration (Brussaard 2004).

Flow cytometry

For prokaryote and VLP enumeration, samples were thawed and diluted 1:100 with Tris-EDTA buffer (pH 8.0, 0.2 µm filtered, 10 mM Tris, 1 mM EDTA) and stained with SYBR Green I (1:500 final dilution commercial stock; Molecular Probes), a nucleic acid-specific dye (Gasol & Del Giorgio 2000; Schapira *et al.* 2009). Samples were then incubated in the dark at 80 ˚C for 10 minutes to optimise VLP counts (Dann *et al.* 2014; Blackburn *et al.* 1998; Brussaard 2004; Marie *et al.* 1999; Schapira *et al.* 2009). Each sample was measured in triplicate to check the precision of the method.

Flow cytometry was performed on a FACSCanto II flow cytometer (BD) using a phosphate-buffered saline (PBS) solution as sheath fluid. Forward-angle light scatter (FSC), right-angle light scatter (SSC) and green fluorescence (SYBR I) were collected for each sample. Each sample was run at a low flow rate setting to obtain less than 1000 events per second. Reference fluorescent yellow beads (1 µm diameter, Molecular Probes) were added to each sample as an internal size and concentration standard with flow cytometer settings normalised to fluorescence and bead concentration (Blackburn *et al.* 1998; Schapira *et al.* 2009). Epifluorescent microscopy was used to ensure bead reliability by confirming a final concentration of

approximately 10^5 beads ml^{-1} per sample (Gasol & Del Giorgio 2000; Roudnew *et al.* 2014; Schapira *et al.* 2009; Smith *et al.* 2015). Flow cytometric data was exported as FCS 3.0 files and prokaryotic and VLP subpopulation enumeration was performed using FlowJo (Tree Star, Inc.) (Smith *et al.* 2015). VLP and prokaryotic subpopulations were discriminated via peaks in monoparametric histograms of SYBR green fluorescence and saturated regions in biparametric cytograms of SYBR green fluorescence and side-scatter (Brussaard 2004; Dann *et al.* 2014; Marie *et al.* 1997, 1999; Roudnew *et al.* 2014).

Microscale distribution analysis

The microscale distributions of prokaryotic and VLP subpopulations at the air- and sediment-water interface were determined via flow cytometric abundance counts in each sample well of the vertical profiles collected. From this, two-dimensional contour plots were constructed using Surfer 10 (Golden Software, Inc.). Background values reported are the median values in the dataset rather than the mean, as inclusion or exclusion of the hotspot values did not overly affect the median values (Wiebe 1970).

For this study, hotspot and background values were determined via rank abundance graphs (Dann *et al.* 2014). Briefly, background values were those that fitted a linear trend and hence were indistinguishable from a random distribution, whilst the hotspot values exceeded this linear fit, exhibiting a steep power law trend, indicating their non-random nature (Dann *et al.* 2014).

Subpopulation correlations

Pearson's correlation coefficients were performed for each vertical profile collected at the air- and sediment-water interface and the α of 0.05 was reduced by sequential Bonferroni (Holm 1979). All possible subpopulation correlations were considered in order to identify potential relationships between the VLP and prokaryotic subpopulations. Two-sample t-tests were employed to identify possible abundance differences between the air- and sediment-water interface.

The VLP to prokaryote ratio (VPR) was used as a potential indicator of VLP and prokaryote interactions. Previous studies used the patchine to investigate the biological dynamics of systems (Roudnew *et al.* 2014). Higher VPRs are typically found in more productive and nutrient-rich ecosystems with the suggestion that these conditions favour maximum prokaryotic growth and productivity (Jacquet *et al.* 2010; Weinbauer 2004; Wommack & Colwell 2000).

Spatial autocorrelation analysis

Spatial autocorrelation analysis was performed to determine the spatial dependence within prokaryote and VLP communities. Originally used by the National Institute of Justice for identifying non-random spatial patterns in crime occurrences (Smith & Bruce 2008), spatial autocorrelation analysis can also be applied to organism distribution and has been used previously to look at the level of spatial complexity within phytoplankton communities (Waters *et al.* 2003); the distribution of cellular nucleic acid signals within floating riverine aggregates (Luef *et al.* 2009), the frequency of particular allozymes in snail populations (Sokal & Oden 1978); and the distribution of particular genes in Australian rat species (Peakall *et al.* 2003). This study used two common spatial autocorrelation analysis statistics; Moran's I and Geary's C, to assess the spatial complexity of the indigenous prokaryote and VLP

communities (Geary 1954; Moran 1950; Sokal & Oden 1978). Moran's I and Geary's C identify whether two-dimensional and multi-directional correlations are present between proximate sample points that have similar values (Geary 1954; Moran 1950; Sokal & Oden 1978). Analysing the spatial autocorrelation in VLP and prokaryotic microscale distributions via Moran's I and Geary's C builds on previous work by Dann *et al.* (2014).

- Moran's I

Moran's I spatial autocorrelation statistics test (CrimeStat 3.3, Ned Levine software) was used to identify the degree of spatial dependence present in the VLP and prokaryotic subpopulations. The equation for Moran's I is:

Eq. 1

$$I = \frac{N \sum i \sum j W_{ij} \left(X_i - \overline{X} \right) \left(X_j - \overline{X} \right)}{\left(\sum i \sum j W_{ij} \right) \sum i \left(X_i - \overline{X} \right)^2}$$

Where N refers to the sample number, X_i and X_j are the variable values at specific locations, i and j (where i ≠ j), $\overline{X}$ is the mean of the variable and W_{ij} is the weight applied to the i and j comparison. A weighted Moran's I test was chosen as this applies a weight value of 1 and 0 to adjacent and non-adjacent sample points respectively.

Moran's I (Moran 1950) is a global statistical test used to identify spatial dependence within a dataset. It is multi-directional as it can use vertical, horizontal and diagonal directional analysis for correlation calculations. Moran's I has a value range from +1 to -1, with +1 indicating perfect clustering where high/low values are proximate, -1 indicating perfect dispersion where high/low values are found far apart and zero being indicative of a random distribution. For Moran's I, the critical cut-off values are

often established through a collation of previous literature using the same or similar scales or values. This information is lacking for prokaryotes and viruses and the critical cut-off values for small-scale spatial studies are unknown.

- Moran correlograms

Moran correlograms were created using CrimeStat 3.3 whereby output values from Moran's I statistic were applied to two pairs of sample values that were separated by a chosen lag distance, in this instance 0.9 cm, the distance between each well in the sampler used.

Significance was determined by the standard error obtained from Moran's I and the degree of spatial autocorrelation was determined via the output values given for each lag distance in the correlograms. From this, the level of spatial dependence could be identified at each sample distance as opposed to the whole sampling area, as was achieved in the Moran's I statistic.

- Geary's C

Geary's C spatial autocorrelation test (Geary 1954) within CrimeStat 3.3 was used to identify deterministic patterning of extreme values and non-spatially related chance phenomena within the dataset. The equation for Geary's C is:

Eq. 2
$$C = \frac{(N-1)\left[\sum i \sum j W_{ij}\left(X_i - X_j\right)^2\right]}{2\left(\sum i \sum j W_{ij}\right)\sum i\left(X_i - \overline{X}\right)^2}$$

All terms are the same as in Eq. 1.

Geary's C is more sensitive to local clustering and can be used in conjunction with Moran's I. The Geary's C statistical test is similar to Moran's I, however spatial

dependence is calculated via the deviation in intensity of each sample value's location compared to one another, whereas Moran's I calculates spatial dependence via the cross-product of the deviations from the mean within the sample values.

Geary's C has a value range from 0 to approximately 2, with no definitive upper limit (Geary 1954). Spatial independence is indicated by a value of 1 whilst positive spatial autocorrelation is indicated by values < 1 and negative spatial autocorrelation is indicated by values > 1. Thus, Geary's C is inversely related to Moran's I (Griffith 1987).

RESULTS

Prokaryotic and VLP abundance

Flow cytometric analysis revealed two prokaryotic subpopulations, referred to as a low- and high-density nucleic acid prokaryote population (LDNA and HDNA), and two virus-like particles (VLP 1 and VLP 2) subpopulations (Fig. 1). From all the samples collected, the mean VLP 1 abundance ranged from 5.4 to 6.7 x 10^7 particles ml^{-1} (95%CI = 7.2 x 10^6 particles ml^{-1}, n = 648) whilst mean VLP 2 abundance ranged from 1.7 to 2.1 x 10^7 particles ml^{-1} (95%CI = 2.3 x 10^6 particles ml^{-1}, n = 648) (Table S1). The mean prokaryotic abundances were lower than the VLP mean abundances, with the mean abundance of the LDNA subpopulation ranging from 1.3 to 1.6 x 10^7 cells ml^{-1} (95%CI = 1.3 x 10^6 cells ml^{-1}, n = 648) and the HDNA subpopulation exhibiting mean abundances that ranged from 0.9 to 1.5 x 10^7 cells ml^{-1} (95%CI = 1.9 x 10^6 cells ml^{-1}, n = 648) (Table S2). This resulted in a VPR ranging from 0.4 to 3.4 (Table S3). In all samples, the VLP 1 abundance was higher than VLP 2 and the

HDNA abundance was equal to or less than the LDNA abundance (Table S1 and S2). There was no significant difference in VLP and prokaryotic mean abundances between the air-water and sediment-water interface ($p = 0.39$).

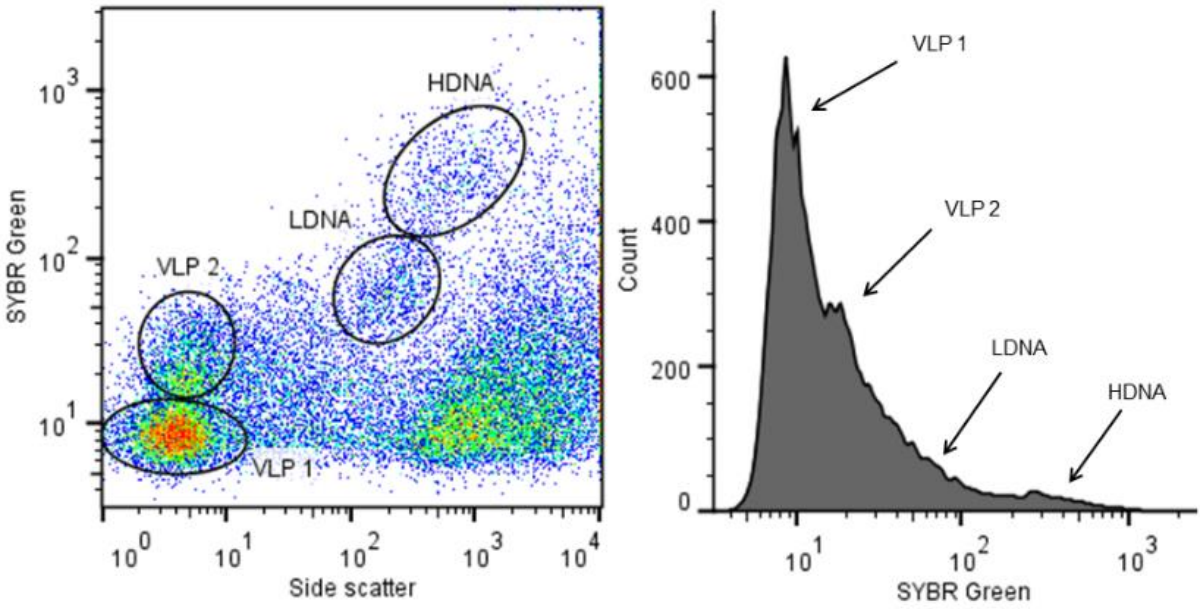


Figure 1. Representative flow cytometric cytogram and histogram. (A) Cytogram of SYBR green fluorescence versus side scatter and (B) histogram of SYBR green fluorescence showing two VLP (VLP 1 and VLP 2) and two prokaryotic (LDNA and HDNA) subpopulations.

Hotspots

Rank abundance graphs of the vertical profiles revealed hotspots within all the prokaryotic and VLP subpopulations at the air- and sediment-water interface. In 19 of 24 samples the background values followed a single linear trend whilst the remainder followed two or three linear trends. Background values followed a linear trend with slopes ranging from -2 x 10^5 and -5 x 10^6, whilst the hotspot values followed a power law trend with exponents ranging from -0.26 to -1.43 (Fig. 2A-D).

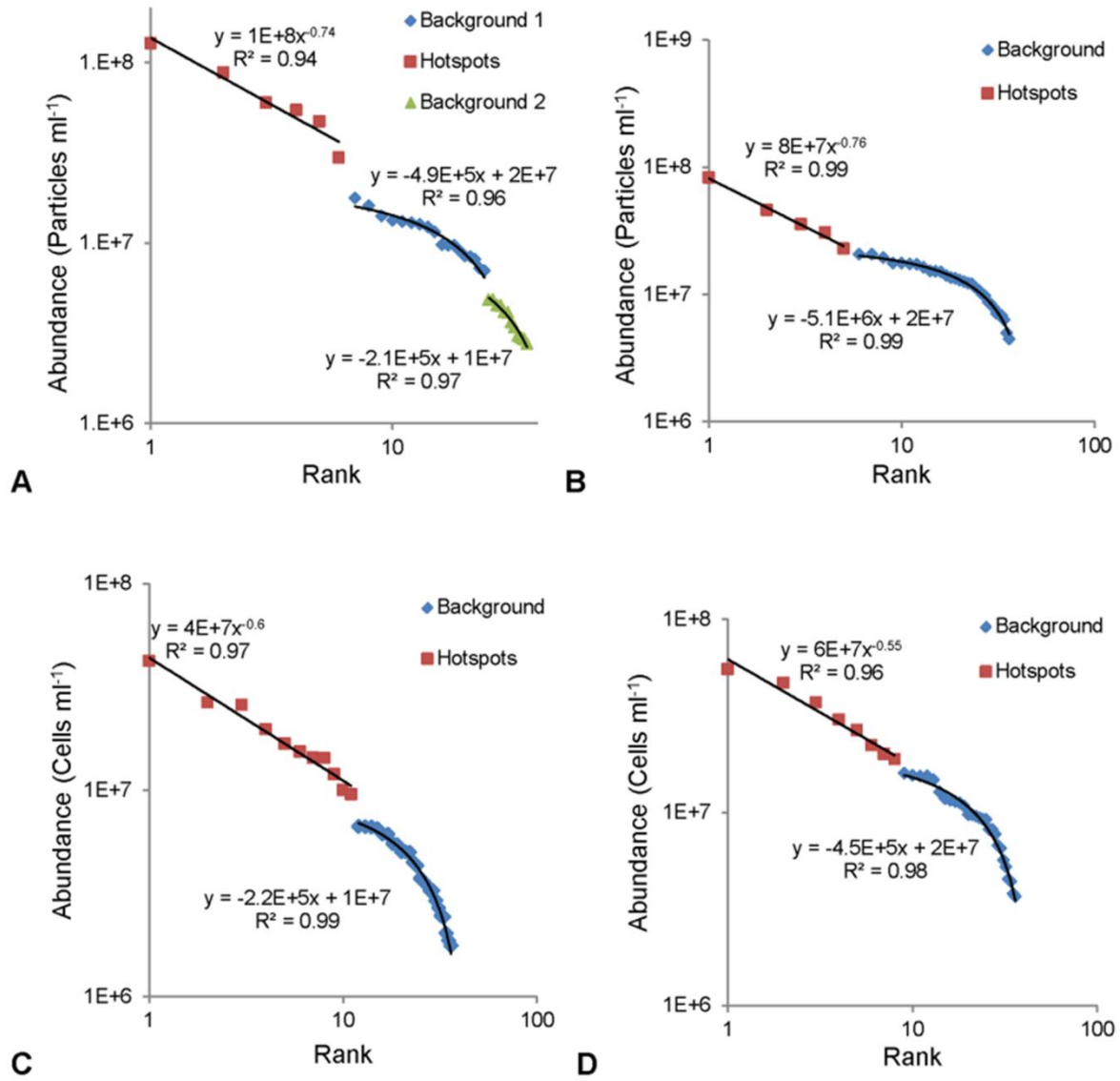


Figure 2. Representative rank abundance graphs for hotspot determination. Background values followed a linear trend, whilst hotspots exhibited a power law trend. (A) VLP 2 at air-water interface. (B) VLP 1 at air-water interface. (C) HDNA at sediment-water interface. (D) LDNA at air-water interface.

In all instances, the maximum abundance hotspots were present at the sediment-water interface, and when analysing the total 1.8 cm x 11.3 cm sampling area, VLP 1 had a maximum hotspot of 8.5 x 10^8 particles ml^{-1} with a background of 3.6 x 10^7 particles ml^{-1}, resulting in a 24 fold increase in heterogeneity (Fig. 3A). Whereas VLP 2 had a lower maximum hotspot value of 2.6 x 10^8 particles ml^{-1} but over a

background of 1 x 10^7 particles ml^{-1} resulting in a 26 fold change in heterogeneity, which was higher than VLP 1 (Fig. 3B).

As the prokaryotic abundances were lower than the VLP abundances, the maximum abundance hotspots found in LDNA and HDNA were lower than VLP 1 and VLP 2. LDNA had a maximum abundance hotspot of 1.3 x 10^8 cells ml^{-1} over a background of 9 x 10^6 cells ml^{-1}, resulting in a 15 fold change in heterogeneity over the 1.8 x 11.3 cm sampling area (Fig. 3C). HDNA had similar values, with a maximum abundance hotspot of 1.4 x 10^8 cells ml^{-1} but over a background of 4.6 x 10^6 cells ml^{-1} resulting in a 31 fold difference in heterogeneity which was higher than LDNA and was the largest change in heterogeneity in all of the subpopulations (Fig. 3D).

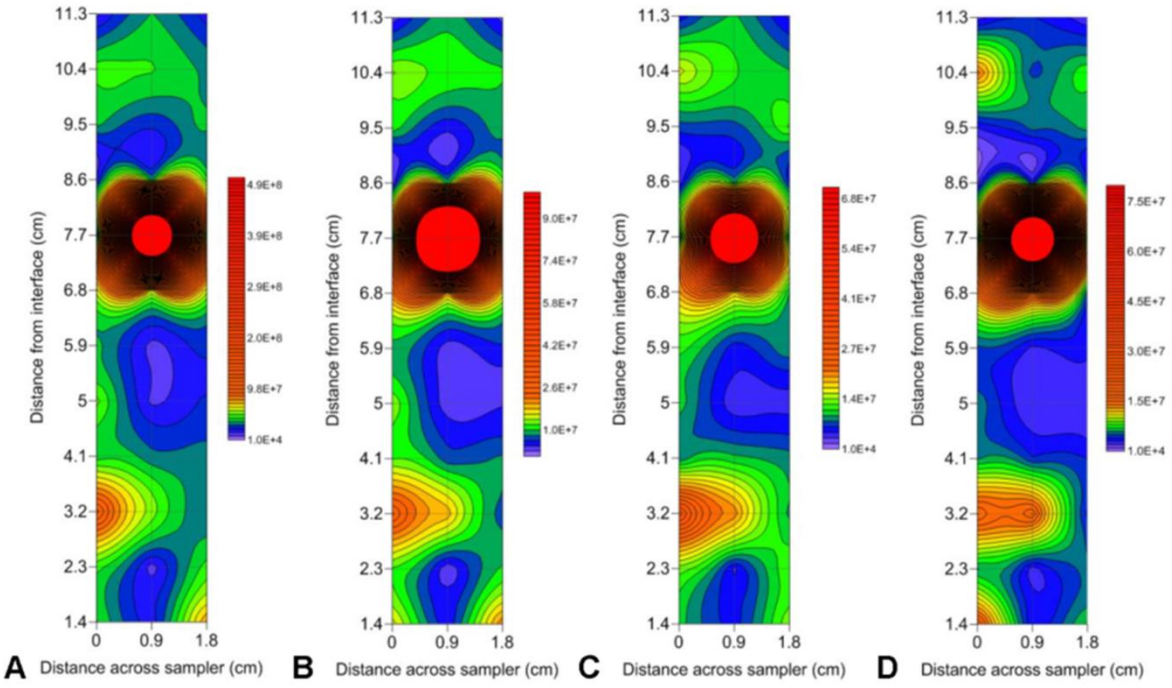


Figure 3. Contour plots showing hotspots in prokaryotic and VLP subpopulations at the sediment-water interface. (A) VLP 1 (B) VLP 2 (C) LDNA (D) HDNA. Abundance levels indicated via a colour intensity scale in units of cells/particles ml^{-1}. Faint gridlines indicate sample intervals. A minimum contour interval value of 10000 was chosen as this was larger than the maximum flow

cytometer machine error observed in blank control samples. Solid red areas indicate abundance points higher than the maximum contour level selected.

However, these maximum abundance hotspot values were not the cause for the largest changes in heterogeneity. The largest fold changes were from one sample point to the next, over a distance of 0.9 cm, rather than across the entire sampling area. This was due to the occurrence of single point hotspots adjacent to low abundance values. VLP 1 showed a maximum 74 fold change in heterogeneity over 0.9 cm, whilst VLP 2 had the highest change in heterogeneity showing a maximum 107 fold change over 0.9 cm (Fig. 3A-B). For the prokaryotic subpopulations, LDNA had a maximum 41.5 fold change over 0.9 cm, whilst HDNA showed a maximum 80.5 fold change over 0.9 cm (Fig. 3C-D).

Subpopulation correlations

Two-dimensional contour plots revealed that the prokaryotic and VLP subpopulations were correlated overall, but that there were locational differences caused by the presence or absence of hotspots in prokaryotic and VLP abundance (Fig. 4).

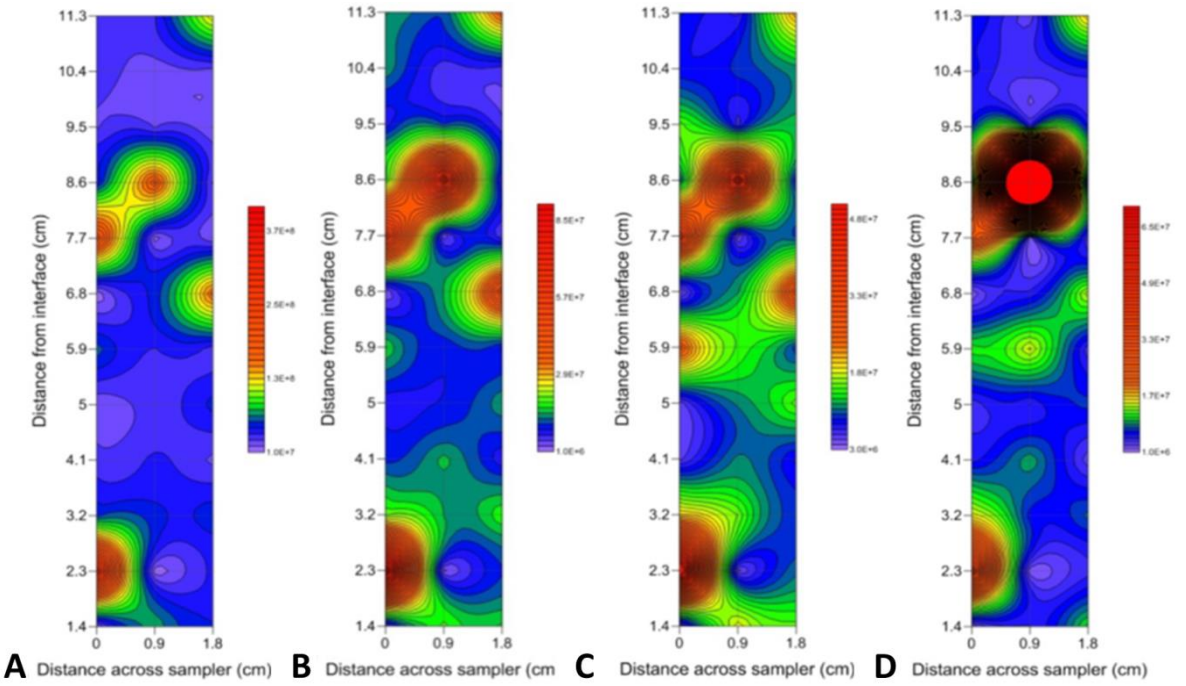


Figure 4. Representative two-dimensional contour plots showing the presence vs. absence of hotspots in prokaryotic and VLP abundance. (A) VLP 1, (B) VLP 2, (C) LDNA and (D) HDNA at the air-water interface. Abundance levels indicated via a colour intensity scale in units of cells/particles ml^{-1}. Faint gridlines indicate sample intervals. A minimum contour interval value of 10000 was chosen as this was larger than the maximum flow cytometer machine error observed in blank control samples. Solid red areas indicate abundance points higher than the maximum contour level selected.

To identify potential relationships between plankton populations, Pearson correlation coefficients were run for each subpopulation pair at the air-water interface and sediment-water interface. From the 108 possible single vertical profile subpopulation correlations between VLP 1, VLP 2, LDNA and HDNA, 92 were significantly correlated with an r value ≥ 0.79 ($p \leq 0.003$, $n = 108$) (Fig. 5A-B). Of these correlated profiles, all of VLP 1 and VLP 2, 16/18 of the VLP 1 and LDNA and VLP 2 and

LDNA, 15/18 of the VLP 2 and HDNA, 14/18 of the LDNA and HDNA and 13/18 of the VLP 1 and HDNA subpopulation profiles were correlated (Table S4).

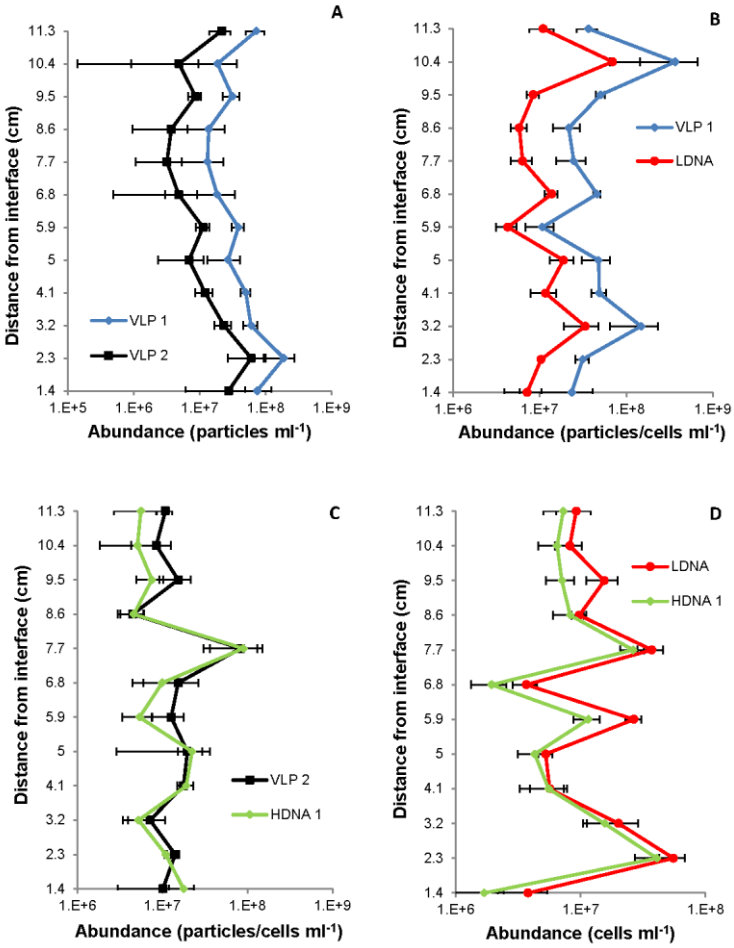


Figure 5. Single vertical profiles showing correlations between VLP and prokaryotic subpopulations. (A) VLP 1 and VLP 2, (B) VLP 1 and LDNA at the sediment-water interface ($r = 0.99$, $p < 0.0001$); and (C) VLP 2 and HDNA 1, (D) LDNA and HDNA 1 at the air-water interface ($r \geq 0.97$, $p < 0.0001$). Error bars represent 95% confidence intervals.

Spatial autocorrelation

- Moran's I and Geary's C

Significant spatial autocorrelation values were seen in 29% of the prokaryotic and VLP subpopulations. Significant Moran's I values were present in VLP 1 and VLP 2 at 1 out of 3 of the sediment-water interface environments. The range of Moran's I values for the VLP subpopulations was -0.07 to 0.05 with significance being seen at I values of 0.05. All prokaryotic subpopulations had non-significant Moran's I values with the range of I values being -0.02 to 0.07 (Table S5 and S6).

Significant Geary's C values were seen in all subpopulations at 1 out of 3 of the sediment-water interface environments, as well as HDNA at 1 out of 3 of the air-water interface environments, whilst their corresponding Moran's I values were non-significant. The range of Geary's C values for the prokaryotic subpopulations was 0.97 to 1.16 with significance seen at C values of ≥ 1.13, whilst the range of Geary's C values for the VLP subpopulations was 0.95 to 1.17 with significance seen at C values of ≥ 1.16 (Table S5 and S6).

- Moran and Geary Correlograms

Significant Moran correlograms were only found for VLP subpopulations at 1 out of 3 of the sediment-water interface environments (Fig. 6A-B). Significant Geary correlograms were seen in all subpopulations for 1 out of 3 of the sediment-water interface environments, as well as HDNA at 1 out of 3 air-water interface environments, whilst their corresponding Moran correlograms were non-significant (Fig. 7A-B). Moran correlograms showed a general trend of positive to negative spatial association whilst Geary correlograms showed alternation between positive and negative spatial autocorrelation with no clear pattern amongst all significant correlograms (Fig. 6 and 7).

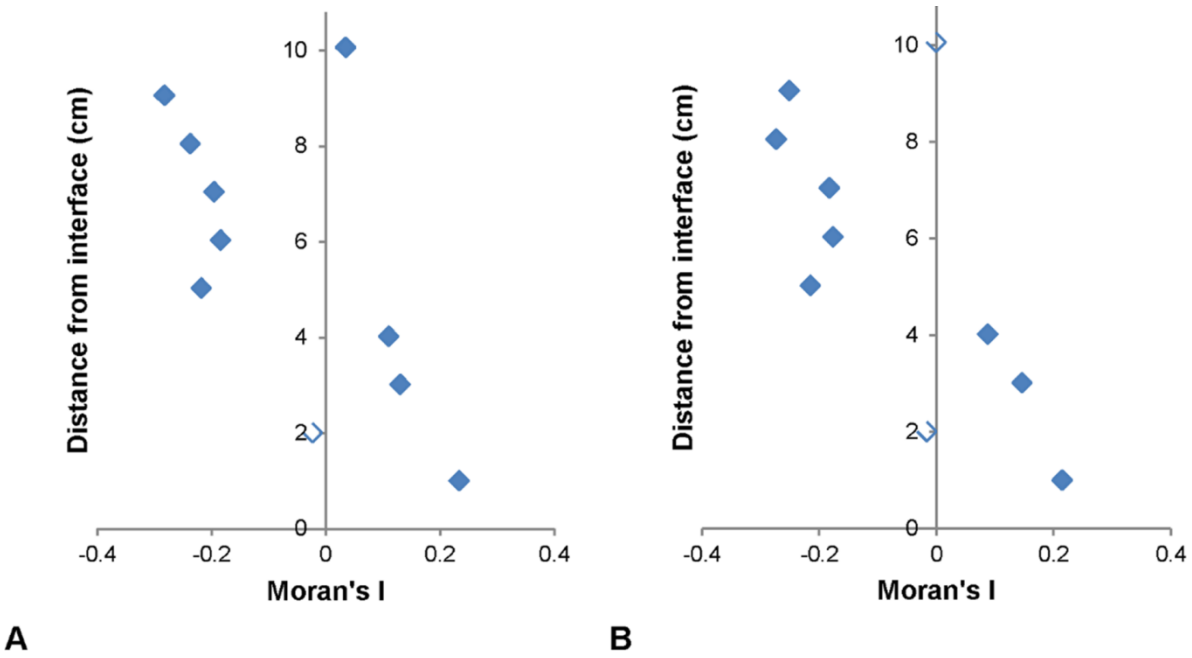


Figure 6. Significant Moran correlograms. (A) VLP 1 and (B) VLP 2 at the sediment-water interface. Unfilled data points indicate non-significance.

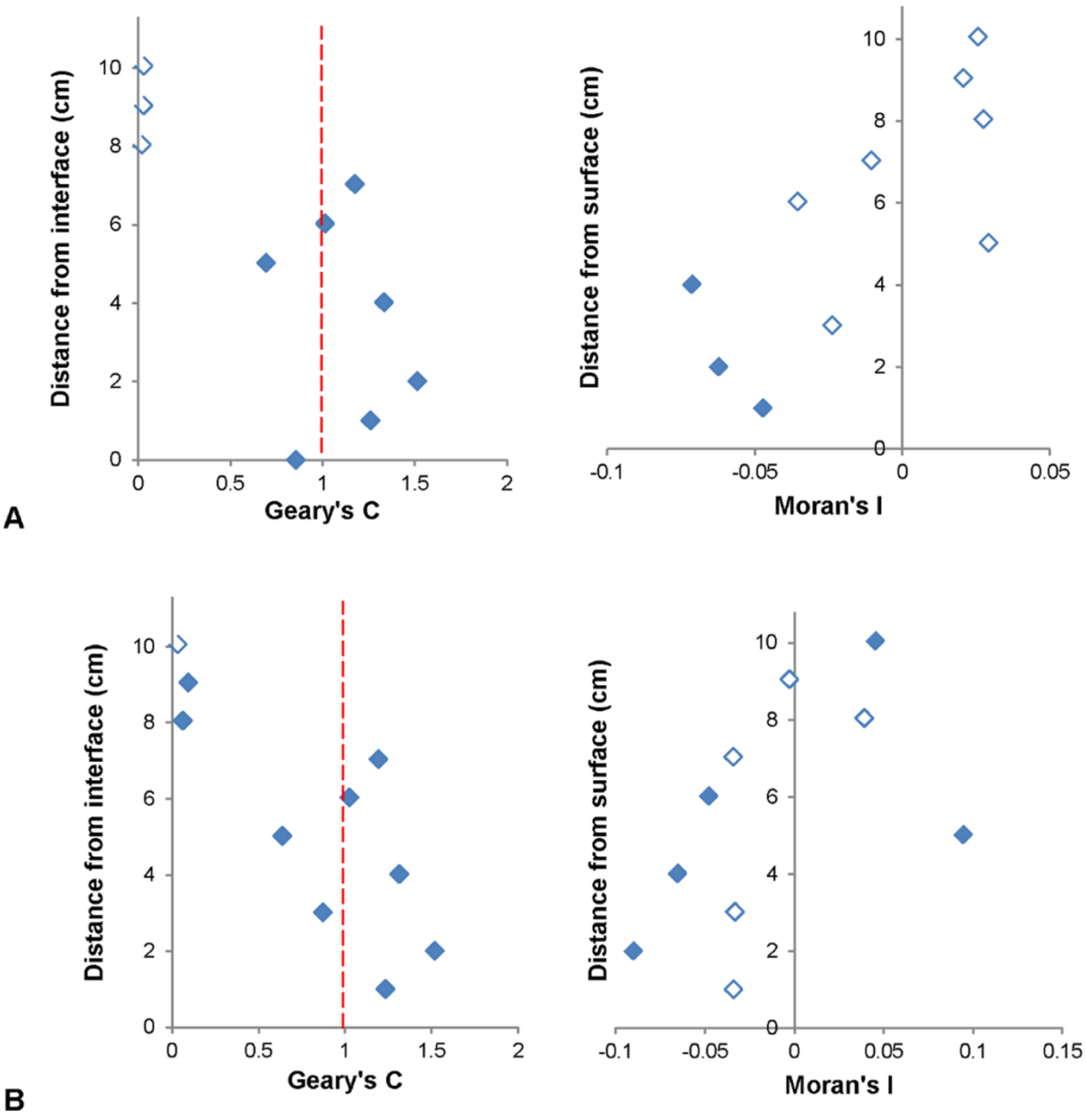


Figure 7. Significant Geary correlograms and corresponding non-significant Moran correlograms. (A) VLP 1 and (B) LDNA at the sediment-water interface. Unfilled data points indicate non-significance. Dashed red lines in Geary correlograms indicate point of spatial independence, with values above indicating negative spatial autocorrelation and values below indicating positive spatial autocorrelation.

DISCUSSION

Microscale microbial patchiness

It is now acknowledged, at least for marine systems, that important microbial processes occur at scales of micrometres to centimetres. Here we tested the hypothesis that microscale microbial communities within freshwater systems exhibit similar patchy abundance distributions as previously observed in marine systems (Blackburn *et al.* 1998; Dann *et al.* 2014; Long & Azam 2002; Margules & Pressey 2000; Seymour *et al.* 2000, 2006, 2008). Our findings showed variations in abundance of 107 and 80.5 fold over 0.9 cm in VLP and prokaryotic subpopulations, indicating heterogeneity consistent with marine systems where it is attributed to mixing and aggregation. However, the magnitude of the gradients was higher than the previously observed maxima of 45 fold variation over 0.9 cm reported for bacterioplankton in marine systems (Dann *et al.* 2014)(Azam 1998)7]. This potentially indicates that microscale patchiness within interface microenvironments may be higher in freshwater rather than marine systems and that microscale heterogeneity exists at scales that directly impact microbial interactions.

However, as flow and shear patterns are often unique within fluvial systems, the microbial patchiness observed may not be the same or similar between other freshwater systems. For instance, less mixing in lakes compared to flowing waters could potentially result in different levels of microscale patchiness, perhaps due to lake snow or more stable phytoplankton cell distributions. Also, as aggregates in riverine systems are smaller in size than those observed in lakes and marine environments, due to their exposure to contact shear force, this could impact on the level of heterogeneity observed in microbial communities (Besemer *et al.* 2005;

Griffith & Douglas 1990; Grossart & Simon 1993, 1998; Luef *et al.* 2007, 2009; Peduzzi & Luef 2008; Simon *et al.* 2002; Zimmermann-Timm 2002). Hydrological and seasonal patterns, such as algal blooms, leaf fall or terrestrial material inflow during flood events, will also impact microbial patchiness due to their observed effects on particle-attached microbial communities (Besemer *et al.* 2005; Luef *et al.* 2009).

Previously, studies of microbial distributions in freshwater systems often relied on bulk phase, large scale sampling where individual samples were separated by metre to kilometre scales (Jezbera *et al.* 2012; Lear *et al.* 2014; Liu *et al.* 2011; Palijan 2012; Pollard & Ducklow 2011). This sampling approach misses the important ecological associations that occur between individual microbial populations at micrometre to centimetre scales. In this study, the vast differences in microbial biomass from one sample point to the next would have been missed had bulk phase sampling been employed, hence influencing estimates in carbon flow. In addition, the results provide further evidence that nutrient exchanges may happen rapidly over very short distances (Azam 1998; Blackburn *et al.* 1998). Here, the ability to collect two-dimensional profiles at a high resolution has allowed discrimination between single point hotspots and adjacent low background values (Fig. 3). These single point hotspots and adjacent low background values were responsible for the large fold changes in this study, and show the utility of this two-dimensional sampling technique to provide enhanced scale resolution, as seen previously in marine systems (Dann *et al.* 2014). Our results suggest that where accurate abundance estimates are needed, large scale sampling may improve accuracy by including a few high resolution samples.

The steep adjacent gradients from one sample point to the next in this study illustrate the presence of microscale heterogeneity primarily due to the presence of hotspots. Hotspots in the microscale distributions of prokaryotes and VLPs are well accepted in marine systems, having been found in coral reefs (Seymour *et al.* 2005b), estuaries (Dann *et al.* 2014; Seymour *et al.* 2000, 2006), seawater aquarium (Blackburn *et al.* 1998), eutrophic coastal waters (Long & Azam 2001; Mitchell & Fuhrman 1989; Seymour *et al.* 2000, 2006, 2008) and oligotrophic open ocean systems (Seymour *et al.* 2006).

The hotspots observed may indicate microbial interactions with suspended particulate matter. For instance, river snow particles can contain high microbial abundances, with these abundances often exceeding the abundance in the water column (Luef *et al.* 2007). These particle associated microbial hotspots can contribute to a significant amount of production and activity (Grifith *et al.* 1990; Grossart & Simon 1993, 1998; Iriberri *et al.* 1987). Viral attachment to particulate matter can lead to viral abundance hotspots due to prolonged survival or increased phage production and transduction (Luef *et al.* 2009; Kapuscinski & Mitchell 1980; Kokjohn *et al.* 1991; Ripp & Miller 1995; Suttle & Chen 1992). This increased viral production on particles can also result in viral abundance hotspots within the free-living portion of the water column (Luef *et al.* 2009).

Nutrient patches may also explain the hotspots in freshwater prokaryote and VLP populations. Nutrient patches are small, being micrometres to centimetres in size and often short-lived, lasting seconds to minutes. These are generated by events such as algal lysis and the sinking of organic particles (Blackburn *et al.* 1998; Stocker & Seymour 2012; Taylor & Stocker 2012; Mitchell *et al.* 1985). As the sampling resolution was larger than the Batchelor scale, chemical gradients could

exist from one sample point to the next. These chemical gradients and nutrient patches attract chemotactic bacteria which form clusters around the high nutrient areas (Stocker & Seymour 2012; Stocker 2012). The bacterial accumulation favours viral infection due to high host density allowing viruses to infect multiple bacteria (Blackburn *et al.* 1998; Taylor & Stocker 2012). The abundance patterns that were observed suggested that bacteria had chemotactically responded to nutrient patches, leading to high bacterial abundance at some locations, followed by viral lysis of some of these bacterial species, and hence associated high viral abundance.

VLP and prokaryote subpopulations

The VLP 1 subpopulation was the most abundant which is consistent with previous studies (Chen *et al.* 2001; Thomas *et al.* 2011) (Fig. 1). Previous studies have also shown VLP 1 to contain bacteriophage (Marie *et al.* 1999; Goddard *et al.* 2005; Payet & Suttle 2008). However, recent work could not rule out the presence of algal and cyanobacterial viruses in the VLP 1 region (Parvathi *et al.* 2012) and additionally identified active lytic bacteriophage, specifically myoviruses, within VLP 2 FCM signatures (Zhong *et al.* 2014).

VLP to prokaryote ratio

The mean VPR found in this study was much lower than previous freshwater studies, with the most similar system being a eutrophic subtropical Australian river, which had a minimum VPR of 3.0 (Pollard & Ducklow 2011; Jacquet *et al.* 2005). However, this minimum VPR was seen during the summer, which is in contrast to this study which sampled during winter. As higher VPRs are typically seen in productive environments, which are suggested to be due to higher nutrient levels favouring maximum prokaryotic growth and productivity rates, the low VPRs obtained in this

study may indicate a less productive system (Jacquet *et al.* 2005, 2010; Maranger & Bird 1995; Meyer *et al.* 2014). The results confirm the mechanisms for large scale viral and bacterial interactions.

VLP and prokaryote relationships

The vertical profiles of 85% of the prokaryotic and VLP subpopulations were correlated and the VPR values at each sample area were consistent, implying a tight-coupling between prokaryotic and VLP communities (Fig. 5). Previous studies have shown positive interdependence between VLP abundance and prokaryotic numbers (Peduzzi & Schiemer 2004) and subpopulation correlations in marine systems for VLP 1 and LDNA and VLP 2 and HDNA leading to the belief VLP populations are the phage of the prokaryotic populations (Blackburn *et al.* 1998; Cochlan *et al.* 1993; Fuhrman 1999; Seymour *et al.* 2005a). The correlations indicate mutual succession of VLPs and prokaryotes, where cell growth is in equilibrium with cell lysis.

The strong correlations may be due to pre-lytic events where bacterial numbers are high and viral numbers are high as was seen in the abundance values (Breitbart 2012; Seymour *et al.* 2006). Also, as sampling occurred in winter when nutrient concentrations are minimal and bacterial productivity is at its lowest, lysis would be low and lysogenic viral activity would dominate, causing bacterial and viral numbers to remain relatively constant (Ortmann *et al.* 2002; Thomas *et al.* 2011). Correlations may also represent phytoplankton bloom demise, where bacterial numbers and productivity are high leading to high viral numbers that can last for days to weeks (Hennes & Simon 1995; Larsen *et al.* 2004; Weinbauer 2004; Yager *et al.* 2001). In

addition, correlations could indicate nutrient patches attracting bacteria and consequently viruses to high host density areas (Stocker 2012).

Microscale prokaryote and VLP distributions showed different hotspot and coldspot patterns. The two-dimensional contour plots at the air-water interface showed prokaryote hotspots but the absence of VLP hotspots 6 cm from the interface surface (Fig. 4). This indicates dynamic differences between prokaryotes and VLPs, perhaps representative of bacterial accumulation around a nutrient source prior to viral attack, or a suspended biofilm particle that is impenetrable to viruses.

Spatial autocorrelation within microbial subpopulations

Spatial autocorrelation analysis revealed 29% of the prokaryotic and VLP subpopulations were non-randomly distributed at 0.9 cm distances on the local and global scale. Previously significant Moran's I values have been observed in marine bacterioplankton, however not for VLP subpopulations (Dann *et al.* 2014). As significant Moran's I values were positive and significant Geary's C values were > 1, this indicates positive and negative spatial autocorrelation and hence regions of clustering and dispersion, which were present as hotspots and coldspots in two-dimensional abundance distributions (Moran 1950). Hotspots may indicate nutrient patches and subsequent bacterial accumulation and viral lysis, whilst coldspots could indicate low nutrient concentration regions where bacterial abundance and productivity is low and lysogeny is favoured (Azam 1998; Patten *et al.* 2006; Seymour *et al.* 2012; Taylor & Stocker 2012; Waters *et al.* 2003; Weinbauer & Suttle 1999).

Skewing was present in the results, with non-significant Moran's I values but significant Geary's C values found for VLP 1, VLP 2, LDNA and HDNA at 1 out of the

3 sediment-water interface environments and HDNA at 1 out of the 3 air-water interface environments (Fig. 7). This indicates, although these subpopulations were spatially dependent on the global scale, on the local scale they showed negative spatial autocorrelation areas (Sokal & Oden 1978). This skewing was due to extreme outlier values, i.e. hotspots, which were almost an order of magnitude higher than the rest of the dataset therefore indicates the importance of using Geary's C in conjunction with Moran's I.

The significant Moran's I values in this study and in Dann *et al.* (2014) and Waters *et al.* (2003) are much lower than the perfect clustering values of +1 indicated by Moran (1950). In Dann *et al.* (2014), who looked at microscale virio- and bacterioplankton distributions in marine habitats, significant Moran's I and Geary's C values ranged from 0.04 to 0.07 and 0.87 to 1.02 compared to 0.02 to 0.07 and 0.95 to 1.17 in this study. In addition, Waters *et al.* (Waters *et al.* 2003), who looked at phytoplankton distributions over 2 cm and 4 cm scales, showed Moran's I values between 0.08 and 0.18 were indicative of clustering. As previous uses of Moran's I related to large scale analyses, the lower Moran's I values obtained in this study could perhaps be characteristic of microscale microbial studies.

CONCLUSION

Here we report microscale patchiness in freshwater microbial communities with abundance variations of 107 and 80.5 fold over 0.9 cm for VLP and prokaryotic subpopulations. This indicates that within freshwater ecosystems microbial interactions are likely to differ markedly at the microscale. The pattern of variation is consistent with observations in marine systems for variation caused by mixing and

aggregation and therefore suggests, as with marine systems, that bulk phase sampling will not provide accurate representation of the dynamics of microbial processes within freshwater systems.

Spatial autocorrelation analysis showed VLP and prokaryotic subpopulation distributions were non-random and spatially dependent due to heterogeneous hotspots and coldspots. This indicates that the dissipation rates within slow flowing river systems allow nutrient patch formation with lifetimes that exceed bacterial chemotaxis rates.

SUPPLEMENTARY INFORMATION

Table S1. Mean VLP abundances at the air- and sediment-water interface. The 95% confidence intervals are included for each mean abundance.

Interface	Subpopulation	Abundance 10^7 x cells ml^{-1} (95%CI, n)
SWI*	VLP 1	5.8 (1.2, 108)
	VLP 2	1.9 (0.4, 108)
	VLP 1	6.7 (1.6, 108)
	VLP 2	2.1 (0.5, 108)
	VLP 1	5.9 (3.1, 108)
	VLP 2	1.7 (0.9, 108)
AWI*	VLP 1	5.4 (1.0, 108)
	VLP 2	1.7 (0.3, 108)
	VLP 1	6.1 (1.7, 108)

	VLP 2	1.8 (0.6, 108)
	VLP 1	5.9 (1.3, 108)
	VLP 2	1.8 (0.4, 108)

*AWI = air-water interface, SWI = sediment-water interface

Table S2. Mean prokaryotic abundances at the air- and sediment-water interface. The 95% confidence intervals are included for each mean abundance. Values indistinguishable from background noise were excluded.

Interface	**Subpopulation**	**Abundance** 10^7 x cells ml^{-1} (95%CI, n)
SWI*	LDNA	1.5 (0.3, 108)
	HDNA 1	0.9 (0.2, 108)
	LDNA	1.6 (0.3, 108)
	HDNA 1	1.0 (0.3, 106)
	LDNA	1.3 (0.5, 108)
	HDNA 1	1.0 (0.6, 108)
AWI*	LDNA	1.5 (0.2, 108)
	HDNA 1	1.5 (0.4, 108)
	LDNA	1.4 (0.3, 108)
	HDNA 1	1.0 (0.4, 108)
	LDNA	1.5 (0.2, 108)
	HDNA 1	1.2 (0.7, 105)

*AWI = air-water interface, SWI = sediment-water interface

Table S3. VLP to prokaryote ratio at the air- and sediment-water interface.

Interface	VPR
SWI*	3.2
	3.4
	3.3
AWI*	2.4
	3.3
	2.9

*AWI = air-water interface, SWI = sediment-water interface

Table S4. Corresponding r and p values for correlated prokaryotic and VLP subpopulations profiles.

Subpopulation profile	R value	P value
VLP 1 and VLP 2	≥ 0.88	≤ 0.0002
VLP 1 and LDNA	≥ 0.85	≤ 0.0005
VLP 2 and LDNA	≥ 0.81	≤ 0.0013
VLP 2 and HDNA	≥ 0.79	≤ 0.0024
LDNA and HDNA	≥ 0.83	≤ 0.0008
VLP 1 and HDNA	≥ 0.79	≤ 0.0025

Table S5. Moran's I and Geary's C values for prokaryotic subpopulations at the air- and sediment-water interface.

Interface	Subpopulation	Moran's I (p-value)	Geary's C (p-value)
SWI*	LDNA	0.02 (n.s)	0.98 (n.s)
	HDNA 1	0.02 (n.s)	0.97 (n.s)

	LDNA	-0.06 (n.s)	0.99 (n.s)
	HDNA 1	-0.06 (n.s)	0.97 (n.s)
	LDNA	-0.04 (n.s)	1.16 (0.0001)
	HDNA 1	-0.04 (n.s)	1.16 (0.0001)
AWI*	LDNA	-0.02 (n.s)	1.00 (n.s)
	HDNA 1	-0.04 (n.s)	0.99 (n.s)
	LDNA	-0.04 (n.s)	1.06 (n.s)
	HDNA 1	-0.04 (n.s)	1.05 (n.s)
	LDNA	-0.07 (n.s)	1.04 (n.s)
	HDNA 1	-0.05 (n.s)	1.13 (0.001)

*AWI = Air-water interface, SWI = Sediment-water interface.

Table S6. Moran's I and Geary's C values for VLP subpopulations at the air- and sediment-water interface.

Interface	**Subpopulation**	**Moran's I** (p-value)	**Geary's C** (p-value)
SWI	VLP 1	0.05 (0.05)	0.95 (n.s)
	VLP 2	0.05 (0.05)	0.96 (n.s)
	VLP 1	-0.07 (n.s)	1.00 (n.s)
	VLP 2	-0.07 (n.s)	1.02 (n.s)
	VLP 1	-0.04 (n.s)	1.17 (0.0001)
	VLP 2	-0.04 (n.s)	1.16 (0.0001)
AWI	VLP 1	-0.03 (n.s)	1.03 (n.s)
	VLP 2	-0.04 (n.s)	1.04 (n.s)
	VLP 1	-0.03 (n.s)	1.05 (n.s)
	VLP 2	-0.04 (n.s)	1.05 (n.s)

	VLP 1	-0.06 (n.s)	1.01 (n.s)
	VLP 2	-0.06 (n.s)	1.04 (n.s)

*AWI = Air-water interface, SWI = Sediment-water interface.

Chapter 5

Microscale Distributions of Planktonic Viruses and Prokaryotes Reflected by Patchiness of their Densities and Bacteria Taxonomic Composition

ABSTRACT

Patchiness in microscale microbial distributions is well established, often observed as abundance hotspots and coldspots. These hotspots and coldspots provide important microenvironments for microbial interactions. However, previous studies are often restricted to abundance estimates alone. Here, at the riverbed of the Murray River, we show 3 categories of microscale microbial heterogeneity: hotspots, coldspots and background regions that represent taxonomically distinct spatial partitions. We report hotspots characterised by discrete genera abundance increases, suggesting hotspots represent increases in particular bacterial groups rather than a general increase in all bacteria. Genera with increased abundances, *Pseudomonas, Parasporobacterium, Lachnospiraceae incertae sedis* and *Bacteroides,* were indicative of human and/or animal inputs and represented up to 14.7% average abundance. Each hotspot contained an abundance of different genera, which led to high dissimilarity among hotspots. Genera exclusivity was higher in background and coldspots, with 54 and 48 genera exclusive compared to 7 and 4 genera in hotspots. This suggests hotspots represent increases in persistent genera, rather than genera appearance. Hotspots were more similar to coldspots, indicating coldspots may represent dying hotspots. Sample category was a better indicator of taxonomic similarity than proximity, indicating hotspots and coldspots contain communities distinct from the background at the sub-centimetre scale. These findings indicate hotspots and coldspots represent distinct spatial taxonomic partition rather than temporary abundance increases or decreases of the overall community. This suggests 300 µl volumes persist together long enough to build up genera numbers and create distinctly different proximate spatial taxonomic partitions, which might have implications for microscale biogeochemical processes.

INTRODUCTION

Bacteria are important for ecosystem function, playing a crucial role in the cycling of carbon, phosphorus and nitrogen (Azam 1998; Long & Azam 2001). It is at the microscale that these nutrient exchanges occur (Azam & Malfatti 2007; Stocker 2008). Previously, microbial distributions were considered homogeneous or to follow patterns of random aggregation, leading to bulk phase sampling being considered representative of microbial microscale processes (Dann et al. 2014 – Chapter 3; Long & Azam 2001; Seymour et al. 2009). However, within aquatic environments, it is now well accepted that microbial abundance and activity differ by orders of magnitude over micrometre to centimetre scales (Azam 1998; Duarte & Vaque 1992; Long & Azam 2001; Mitchell & Fuhrman 1989; Seymour et al. 2004, 2005a, 2006), with previous studies showing 45 and 2584 fold cm^{-1} variation in prokaryotic and viral abundance (Dann et al. 2014 – Chapter 3). In addition, recent work showed this patchiness also exists in freshwater systems with river prokaryotic populations exhibiting up to an 80 fold change in abundance per 0.9 cm (Dann et al. 2016 – Chapter 4).

Patchiness in microscale microbial distributions is characterised typically by abundance hotspots or coldspots. Hotspots are regions of elevated bacterial abundance resulting from microbial accumulation around high nutrient areas via chemotaxis, particle aggregation or disintegration, small scale water mixing or grazing events in proximate areas. These high bacterial abundance regions have been suggested to result in increased viral production via lysis due to increased host density (Barbara & Mitchell 2003; Blackburn et al. 1998; Mitchell 2002; Seymour et al. 2000, 2005a, 2006).

Coldspots are believed to result from removal mechanisms, such as lysis or grazing events (Dann et al. 2014 – Chapter 3; Seymour et al. 2004, 2006; Waters et al. 2003). Whether grazing or lysis dominates as the main source of prokaryotic mortality will determine the fate of energy within a system (Tsai et al. 2013). For instance, grazing of prokaryotic cells reduces prokaryotic production by releasing energy to higher trophic levels, whilst lysis diverts energy from higher trophic levels where it is made available to other non-infected prokaryotes, with the suggestion that this stimulates prokaryotic production (Suttle 2005, 2007; Tsai et al. 2013; Weinbauer & Höfle 1998). Grazing can be taxonomy- or size-selective and lysis is typically favoured towards high density populations, therefore leading to coldspots in selected bacterial populations, which will assist the enrichment, and hence hotspot formation, of other non-selected bacteria (Chow et al. 2014; Thingstad 2000; Weinbauer et al. 2006). However, as previous studies of microscale prokaryotic distributions in freshwater systems have been limited to abundance estimates alone, whether hotspots and coldspots represent discrete spatial taxonomic partition, and hence whether heterogeneous species richness and composition exists in freshwater, have remained unknown.

Within marine systems, Long and Azam (2001) analysed 16s rRNA gene diversity and richness using microlitre samples of seawater to test for small-scale patchiness in bacterial species. Their results indicated microscale variation in the bacterial community richness of the ocean, yet no similar investigation exists for freshwater. Therefore, the aim of this study was to determine whether microscale variation in species richness and composition occurs in freshwater systems and confirm whether microscale patches represent discrete communities. We hypothesise that hotspots, coldspots and background represent discrete spatial taxonomic partitions where

hotspots are representative of discrete abundance increases of individual taxa. Here, we test this hypothesis by analysing the taxonomic composition of multiple hotspots, coldspots and background samples. From this, potential taxonomic patterns within these microenvironments can be determined.

MATERIALS AND METHODS

Sample collection

Samples were collected from the Murray River at Murray Bridge, South Australia (35°07'7"S, 139°16'56"E) on July 2nd, 2014. At the time of sampling the water temperature was 13.6˚C, pH was 7.8, total dissolved solids was 445 mg L^{-1} and electrical conductivity was 445 µS/cm and conductivity was 2.24 KΩ*cm. All environmental parameters were measured using a HydroLab DataSonde probe.

A two-dimensional sampling device (Dann et al. 2014 – Chapter 3) was employed at the sediment-water interface of the river bed enabling the collection of 8 vertical profiles with 0.9 cm separation and containing 12 sample points from 1.4 cm to 11.3 cm from the sediment-water interface. This 96-well microplate collected approximately 300 µl per sample well and sampling occurred at the sediment-water interface of the river benthos.

For taxonomic analysis, 100 µl samples were collected from each microplate well, aliquoted into cryovials, immediately snap frozen in liquid nitrogen and stored at -80°C until further analyses. For flow cytometric enumeration of prokaryotic and viral abundances, 200 µl samples were collected from each microplate well, immediately fixed with glutaraldehyde (0.5% final concentration) and stored at 4 ˚C in the dark for

15 minutes. Samples were snap frozen in liquid nitrogen and stored at -80˚C until required for further analyses. Samples were processed within one week of storage to avoid potential deterioration (Brussaard 2004).

Flow cytometry bacterial and viral abundance sample preparation, processing and analysis

Triplicate flow cytometry samples were prepared as described previously (Dann et al. 2014 – Chapter 3). Briefly, thawed samples were diluted 1:100 in Tris-EDTA buffer (0.2 µm filtered, pH 8.0, 10 mM Tris, 1 mM EDTA) and stained with SYBR Green I nucleic acid dye (1:20,000 final dilution; Molecular Probes). To optimise viral counts, samples were incubated at 80 ˚C in the dark for 10 minutes (Brussaard 2004). Each sample was run in triplicate to check method precision. Reference beads of 1 µm diameter (Molecular Probes) were added to act as an internal size and concentration standard with a final concentration of approximately 10^5 beads ml^{-1} per sample. Bead fluorescence and concentration were used to calibrate sample volume and target size (Brussaard 2004).

A FACSCanto II cytometer equipped with blue laser (488 nm, 20 mW, air-cooled) and a sheath fluid of phosphate-buffered saline (PBS) solution was used for sample analysis (Dann et al. 2014 – Chapter 3). Samples were run for 2 minutes on a low flow rate to obtain less than 1000 events per second. Green fluorescence (SYBR I), right-angle light scatter (SSC) and forward-angle light scatter (FSC) were acquired for each sample. Triplicate blank control samples containing filtered Tris-EDTA buffer (0.2 µl filtered) stained with SYBR Green I were prepared following the same sample preparation aforementioned and run during each flow cytometry session to eliminate potential background noise created via flow cytometer artifacts or sample preparation (Dann et al. 2014 – Chapter 3).

Flow cytometric histograms and cytograms were exported as FCS 3.0 files and analysed via FlowJo (Tree Star, Inc.) to enumerate the prokaryotic and viral populations present (Dann et al. 2016 – Chapter 4; Smith et al. 2015). Viral and prokaryotic populations were determined via their position in biparametric cytograms of SYBR Green fluorescence and side scatter (SSC) and the presence of discrete peaks in monoparametric histograms of SYBR Green fluorescence. Side-scatter and SYBR Green fluorescence were used as indicators of cell size and nucleic acid content (Brussaard 2004).

Flow cytometry data analysis and representation

Rank abundance graphs were used to determine hotspot, coldspot and background values (Dann et al. 2014 – Chapter 3, 2016 – Chapter 4). Background values were determined according to Weibe (1970) where the median value within the dataset was used as the background due to inclusion or exclusion of hotspots or coldspots not overly affecting the values obtained. Background values exhibited a linear trend and hence were indistinguishable from a random distribution, whilst the hotspots were identified as the sample points that exceeded this linear fit (Fig. S1). Lastly, the coldspot values were identified as sample points that fit a linear trend but exhibited gentler slopes than what were observed in the background values due to small differences between adjacent sample points (Fig. S1). From this, three of the highest abundance hotspots (H1, H2 and H3), lowest abundance coldspots (C1, C2 and C3) and median background samples (B1, B2 and B3) were selected for taxonomic analysis. For the spatial relationship of these hotspot, coldspot and background samples, refer to Figure 1C. Subpopulation correlations were determined via Pearson's coefficient. Two-dimensional contour plots were constructed using Surfer 10 (Golden Software, Inc.). When constructing these plots, a minimum contour

interval value ≥ 1000 events ml^{-1} was chosen as this was higher than the maximum flow cytometric error, i.e. the background noise within triplicate blank control samples. This value is conservative due to the maximum flow cytometric error being < 24 events ml^{-1}.

Quantitative PCR

The use of appropriate PCR reagents and sample volumes were trialled. To ensure sufficient prokaryotic amplification, different DNA sample volumes (1, 2, 3, 5, 10 and 23 µl) were analysed via quantitative PCR (qPCR). Briefly, 16s region specific forward 27F (5'-AGRGTTTGATCMTGGCTCAG -3') and reverse primers 519R (5'-GTNTTACNGCGGCKGCTG -3') were added to 1 µl DNA samples stained with Universal KAPA SYBR Fast qPCR Master Mix 2x (KAPA Biosystems) and run through 42 qPCR cycles on a Rotor-Gene. SYBR quantitation reports using Rotor-Gene Real-Time Analysis Software (6.0.27) showed amplification in the 1 µl DNA samples whilst the 2, 3, 5, 10 and 23 µl samples were inhibited perhaps due to the high level of tannins in the water samples (Kontanis & Reed 2006). To ensure 1 µl samples were representative, ten 1 µl replicates were taken for the highest hotspot, lowest coldspot and mean background, whilst five 1 µl replicates were taken for the remaining two samples, with taxonomic analysis performed on each replicate.

Direct PCR amplification of bacterial communities

Direct PCR amplification was performed on 1 µl samples in 25 µl PCR reactions that consisted of KAPA Taq Ready Mix 2x (KAPA Biosystems), and forward (27F) and reverse (519R) primers specific to the 16s sequence region. Samples were run on a Veriti 96 well Thermal Cycler (Applied Biosystems) for 42 cycles consisting of three stages. Of the resulting PCR products, 5 µl was run on an electrophoresis agarose

gel to ensure adequate DNA amplification and 20 µl was used for 16s rRNA sequencing. Negative controls were run with sterile filtered water in place of template. For sequencing, the variable 16s rRNA gene region PCR primers with forward primer barcodes were used in a 5 cycle PCR using HotStarTaq Plus Master Mix Kit (Qiagen, USA) with the following conditions: 94°C for 3 minutes, followed by 28 cycles at 94°C for 30 seconds, 53°C for 40 seconds, 72°C for 1 minute and a final elongation step at 72°C for 5 minutes. PCR products were run on a 2% agarose gel to determine amplification success and relative band intensity. Barcoded samples were pooled together in equal proportions based on molecular weight and DNA concentrations and then purified via calibrated Ampure XP beads. Sequencing was performed at Molecular Research (Shallowater, TX, USA) on an Illumina MiSeq following the manufacturer's guidelines.

Taxonomic analysis

Sequenced bacterial DNA was quality filtered and length truncated with reads < 250 bp or containing > 0.5% expected errors for all bases discarded. Sequences lacking a recognisable barcode or forward PCR primer were also discarded. The USEARCH pipeline was employed for full length dereplication and abundance sorting using a minimum size of 2 to ensure singletons removal (Edgar 2010). Removal of singletons ensured spurious OTUs resulting from PCR errors and/or sequencing artifacts were discarded (Edgar 2013; Liu et al. 2015). OTUs clustering at 97% identity was performed via the cluster_otus command using UPARSE in USEARCH v8 (Edgar 2013). Reference-based chimeric filtering was carried out using the gold database via UCHIME (Edgar et al. 2011; Pylro et al. 2014). Reads were then globally mapped to the OTUs using 97% identity threshold and taxonomic assignment was performed using RDP Classifier via the utax command in

USEARCH v8 (Edgar 2013; Wang et al. 2007). OTU tables were constructed using python scripts and the resulting OTU tabbed text file was used to determine average OTU abundances.

OTU tables were analysed in PRIMER (version 7) where square root overall transformed data was used for similarity percentage (SIMPER) analysis (Clarke & Gorley 2006). SIMPER was used to determine what is driving Bray-Curtis dissimilarity within and between sample values. Bray-Curtis resemblance was performed to construct dissimilarity matrices for PERMdisp, which was used to determine the level of dispersion between samples (Clarke 1993). Bray-Curtis resemblance of presence vs. absence was used to determine sample exclusivity of bacterial sequences. Metric multidimensional scaling (MDS) using Bootstrap Average analysis was performed to determine the level of spread between samples and produce smoothed 95% bootstrap regions for each sample type. Metric MDS ordination employed 500 bootstrap averages of the centroid of each sample to show where 95% of the centroid averages lie within multivariate space.

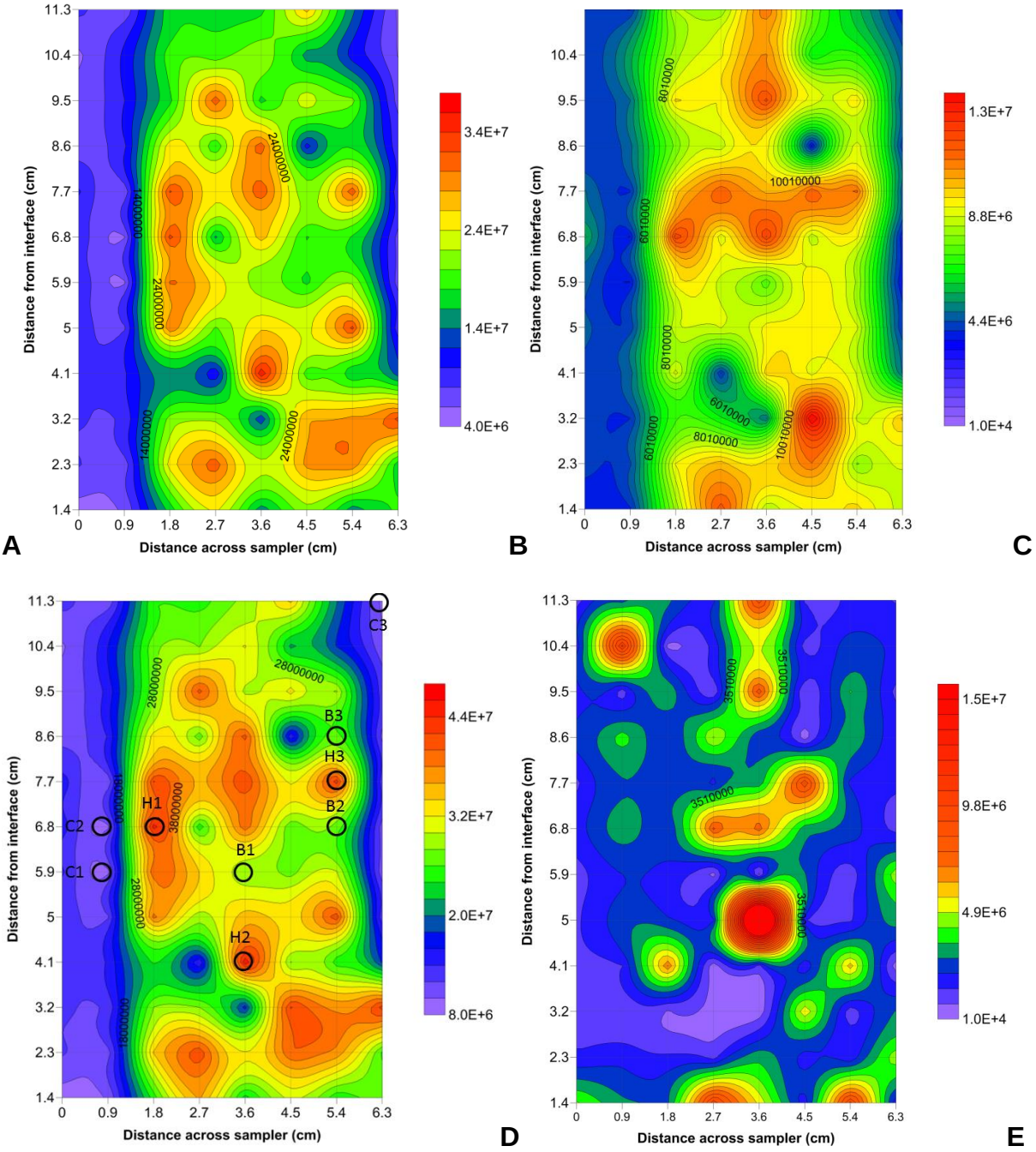
RESULTS

Prokaryotic and viral enumeration and two-dimensional distributions

Flow cytometric analysis revealed two viral (V1 and V2) and two prokaryotic subpopulations, Low DNA (LDNA) and high DNA (HDNA) via dense regions in biparametric cytograms of SYBR Green fluorescence and side-scatter (Fig. S2). The VLP and prokaryotic subpopulations are similar to those observed previously at Murray Bridge, South Australia (Dann et al. 2016 – Chapter 4). Viral abundances ranged from 2.5 x 10^5 to 2.4 x 10^7 particles ml^{-1} (n = 95, 95%CI = 6.4 x 10^5) for the V1 subpopulation, from 4.5 x 10^4 to 8.6 x 10^6 particles ml^{-1} (n = 95, 95%CI = 2.1 x

10^5) for the V2 subpopulation and from 2.9 x 10^5 to 3.3 x 10^7 particles ml^{-1} (n = 95, 95%CI = 5.6 x 10^5) for total viruses. Prokaryotic abundances ranged from 5.6 x 10^6 to 3.6 x 10^7 cells ml^{-1} (n = 95, 95%CI = 1.8 x 10^6) for the LDNA subpopulation, from 3.4 x 10^6 to 1.4 x 10^7 cells ml^{-1} (n = 95, 95%CI = 5.6 x 10^5) for the HDNA subpopulation and from 9.2 x 10^6 to 4.6 x 10^7 cells ml^{-1} (n = 95, 95%CI = 2.1 x 10^5) for total prokayotes.

The two-dimensional distributions of prokaryotes and viruses were characterised by hotspots and coldspots in abundance (Fig. 1 and S2). V1 had a maximum hotspot of 2.4 x 10^7 particles ml^{-1} (n = 3, 95%CI = 4.3 x 10^7) and a minimum coldspot of 2.5 x 10^5 particles ml^{-1} (n = 3, 95%CI = 5.0 x 10^4) resulting in an overall 97.5-fold change in abundance over the sampling area. V2 had a maximum hotspot of 8.8 x 10^6 particles ml^{-1} (n = 3, 95%CI = 1.5 x 10^7) and a minimum coldspot of 4.5 x 10^4 particles ml^{-1} (n = 3, 95%CI = 5.6 x 10^4) resulting in an overall 195.4-fold change in abundance over the sampling area. LDNA had a maximum hotspot of 3.6 x 10^7 cells ml^{-1} (n = 3, 95%CI = 5.6 x 10^6) and a minimum coldspot of 3.7 x 10^6 cells ml^{-1} (n = 3, 95%CI = 3.5 x 10^6) resulting in a 9.7-fold change in abundance over the sampling area. HDNA had a maximum hotspot of 1.4 x 10^7 cells ml^{-1} (n = 3, 95%CI = 1.8 x 10^6) and a minimum coldspot of 2.7 x 10^6 cells ml^{-1} (n = 3, 95%CI = 2.5 x 10^6) resulting in a 5.2-fold change in abundance over the sampling area. From one sampling well to the next, the largest change in abundance for VLP 1 was 84.2-fold, V2 was 115-fold, LDNA was 5.9-fold, HDNA was 3.7-fold, total viruses was 90.6-fold and total prokaryotes was 4.5-fold per 0.9 cm. In addition, vertical gradients were present within the prokaryotic subpopulations with low abundance values in one of the bordering vertical profiles within the LDNA, HDNA and total prokaryotes (Fig. 1).


Distance from interface (cm)
Distance across sampler (cm)
3.4E+7
2.4E+7
1.4E+7
4.0E+6
14000000
24000000
A
8010000
10010000
6010000
1.3E+7
8.8E+6
4.4E+6
1.0E+4
B
C
C3
B3
H3
B2
H1
C2
C1
B1
H2
28000000
38000000
18000000
4.4E+7
3.2E+7
2.0E+7
8.0E+6
D
3510000
1.5E+7
9.8E+6
4.9E+6
1.0E+4
E

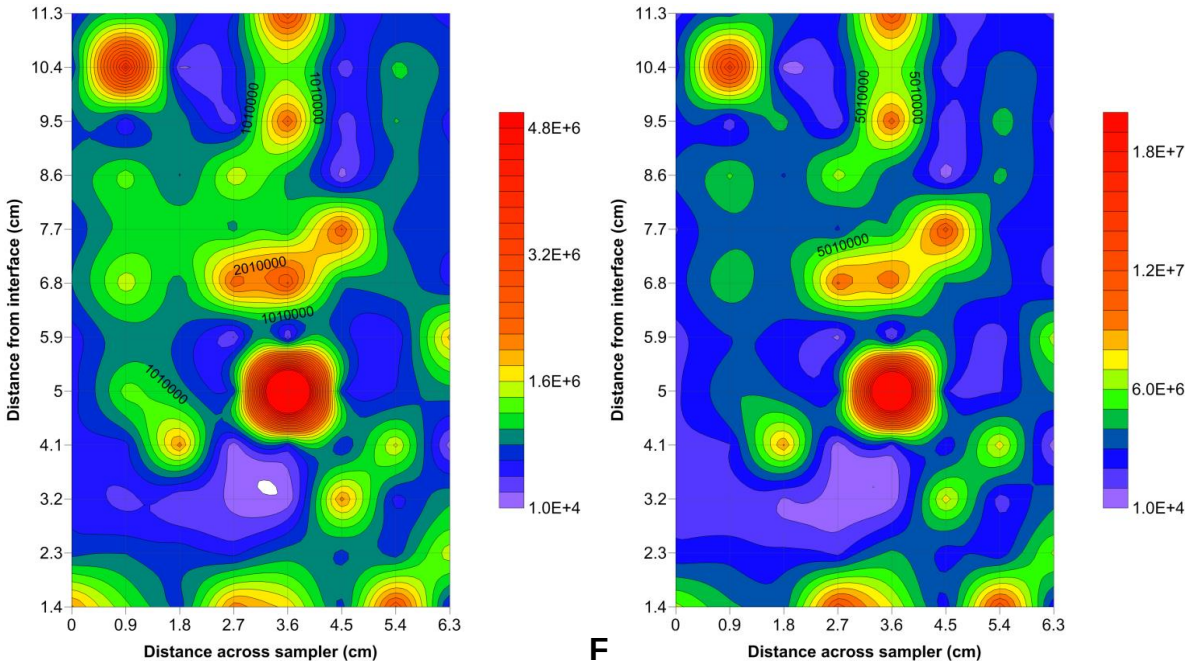


Figure 1. Two-dimensional contour plots showing the presence of hotspots, coldspots and background in **A** LDNA, **B** HDNA, **C** Total prokaryotes, **D** V1, **E** V2 and **F** Total viruses. Faint gridlines indicate sampling interval. Transects separated by 0.9 cm. Colour intensity scale in cells ml^{-1}. Circles in **C** indicate sample location and corresponding label.

Correlations

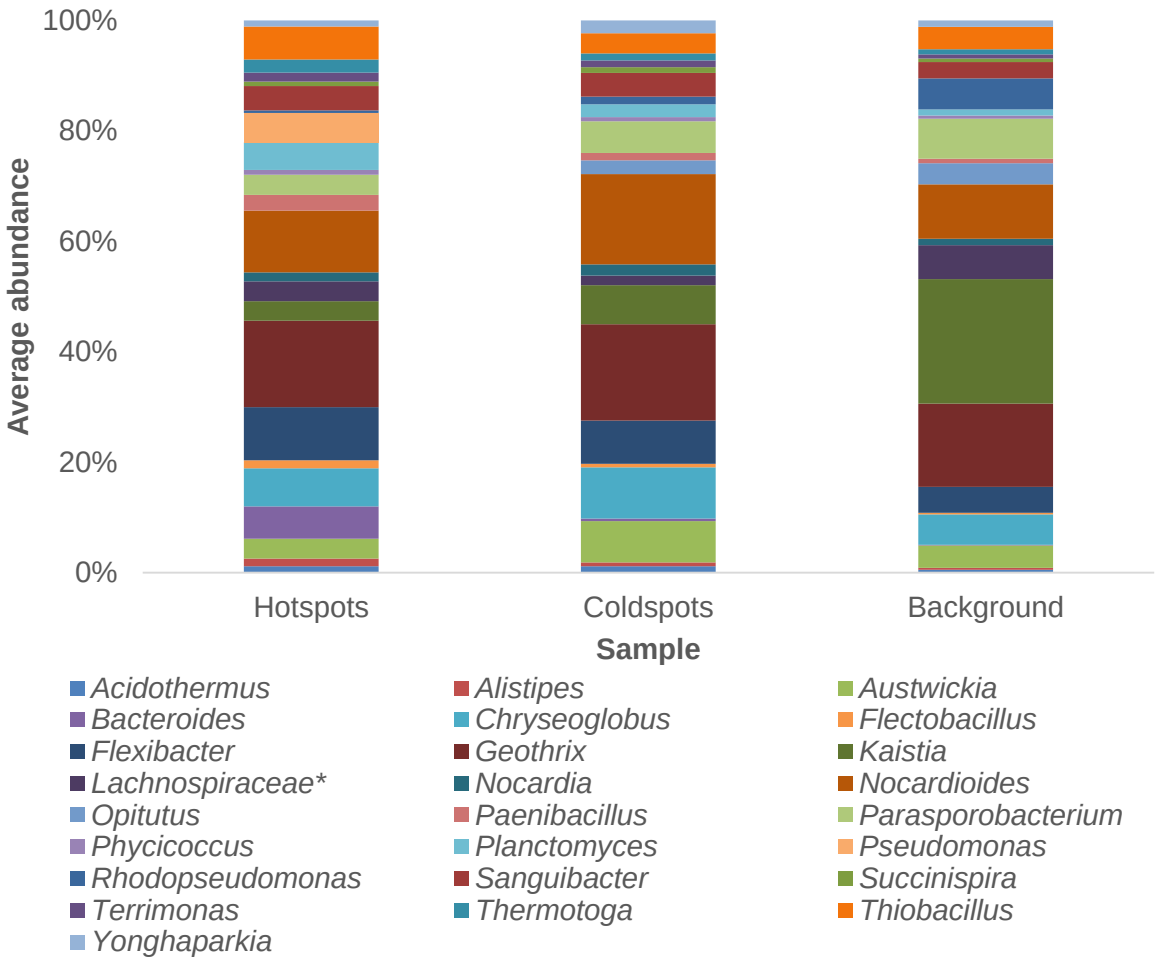
The single vertical profiles of V1 and V2 were positively correlated to one another in all instances (r^2 = ≥ 0.94, n = 12, p < 0.0001). Whilst five of the single vertical profiles of LDNA and HDNA were positively correlated (r^2 = ≥ 0.63, n = 12, p ≤ 0.03). Two single vertical profiles of V1 were negatively correlated with LDNA (r^2 = ≤ -0.6, n = 12, p ≤ 0.04) and two were positively correlated with HDNA (r^2 = ≥ 0.59, n = 12, p ≤ 0.04). V2 was negatively correlated with LDNA and HDNA in 1 vertical profile (r^2 = ≤ -0.6, n = 12, p ≤ 0.04). One of the single vertical profiles of total viruses and total prokaryotes were negatively correlated (r^2 = -0.6, n = 12, p =0.04).

Taxonomic profiles

- Hotspots

Phylogenetic average abundance analysis of the hotspots showed a high occurrence of Proteobacteria 28.9%, Actinobacteria, 21.9%, Bacteroidetes 20.0% and Acidobacteria, 10.1% (Fig. S3, Table S1). Also present were Firmicutes 7.8%, Planctomycetes 4.8%, Verrucomicrobia 4.0% and Thermotogae 1.5% (Fig. S3, Table S1). Within the averaged taxonomic profiles, phyla dominance was less obvious in the hotspots, as Proteobacteria had an average abundance of 7.0% and 8.9% above the second and third most abundant phyla, Actinobacteria and Bacteroidetes (Fig. S3A, Table S1).

At the genus level, hotspots contained an abundance of *Geothrix* 12.7%, *Nocardioides* 9.1%, *Flexibacter* 7.8%, *Chryseoglobus* 5.6%, *Thiobacillus* 4.9%, *Bacteroides* 4.8%, *Pseudomonas* 4.3% and *Planctomyces* 4.1% (Fig. 2, Table S2). SIMPER analysis of the hotspots at the genus level revealed a similarity of 55.4, with *Geothrix*, *Nocardioides* and *Chryseoglobus* the main contributors to similarity, contributing to > 5% similarity. PERMdisp analysis showed an average dispersion of 30.3 (SE = 2.7).


100%
80%
60%
40%
20%
0%
Average abundance
Hotspots
Coldspots
Background
Sample
Acidothermus
Alistipes
Austwickia
Bacteroides
Chryseoglobus
Flectobacillus
Flexibacter
Geothrix
Kaistia
Lachnospiraceae*
Nocardia
Nocardioides
Opitutus
Paenibacillus
Parasporobacterium
Phycicoccus
Planctomyces
Pseudomonas
Rhodopseudomonas
Sanguibacter
Succinispira
Terrimonas
Thermotoga
Thiobacillus
Yonghaparkia

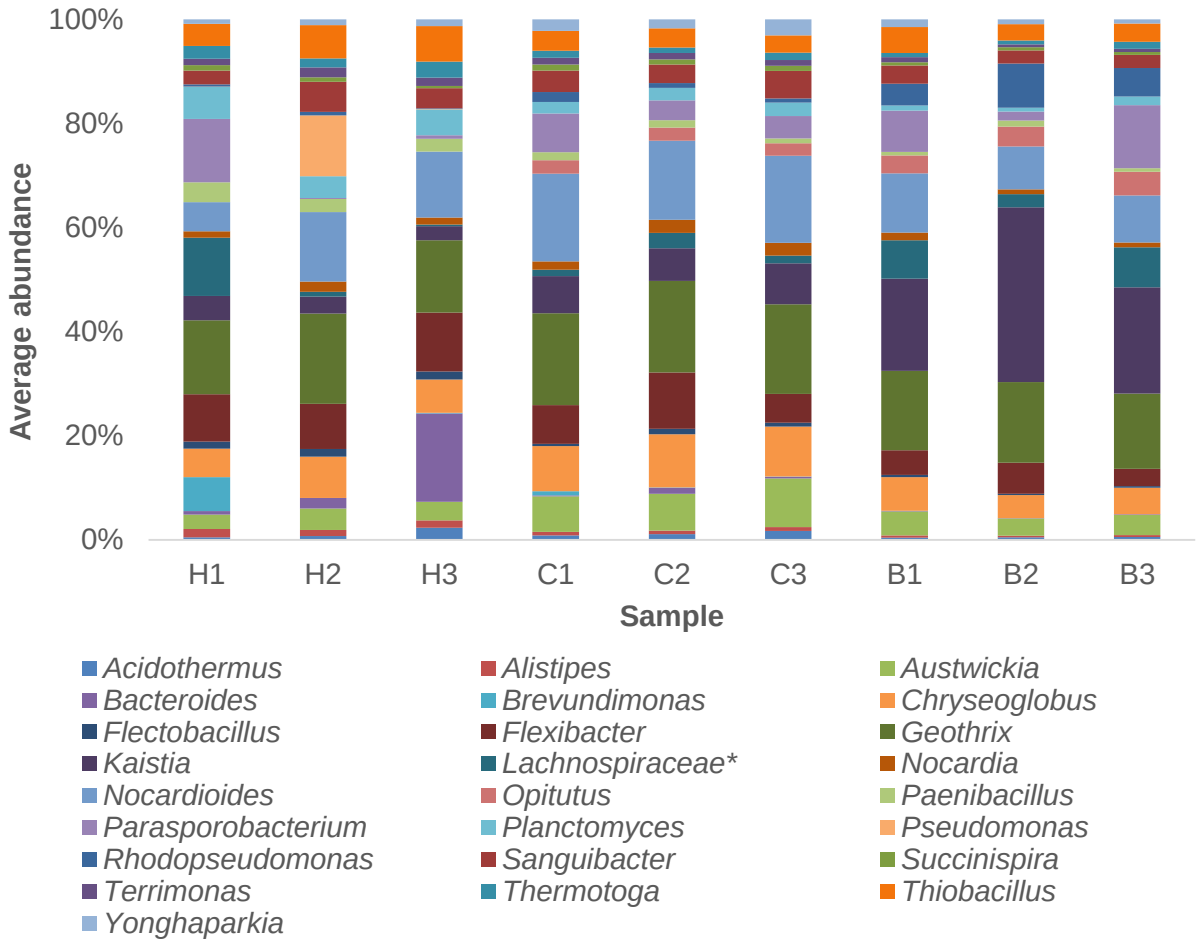


Figure 2. A Averaged taxonomic profiles of hotspots (n = 20), coldspots (n = 20) and background (n = 20) samples at the genus level. **B Figure 4.** Taxonomic profiles at the genus level of the individual hotspot, coldspot and background samples.

For clarity, genera representing < 2% and < 6% total average abundance were omitted. * *incertae sedis*.

Phylogenetic profiles of individual hotspots showed a higher average abundance of Proteobacteria 33.0% in H2, Firmicutes 20.2% in H1 and Bacteroidetes 29.2% in H3 (Fig. S4). This abundance of Proteobacteria in H2 was primarily due to *Pseudomonas* 9.4%, Firmicutes in H1 was primarily due to *Parasporobacterium* 10.2% and *Lachnospiraceae incertae sedis* 9.4% and Bacteroidetes in H3 was

primarily due to *Bacteroides* 14.2% (Fig. 4, Table S3). At the genus level, SIMPER analysis of the individual samples revealed similarities of 61.2 for H3, 53.3 for H2 and 57.5 for H1. Between sample positions there was lower dissimilarity between H1 and H3 than H3 and H2 and H1 and H2, with dissimilarities of 42.9, 43.4 and 46.5, respectively. The main contributor to dissimilarity between the hotspot samples were *Bacteroides* and *Parasporobacterium*, which contributed to > 3% dissimilarity. SIMPER analysis of presence vs. absence revealed an average similarity of 56.5. Between H1 and H3, 35 genera were exclusive with 25 exclusive to H1 and 10 exclusive to H3 (Table S4). Between H3 and H2, 15 genera were exclusive with 9 exclusive to H3 and 6 exclusive to H2 (Table S5). Between H1 and H2, 24 genera were exclusive with 16 exclusive to H1 and 8 exclusive to H2 (Table S6).

- Coldspots

Phylogenetic average abundance analysis of the coldspots, showed an abundance of Actinobacteria 37.0%, Proteobacteria 19.3%, Acidobacteria 13.8% and Bacteroidetes 13.8% (Fig. S3, Table S1). Also present were Firmicutes 9.6%, Plantomycetes 2.6%, Verrucomicrobia 2.11% and Thermotogae 1.0% (Fig. S3, Table S1). The most dominant phyla in the coldspots, Actinobacteria, represented 17.7% more than the second most abundant phyla, respectively (Fig. S3A).

At the genus level, coldspots contained an abundance of *Geothrix* 15.7%, *Nocardioides* 14.7%, *Chryseoglobus* 8.3%, *Flexibacter* 7.0%, *Austwickia* 6.7%, *Kaistia* 6.3% and *Parasporobacterium* 5.1% (Fig. 2, Table S2). SIMPER analysis at the genus level revealed a similarity of 69.8, with *Nocardioides* and *Geothrix* the main contributors to similarity, contributing to > 5% similarity. PERMdisp analysis showed an average dispersion of 20.7 (SE = 1.1).

At the genus level, SIMPER analysis of the individual samples revealed similarities of 70.3 for C2, 71.6 for C1 and 66.0 for C3, with a lower dissimilarity between C1 and C2 than C1 and C3 and C2 and C3, with dissimilarities of 29.2, 31.4 and 31.7 respectively. *Parasporobacterium* was the main contributor to dissimilarity between coldspot samples, accounting for ≥ 1.6% dissimilarity. SIMPER analysis of presence vs. absence revealed an average similarity of 75.2.

- Background

Phylogenetic average abundance analysis of the background values showed an abundance of Proteobacteria 37.2%, Actinobacteria 22.6%, Firmicutes 13.3% and Acidobacteria 12.2% (Fig. S3, Table S1). Also present were Bacteroidetes 8.7%, Verrucomicrobia 3.3% and Plantomycetes 1.2% (Fig. S3). The most abundant phyla, Proteobacteria, represented 14.6% more than the next abundant phyla (Fig. S3A).

At the genus level, background samples had an abundance of *Kaistia* 19.9%, *Geothrix* 13.2%, *Nocardioides* 8.7%, *Parasporobacterium* 6.4%, *Lachnospiraceae incertae sedis* 5.4%, *Chryseoglobus* 4.9% and *Flexibacter* 4.1% (Fig. 2, Table S2). SIMPER analysis at the genus level, revealed a similarity of 74.3, with *Kaistia* and *Geothrix* the main contributors to similarity, accounting for 6.2% and 5.3% similarity. PERMdisp analysis at the genus level revealed an average dispersion of 17.7 (SE = 0.4).

At the genus level, SIMPER analysis of the individual samples revealed similarities of 76.2 for B1, 74.6 for B2 and 74.1 for B3, with a lower dissimilarity between B1 and B3 than B2 and B3 and B2 and B1, with dissimilarities of 25.5, 26.6 and 27.2 respectively. *Lachnospiracea incertae sedis* was the main contributor to dissimilarity between the background samples, accounting for ≥ 3.7% dissimilarity. SIMPER

analysis of presence vs. absence revealed an average similarity of 77.3. In all instances, hotspots, coldspots and background, the most abundant Proteobacteria was Alpha-, followed by Betaproteobacteria (Fig. S4). No archaea were identified in the taxonomic profiles. Sequence quality filtering, chimera detection, read lengths and OTU classification results provided in supplementary information.

- Sample type comparisons

For the hotspots and coldspots, SIMPER analysis at the genus level revealed an average dissimilarity of 47.2, with *Parasporobacterium* and *Opitutus* contributing to > 1.7% of the dissimilarity. SIMPER analysis of presence vs. absence revealed an average dissimilarity of 55.4 with 52 genera exclusive of which 48 were exclusive to the coldspots and 4 were exclusive to the hotspots (Table S7). For the hotspots and background, SIMPER analysis at the genus level revealed an average dissimilarity of 51.0, with *Kaistia*, *Thiobacter* and *Thiobacillus* contributing > 2.2% of this dissimilarity. SIMPER analysis of presence vs. absence revealed an average dissimilarity of 56.9 with 61 genera exclusive of which 54 were exclusive to background and 7 were exclusive to the hotspots (Table S8). For the coldspots and background, SIMPER analysis at the genus level revealed an average dissimilarity of 39.6, with *Kaistia*, *Thiobacillus* and *Thiobacter* contributing to > 2.5% of this dissimilarity. SIMPER analysis of presence vs. absence revealed an average dissimilarity of 36.7 with 74 genera exclusive of which 39 were exclusive to coldspots and 35 were exclusive to background (Table S9). Metric MDS analysis showed a difference between sample types with higher spread between hotspots rather than coldspots and background (Fig. 3).

Figure. 3 Metric multidimensional scaling (MDS) analysis of sample type, hotspot, coldspot or background.

- Proximate samples

Proximate samples were compared to determine whether a spatial component was present within the taxonomic profiles. At the genus level, SIMPER analysis of proximate coldspots and hotspots revealed an average dissimilarity of 29.2 between sample C2 and sample C1, 47.2 between sample C2 and sample H1 and 46.1 between sample C1 and sample H1. *Parasporobacterium* contributed to > 2.5% dissimilarity between the coldspot samples and Parasporobacterium, *Lachnospiracea incertae sedis* and *Nocardiodes* contributed to > 1.9% dissimilarity between coldspot and hotspot samples. SIMPER analysis of presence vs absence data showed a dissimilarity of 24.4 between sample C1 and sample C2, 53.3 between sample C1 and sample H1 and 54.3 between sample C2 and sample H1. Between sample C1 and sample H1, 94 genera were exclusive, with 90 found only in sample C1 and 4 found only in sample H1 (Table S10). Between sample C2 and sample H1, 103 genera were exclusive, with 98 found only in sample C2 and 5 found only in sample H1 (Table S11). Between sample C1 and sample C2, 17 genera were exclusive, with 10 exclusive to sample C1 and 7 exclusive to sample C2 (Table S12).

At the genus level, SIMPER analysis of the proximate background and hotspot samples showed an average dissimilarity of 50.0 between sample B2 and sample H3, 50.9 between sample B3 and sample H3, and 27.17 between sample B2 and sample B1. *Lachnospiraceae incertae sedis* and *Kaistia* accounted for ≥ 3.6% dissimilarity between the background and *Kaistia* and *Bacteroides* contributed to ≥ 3.4% dissimilarity between background and hotspot samples.

SIMPER analysis of presence vs absence data showed a dissimilarity of 53.7 between sample H3 and sample B2 and 81 genera were exclusive, with 72 found only in sample B2 (Table S13). There was a dissimilarity of 56.2 between sample H3 and sample B3 and 115 genera were exclusive, with 106 found only in sample B3 (Table S14). Lastly, there was a dissimilarity of 24.9 between sample B1 and sample B2 and 18 genera were exclusive, with all 18 found only in sample B1 (Table S15).

DISCUSSION

Microscale viral and prokaryotic distributions – hotspots, coldspots and surface gradients

Hotspots and coldspots were present in prokaryotic and viral microscale distributions near to the sediment water interface of the river bed (Fig. 1). These patchy distributions in prokaryotic and viral abundance appear ubiquitous, having been found in a number of studies previously (Dann et al. 2014 – Chapter 3; Gharasoo et al. 2014; Seymour et al. 2000, 2004, 2005a, 2005b, 2006, 2007; Stocker 2008). In addition to hotspots and coldspots, surface abundance gradients were present within the prokaryotic subpopulations (Fig. 1). Unlike previous observations, these gradients vertical to the sediment-water interface were characterised by lowered rather than increased abundance (Dann et al. 2014 – Chapter 3; Seymour et al. 2007). Increased abundance gradients relate to organic matter sinking and its incorporation into the benthos allowing its degradation and transformation by bacteria, therefore resulting in high nutrient concentrations at the sediment-water interface and hence high microbial abundance (Stocker & Seymour 2012). The lowered abundance in these observed gradients implies a gradient of a different

mechanism, perhaps extensive lysis via phage, or non-selective grazing via protists, with the latter being favoured due to a lack of correlation between viral and prokaryotic abundances within this gradient (Weinbauer & Höfle 1998).

A maximum abundance variation of 115-fold over 0.9 cm was observed in V2, this is larger than a maxima of 107-fold over 0.9 cm observed recently in the same river environment, therefore indicating viral spatial heterogeneity within rivers is higher than was assumed (Dann et al. 2016). A maximum 5.9-fold change per 0.9 cm in abundance observed in prokaryotic subpopulations is much lower than previous reported maxima of 80.5-fold per 0.9 cm (Dann et al. 2016 – Chapter 4), perhaps due to the surface gradients within the prokaryotic microscale distributions as these would lead to a more gradual change in abundance over the sampling area.

Viral subpopulations were positively correlated with one another in all instances and the majority of prokaryotic subpopulations were positively correlated to each other (Fig. 1). In contrast, the prokaryotic and viral subpopulations showed minimal correlations, and when correlated, the relationship was negative, which is suggestive of viral lysis leading to increased progeny and reduced prokaryote hosts. Prokaryotic subpopulation correlations could relate to stimuli responses, such as to particulate organic matter, nutrient gradients or hydrological conditions, leading to the formation of high abundance areas (Stocker et al. 2008). Whilst viral correlations could relate to lysis events where viral progeny are released and their hosts are destroyed (Seymour et al. 2006). The general lack of strong negative and positive correlations between viral and prokaryotic subpopulations may be due to temporal mismatch of the populations. However, it also suggests that grazers may be the dominant source of prokaryotic mortality, therefore releasing energy to higher trophic levels, and hence being the prevailing removal process for coldspot formation (Tsai et al. 2013;

Weinbauer & Höfle 1998). In addition, observed viral subpopulations may have eukaryotic hosts, as has been found in previous work by Parvathi et al. (2012), with the possible presence of algal viruses in the V1 region that do not respond to prokaryote population fluctuations. The lack of correlations might also indicate the ephemerality of heterogeneity where the diffusivity of viral progeny is low in comparison to potential bacterial motility rates, leading to high viral abundances in areas of past nutrient patches and prokaryotic hotspots (Seymour et al. 2006).

The viral concentrations were low compared to **a previous study conducted in this environment (Dann et al. 2016 – Chapter 4), despite both studies sampling during winter. However, during** 2013 and 2014, the delivery of environmental water via flows and releases from locations upstream, resulted in higher flows to the South Australian portion of the Murray River sampled here (DEWNR 2014; DPI 2014).This potentially indicates that within this highly regulated river, viral abundance may be more dependent on annual river flow dynamics rather than typical seasonal variations typically observed in other freshwater studies (Maheshwari *et al.* 1995; Walker & Thoms 1993).***Hotspots and coldspots as distinct spatial taxonomic partitions***

Since early studies showing patchiness within plankton distributions (Bainbridge 1957; Bérzins 1958; Birge 1897; Langford 1938; Malone & McQueen 1983; Riley 1963; Sutcliffe et al. 1963), to recent microscale work confirming patchiness in bacterial, viral and phytoplankton communities (Dann et al. 2014 – Chapter 3; Doubell et al. 2006; Seuront et al. 2007; Seymour et al. 2004, 2005ab, 2006, 2007, 2008; Waters et al. 2003), whether this patchiness represented temporary abundance increases or discrete communities remained unknown.

Here, we confirm microscale microbial hotspots, coldspots and background regions represent distinct spatial taxonomic partitions (Fig. 3.). Hotspots were characterised by discrete genera abundance increases, specifically *Parasporobacterium*, *Lachnospiraceae incertae sedis*, *Pseudomonas* and *Bacteroides* (Fig. 4). Genera abundance increases differed between hotspot samples, leading to high dissimilarity to coldspots and background as well as other hotspot samples. Increases in specific genus abundance indicates exploitation of favourable environmental conditions, for instance fast moving motile bacteria taking advantage of ephemeral nutrient patches, which other bacterial species may not have the speed to encounter in time (Seymour et al. 2006; Stocker et al. 2008). Success in competition could be coupled with the decline in other species, which would be likely due to selective grazing by protists or lysis by the phage of these non-successive species. Stocker et al. (2008) showed that some species of bacteria are capable of better exploiting ephemeral nutrient patches due to their motility rates, referred to as “opportunitrophs”. This exploitation is possible for *Pseudomonas,* as well as some species of *Bacteroides,* as they are capable of motility (McBride & Zhu 2013; Shrout et al. 2006).

In addition, specific genus abundance increases may represent microbial associations with suspended particulate matter. For instance, within riverine systems, organic and inorganic suspended particles produced via auto- and allochthonous sources can harbour a range of microbial communities (Kernegger et al. 2009). Specifically, Bacteroidetes have been shown to colonise and dominate large particles due to their ability to grow whilst attached to particles (Fernández-Gómez et al. 2013). Bacteroidetes have been shown to contain many adaptations for particle attachment, such as adhesion proteins and glycoside hydrolase and peptidase enzymes, which are involved in polysaccharide degradation (Fernández-

Gómez et al. 2013). These particle-associated microbial communities contribute to a high proportion of activity and production (Griffith et al. 1990; Grossart & Simon 1993, 1998; Iriberri et al. 1987). Abundance increases of *Pseudomonas* may suggest microenvironments where this genus is able to out compete other genera for a nutrient source, hence leading to its increased abundance. Pseudomonads are strict aerobic, motile bacteria which have the capacity to live anaerobically when in biofilms, such as has been observed in *Pseudomonas aeruginosa* (Hassett et al. 2002). *Pseudomonas* can act as opportunistic pathogens to plants, fish or humans, or can be beneficial to their host, sometimes being found to promote plant growth. *Pseudomonas* has also been identified as a symbiont of freshwater sponges (Keller-Costa et al. 2014). Therefore, the abundance increases of *Pseudomonas* could indicate either biofilm aggregates or the exploitation of nutrient patches due to its motility.

Bacteroides are anaerobic and constitute the majority of mammalian gastrointestinal flora, playing an important role in the intestines. For this reason, *Bacteroides* have been used as faecal indicators in river systems hence indicating the presence of mammalian faecal-associated particulate matter (Okabe et al. 2007). As the Murray River is impacted by human activity, the presence of mammalian faecal bacteria is expected.

The specific genera abundance increases in the hotspots, *Parasporobacterium* and *Lachnospiraceae incertae sedis,* are members of the Firmicutes, within the class Clostridia. *Parasporobacterium* are anaerobic bacteria typically found in freshwater sediments, and certain species, such as *Parasporobacterium paucivorans* are responsible for methanethiol and dimethyl sulphide formation via sulphide methylation (Lomans et al. 2001). Therefore, this increase in *Parasporobacterium*

abundance could indicate sediment biofilm aggregates that had been resuspended into the water column. *Lachnospiraceae incertae sedis* are also anaerobic and are involved in biohydrogenation in rumens, indicating the presence of faecal particles from ruminant animals, such as sheep and cattle, and therefore the potential impact of ruminant animals on the microbial river biota (Huws et al. 2011).

Hotspots and coldspots had higher similarity and lower exclusivity than hotspots and background, suggesting coldspots are the result of dying hotspots or hotspots are the prolific growth of coldspots. For instance, coldspots may indicate the promotion of specific genera leading to the distinct genus abundance increases observed within the hotspots. Specific genus abundance increases and overall abundance increases may also occur where *Kaistia* has a lowered abundance. This genus was found in high abundances within background, but lowered abundance in coldspots and hotspots (Fig. 4). Therefore hotspots and coldspots could indicate a reduction in the persistent and most abundant genera within background environments.

Here we identified 52, 61 and 74 bacterial genera exclusive between hotspots and coldspots, hotspots and background and background and coldspots respectively (Table S7-S9). Previously Dann et al. (submitted) found 117 bacteria exclusive between an upstream site and 3.3 km downstream within the same river environment, therefore indicating larger scale mechanisms, such as the flow of genes through fluvial systems, may lead to higher exclusivity between samples (Leff et al. 1992). Coldspots and background contained more exclusive genera than hotspots indicating hotspots allow specific genera enrichment at the expense of genera diversity. As the genera exclusive to the hotspots were not the genera with increases abundance causing the high dissimilarity between hotspots, coldspots and background this indicates hotspots are not exclusive representations of genera

appearance and further supports the notion that hotspots represent the enrichment of persistent genera. Therefore, it is the dominance or whether genera succeed within each microenvironment, rather than appearance or loss, which leads to the main differences between hotspots, coldspots and background.

By looking at adjacent samples, potential proximate relationships were determined. Lower similarity between proximate samples suggested sample proximity is not indicative of taxonomic and abundance similarity. Two coldspots proximate to a hotspot showed the highest similarity due to sample type, with higher similarity between the coldspots. The main driver for dissimilarity between the coldspots and hotspot was *Parasporobacterium* and *Lachnospiracea incertae sedis*, which had heightened abundance within hotspots, and *Nocardiodes*, which was more abundant in coldspots. Higher dissimilarity between coldspots and hotspots were indicated by higher exclusivity in the coldspots suggesting coldspots are high diversity, yet low abundance microenvironments, whereas hotspots contain heightened abundance of bacteria that are common to coldspot samples.

Proximate background and hotspot samples also showed a higher similarity between samples of the same type rather than proximate samples of different type. A greater dissimilarity between sample types was due to *Kaistia* and *Bacteroides,* with the former more abundant within the background and the latter showing heightened abundance in hotspots. In addition to the heightened *Bacteroides* abundance was its low abundance in the background, which indicated the background does not provide an environment allowing *Bacteroides* enrichment. This could perhaps be due to competing bacterial genera, for instance *Kaistia* as this genus had heightened abundance within the background but a lowered abundance in the hotspot. These proximate relationships indicate sample types are a better indicator for taxonomic

relatedness rather than sample proximity hence showing background, hotspots and coldspots contain discrete taxonomic profiles (Fig. 3).

Common background within microbial taxonomic profiles

Phylogenetic analysis of the hotspots, coldspots and background revealed a dominance of Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes and Firmicutes (Fig. S3). Previous studies have shown Proteobacteria, Actinobacteria and Cytophaga-Flavobacterium-Bacteroides (CFB) are common freshwater microbial phylotypes (Andersson et al. 2008; Ballesté & Blanch 2010; Newton et al. 2006; Wu et al. 2007; Zwart et al. 2002). Dann et al. (submitted) also found Actinobacteria, Proteobacteria and Bacteroidetes, filamentous bacteria associated with the human gut microbiota, to be abundant within the same river environment. Acidobacteria are less common in freshwater systems, yet Dann et al. (submitted) identified an abundance of Acidobacteria during winter sampling of the Murray River (Zwart et al. 2002). At the genus level, hotspot, coldspot and background samples contained an increased presence of *Geothrix* and *Nocardioides*, as well as *Kaistia* within the background samples (Fig. 2). *Geothrix*, *Nocardioides* and *Kaistia* are typical freshwater sediment bacteria suggesting their reincorporation/resuspension into the water column from the benthos (Coates et al. 1999; Jin et al. 2012; Topp et al. 2000). In addition, *Geothrix* species are associated with the breakdown of organic material and Fe(III)-reduction, and can assist in the oxidation of organic contaminants and therefore influence water quality, as has been observed in aquifers (Nevin & Lovley 2002). Species of the *Nocardioides* degrade numerous toxic organic pollutants, such as insecticides used in agricultural farming, therefore implying a potential anthropogenic effect that persists at background levels (Topp et al. 2000).

Long and Azam (2001) hypothesised that using 1 µl samples in heterogeneous environments would enable better detection of rare phylotypes that would otherwise be missed if bulk phase sampling was used. Here we support this hypothesis by obtaining > 1244 OTUs by using 1 µl sample volumes compared to a previous study within the same environment obtaining 613 OTUs by using 5 L sample volumes (Dann et al. submitted). This therefore indicates taxonomic studies within heterogeneous environments, particularly those concerned with the rare biosphere, require small scale sampling to achieve adequate sensitivity in taxonomic profiling.

CONCLUSION

Here we report microscale microbial hotspot, coldspot and background regions represent discrete spatial taxonomic partitions. Hotspots were characterised by distinct genera abundance increases, with these genera differing between hotspots. Genera showing heightened abundance, specifically *Parasporobacterium, Lachnospiraceae incertae sedis* and *Bacteroides,* were indicative of animal and/or human inputs indicating a potential anthropogenic effect on microbial patchiness. Genera exclusivity was highest in coldspots and background, indicating hotspots do not represent the gain of genera but rather enrichment of particular persistent bacteria. Higher similarity between hotspots and coldspots suggested coldspots may represent dying hotspots caused by removal processes with grazing favoured as the likely mechanism due to a lack of strong correlation between prokaryotic and viral populations. Sample proximity did not determine taxonomic likeness, rather sample type, indicating hotspots, coldspots and background represent exclusive microenvironments. These findings support the concept that heterogeneous genus

richness and composition exists within this river system, which will therefore have a significant impact on local microbial biogeochemical cycling.

SUPPLEMENTARY INFORMATION

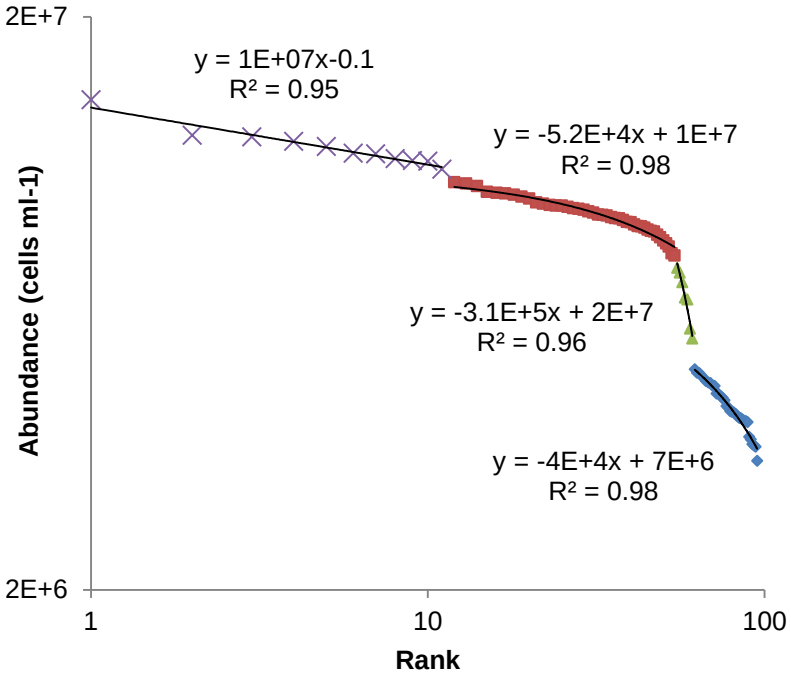


Figure S1. Representative rank abundance graph used for hotspot, background and coldspot determination.

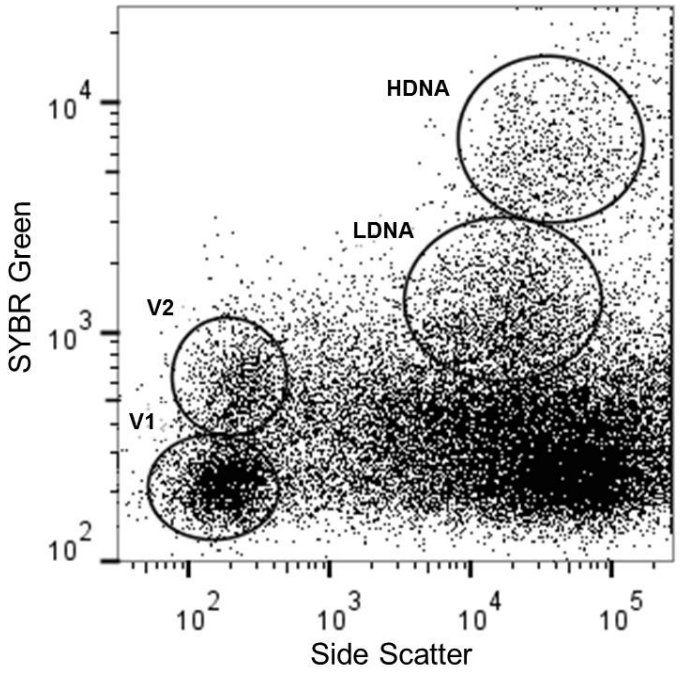
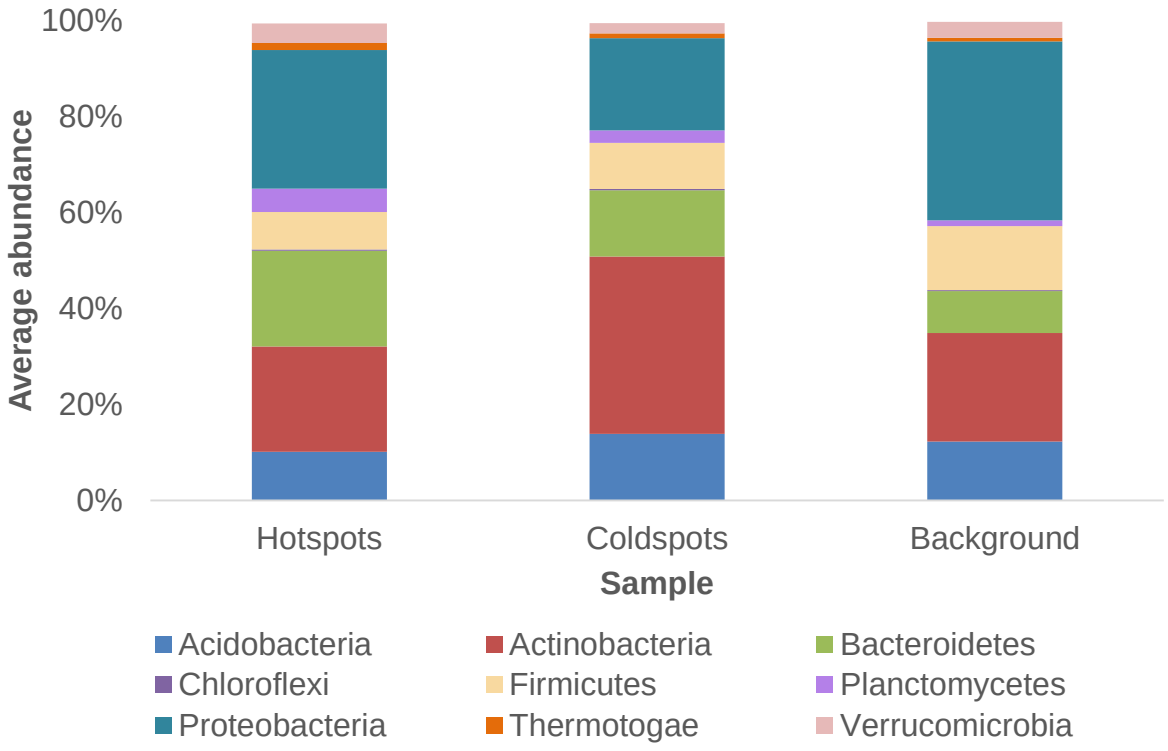


Figure S2. Flow cytometric determination of bacterial and viral subpopulations in biparametic cytogram of SYBR Green fluorescence vs. side scatter (size) showing the presence of prokaryotic (LDNA and HDNA) and virus (V1 and V2) subpopulations.



A

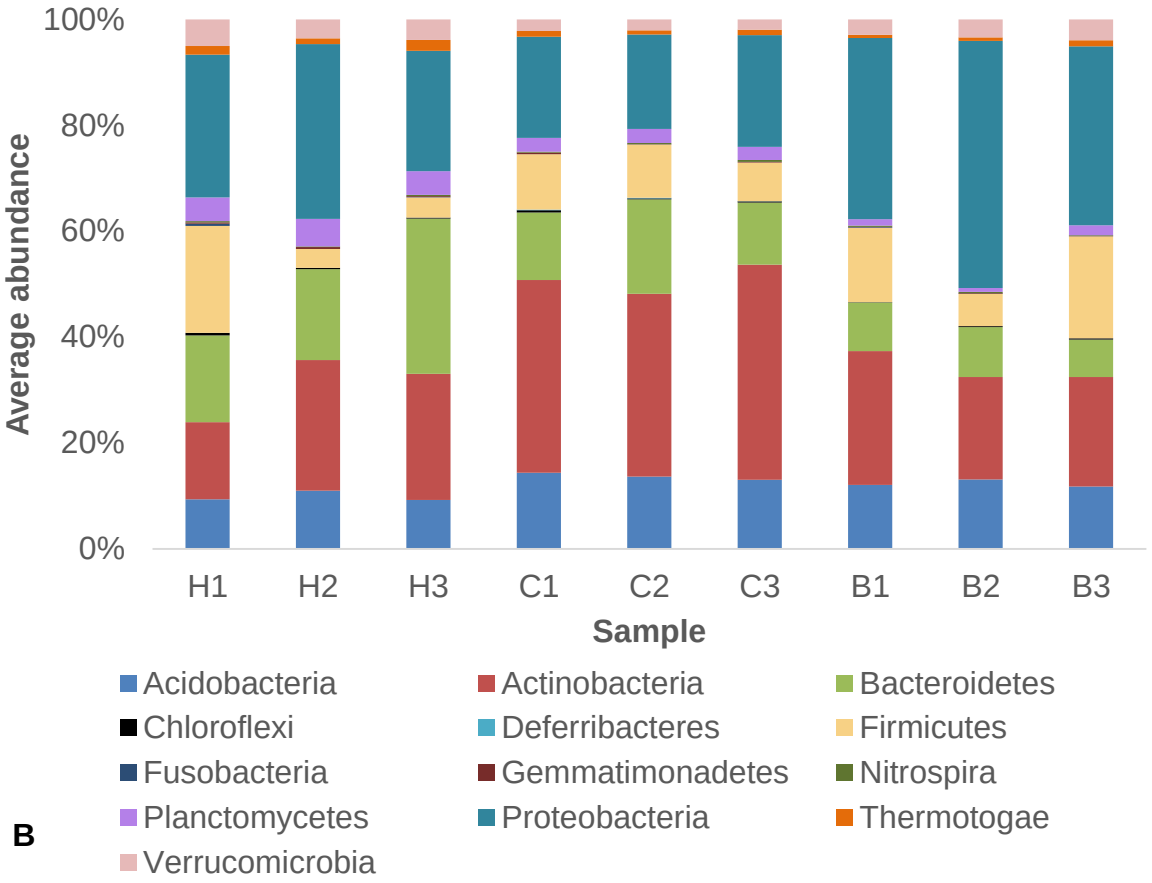


Figure S3. A Averaged phylogenetic profiles of hotspot (n = 20), coldspot (n = 20) and background (n = 20) samples. **B** Phylogenetic profiles of individual background, coldspot and hotspot samples. For clarity, phyla representing < 0.1% average abundance were omitted.

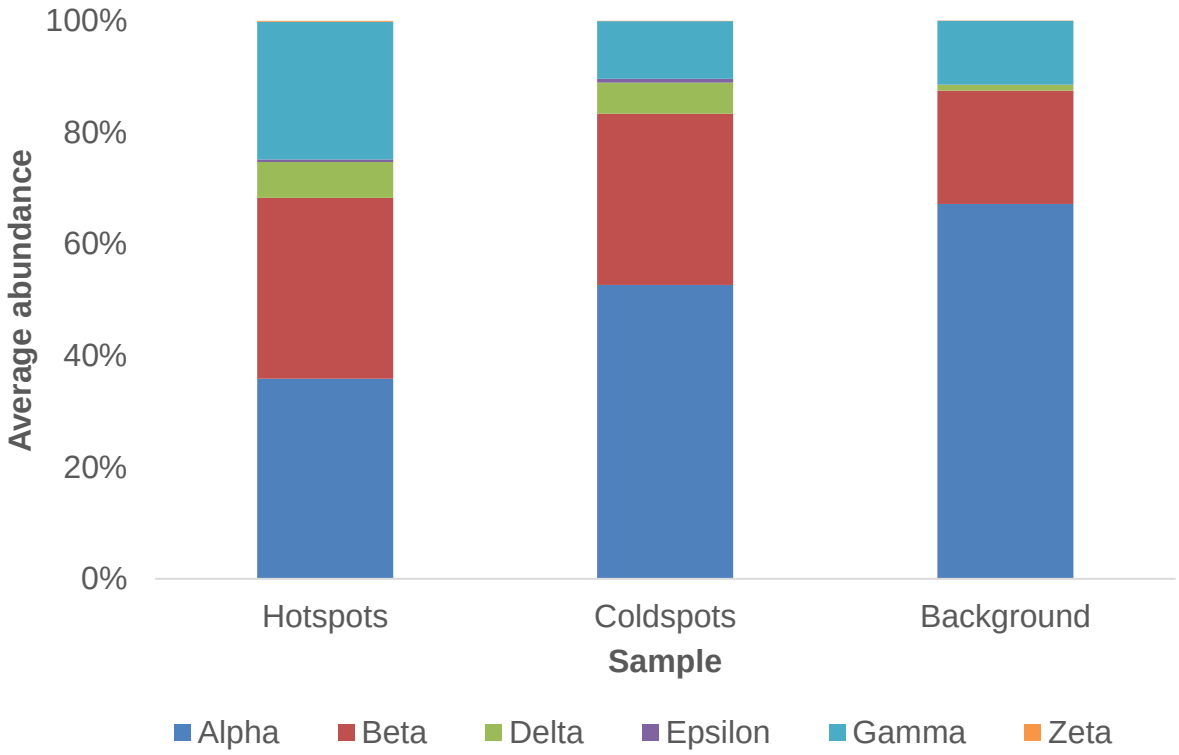


Figure S4. Average abundance of Proteobacteria at the class level in averaged hotspot (n = 20), coldspot (n = 20) and background (n = 20) samples.

Table S1. Average abundances at the phyla level. n = 20. 95% confidence intervals provided for each abundance.

Phyla	**Average Abundance (%)**		
	Hotspots	**Coldspots**	**Background**
Proteobacteria	28.9 (0.02)	19.3 (0.02)	37.2 (0.03)
Actinobacteria,	21.9 (0.04)	37.0 (0.02)	22.6 (0.02)
Bacteroidetes	20.0 (0.04)	13.8 (0.02)	8.7 (0.01)
Acidobacteria	10.1 (0.02)	13.8 (0.02)	12.2 (0.01)
Firmicutes	7.8 (0.05),	9.6 (0.03)	13.3 (0.04)
Planctomycetes	4.8 (0.01),	2.6 (0.01)	1.2 (0.004)
Verrucomicrobia	4.0 (0.01)	2.1 (0.003)	3.3 (0.005)

Thermotogae	1.5 (0.003)	1.0 (0.002)	

Table S2. Average abundances at the genus level. n = 20. 95% confidence intervals provided for each abundance.

Genus	**Average Abundance (%)**		
	Hotspots	**Coldspots**	**Background**
Geothrix	12.7 (0.02)	15.7 (0.02)	13.2 (0.01)
Nocardioides	9.1 (0.02)	14.7 (0.01)	8.7 (0.01)
Flexibacter	7.8 (0.01),	7.0 (0.01)	4.1 (0.01)
Chryseoglobus	5.6 (0.005)	8.3 (0.01)	4.9 (0.005)
Thiobacillus	4.9 (0.01)		
Bacteroides	4.8 (0.04)		
Pseudomonas	4.3 (0.07)		
Planctomyces	4.1 (0.01)		
Austwickia		6.7 (0.01)	
Kaistia		6.3 (0.01)	19.9 (0.03)
Parasporobacterium		5.1 (0.03)	6.4 (0.03)
Lachnospiraceae incertae sedis			5.4 (0.03)

Table S3. Average abundances of phyla and genera showing heightened abundance in hotspot samples. For H2, n = 10. For H1 and H3, n = 5. 95% confidence intervals provided for each abundance.

	Average Abundance (%)		
	H2	**H1**	**H3**
Phyla	Proteobacteria	Firmicutes 20.2 (0.15)	Bacteroidetes 29.2 (0.10)
Genus	*Pseudomonas* 9.4	*Parasporobacterium* 10.2	*Bacteroides* 14.2 (0.14)

	(0.15)	(0.06)	
		Lachnospiraceae incertae sedis 9.4 (0.002)	

Table S4. Genera exclusive between H3 and H1 as determined via SIMPER presence vs. absence data analysis.

H3	**H1**	
Phascolarctobacterium	*Geobacter*	*Lechevalieria*
Faecalibacterium	*Devosia*	*Enterococcus*
Mucilaginibacter	*Solirubrobacter*	*Nitrobacter*
Psychrobacter	*Anaplasma*	*Desulfobulbus*
Levilinea	*Aeromonas*	*Actinoplanes*
Rothia	*Geosporobacter*	*Altererythrobacter*
Sutterella	*Bellilinea*	*Campylobacter*
Dyadobacter	*Cytophaga*	*Hydrogenoanaerobacterium*
Joostella	*Laribacter*	*Kytococcus*
Butyricicoccus	*Gluconacetobacter*	*Leptolinea*
	Hyphomicrobiaceae	*Leptospira*
	Kineococcus	*Microlunatus*
	Pandoraea	

Table S5. Genera exclusive between H3 and H2 as determined via SIMPER presence vs. absence data analysis.

H3		**H2**
Phascolarctobacterium	*Neisseria*	*Singulisphaera*
Brevundimonas	*Parvibaculum*	*Nitrobacter*
Psychrobacter	*Rivibacter*	*Lechevalieria*
Rothia	*Sphingobacterium*	*Zoogloea*
Sutterella	*Thermodesulfobium*	*Kerstersia*

Dyadobacter	*Corynebacterium*	*Sulfurimonas*
Tetrasphaera	*Dictyoglomus*	*Geobacter*
Butyricicoccus	*Mycoplasma*	*Comamonas*

Table S6. Genera exclusive between H1 and H2 as determined via SIMPER presence vs. absence data analysis.

H1	**H2**
Brevundimonas	*Singulisphaera*
Aeromonas	*Paludibacter*
Geosporobacter	*Rhodobacter*
Cytophaga	*Zoogloea*
Gluconacetobacter	*Mucilaginibacter*
Fusobacterium	*Comamonas*
Kineococcus	
Pandoraea	
Enterococcus	

Table S7. Genera exclusive between background and hotspot samples as determined via SIMPER presence vs. absence data analysis.

Background		
Agaricicola	Rhodobacteraceae	*Methylophilus*
Cupriavidus	*Thermosediminibacter*	*Melitea*
Ectothiorhodosinus	*Pedomicrobium*	*Methylosarcina*
Inhella	*Nevskia*	*Psychrosinus*
Methylibium	*Chryseobacterium*	*Serpens*
Natronincola	*Lysobacter*	*Rhodospirillum*
Pelomonas	*Thermolithobacter*	*Psychromonas*
Perlucidibaca	*Dysgonomonas*	*Variovorax*
Pseudorhodoferax	*Porphyromonas*	*Cellulomonas*
Sediminibacterium	*Chitinibacter*	*Acidocella*
Thermotoga	*Staphylococcus*	*Luteimonas*

Verrucomicrobium	*Caldanaerobius*	*Cellvibrio*
Desulfobacterium	*Mesorhizobium*	*Cellulomonas*
Thiomicrospira	*Sulfurovum*	*Prosthecobacter*
Megamonas	*Caldanaerobacter*	*Burkholderia*
Bowmanella	*Meniscus*	*Rugamonas*
Gemmobacter	*Rhizobium*	*Labrys*
Cerasicoccus	*Sphingosinicella*	*Azonexus*
Rubrobacter	*Rheinheimera*	
Hotspots		
Thermovenabulum	*Janthinobacterium*	*Vibrio*
Xanthomonas	*Sediminobacterium*	*Bauldia*

Table S8. Genera exclusive between coldspot and background samples as determined via SIMPER presence vs. absence data analysis.

Coldspots	Background
Niastella	*Dechloromonas*
Desulfobacterium	*Tistlia*
Microvirga	*Ectothiorhodosinus*
Thermocrispum	*Nevskia*
Eubacterium	*Niabella*
Enterococcus	*Pelomonas*
Sporobacter	*Pseudorhodoferax*
Lactobacillus	*Chitinibacter*
Thermovenabulum	*Desulfobacterium*
Vibrio	*Megamonas*
Xanthomonas	*Pseudoxanthobacter*
Acidiphilium	*Cerasicoccus*
Capnocytophaga	*Thiorhodospira*
Catenulispora	*Cellulomonas*
Anaerofilum	*Parvibaculum*
Methylocapsa	*Mesorhizobium*
Paucimonas	*Devosia*
Terracoccus	*Thermolithobacter*
Citrobacter	*Thioflavicoccus*
Larkinella	*Undibacterium*

Limnobacter	*Psychrosinus*
Laribacter	*Serpens*
Comamonas	*Thermodesulfovibrio*
Halochromatium	*Psychrobacter*
Phascolarctobacterium	*Psychromonas*
Runella	*Desulfomicrobium*
Tepidimonas	*Spirillum*
Pseudaminobacter	*Pandoraea*
Saprospira	*Zoogloea*
Zunongwangia	*Fontibacter*
Anaerovorax	*Shinella*
Arenibacter	*Neisseria*
Schwartzia	*Melitea*
Papillibacter	*Wautersia*
Zhangella	*Sediminitomix*
Heliorestis	
Dyadobacter	
Coenonia	
Segetibacter	

Table S9. Genera exclusive between coldspot and hotspot samples as determined via SIMPER presence vs. absence data analysis.

Coldspots		
Desulfobacterium	*Mucispirillum*	*Thermotoga*
Inhella	*Catenulispora*	*Larkinella*
Sediminibacterium	*Anaerofilum*	*Longilinea*
Staphylococcus	*Gemmobacter*	*Variovorax*
Acidocella	*Novosphingobium*	*Methylibium*
Agaricicola	*Dysgonomonas*	*Anaerovorax*
Acidiphilium	*Citrobacter*	*Arenibacter*
Capnocytophaga	*Rugamonas*	*Cellulomonas*
Rhodopirellula	*Sulfurovum*	*Cellvibrio*
Perlucidibaca	*Caldanaerobacter*	*Zhangella*
Burkholderia	*Thiomicrospira*	*Roseibacillus*
Bowmanella	*Natronincola*	*Rhodoferax*
Pedomicrobium	*Spirochaeta*	*Halochromatium*
Tepidimonas	*Meniscus*	*Sphingosinicella*
Cupriavidus	*Schwartzia*	*Sporobacter*
Thermocrispum	*Sulfuricella*	*Rubrobacter*
Thermosediminibacter		
Hotspots		

Thiorhodospira	*Sediminobacterium*	Rhodobacteraceae
Niabella		

Table S10. Genera exclusive between C1 and H1 as determined via SIMPER presence vs. absence data analysis.

C1			
Desulfobacterium	*Bosea*	*Rothia*	*Acidocella*
Frateuria	*Citrobacter*	*Sediminibacterium*	*Novosphingobium*
Herbaspirillum	*Burkholderia*	*Singulisphaera*	*Rugamonas*
Ilumatobacter	*Perlucidibaca*	*Sphaerotilus*	*Thioalkalivibrio*
Inhella	*Veillonella*	*Sporomusa*	*Paludibacter*
Methylomicrobium	*Runella*	*Zavarzinella*	*Sediminicola*
Niastella	*Dysgonomonas*	*Corynebacterium*	*Lactobacillus*
Opitutus	*Halochromatium*	*Mucispirillum*	*Rhodoferax*
Parabacteroides	*Tepidimonas*	*Thermocrispum*	*Meniscus*
Paracraurococcus	*Nitrosococcus*	*Ideonella*	*Faecalibacterium*
Beijerinckia	*Thiomicrospira*	*Microvirgula*	*Cellvibrio*
Naxibacter	*Ahrensia*	*Pedomicrobium*	*Phaeospirillum*
Agaricicola	*Rubrobacter*	*Eubacterium*	*Dyadobacter*
Staphylococcus	*Desulforhabdus*	*Erysipelothrix*	*Pseudoflavonifractor*
Rhodopirellula	*Heliobacterium*	*Sporobacter*	*Dehalobacter*
Acidiphilium	*Bowmanella*	*Rhizobacter*	*Zhangella*
Anaerofilum	*Arenibacter*	*Thermotoga*	*Succiniclasticum*
Capnocytophaga	*Natronincola*	*Odoribacter*	*Gemmobacter*
Catenulispora	*Roseibacillus*	*Saprospira*	*Sulfurovum*
Propionibacterium	*Caldanaerobacter*	*Joostella*	*Leptothrix*
Filimonas	*Beggiatoa*	*Cupriavidus*	*Sphingosinicella*
Paucimonas	*Zunongwangia*	*Comamonas*	*Prevotella*
Variovorax	*Phascolarctobacteri*	*Butyricicoccus*	

	um		
H1			
Rhodobacteraceae	*Cyclobacterium*	*Thiorhodospira*	*Niabella*

Table S11. Genera exclusive between C2 and H1 as determined via SIMPER presence vs. absence data analysis.

C2			
Acidiphilium	*Nitrosococcus*	*Pseudoflavonifractor*	*Caedibacter*
Acidocella	*Rikenella*	*Succiniclasticum*	*Cellulomonas*
Agaricicola	*Schwartzia*	*Sulfuricella*	*Longilinea*
Anaerofilum	*Acidaminococcus*	*Sulfurovum*	*Pedomicrobium*
Azospirillum	*Anaeromyxobacter*	*Thermodesulfatator*	*Thermosediminibacter*
Bowmanella	*Cellvibrio*	*Thiomicrospira*	*Magnetospirillum*
Butyricicoccus	*Comamonas*	*Rubrobacter*	*Roseibacillus*
Capnocytophaga	*Desulfonauticus*	*Lactobacillus*	*Variovorax*
Citrobacter	*Desulfonema*	*Natronincola*	*Propionibacterium*
Corynebacterium	*Eubacterium*	*Rhizobacter*	*Staphylococcus*
Desulfobacterium	*Labrys*	*Salinihabitans*	*Catenulispora*
Frateuria	*Paracoccus*	*Caldanaerobacter*	*Ahrensia*
Gemmobacter	*Saprospira*	*Cupriavidus*	*Beijerinckia*
Heliobacterium	*Sediminicola*	*Dysgonomonas*	*Desulforhabdus*
Herbaspirillum	*Segetibacter*	*Filimonas*	*Rhodopirellula*
Ideonella	*Thermotoga*	*Larkinella*	*Rugamonas*
Ilumatobacter	*Zunongwangia*	*Prevotella*	*Sphingosinicella*
Inhella	*Burkholderia*	*Sandarakinotalea*	*Anaeroplasma*
Leptothrix	*Paludibacter*	*Sporobacter*	*Mucispirillum*
Methylomicrobium	*Parabacteroides*	*Coenonia*	*Papillibacter*

Microvirgula	*Paracraurococcus*	*Joostella*	*Sporomusa*
Naxibacter	*Paucimonas*	*Meniscus*	*Tepidimonas*
Niastella	*Perlucidibaca*	*Thioalkalivibrio*	*Spirochaeta*
Novosphingobium	*Rothia*	*Veillonella*	*Sphaerotilus*
Opitutus	*Sediminibacterium*	*Zavarzinella*	
H1			
Rhodobacteraceae	*Aeromonas*	*Cyclobacterium*	*Thiorhodospira*

Table S12. Genera exclusive between C1 and C2 as determined via SIMPER presence vs. absence data analysis.

C1	**C2**
Chthonomonas/Armatimonadetes_gp3	*Desulfonema*
Odoribacter	*Rhodocista*
Agrococcus	*Marinobacterium*
Beggiatoa	*Lysobacter*
Bilophila	*Syntrophus*
Geoalkalibacter	*Anaeroarcus*
Pleomorphomonas	*Desulfocurvus*
Robiginitalea	
Gaetbulibacter	
Rhizobium	

Table S13. Genera exclusive between B2 and H3 as determined via SIMPER presence vs. absence data analysis.

B2			
Actinoplanes	*Ahrensia*	*Lechevalieria*	*Methylibium*
Agaricicola	*Caldanaerobius*	*Sulfurimonas*	*Microlunatus*

Bradyrhizobium	*Cellulomonas*	*Derxia*	*Natronincola*
Cerasicoccus	*Desulforhabdus*	*Alkaliphilus*	*Naxibacter*
Cupriavidus	*Pseudoxanthobacter*	*Staphylococcus*	*Thiomicrospira*
Dechloromonas	*Geobacter*	*Solirubrobacter*	*Verrucomicrobium*
Desulfobacterium	*Caldanaerobacter*	*Rhizobium*	*Wandonia*
Ectothiorhodosinus	*Leptolinea*	*Wautersia*	*Methylosarcina*
Frateuria	*Paracoccus*	*Bellilinea*	*Melitea*
Gemmobacter	*Ralstonia*	*Variovorax*	*Nitrobacter*
Gluconacetobacter	*Roseiflexus*	*Tistlia*	*Pelomonas*
Herbaspirillum	*Rugamonas*	*Caedibacter*	*Perlucidibaca*
Ideonella	*Acidocella*	*Hahella*	*Pseudorhodoferax*
Inhella	*Burkholderia*	*Patulibacter*	*Pandoraea*
Kerstersia	*Cellulomonas*	*Sarcina*	*Labrys*
Lutispora	*Megamonas*	*Prosthecobacter*	*Thermosediminibacter*
Spirillum	*Microvirgula*	*Rhizobacter*	*Thioflavicoccus*
Porphyrobacter	*Thermotoga*	*Devosia*	*Cellvibrio*
Sediminibacterium			
H3			
Sediminobacterium	*Vibrio*	*Xanthomonas*	*Faecalibacterium*
Thermovenabulum	*Janthinobacterium*	*Phascolarctobacterium*	*Rhodobacteraceae*
Bauldia			

Table S14. Genera exclusive between B3 and H3 as determined via SIMPER presence vs. absence data analysis.

B3			
Acidocella	*Solirubrobacter*	*Algibacter*	*Prosthecobacter*
Agaricicola	*Staphylococcus*	*Anaerovirgula*	*Succiniclasticum*
Ahrensia	*Sulfurovum*	*Cellvibrio*	*Aeromonas*

Aquaspirillum	*Thiomicrospira*	*Sphingosinicella*	*Cellulomonas*
Bowmanella	*Tistlia*	*Spiroplasma*	*Rugamonas*
Bradyrhizobium	*Rhodoferax*	*Caldanaerobacter*	*Rhodoblastus*
Cerasicoccus	*Chondromyces*	*Corallococcus*	*Desulfomicrobium*
Cupriavidus	*Duganella*	*Maricaulis*	*Desulforhopalus*
Dechloromonas	*Bosea*	*Marivirga*	*Devosia*
Ectothiorhodosinus	*Fontibacter*	*Chryseobacterium*	*Rhodospirillum*
Frateuria	*Pandoraea*	*Lysobacter*	*Nevskia*
Gaetbulibacter	*Zoogloea*	*Spirochaeta*	*Paracoccus*
Geobacter	*Hahella*	*Zimmermannella*	*Pseudoxanthobacter*
Gluconacetobacter	*Herbaspirillum*	*Alkaliphilus*	*Ralstonia*
Labrys	*Ideonella*	*Azonexus*	*Actinoplanes*
Lechevalieria	*Inhella*	*Desulforhabdus*	*Gemmobacter*
Luteimonas	*Kerstersia*	*Dysgonomonas*	*Mesorhizobium*
Lutispora	*Leptolinea*	*Geosporobacter*	*Methylosarcina*
Megamonas	*Thermosediminibacter*	*Haliea*	*Psychromonas*
Methylibium	*Thermolithobacter*	*Phaeospirillum*	*Undibacterium*
Microlunatus	*Thermotoga*	*Pseudorhodoferax*	*Meniscus*
Microvirgula	*Thioflavicoccus*	*Psychrosinus*	*Naxibacter*
Natronincola	*Variovorax*	*Roseiflexus*	*Porphyromonas*
Nitrobacter	*Verrucomicrobium*	*Sediminibacterium*	*Burkholderia*
Pelomonas	*Wandonia*	*Serpens*	*Cellulomonas*
Perlucidibaca	*Anaplasma*	*Pedomicrobium*	*Desulfobacterium*
Bellilinea	*Melitea*	*Steroidobacter*	
H3			
Bauldia	*Sediminobacterium*	*Faecalibacterium*	*Thermovenabulum*
Xanthomonas	*Phascolarctobacterium*	Rhodobacteraceae	

Table S15. Genera exclusive between B2 and B1 as determined via SIMPER presence vs. absence data analysis.

B1		
Steroidobacter	*Chitinimonas*	*Bacteriovorax*
Azonexus	*Rhodoferax*	*Desulfobacca*
Acinetobacter	*Parasutterella*	*Odoribacter*
Sulfurovum	*Dokdonella*	*Thermomonas*
Desulforhopalus	*Yersinia*	*Heliobacterium*
Lacibacter	*Pectinatus*	*Pigmentiphaga*
Frigoribacterium		

Additional taxonomic analysis results

For the hotspots, a total of 2,127,612 primer matched sequences were quality filtered to yield 1,795,425 (84.4%) passed reads, 4,845 short reads discarded (< 250 bp) and 327,342 low quality records discarded (expected errors > 0.5). Dereplication resulted in 442,770 unique reads and 371,806 singletons. OTU clustering using 97% identity produced 1,601 OTUs and 2,945 chimeras (4.1%). Out of the 1,601 OTUs, reference-based chimera detection identified 47 chimeric reads. Taxonomy assignment via the RDP Classifier produced a total of 1,527 OTUs which contained 1,637,388 total reads. Samples H2 contained between 38,357 and 115,285 reads (mean = 68,426, std. dev. = 20,971), H1 contained between 46,545 and 114,499 reads (mean = 87,493, std, dev, = 26,146.7) and H3 contained between 55,989 and 146,072 reads (mean = 103,492, std. dev. = 31,990.4).

For the coldspots, a total of 1,366,587 primer matched sequences were quality filtered to yield 969,705 (71%) passed reads, 468 short reads discarded (< 250 bp)

and 396,414 low quality reads discarded (expected errors > 0.5). Dereplication resulted in 273,700 unique reads and 233,597 singletons. OTU clustering at 97% identity produced 1,322 OTUs and 1,121 chimeras (2.8%). Of the 1,322 OTUs, reference-based chimera detection revealed 46 were chimeric reads. Taxonomy assignment via the RDP Classifier produced a total of 1,264 OTUs which contained 879,823 total reads. Samples C1 contained between 18,612 and 52,884 reads (mean = 36,280, std. dev. = 10,679.7), C2 contained between 34,271 and 63,451 reads (mean = 46,071, std. dev. = 11,920.3) and C3 contained between (mean = 57,333, std. dev. = 41,730.9).

For the background, a total of 1,300,203 primer matched sequences were quality filtered producing 1,007,474 passed reads, 819 short reads discarded (< 250 bp) and 291,910 low quality reads discarded (expected errors > 0.5). Dereplication resulted in 305,310 unique reads and 264,902 singletons. OTU clustering at 97% identity produced 1,299 OTUs and 2,131 chimeras (5.3%). Of these 1,299 OTUs, reference-based chimera detection revealed 46 were chimeric reads. Taxonomy assignment via the RDP Classifier produced a total of 1,244 OTUs which contained 889,419 total reads. Samples B2 contained between 15,399 and 31,150 reads (mean = 23,849, std. dev. = 6,003.8), samples B3 contained between 27,283 and 65,631 reads (mean = 50,699, std. dev. = 15,027.6) and samples B1 contained between 18,344 and 99,216 reads (mean = 68,177, std. dev. = 46,734.9).

Chapter 6

The internal structure of microscale microbial hotspots

ABSTRACT

Microscale microbial distributions are known to be patchy, with hotspots and coldspots often observed over distances of micrometres to centimetres. Hotspots have been discussed previously and here we provide the internal structure for these microscale hotspots. Microscale hotspots, coldspots and background regions showed heterogeneity between 1 µl subsamples taken from each region. This heterogeneity was due to heightened abundance of specific genera with this increase most extreme within hotspot subsamples. Specifically, *Parasporobacterium* and *Lachnospiraceae incertae sedis*, were common amongst all region types whilst *Pseudomonas* and *Bacteroides* were only dominant within subsamples from hotspot regions. Furthermore, *Pseudomonas* and *Bacteroides* were responsible for a shift in the most dominant genera in hotspot subsamples potentially indicating exploitation by these genera. Subsamples lacking these specific genera increases showed similarities in abundant genera, namely *Geothrix*, *Nocardioides*, *Flexibacter*, *Kaistia*, *Austwickia* and *Chryseoglobus*, inferring a common taxonomic background and hence indicating taxonomic groups capable of persistence. Subsample heterogeneity revealed two distinct taxonomic patterns within hotspot subsamples: overall increases in genera common to most subsamples or the dominance of specific individual genera. Genera exhibiting heightened abundance between subsamples were characterised by exponential trends indicating highly structured spatial heterogeneity, whereas genera common to most subsamples exhibited either linear or logarithmic trends indicating random, or a combination of random and active, processes influencing their distributions. Shifts in the most dominant genera of some hotspot subsamples caused distinct taxonomic profiles compared to other subsamples as the whole taxonomic profile was altered by this dominance. These

findings suggest microscale microbial hotspots represent discrete microenvironments within this river system, which will have important implications for local nutrient exchange and cellular interactions and hence system function.

INTRODUCTION

Prokaryotic communities are important for the cycling of crucial nutrients, such as nitrogen, carbon and phosphorus, which are required for proper ecosystem function (Azam 1998). Most of these nutrient exchanges and interactions occur over micrometre to centimetre scales (Azam & Malfatti 2007; Stocker *et al.* 2008). However, microbial distributions are often measured using bulk phase sampling in what is termed as the "mean field approach", whereby large volume samples are collected and an average is taken as a representation of microbial processes for larger samples or volumes (Dann *et al.* 2014 – Chapter 3; Hillmer *et al.* 2008; Liu *et al.* 2002; Seymour *et al.* 2004). There is growing awareness of the deficiencies of this method with microbial abundance and activity differing by orders of magnitude over micrometres to centimetres (Blackburn *et al.* 1998; Long & Azam 2001; Mitchell & Fuhrman 1989; Seymour *et al.* 2000, 2006). Prokaryotic abundance patches, often referred to as hotspots and coldspots, vary by up to 90 fold per centimetre (Dann *et al.* 2014 – Chapter 3).

These hotspots in prokaryotic abundance may represent particle aggregation or disintegration, small scale water mixing, grazing events in proximate areas or bacterial chemotactic response to nutrient patches (Azam 1998; Barbara & Mitchell 2003; Blackburn *et al.* 1998; Fenchel 2001; Matz & Jürgens 2005; Mitchell 2002). Previous work has shown bacteria are capable of exploiting ephemeral nutrient

patches (Stocker et al. 2008). In addition, specific prokaryotic species are better equipped at exploiting these areas of high nutrient concentration due to having higher chemotactic advantage (Stocker et al. 2008). The coldspots in prokaryote distributions are indicative of removal processes such as viral lysis or protistan grazing. Grazing and/or lysis can lead to coldspots of specific bacterial species as grazing can be taxonomy- or size-selective whilst lysis is typically favoured towards high density populations. These mechanisms therefore assist in the enrichment, and hence hotspot formation, of other non-selected bacteria (Azam 1998; Barbra & Mitchell 2003; Blackburn *et al.* 1998; Fenchel 2001; Long & Azam 2001; Mitchell & Fuhrman 1989; Seymour *et al.* 2000, 2004, 2006).

Long and Azam (2001) identified taxonomic differences between 1 µl seawater samples emphasizing the importance of using small sample volumes within heterogeneous environments to better detect rare phylotypes. In addition, Meckenstock et al. (2014) found heterogeneity between 1 µl subsamples of oil droplets, suggesting actively degrading microbial populations are distributed heterogeneously within small water-filled cavities. Recently, Dann *et al.* (2016 – Chapter 4) found high dissimilarity between the taxonomic profiles of hotspots, coldspots and background regions with high dissimilarity between hotspots due to different genera enrichment. This study aims to investigate whether heterogeneity exists between 1 µl subsamples from hotspots, coldspots and background regions in prokaryotic populations. We hypothesise taxonomic heterogeneity between 1 µl subsamples from hotspot, coldspot and background regions with higher heterogeneity between hotspot subsamples. We also hypothesise the internal structure of hotspots is not numerically uniform for all genera present. To test this,

taxonomic analysis was performed on 1 µl subsamples from microscale hotspots, coldspots and background.

MATERIALS & METHODS

Sample collection

Samples were collected from the Murray River at Murray Bridge, South Australia (-35.1, 139.3) on July 2nd, 2014. At the time of sampling the pH was 7.8, water temperature was 13.6˚C, total dissolved solids was 445 mg/L and electrical conductivity was 445 µS/cm. A HydroLab DataSonde probe was used to measure environmental parameters. Sample collection was achieved using a two-dimensional micro-titre plate trialled previously (Dann *et al.* 2014 – Chapter 3, 2016 – Chapter 4). This sampler allowed the collection of 8 x 12 vertical profiles, each separated by 0.9 cm and ranging from 1.4 cm to 11.3 cm from the sediment-water interface. Approximately 300 µl per sample well was collected and sampling occurred at the sediment-water interface of the river benthos. Specific permission to access the sampling site was not required. The field study did not involve endangered or protected species.

For taxonomic identification, 100 µl samples were collected from each microplate well, aliquoted into cryovials and frozen at -80°C until further analyses. For enumeration, 200 µl samples were collected from each microplate well, aliquoted into cryovials containing glutaraldehyde (0.5% final concentration) to fix bacterial cells and viral particles and stored in the dark at 4 ˚C for 15 minutes. Samples were then quick frozen in liquid nitrogen and stored at -80˚C until required for further

analyses. Flow cytometric enumeration was performed within one week of storage to avoid potential sample deterioration (Brussaard 2004).

Flow cytometry

For flow cytometric enumeration, triplicate samples were prepared according to established methods (Dann *et al.* 2016 – Chapter 4; Roudnew *et al.* 2012, 2014). Briefly, samples were thawed, diluted 1:100 in Tris-EDTA buffer (0.2 µm filtered, pH 8.0, 10 mM Tris, 1 mM EDTA) and stained with SYBR Green I nucleic acid dye (1:20,000 final dilution; Molecular Probes). Virus-like particle (VLP) counts were optimised by incubating samples in the dark at 80 ˚C for 10 minutes (Brussaard 2004; Dann *et al.* 2014 – Chapter 3; Seymour *et al.* 2007). Reference beads of 1 µm diameter (Molecular Probes) were used as an internal concentration and size standard with a final concentration per sample of approximately 10^5 beads ml^{-1}. Measured flow cytometry parameters were normalised to bead concentration and fluorescence (Brussaard 2004; Dann *et al.* 2014 – Chapter 3; Marie *et al.* 1997).

A FACSCanto II cytometer equipped with a blue (488 nm, 20 mW, air-cooled), red (633 nm, 17 mW), and violet (405 nm, 30 mW) laser and a phosphate-buffered saline (PBS) solution sheath fluid was employed to analyse the samples (Dann *et al.* 2014 – Chapter 3). Samples were run on a low flow rate for 2 minutes. For each sample, green fluorescence (SYBR I), right-angle light scatter (SSC) and forward-angle light scatter (FSC) were recorded. Triplicate blank control samples containing filtered Tris-EDTA buffer (0.2 µl filtered) stained with SYBR Green I were prepared following the same sample preparation aforementioned and processed in each flow cytometry session to eliminate possible background noise from sample preparation or flow cytometer artifacts (Dann *et al.* 2014 – Chapter 3).

Flow cytometric cytograms and histograms were exported as FCS 3.0 files and analysed via FlowJo (Tree Star, Inc.) to enumerate prokaryotic and VLP populations (Dann *et al.* 2016; Smith *et al.* 2015). These populations were determined via the presence of discrete peaks in SYBR Green fluorescence histograms and their position in SYBR Green fluorescence and side scatter (SSC) cytograms, with side-scatter being indicative of cell size and SYBR Green fluorescence being indicative of nucleic acid content (Brussaard 2004; Dann *et al.* 2014; Marie *et al.* 1997, 1999). Flow cytometry was used for enumeration due to its precision and rapidity (Brussaard 2004; Dann *et al.* 2014; Marie *et al.* 1997).

Data analysis

Hotspots, coldspots and background regions were defined using prokaryote rank abundance graphs (Dann *et al.* 2014 – Chapter 3, 2016 – Chapter 4). Background values exhibited a linear trend and were therefore indistinguishable from a random distribution, whilst the hotspots were identified as the sample points that exceeded this linear fit. The coldspots were identified as sample values that fit a linear trend but exhibited gentler slopes than those observed in the background values due to small differences between adjacent sample points (Dann *et al.* 2014 – Chapter 3, 2016 – Chapter 4). Median abundance samples were selected as background rather than the mean, as inclusion or exclusion of the hotspot and/or coldspot values did not overly affect the median values (Wiebe 1970). From this, three of the highest abundance hotspots (H1, H2 and H3), lowest abundance coldspots (C1, C2 and C3) and median background samples (B1, B2 and B3) were chosen for taxonomic analyses, with five to ten subsamples analysed from each hotspot, coldspot and background sample. For the spatial relationship of these hotspot, coldspot and background samples, refer to figure 1. Two-dimensional contour plots were created

via Surfer 10 (Golden Software, Inc.). Contour plots used a conservative minimum contour interval value of ≥ 1000 events μl^{-1}, which was higher than the maximum flow cytometric error of < 24 events μl^{-1}.

Quantitative PCR

Microplate well samples were analysed via quantitative PCR (qPCR) following methods described previously (Dann *et al.* submitted). Briefly, ten 1 µl subsamples (R1-R10) were taken for the highest hotspot, lowest coldspot and median background value and five 1 µl subsamples (R1-R5) were taken for the next two highest hotspots, lowest coldspots and median background values. 16S rRNA region-specific forward (27F) and reverse primers (519R) were added to 1 µl DNA samples and stained with Universal KAPA SYBR Fast qPCR Master Mix 2x (KAPA Biosystems). Stained samples were run through 42 cycles on a Rotor-Gene to determine whether sufficient DNA amplification was achieved prior to direct PCR.

Direct PCR

Direct PCR amplification was performed as described previously (Dann *et al.* submitted). Briefly, 1 µl DNA samples were used in each 25 µl PCR reaction consisting of KAPA Taq Ready Mix 2x (KAPA Biosystems) and 16S rRNA sequence region-specific forward (27F) and reverse (519R) primers. Samples were run for 42 cycles on a Veriti 96 well Thermal Cycler (Applied Biosystems). From each reaction, 20 µl of 16S rRNA amplicon DNA was sequenced whilst the remaining 5 µl was run on an electrophoresis agarose gel to check DNA quality and amplicon size. For sequencing, the variable 16s rRNA gene region PCR primers with forward primer barcodes were used in a 5 cycle PCR using HotStarTaq Plus Master Mix Kit (Qiagen, USA). The following stages were applied: 94°C for 3 minutes, followed by

28 cycles at 94°C for 30 seconds, 53°C for 40 seconds, 72°C for 1 minute and a final elongation step at 72°C for 5 minutes. PCR products were then run on an agarose gel (2%) to determine relative band intensity and amplification success. Barcoded samples were pooled into equal proportions based on DNA concentrations and molecular weight and were then purified with calibrated Ampure XP beads. Sequencing occurred on an Illumina MiSeq at Molecular Research (Shallowater, TX, USA) following the manufacturer's guidelines.

Taxonomic analyses

The resulting sequenced bacterial DNA was quality filtered and length truncated with sequences discarded if they contained > 0.5% expected errors for all bases or < 250 bp. Sequences were also discarded if they lacked a recognisable barcode or forward PCR primer. Full length dereplication and abundance sorting using a minimum size of 2 to ensure singletons removal were performed via the USEARCH pipeline (Edgar 2010). Singletons removal ensured potential spurious OTUs from sequencing artifacts and/or PCR errors were discarded (Edgar 2013; Liu *et al.* 2015). The cluster_otus command within the UPARSE pipeline in USEARCH v8 was used for OTUs clustering at 97% identity (Edgar 2013). Reference-based chimeric filtering using the gold database was performed via UCHIME (Edgar 2011; Pylro *et al.* 2014). Reads were then globally mapped to the OTUs using 97% identity threshold. The utax command in USEARCH v8 was used for taxonomic assignment via the RDP Classifier (Edgar 2013; Wang *et al.* 2007). OTU tables were created via python scripts and the resulting OTU tabbed text file was used to determine average abundances of OTUs. Statistical analyses were performed in PRIMER (version 7) where OTU tables were square root overall transformed and used for similarity percentage (SIMPER) analysis (Clarke & Gorley 2006). SIMPER analyses

determined what was driving Bray-Curtis dissimilarity within and between samples. Bray-Curtis resemblance was performed to create dissimilarity matrices for PERMdisp, which determined the level of dispersion between subsamples (Clarke 1993). Presence vs. absence overall transformed data with Bray-Curtis resemblance was employed to determine subsample exclusivity of bacterial sequences. Rank abundance graphs at the phyla and genus level were created incorporating all hotspot, coldspot and background regions to determine the function that best described each taxonomic group. Sequences were uploaded to the Harvard Dataverse Network and can be accessed at http://dx.doi.org/10.7910/DVN/SITGAO.

RESULTS

Flow cytometric abundance

Flow cytometric analyses of the microplate wells revealed the total prokaryotic abundance within the hotspots was 4.5 x 10^7 cells ml^{-1} in H1, 4.6 x 10^7 cells ml^{-1} in H2 and 4.3 x 10^7 cells ml^{-1} in H3. Within the coldspots the total prokaryotic abundance was 0.92 x 10^7 cells ml^{-1} in C1, 0.94 x 10^7 cells ml^{-1} in C2 and C3. In the background samples, prokaryotic abundance was 2.7 x 10^7 cells ml^{-1} in B1, B2 and B3 (Fig. 1, S1 Table).

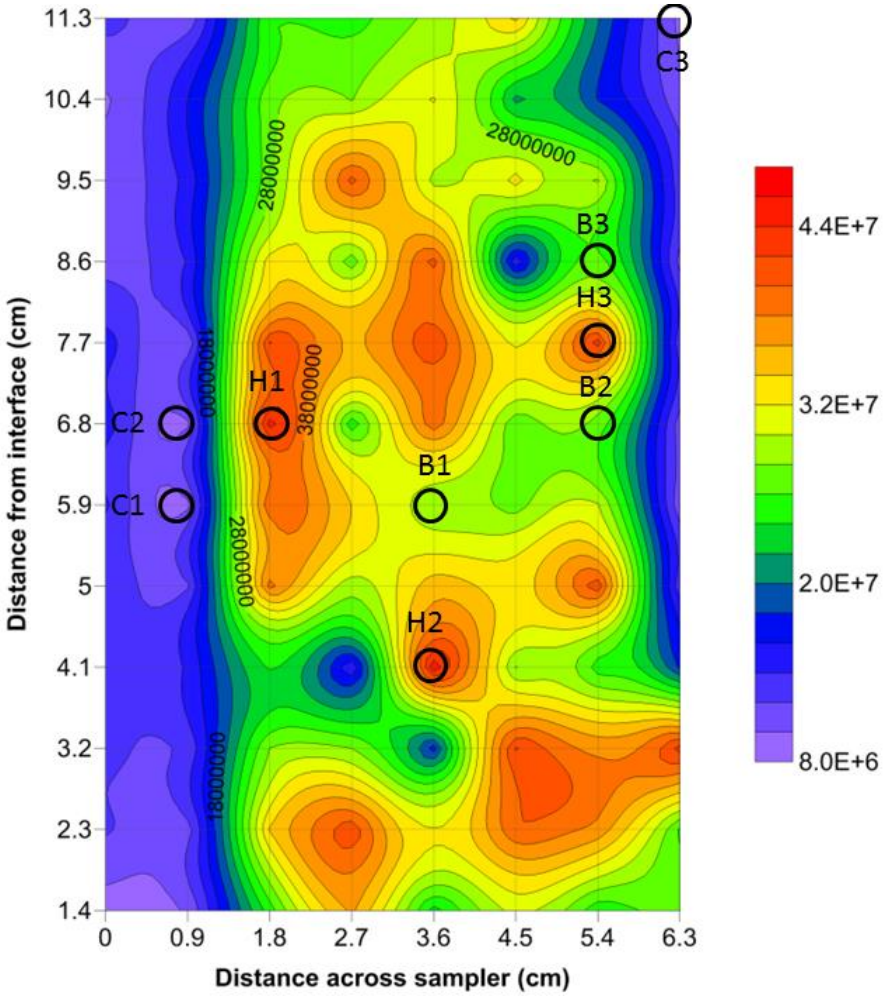


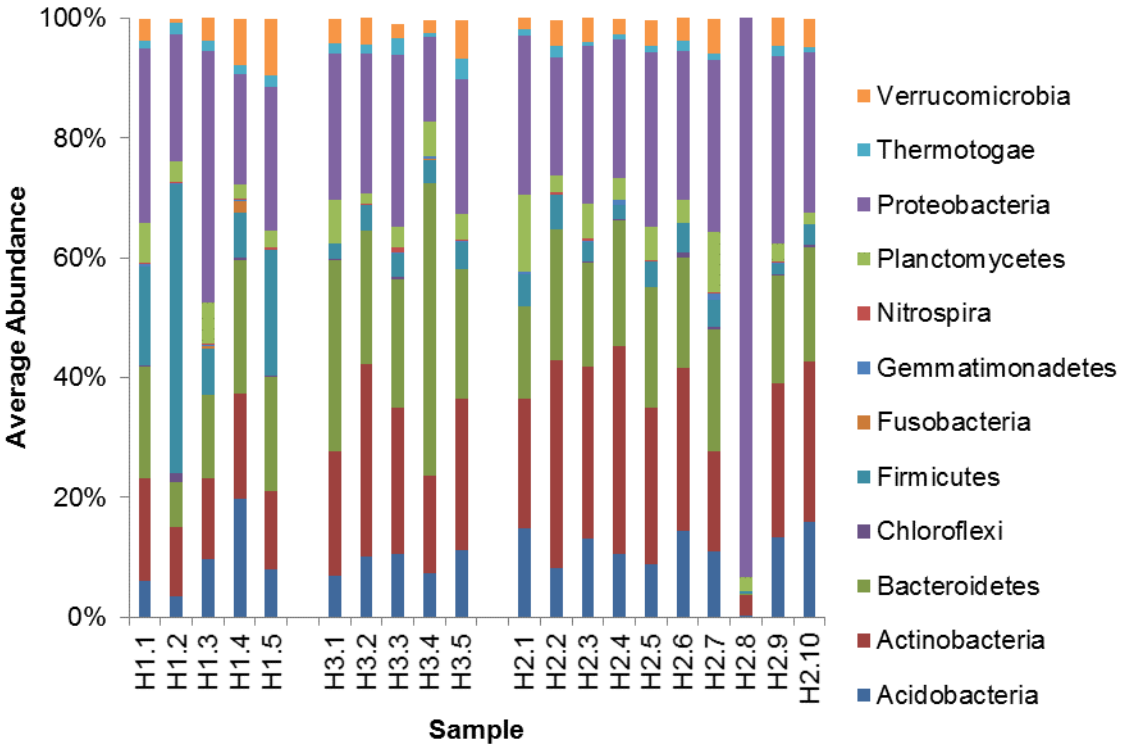
Figure 1. Representative two dimensional contour plot of total prokaryotic abundance showing hotspots, coldspots and background. Circles indicate sample location and corresponding label. Faint gridlines indicate sampling intervals. Colour intensity scale in cells ml^{-1}.

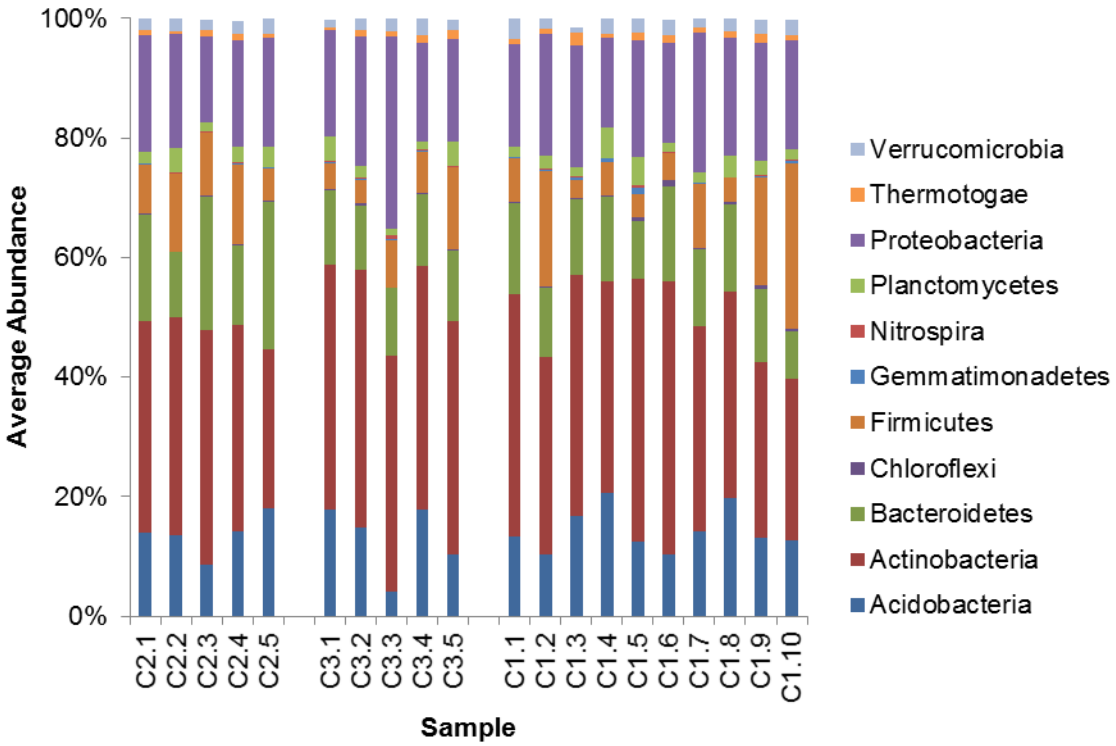
Taxonomic profiles

Phylogenetic profiles of the hotspots revealed the most dominant phyla was Proteobacteria in 9 out of the 20 subsamples, Actinobacteria in 7/20 subsamples, Bacteroidetes in 3/20 subsamples and Firmicutes in 1/20 subsamples (Fig. 2A). At the genus level, 9/20 subsamples were dominated by *Geothrix*, 4/20 subsamples were dominated by *Nocardioides*, 2/20 subsamples were dominated by *Bacteroides*

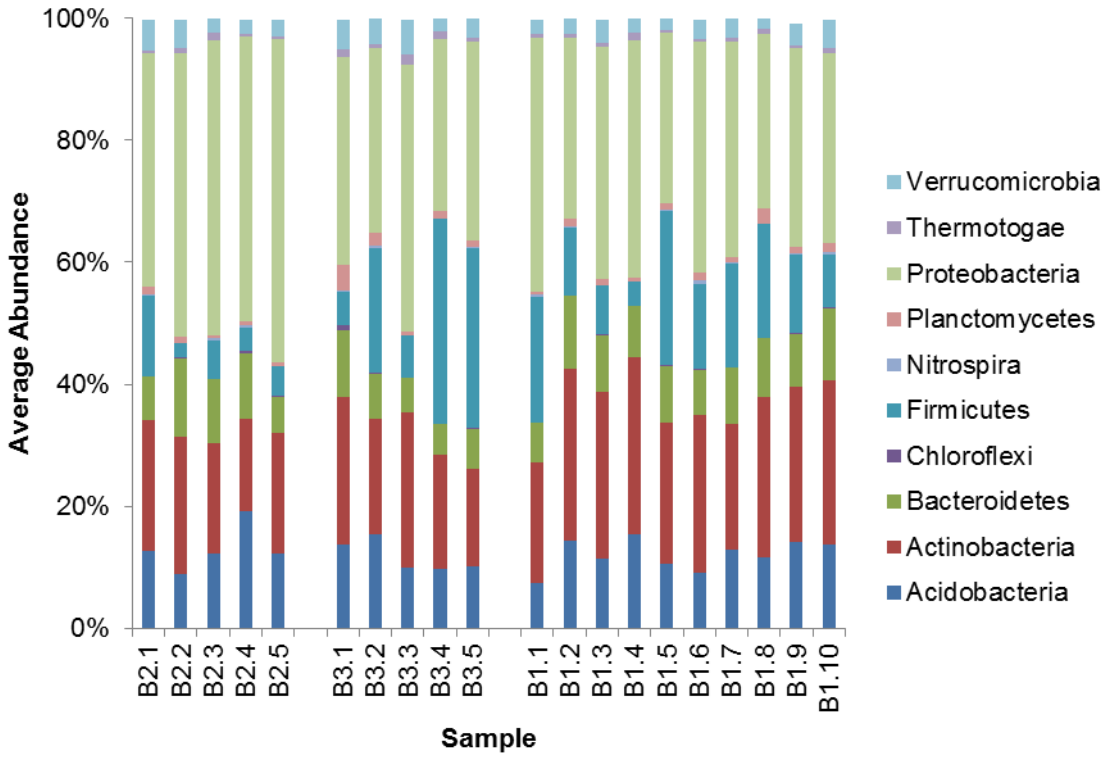
or *Parasporobacterium* whilst *Brevundimonas*, *Lachnospiraceae incertae sedis* and *Pseudomonas* each dominated a single hotspot subsample (Fig. 3). SIMPER analysis revealed an average similarity of 57.5 for H1, 61.2 for H3 and 53.3 for H2. The main driver for similarity within H1, H3 and H2 was *Geothrix*, which contributed to ≥ 5.5% similarity. Comparisons between hotspots revealed a similarity of 57.11 between H1 and H3, 53.52 between H1 and H2 and 56.6 between H2 and H3. The main drivers for dissimilarity were different between hotspots, with *Bacteroides* and *Parasporobacterium* contributing to ≥ 3.2% dissimilarity between H1 and H3, *Parasporobacterium* and *Lachnospiraceae incertae sedis* contributing to ≥ 2.6% dissimilarity between H1 and H2 and *Bacteroides* and *Pseudomonas* contributing to ≥ 2.2% dissimilarity between H3 and H2. PERMdisp analysis at the genus level revealed differences in multivariate dispersion between the hotspots, H1 = 26.9 (SE 1.6), H3 = 24.6 (SE 0.8) and H2 = 30.7 (SE 5.1).

A


100%
80%
60%
40%
20%
0%
Average Abundance
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample
Verrucomicrobia
Thermotogae
Proteobacteria
Planctomycetes
Nitrospira
Gemmatimonadetes
Fusobacteria
Firmicutes
Chloroflexi
Bacteroidetes
Actinobacteria
Acidobacteria


100%
80%
60%
40%
20%
0%
Average Abundance
C2.1
C2.2
C2.3
C2.4
C2.5
C3.1
C3.2
C3.3
C3.4
C3.5
C1.1
C1.2
C1.3
C1.4
C1.5
C1.6
C1.7
C1.8
C1.9
C1.10
Sample
Verrucomicrobia
Thermotogae
Proteobacteria
Planctomycetes
Nitrospira
Gemmatimonadetes
Firmicutes
Chloroflexi
Bacteroidetes
Actinobacteria
Acidobacteria

B

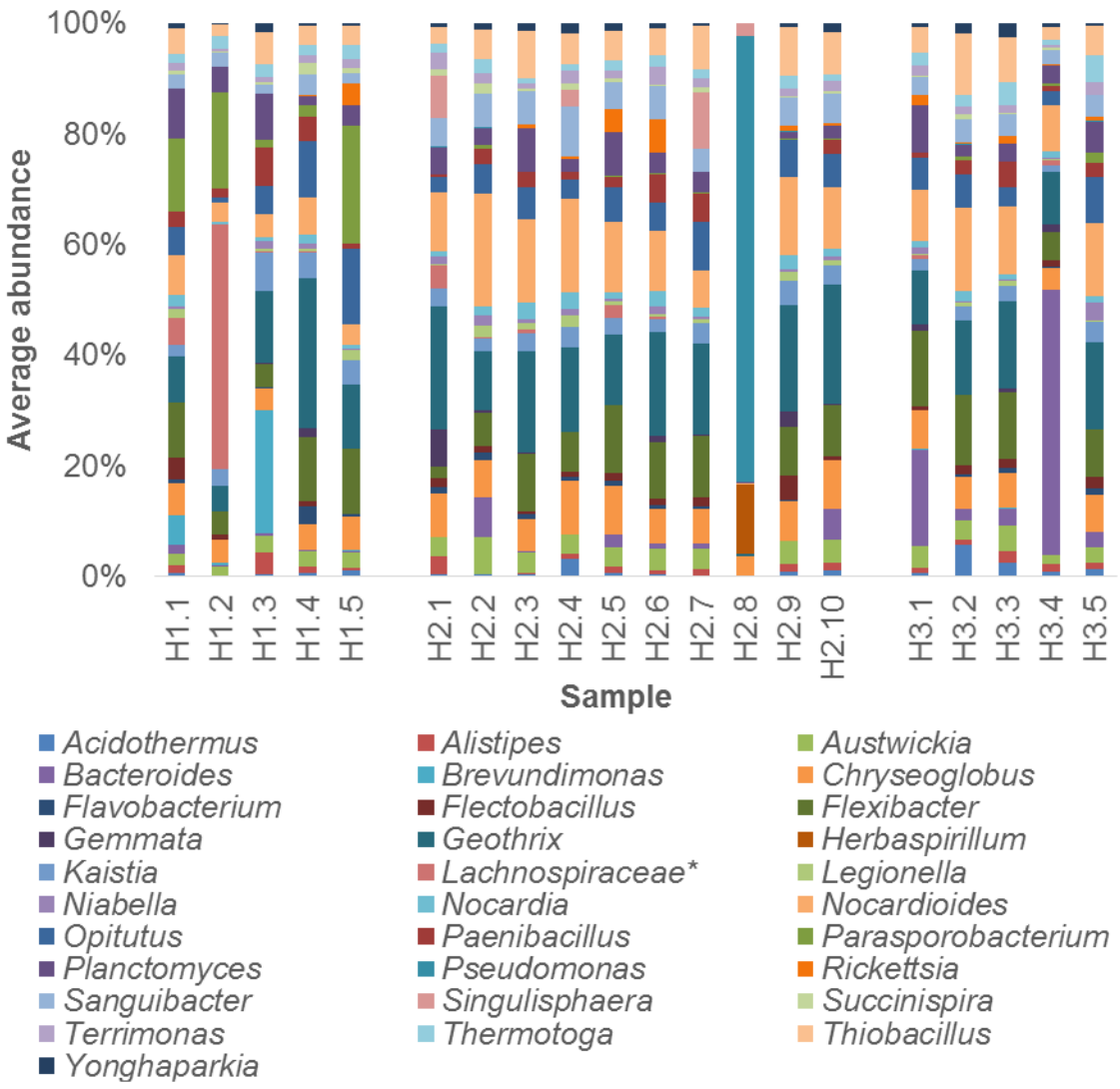


C

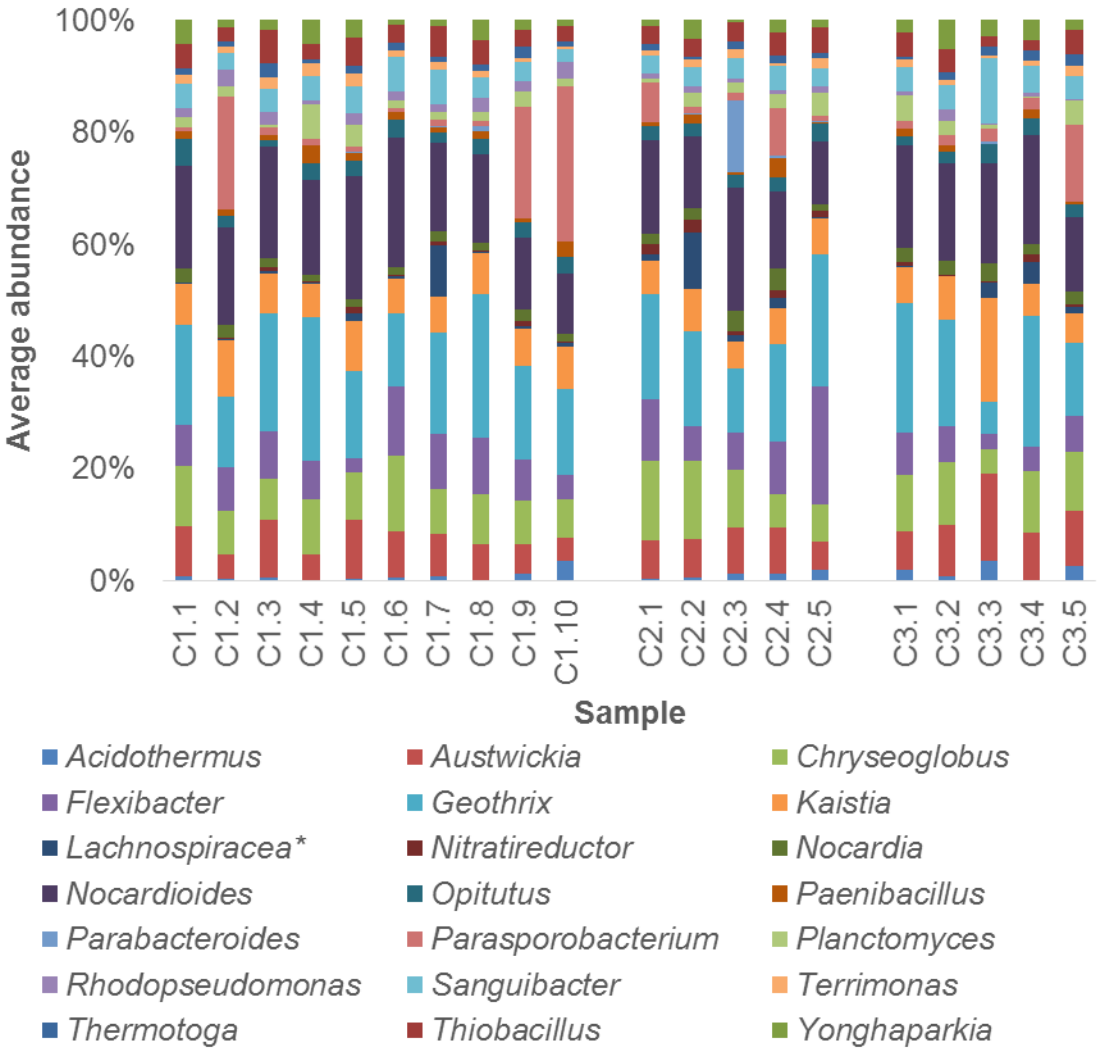
Figure 2. Phylogenetic profiles of **A** hotspot, **B** coldspots and **C** background subsamples showing the heightened abundance of Firmicutes, Proteobacteria and Bacteroidetes. For clarity, only phyla representing > 2% average abundance are shown. OTUs determined via RDP Classifier within the UPARSE pipeline (Stocker et al. 2008)2, 36].

For the coldspots, phylogenetic profiles showed a dominance of Actinobacteria in 19 out of the 20 coldspot subsamples and Firmicutes in 1/20 subsamples (Fig. 2B). At the genus level, 11/20 coldspot subsamples were dominated by *Geothrix*, 4/20 subsamples were dominated by *Nocardioides* or *Parasporobacterium*, whilst 1/20 subsamples were dominated by *Kaistia* (Fig. 3). SIMPER analysis revealed an average similarity of 71.6 for C1, 70.3 for C2 and 66.0 for sample C3. The main drivers for similarity within C1, C2 and C3 were *Geothrix* and *Nocardioides*, which contributed to ≥ 4.9% similarity. Comparisons between coldspots revealed an

average similarity of 70.8 between C2 and C1, 68.3 between C2 and C3, and 68.6 between C1 and C3. The main driver for dissimilarity between C1, C2 and C3 was *Parasporobacterium*, which contributed to ≥ 1.6% dissimilarity. PERMdisp analysis at the genus level revealed dispersions of 18.8 (SE 0.8) for C2, 19.1 (SE 0.3) for C1 and 21.4 (SE 3.3) for C3.



A


100%
80%
60%
40%
20%
0%
Average abundance
C1.1
C1.2
C1.3
C1.4
C1.5
C1.6
C1.7
C1.8
C1.9
C1.10
C2.1
C2.2
C2.3
C2.4
C2.5
C3.1
C3.2
C3.3
C3.4
C3.5
Sample
Acidothermus
Austwickia
Chryseoglobus
Flexibacter
Geothrix
Kaistia
Lachnospiracea*
Nitratireductor
Nocardia
Nocardioides
Opitutus
Paenibacillus
Parabacteroides
Parasporobacterium
Planctomyces
Rhodopseudomonas
Sanguibacter
Terrimonas
Thermotoga
Thiobacillus
Yonghaparkia

B

C

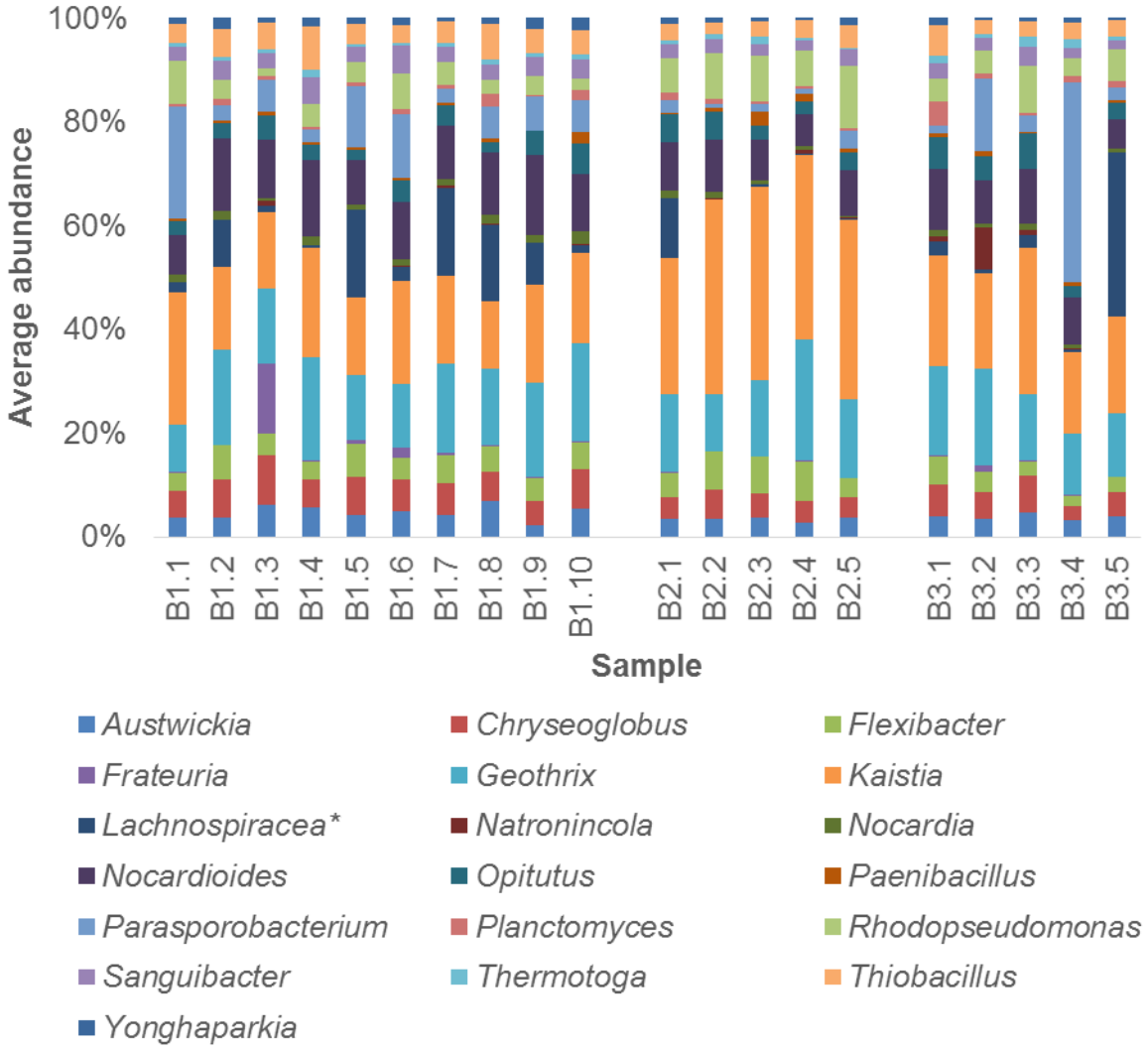


Figure 3. Genera profiles of **A** hotspot, **B** coldspot and **C** background subsamples. For clarity, only genera with average abundances ≥ 2% are shown. **incertae sedis*. OTUs and associated abundance percentages determined via the RDP Classifier within the UPARSE pipeline (Stocker et al. 2008)2, 36].

For the background, phylogenetic analyses indicated 19 out of the 20 subsamples were dominated by Proteobacteria whilst 1/20 subsamples were dominated by Firmicutes (Fig. 2C). At the genus level, 11/20 background subsamples were dominated by *Kaistia*, 5/20 subsamples were dominated by *Geothrix*, 3/20

subsamples were dominated by *Lachnospiraceae* and 1/20 subsamples were dominated by *Parasporobacterium* (Fig. 3). SIMPER analysis revealed an average similarity of 74.6 for B2, 74.1 for B3 and 76.2 for B1. The main drivers for similarity within B2, B3 and B1 were *Kaistia*, *Geothrix* and *Nocardioides*, which contributed to ≥ 4.2% similarity. Comparisons between background samples revealed average similarities of 73.4 between B2 and B3, 74.5 between B1 and B3, and 72.8 between B2 and B1. The main drivers of dissimilarity between B1, B3, and B2 was *Lachnospiraceae incertae sedis*, which contributed to ≥ 3.7% dissimilarity. PERMdisp analysis at the genus level revealed dispersions of 16.0 (SE 0.2) for B2, 16.4 (SE 0.4) for B3 and 15.9 (SE 0.5) for B1 between the background. No archaea were identified in the taxonomic profiles.

- Heterogeneity in taxonomic profiles

Subsample heterogeneity, which was most evident in the hotspot subsamples, was due to heightened abundance of specific genera. For instance, subsample H2.8 contained an increased abundance of Proteobacteria representing 93.4% of its average abundance whilst the other hotspot subsamples from H2 had average abundances ≤ 31.4% (Fig. 2A). Subsample H1.3 also had an abundance of Proteobacteria with an average abundance of 42.0% whilst the other subsamples within H1 had average abundances ≤ 29.1% (Fig. 2A). Firmicutes had increased abundance in subsample H1.1, H1.2 and H1.5, having average abundances of 16.5%, 48.1% and 21.2% respectively, whilst the remaining subsamples in H1 had average abundances ≤ 7.8% (Fig. 2A). Bacteroidetes was abundant in subsample H3.4 and H3.1, representing 48.8% and 32.0% average abundance whilst other subsamples within H3 contained average abundances ≤ 22.3% (Fig. 2A).

Firmicutes and Proteobacteria also exhibited heightened abundance within the background and coldspot subsamples. For instance, Firmicutes within subsample B3.2, B3.4 and B3.5 contained 20.4%, 33.7% and 29.6% average abundance whilst the remaining subsamples of B3 contained ≤ 7.0% average abundance (Fig. 2C). In addition, Firmicutes within subsample C1.2, C1.9, and C1.10 contained 19.3%, 18.1% and 27.8% average abundance whilst the remaining subsamples of C1 contained average abundances ≤ 10.8% (Fig. 2B). For the Proteobacteria, subsample B3.3 contained 43.8% average abundance whilst the remaining B3 subsamples contained average abundances ≤ 34.0% (Fig. 2C). Proteobacteria in subsample C3.3 had an average abundance of 32.3% whilst the remaining C3 subsamples had average abundances ≤ 21.8% (Fig. 2B).

Genera analyses of Proteobacteria revealed the dominant genera within subsample H2.8 were *Pseudomonas* (80.6%) and *Herbaspirillum* (12.6%) (Fig. 3A). Whereas the other H2 subsamples contained multiple Proteobacteria genera at low abundances, with *Kaistia* and *Thiobacillus,* the most abundant, representing ≥ 1.8% and 2.3% average abundance (Fig. 3A). The heightened Proteobacteria abundance in subsample H1.3 was primarily due to *Brevundimonas*, which had an average abundance of 22.6% whilst the other subsamples within H1 had average abundances ≤ 5.6% and were also dominated by *Kaistia* and *Thiobacillus* ≥ 1.6% and ≥ 1.5% average abundance (Fig. 3A). In the coldspot subsample C3.3, the heightened Proteobacteria abundance was due to an increase in *Kaistia*, which had an average abundance of 18.1% compared to ≤ 7.5% in the remaining subsamples (Fig. 3B). The heightened abundance of Bacteroidetes within subsample H3.4 and H3.1 was due to *Bacteroides*, which represented 47.3% and 16.9% average abundance compared to ≤ 3.0% in the remaining subsamples (Fig. 3A).

Genera analyses of the heightened Firmicutes abundance revealed H1.2 was dominated by *Lachnospiraceae incertae sedis* (43.8%), whilst H1.1 and H1.5 were dominated by *Parasporobacterium* (9.9% and 15.0%). Subsample H1.2, also contained heightened abundance of *Parasporobacterium* (13.3%), whilst H1.1 also contained *Lachnospiraceae incertae sedis* (3.7%) (Fig. 3A). The remaining subsamples lacking heightened Firmicutes abundance, H1.3 and H1.4, were dominated by *Paenibacillus* (5.5% and 3.2%) (Fig. 3A). The heightened Firmicutes abundance in the background, B3.2 was due to *Parasporobacterium* (13.6%) and *Natronincola* (7.8%), B3.4 was also due to *Parasporobacterium* (37.6%) B3.5 was due to *Lachnospiraceae incertae sedis* (31.1%) and coldspots C1.2, C1.9, and C1.10 was due to *Parasporobacterium* which represented 19.6%, 19.2% and 27.1% (Fig 3B-C).

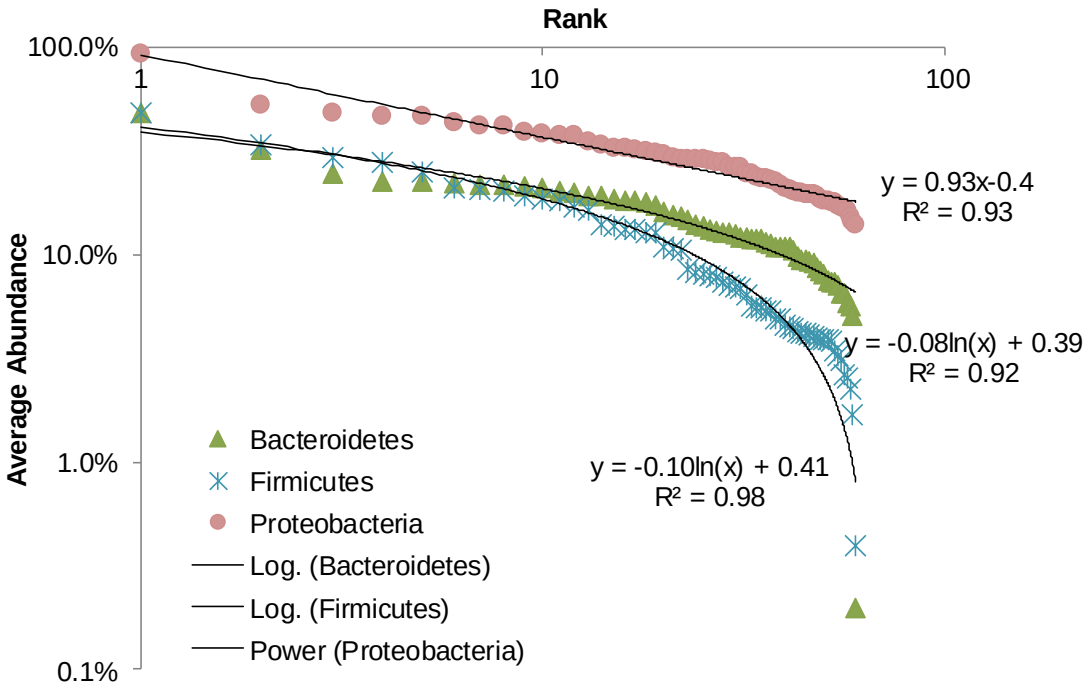
- Homogeneity in taxonomic profiles

Acidobacteria, Actinobacteria, Chloroflexi, Nitrospira, Planctomycetes, Thermotogae and Verrucomicrobia were present in all subsamples and had relatively consistent average abundances between subsamples (Fig. 2). In addition, Gemmatimonadetes was present in the hotspots and coldspots and Fusobacteria was present in the hotspots (Fig. 2). Genera analyses of Acidobacteria revealed *Geothrix* was the most abundant genus within all subsamples (Fig. 3). Genera analyses of Actinobacteria revealed *Nocardioides*, *Chryseoglobus*, *Sanguibacter* and *Austwickia* were the most dominant genera within all subsamples (Fig. 3). Chloroflexi was represented by *Caldilinea* whilst Nitrospira corresponded to the genus *Nitrospira*, Genera analyses of Plantomycetes revealed *Planctomyces* was the most abundant genus in the subsamples (Fig. 3). Genera analyses of Thermotogae revealed *Thermotoga* was the most dominant genus whilst *Opitutus* represented the most abundant

Verrucomicrobia genus (Fig. 3). SIMPER analysis revealed genera with abundances that were common amongst subsamples, *Geothrix*, *Nocardioides*, *Flexibacter*, *Kaistia*, *Austwickia* and *Chryseoglobus* were the main drivers for similarity between subsamples, hence indicating a common background between subsamples.

- Rank abundance analyses

Rank abundance graphs of phylogenetic groups were created using all hotspot, coldspot and background regions (Fig. 4). Phyla exhibiting discrete individual abundance increases between subsamples revealed a logarithmic trend within Firmicutes and Bacteroidetes (Fig. 4A), whilst Proteobacteria showed a power law trend (Fig. 4A). Within the rank abundance graphs of the phyla with common abundances between subsamples, a linear relationship was present in Acidobacteria and Actinobacteria (Fig. 4B); whereas a logarithmic relationship was present in Planctomycetes, Verrucomicrobia, Chloroflexi, Thermotogae and Nitrospira (Fig. 4B). Subsample H2.8 was an outlier in 7 of the 10 phyla identified, with the highest abundance of Proteobacteria and lowest abundance of Firmicutes, Bacteroidetes, Actinobacteria, and Acidobacteria (Fig. 4A-B). At the genus level, rank abundance analyses of the most abundant genera with common abundances between subsamples revealed a logarithmic or linear trend, whilst the genera showing specific abundance increases between subsamples exhibited an exponential trend (Fig. 5A-B).


Rank
100.0%
10.0%
1.0%
0.1%
1
10
100
Average Abundance
y = 0.93x-0.4
R² = 0.93
y = -0.08ln(x) + 0.39
R² = 0.92
y = -0.10ln(x) + 0.41
R² = 0.98
Bacteroidetes
Firmicutes
Proteobacteria
Log. (Bacteroidetes)
Log. (Firmicutes)
Power (Proteobacteria)

A

B

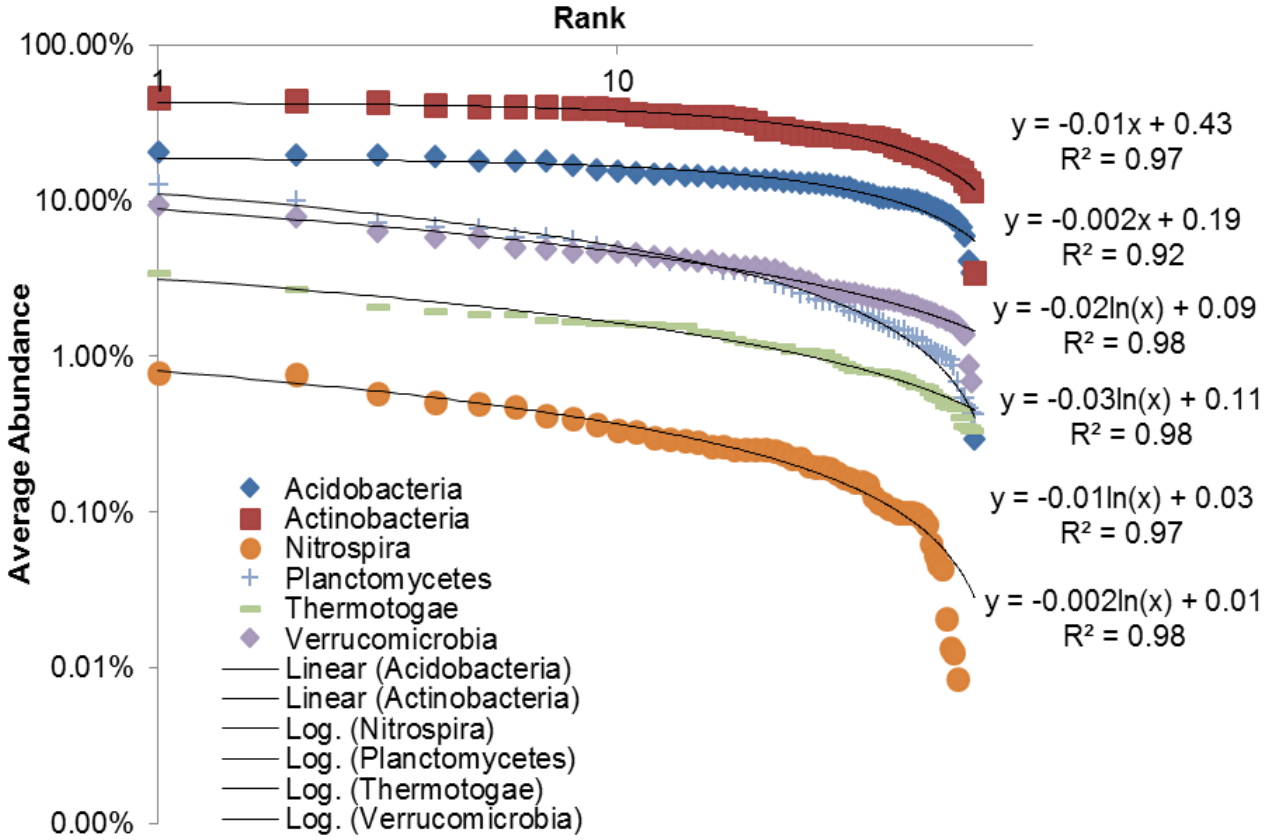
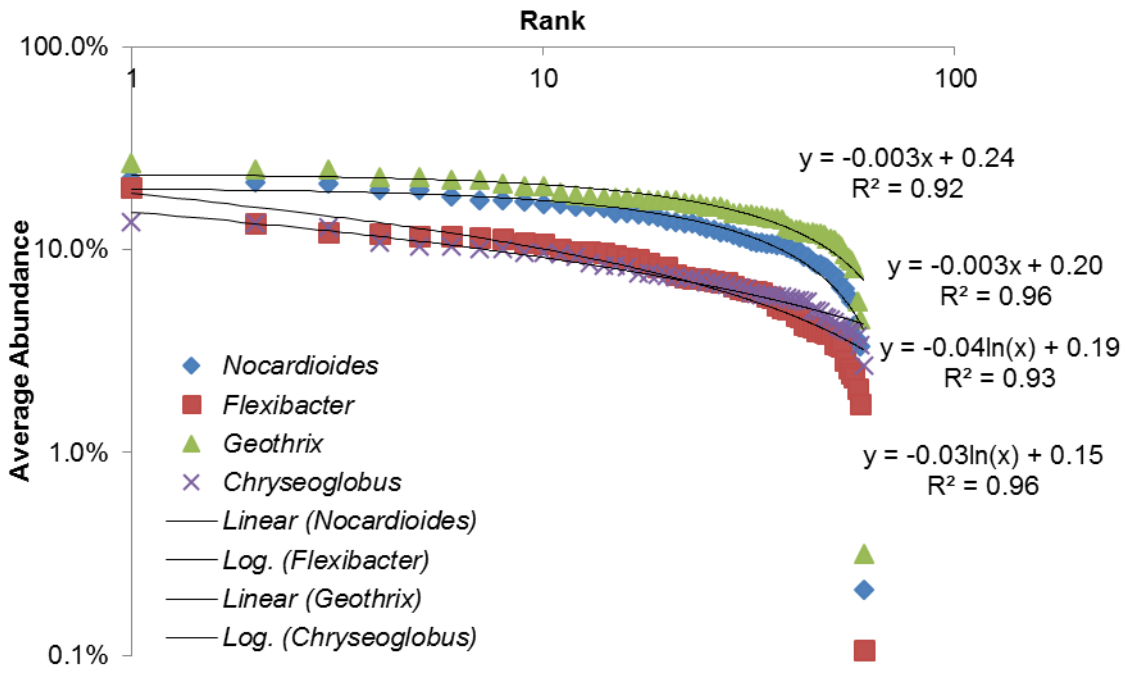
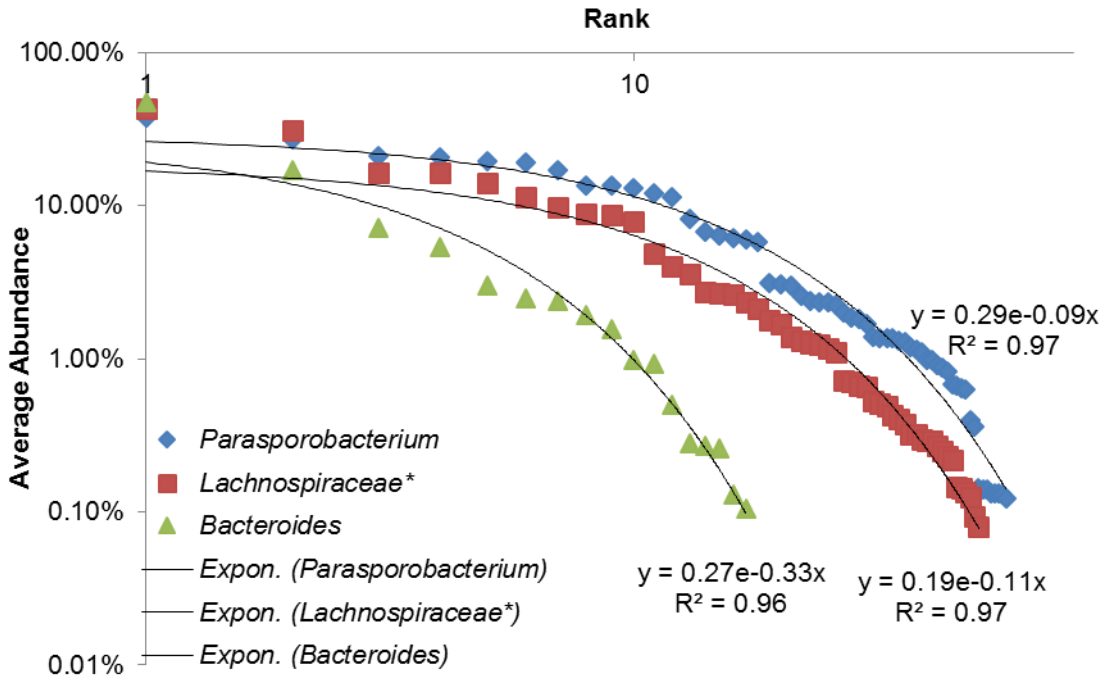

Rank
100.00%
10.00%
1.00%
0.10%
0.01%
0.00%
1
10
Average Abundance
y = -0.01x + 0.43
R² = 0.97
y = -0.002x + 0.19
R² = 0.92
y = -0.02ln(x) + 0.09
R² = 0.98
y = -0.03ln(x) + 0.11
R² = 0.98
y = -0.01ln(x) + 0.03
R² = 0.97
y = -0.002ln(x) + 0.01
R² = 0.98
Acidobacteria
Actinobacteria
Nitrospira
Planctomycetes
Thermotogae
Verrucomicrobia
Linear (Acidobacteria)
Linear (Actinobacteria)
Log. (Nitrospira)
Log. (Planctomycetes)
Log. (Thermotogae)
Log. (Verrucomicrobia)

Figure 4. Representative rank abundance graphs of **A** phyla showing heightened abundance and **B** phyla with abundance common between subsamples.



A



B

Figure 5. Representative rank abundance graphs of **A** genera with abundance common between subsamples and **B** genera with heightened abundance.

- *Background taxonomic profiles*

To determine whether a common background existed between subsamples of hotspot, coldspot and background regions, genera exhibiting heightened abundance between subsamples, *Pseudomonas*, *Parasporobacterium*, *Lachnospiraceae incertae sedis* and *Bacteroides*, were removed. SIMPER analysis showed an increase in average similarities within and between samples, in particular a similarity of 61.6 for H3, 58.2 for H1 and 53.9 for H2 (S2 & S3 Tables). The average dissimilarity between H3 and H1 was 41.8, H3 and H2 was 42.7 and H1 and H2 was 42.7. Whereas for the hotspot samples, complete removal of the subsamples that showed the most extreme genera abundance increases, i.e. H2.8, H1.2 and H3.4, provided higher average similarities of 60.4 within H1 and 63.84 within H2 but a slightly lower similarity in H3 of 61.1. Comparisons between the hotspots showed an average dissimilarity of 41.1 between H1 and H3, 38.1 between H2 and H3 and 40.2 between H2 and H1.

SIMPER analysis between all subsamples indicated the main drivers for similarity were *Geothrix*, *Flexibacter Nocardioides*, *Austwickia* and *Chryseoglobus*, with the former most dominant for hotspots and coldspots, as well as *Kaistia*, which dominated for background, suggesting these genera constitute a common background amongst subsamples.

DISCUSSION

Here we report heterogeneity between the taxonomic profiles of 1 µl subsamples from hotspots, coldspots and background regions (Figs. 2 & 3). This heterogeneity was due to increases in specific genera abundance, with this heightened abundance most extreme within hotspot subsamples, therefore supporting our first hypothesis. This led to highest dissimilarity and dispersion between hotspot, rather than coldspot or background, subsamples. Specific individual genera increases were not present in all hotspot subsamples, with two distinct taxonomic patterns present: overall increases in genera common to most subsamples or the dominance of specific individual genera, therefore supporting our second hypothesis. Overall increases in genera common to most subsamples may indicate equal genera succession and exploitation, where all bacterial genera have the ability to take advantage of growth inducing signals such as nutrient patches (Blackburn & Fenchel 1999; Stocker et al. 2008). Whereas, exclusive increases in individual genera may indicate environmental conditions that only favour specific genera, for instance selective removal processes such as viral lysis or protistan grazing, which do not select for these specific individual genera exhibiting heightened abundance, or nutrient source availability or exploitation that is optimal or achievable for these increased genera (Chow *et al.* 2014; Stocker & Seymour 2012; Thingstad 2000).

Increased heterogeneity within 1 µl hotspot subsamples corresponds to previous findings by Long and Azam (2001) whom found heterogeneous species richness and composition within the ocean by using 1 µl seawater samples, as well as Meckenstock (2014) whom found heterogeneity within the OTU profiles of 1 µl oil

droplet subsamples. Here we confirm hotspots in microbial abundance are microenvironments of particular importance for subsample heterogeneity and hence require high resolution sampling.

Proteobacteria, Firmicutes and Bacteroidetes were the primary phyla exhibiting heterogeneous abundance increases between subsamples (Fig. 2). As Proteobacteria, Firmicutes and Bacteroidetes are arcotypal freshwater bacteria their ability to form regions of heightened abundance may relate to their ability to flourish/sustain within freshwater systems. More specifically, the heightened abundance may indicate microenvironments where these particular bacteria are able to grow faster than the dilution rate and simultaneously avoid viral lysis and protistan grazing, which is an indication of success in competition and defense (Thingstad 2000; Winter *et al.* 2010). Heightened Proteobacteria abundance corresponds with previous freshwater studies showing a high incidence of Proteobacteria attachment on suspended particles and biofilms (Carrias & Sime-Ngando 2009; Newton *et al.* 2011). Proteobacteria represent 20-50% total bacterial abundance within suspended river particles and biofilms, whilst their abundance is much lower in free-living bacterial communities (Carrias & Sime-Ngando 2009). This coincides with previous work by Carrias & Sime-Ngando (2009), which showed living attached to particles offers a variety of advantages for bacteria, such as protection from harmful environmental variables such as predatory attack, UV radiation, desiccation and toxic substances. Particle attachment was also responsible for enhanced interactions and higher nutrient availability due to close cell proximity therefore enabling higher productivity levels (Carrias & Sime-Ngando 2009).

Heightened Firmicutes abundance has been identified previously, being indicators of human faecal matter due to their association with the human gut microbiota, and

therefore serving as important faecal-indicators in environmental water systems (Zheng *et al.* 2009). Whilst increased Bacteroidetes abundance corresponds to previous studies showing this phylum is commonly found associated with particles, and due to their association with the gut microflora, may also indicate the presence of faecal particles within the river system (Carrias & Sime-Ngando 2009; Lamendella *et al.* 2007; Mariat *et al.* 2009). As the Murray River is subject to anthropogenic effects, this perhaps allows high abundance, as well as persistence, of faecal-associated genera within this system (Gell *et al.* 2007). As Bacteroidetes are anaerobic bacteria responsible for specialised degradation of complex macro-molecules, they are typically found in the anoxic regions of microbial biofilms as these contain refractory materials (Burns & Walker 2000; Carrias & Sime-Ngando 2009). Therefore the heightened Bacteroidetes abundance may be due to their ability to form biofilms and could indicate shedding, detachment or disruption of cells of a biofilm containing a high concentration of this particular phylum (Carrias & Sime-Ngando 2009; Donlan 2002; Telgmann *et al.* 2004).

The phyla with common abundances between subsamples, Actinobacteria*,* Acidobacteria*,* Chloroflexi*,* Plantomycetes and Verrucomicrobia*,* indicate phyla that, at least for this environment, are unable/less able to proliferate to abundances above or the same as the phyla showing specific abundance increases between subsamples (Fig. 2). Such phyla may indicate those able to persist but at low numbers such as slow growing defensive strains (Thingstad 2000; Winter *et al.* 2010). However, despite the low abundance of these phyla, their presence is vital to system function, with these minor freshwater phyla being found capable of discrete functions (Newton *et al.* 2011; Wagner & Horn 2006).

Exclusive genera were responsible for the heightened abundance of Proteobacteria in H2.8 and H1.3 (Fig. 2A). Discrete abundance increases of Proteobacteria were restricted to one subsample in H2, with *Pseudomonas* representing 93.35% average abundance in H2.8 (Fig. 3A). *Pseudomonas* belongs to the Gamma Proteobacteria clade and is thought to originate from zoonotic or anthropogenic sources as it is typically a transient member of freshwater systems (Newton *et al.* 2011). *Pseudomonas* is found in many vast environments and can be pathogenic to animals or plants (Palleroni 1981). In contrast, the heightened abundance of Proteobacteria in H1.3 was due to *Brevundimonas* (Fig. 3A). *Brevundimonas* is an aquatic chemo-organotroph known mainly for its characteristic dimorphic reproduction whereby two atypical cells are produced; one is a motile cell with a polar flagellum known as a swarmer and the other containing a stalked prostheca with a holdfast at its distal tip (Tsubouchi *et al.* 2013). Heightened Firmicutes abundance was exclusive to H1, therefore representing a uniform hotspot, with *Lachnospiraceae incertae sedis* dominant in H1.2 whilst *Parasporobacterium* dominated H1.1 and H1.5. In addition, H1.2 also contained increased abundance of *Parasporobacterium*. Differences in the genera showing the highest abundance increases of these subsamples may indicate the microenvironment within H1.2 was most suited to exploitation by *Lachnospiraceae incertae sedis,* whilst this sample as well as H1.1 and H1.5 were more favourable for *Parasporobacterium* succession. This could be due to a multitude of factors, such as a lack of competition from the other genera present, limited/no phage or protists present to prey on these genera or the availability of a particular nutrient source that is better utilised by these genera (Chow *et al.* 2014; Stocker et al. 2008). As *Parasporobacterium* is an anaerobic bacterium typically isolated from freshwater benthos and involved in methanethiol and dimethyl sulphide

formation via sulphide methylation, their increased abundance indicates reincorporation of sediment biofilm aggregates into the water column (Lomans *et al.* 2001). *Lachnospiraceae incertae sedis* is known to be anaerobic and involved in rumen biohydrogenation hence indicating their origin from the faecal particles of ruminant animals (Huws *et al.* 2011). Specific abundance increases of Bacteroidetes within H3.4 was due to a single genus, *Bacteroides.* As *Bacteroides* is commonly associated with the mucosal surfaces of digestive regions of humans, and several genera are used as faecal-indicators, its heightened abundance in H3.4 may indicate the presence of faecal particles (Nava & Stappenback 2011; Ponce-Terashima *et al.* 2014). Analyses of the phyla with common abundances between subsamples revealed Acidobacteria was primarily represented by *Geothrix*, Actinobacteria contained *Nocardioides*, *Chryseoglobus*, *Austwickia* and *Sanguibacter*, Chloroflexi was represented by *Caldilinea*, Verrucomicrobia was primarily represented by *Opitutus* and Plantomycetes had *Planctomyces* as the most abundant genus in all samples. In addition, genera with common abundances between subsamples from phyla showing heightened abundances, for instance *Kaistia* of the Proteobacteria and *Flexibacter* of the Bacteroidetes, made up substantial abundances within all subsamples. The commonality of these genera between samples implies this environment provides favourable conditions allowing these specific genera to persist in the system. As in order for bacteria to persist their growth rate must exceed the dilution rate, that is removal processes such as grazing or lysis, these genera must be able to maintain suitable growth rates and avoid significant mortality (Thingstad 2000; Winter *et al.* 2010).

Different genera dominance patterns were observed between subsamples showing specific individual genera abundance increases and genera with common

abundances between subsamples. For instance, in H3.4, the most abundant genus within Bacteroidetes was *Bacteroides* whilst the most abundant genus within the remaining subsamples was *Flexibacter*. These shifts in the most abundant genus indicate these samples are providing microenvironments that allow these genera to dominate. This perhaps indicates the specific individual genera abundance increases represent genera able to exploit favourable conditions and may be able to do so when the usual abundant genera are present at lower abundances. This is further supported by the presence of these specific individual genera showing heightened abundance in the other samples but at a lower abundance. For this to be possible, according to community assembly theory, these less abundant species can avoid local extinction and out compete other more abundant species if they are able to have a competitive advantage, become replenished/supplemented from surrounding sources or access a different ecological niche (Costello *et al.* 2012).

Taxonomic analyses revealed H2.8 was an outlier with a distinct taxonomic profile, showing the highest abundance of Proteobacteria but the lowest abundance of other abundant phyla, Firmicutes, Bacteroidetes, Actinobacteria and Acidobacteria (Fig. 2A). This implies almost complete succession of Proteobacteria, therefore hindering the enrichment of other phyla. However, H2.8 was not an outlier in the less abundant phylogenetic groups of Plantomycetes, Chloroflexi, Nitrospira, Thermotogae and Verrucomicrobia, indicating these phyla may be able to survive in microenvironments dominated by Proteobacteria. This suggests these phyla exploit different nutrient sources or have predators, such as protists and viruses, which are found in low abundance when the predators of Proteobacteria are also low.

Rank abundance graphs of the phyla showing heightened abundance between subsamples revealed a logarithmic relationship within Bacteroidetes and Firmicutes

implying a combination of underlying processes favouring homogeneity and hence a random distribution as well as processes favouring structure and hence a non-random distribution (Dann *et al.* 2014 – Chapter 3) (Fig. 4A). Specifically this may represent the genera exhibiting heterogeneous and homogeneous abundance distributions within these phyla. A power law relationship within Proteobacteria implies this is a highly structured phyla influenced by active processes, such as aggregation, as there is no true background (Fig. 4A). The rank abundance graphs of the phyla that had common abundances between most subsamples showed a linear relationship in Acidobacteria and Actinobacteria indicating a homogeneous distribution via random processes, whilst a logarithmic relationship in Plantomycetes, Verrucomicrobia and Chloroflexi, again indicated a mixture of random and active processes (Dann *et al.* 2014 – Chapter 3) (Fig. 4B). Plantomycetes, contains several bacterial genera capable of anaerobic ammonia oxidation, therefore the active process that could be occurring within this specific phyla may be related to nitrogen distributions within the system (Sonthiphand *et al.* 2014). Heterogeneity may be present in the phyla showing heightened abundance due to the influence of particular processes that encourage a heterogeneous distribution, such as chemotactic response to nutrient patches or particle aggregation/biofilm formation to avoid unfavourable environmental conditions, such as predatory attack (Carrias & Sime-Ngando 2009; Stocker et al. 2008; Stocker & Seymour 2012). The phyla that had common abundance between most subsamples showed a homogeneous relationship, which may indicate the influence of random processes encouraging a homogeneous distribution (Dann *et al.* 2014 – Chapter 3).

A common taxonomic background could be better identified after the complete removal of hotspot subsamples containing specific individual genera abundance

increases, rather than the removal of these genera from all sub-samples (S2 & S3 Tables). *Geothrix*, *Nocardioides*, *Flexibacter*, *Kaistia*, *Austwickia* and *Chryseoglobus* were the main drivers for similarity between subsamples, and hence constituted the common taxonomic profile, suggesting an environment that favours these particular genera hence allowing their persistence (Hibbing *et al.* 2010). No homogeneous background was identified amongst hotspot subsamples with specific individual genera abundance increases as there was a distinct pattern within the other genera present, which was not similar between other hotspots and/or subsamples. In the latter case, this implies genera interactions are more complex than single genera enrichment matched with homogeneous changes in genera composition and abundance and rather the type of genera enrichment will determine the patterns within the other genera present. This indicates hotspots with increased abundance of specific individual genera contain discrete taxonomic spatial partitions, which are influenced by the type, combination and abundance of genera present.

Average similarities between hotspot subsamples were lower than those between coldspot and background subsamples. Hotspots may not exhibit as high similarities due to sensitive correlations between genera abundance fluctuations leading to higher dissimilarities in their background taxonomy therefore leading to higher subsample heterogeneity within hotspot microenvironments.

CONCLUSION

Here we report heterogeneity between 1 µl subsamples from microbial hotspots, coldspots and background regions. This heterogeneity was due to heightened abundance of specific genera, which was most extreme within hotspot subsamples.

Heterogeneity between subsamples revealed hotspot regions were characterised by two distinct taxonomic patterns: overall increases in genera common to most subsamples or the dominance of specific individual genera, hence revealing a complex structure within the hotspots we observe in microbial microscale distributions. The genera exhibiting heightened abundance were common, belonging to distinct phylogenetic groups of Firmicutes, Proteobacteria and Bacteroidetes, indicating the ability these groups have to succeed in heterogeneous microenvironments. Phyla with common abundance between most subsamples, Acidobacteria, Actinobacteria, Chloroflexi, Planctomycetes and Verrucomicrobia represented those capable of persistence.

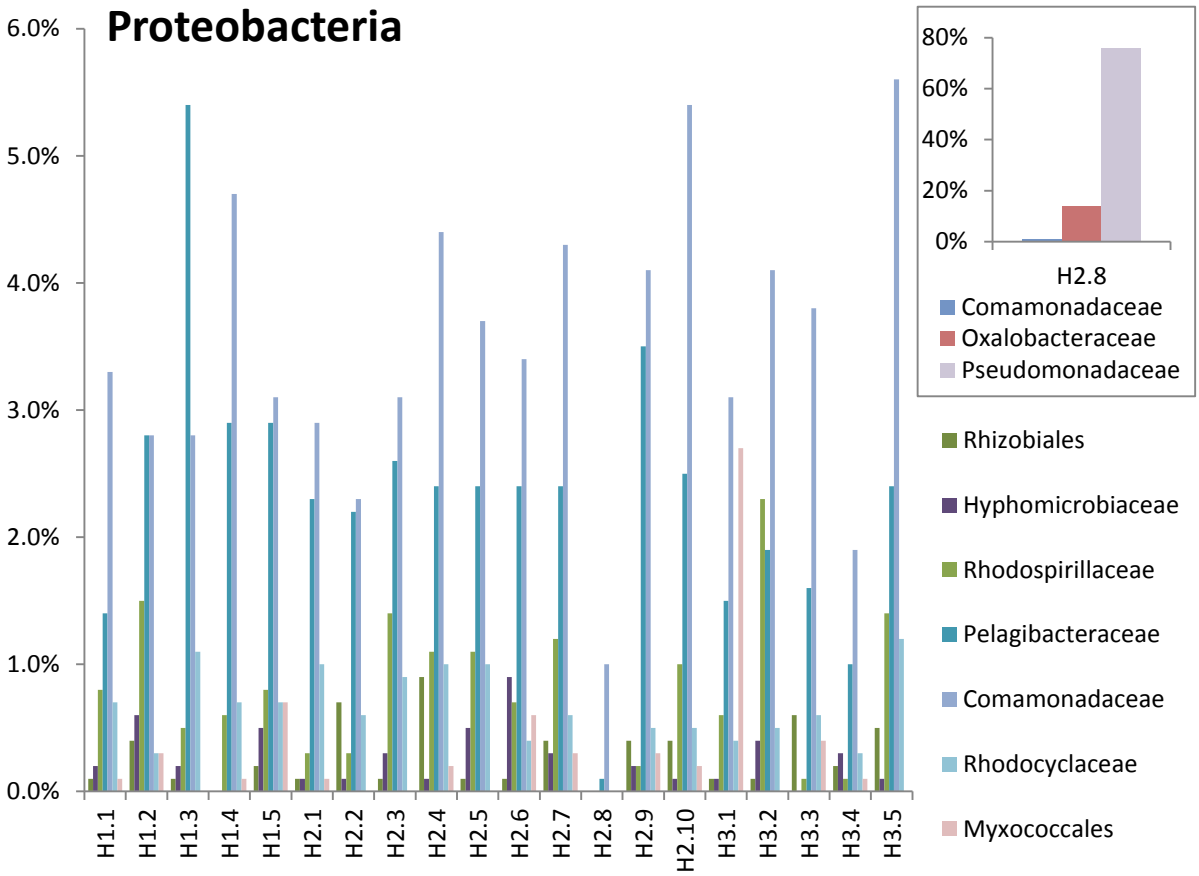
Rank abundance analyses revealed exponential functions in genera exhibiting heightened abundance between subsamples, indicating highly structured spatial heterogeneity perhaps influenced by active processes such as swarming or aggregation. In contrast, the genera with abundances that were common between most subsamples exhibited a linear or logarithmic trend indicating the presence of random processes or a combination of random and active processes. These different functions highlight the ecological differences between these genera and hence the different processes influencing their distributions.

Shifts in the most dominant genera of some hotspot subsamples caused distinct taxonomic profiles compared to other subsamples as the whole taxonomic profile was altered by this dominance. This indicated hotspots contain discrete taxonomic microenvironments, which are influenced by the type, combination and abundance of genera present. In addition, the most abundant genera within subsamples lacking specific individual genera abundance increases were comparable, indicating a common taxonomic background. As microbial hotspots and coldspots are important

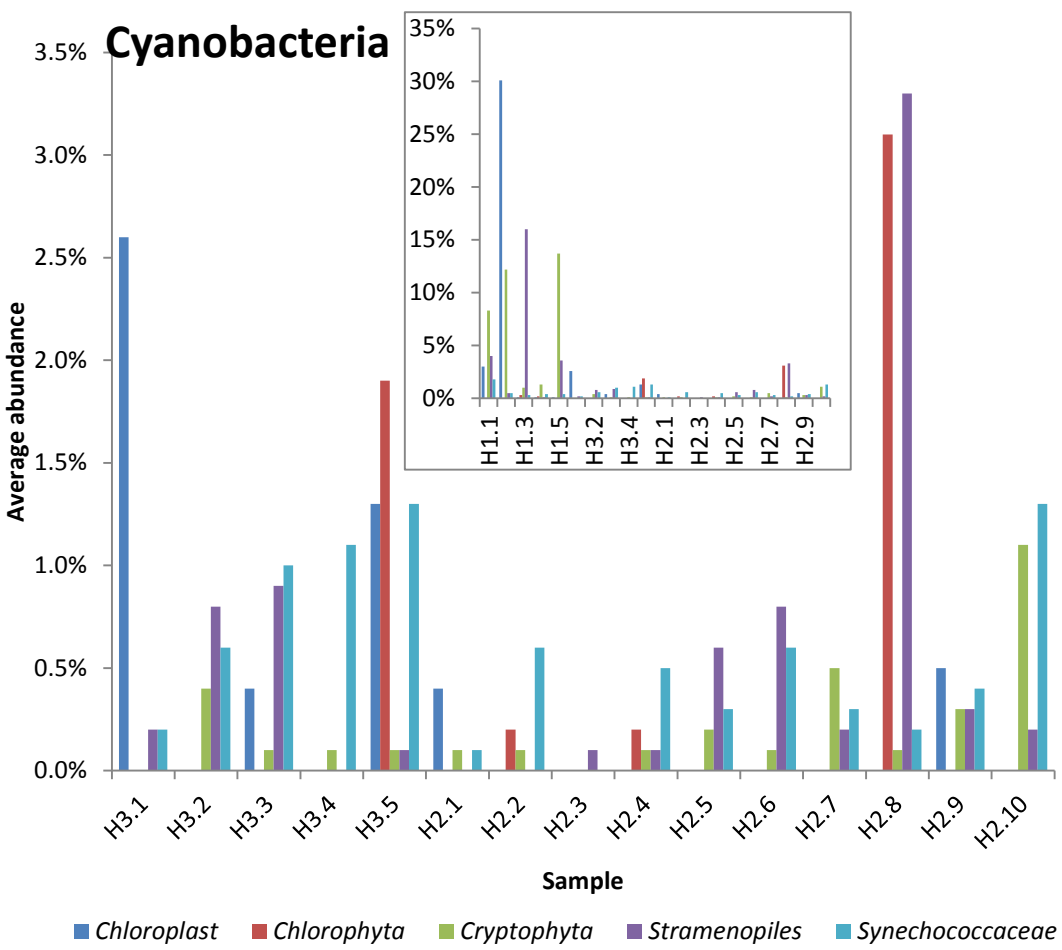
microenvironments for nutrient exchange and cellular interactions, understanding their taxonomic makeup will aid further understanding in environmental heterogeneity, microbial diversity and system function.

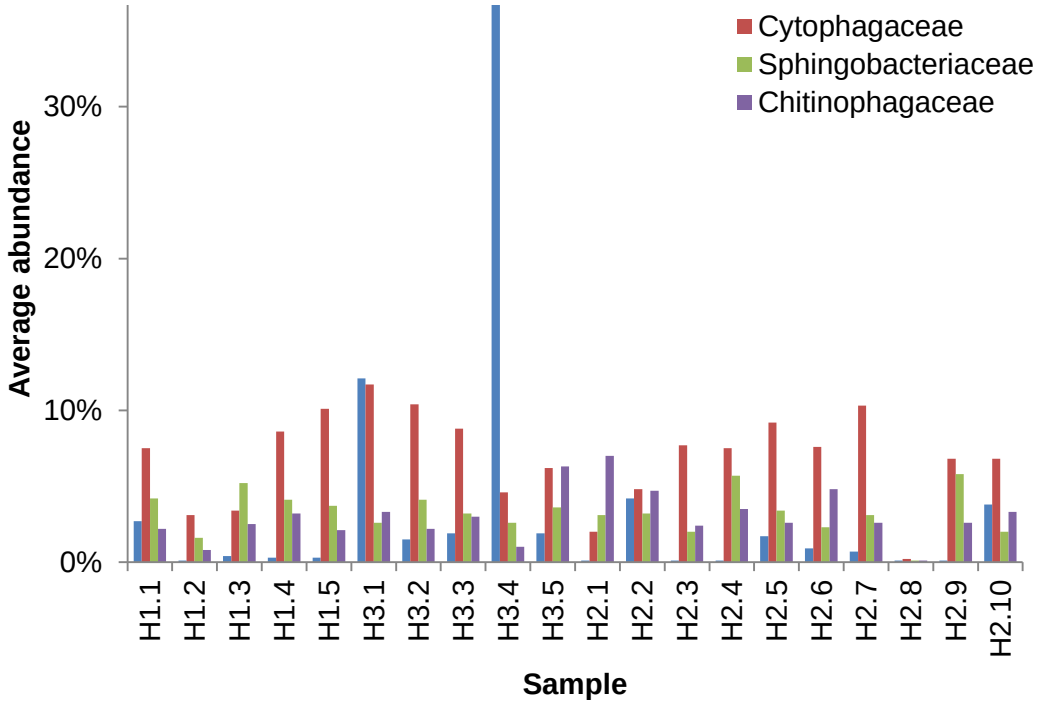
SUPPLEMENTARY INFORMATION

A



B

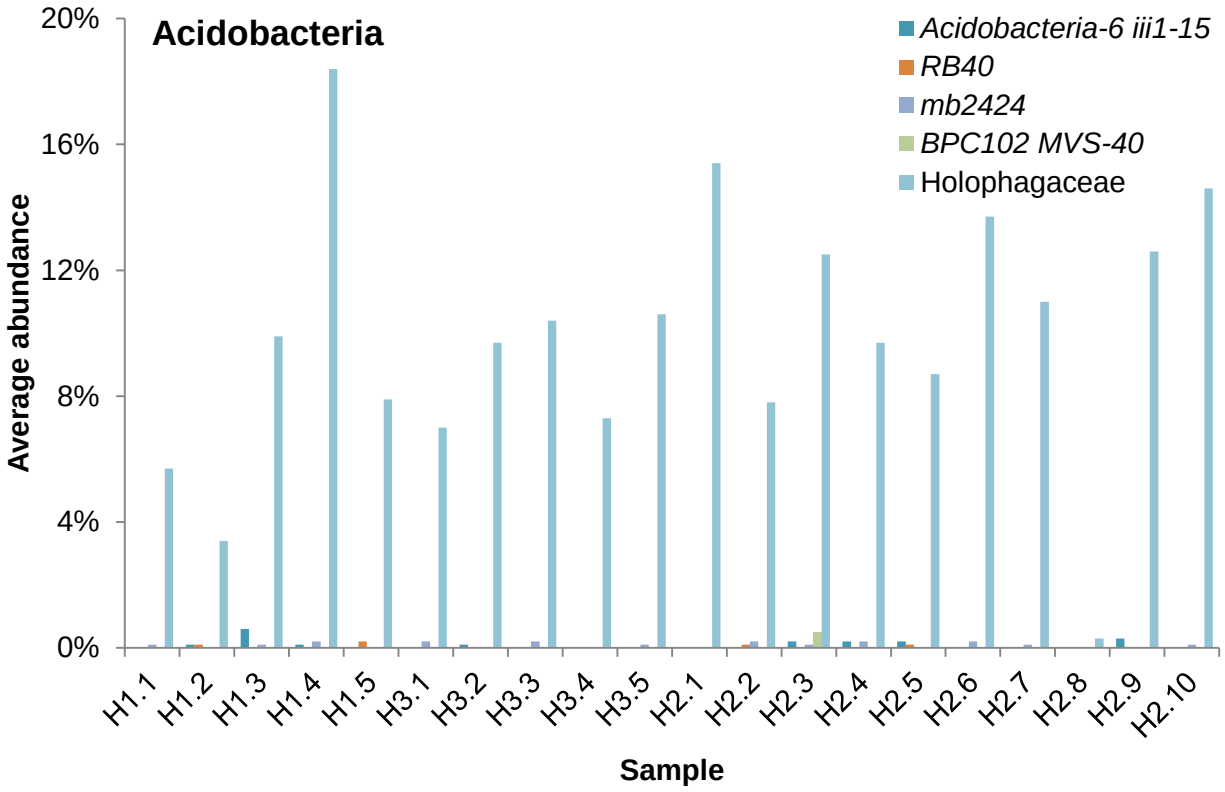

Cyanobacteria
Average abundance
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
35%
30%
25%
20%
15%
10%
5%
0%
H1.1 H1.3 H1.5 H3.2 H3.4 H2.1 H2.3 H2.5 H2.7 H2.9
H3.1 H3.2 H3.3 H3.4 H3.5 H2.1 H2.2 H2.3 H2.4 H2.5 H2.6 H2.7 H2.8 H2.9 H2.10
Sample
Chloroplast
Chlorophyta
Cryptophyta
Stramenopiles
Synechococcaceae


Cytophagaceae
Sphingobacteriaceae
Chitinophagaceae
Average abundance
30%
20%
10%
0%
H1.1 H1.2 H1.3 H1.4 H1.5 H3.1 H3.2 H3.3 H3.4 H3.5 H2.1 H2.2 H2.3 H2.4 H2.5 H2.6 H2.7 H2.8 H2.9 H2.10
Sample

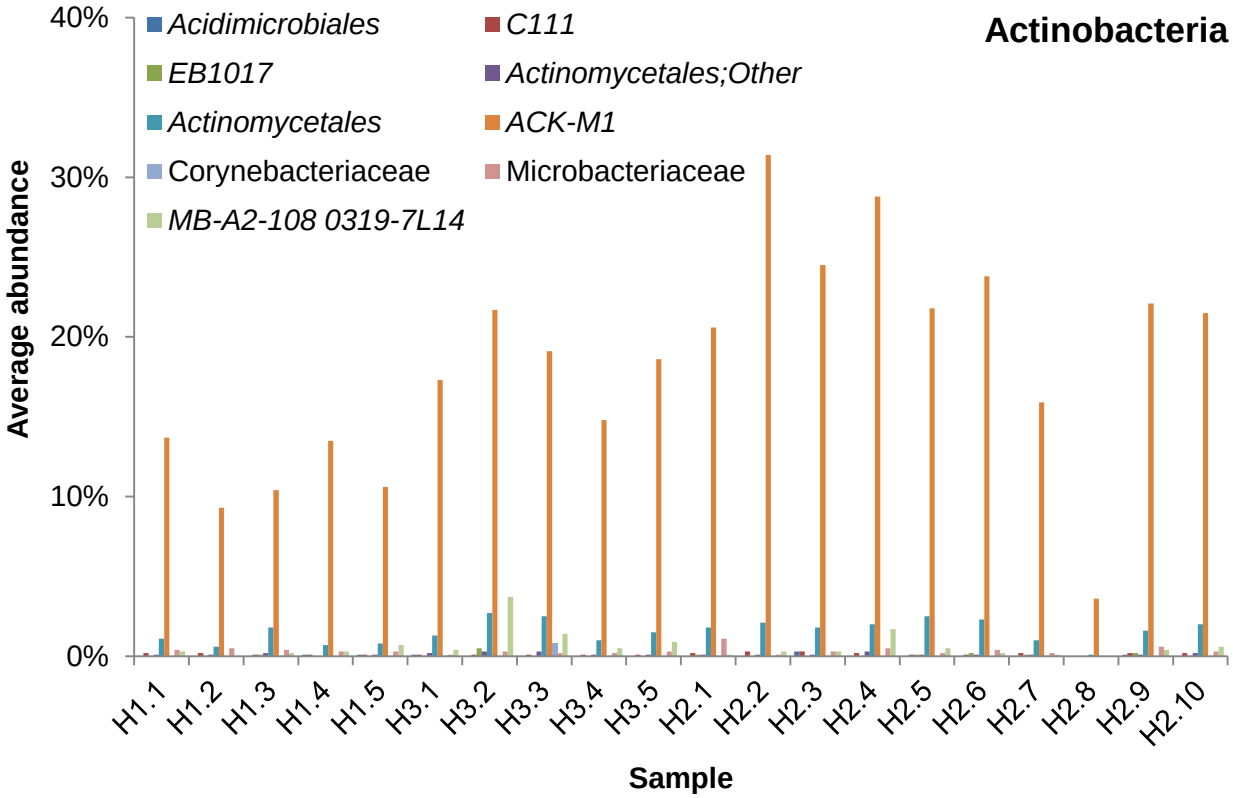
C

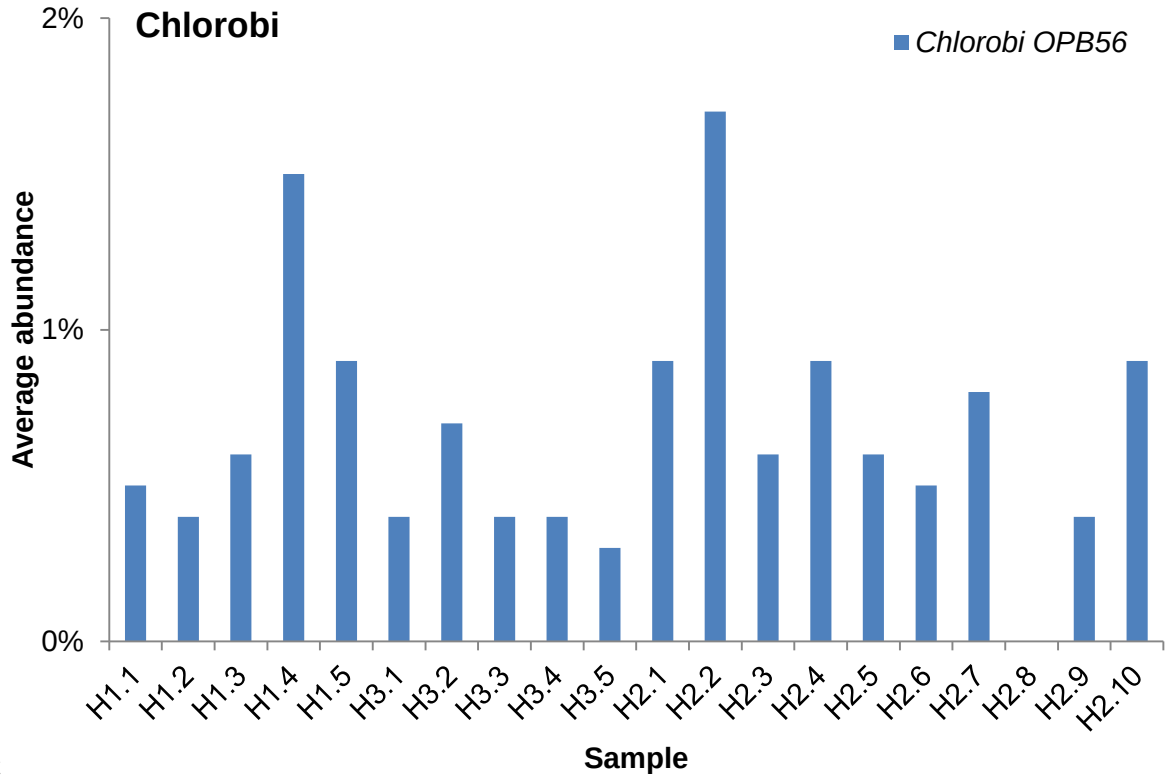
Figure S1. Familial profiles of the clustered phyla **A** Proteobacteria, inset shows inclusion of H2.8 with an abundance of Pseudomonadaceae (75.7%) and Oxalobacteraceae (13.8%), **B** Cyanobacteria, inset shows inclusion of clustered H1 samples, and **C** Bacteroidetes, within the hotspot replicates.

A



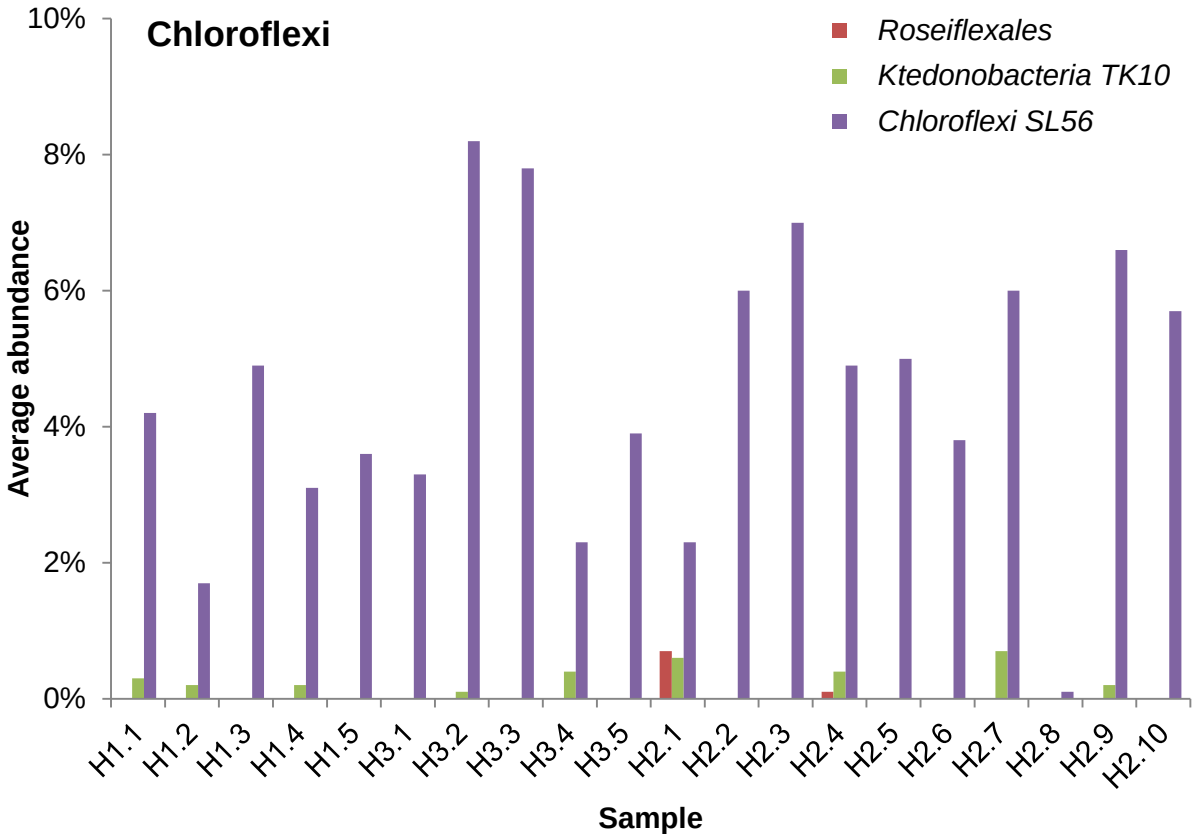
B


Actinobacteria
Acidimicrobiales
C111
EB1017
Actinomycetales;Other
Actinomycetales
ACK-M1
Corynebacteriaceae
Microbacteriaceae
MB-A2-108 0319-7L14
40%
30%
20%
10%
0%
Average abundance
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample

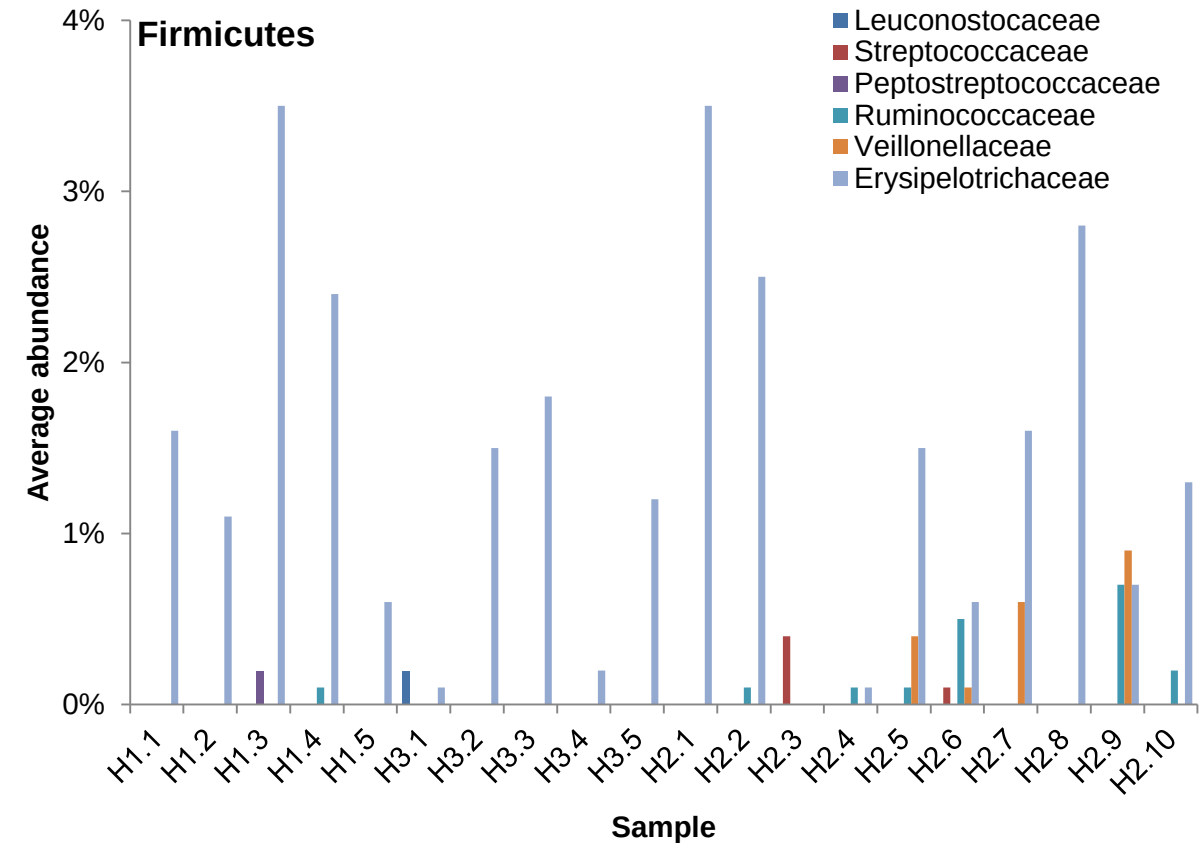

Chlorobi
Chlorobi OPB56
2%
1%
0%
Average abundance
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample

C

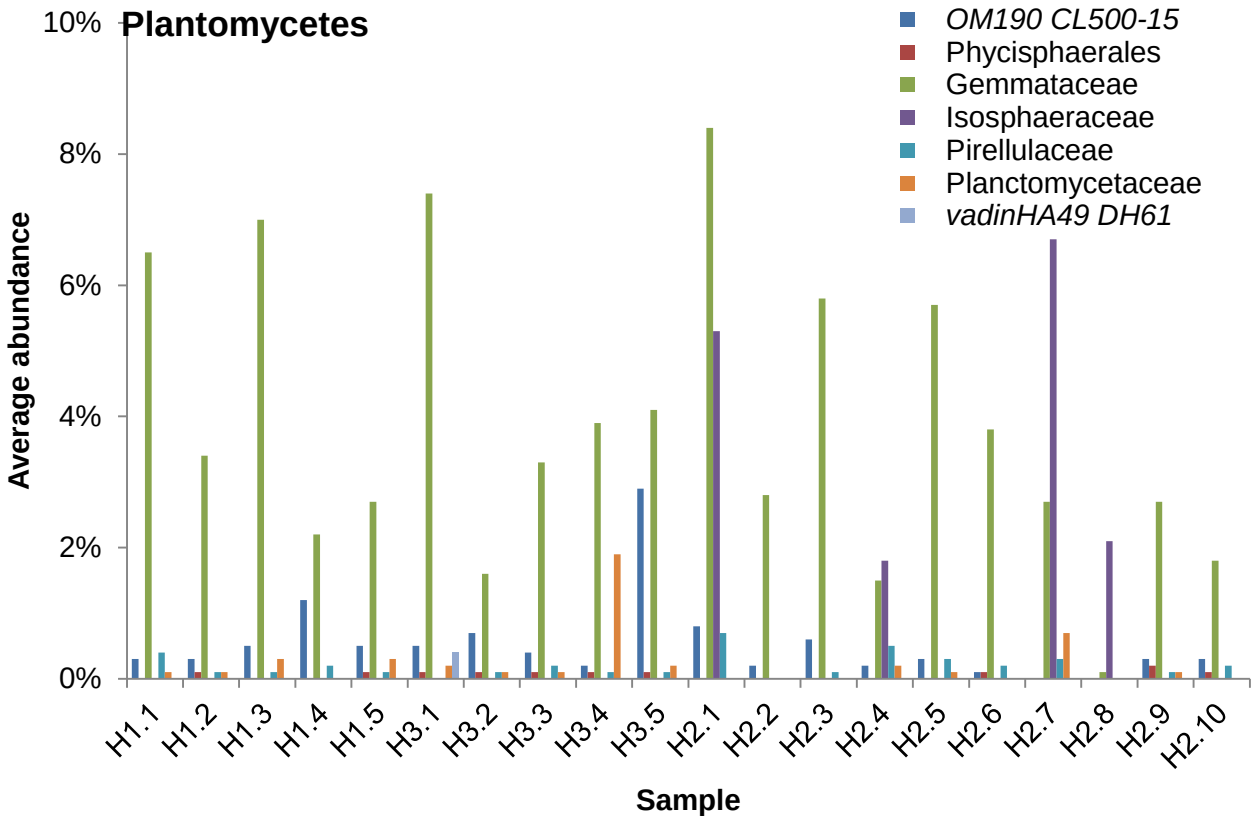
D

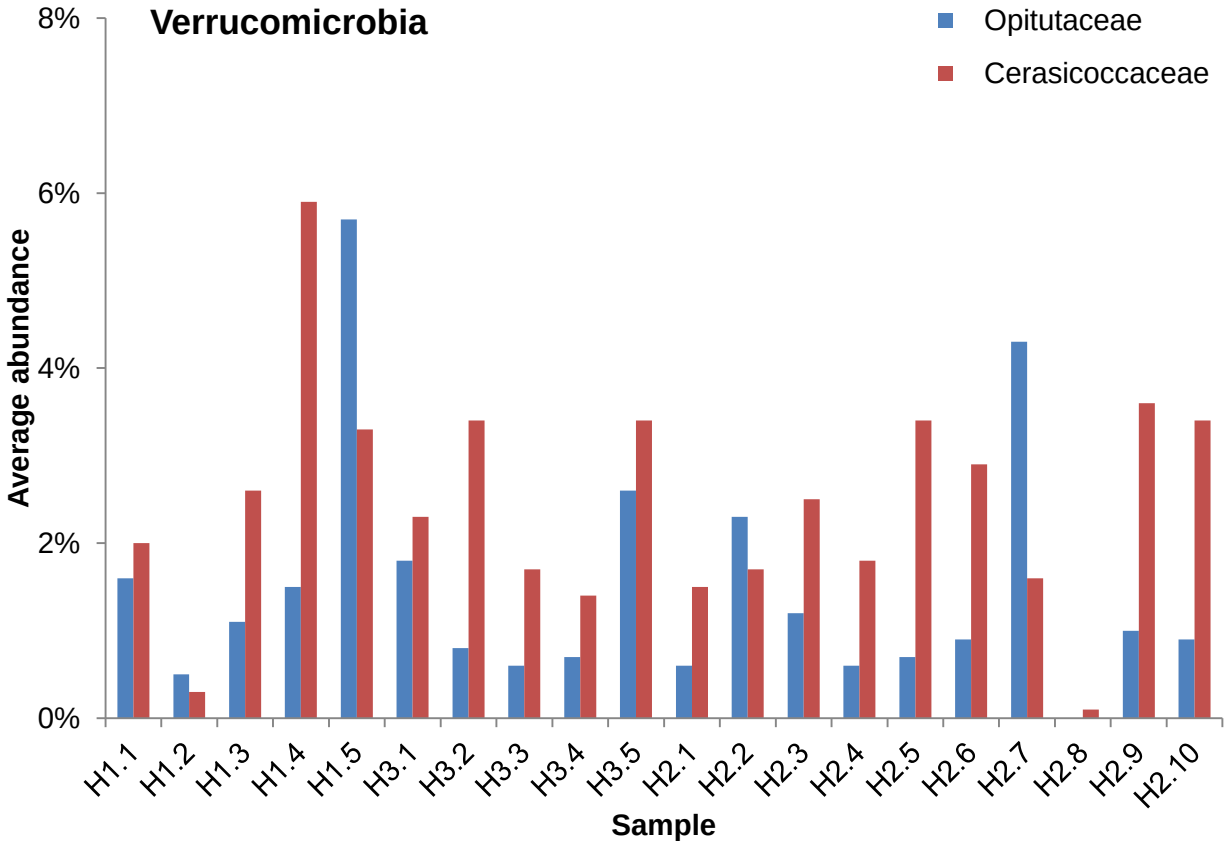

Chloroflexi
Roseiflexales
Ktedonobacteria TK10
Chloroflexi SL56
Average abundance
10%
8%
6%
4%
2%
0%
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample

E


Firmicutes
Leuconostocaceae
Streptococcaceae
Peptostreptococcaceae
Ruminococcaceae
Veillonellaceae
Erysipelotrichaceae
Average abundance
4%
3%
2%
1%
0%
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample

F


Plantomycetes
10%
8%
6%
4%
2%
0%
Average abundance
OM190 CL500-15
Phycisphaerales
Gemmataceae
Isosphaeraceae
Pirellulaceae
Planctomycetaceae
vadinHA49 DH61
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample


Verrucomicrobia
8%
6%
4%
2%
0%
Average abundance
Opitutaceae
Cerasicoccaceae
H1.1
H1.2
H1.3
H1.4
H1.5
H3.1
H3.2
H3.3
H3.4
H3.5
H2.1
H2.2
H2.3
H2.4
H2.5
H2.6
H2.7
H2.8
H2.9
H2.10
Sample

G

Figure S2. Familial profiles of the remaining identified phyla, **A** Acidobacteria, **B** Actinobacteria, **C** Chlorobi, **D** Chloroflexi, **E** Firmicutes, **F** Plantomycetes and **G** Verrucomicrobia, which accounted for ≥ 1.70% of the average abundance within the hotspot samples.

Table S1. Total prokaryotic abundances determined via flow cytometry.

Sample	**Abundance** x 10^7 cells ml^{-1} (95%CI, n)
H1	4.5 (1.9 x 10^7)
H2	4.6 (0.7 x 10^7)
H3	4.3 (0.8 x 10^7)
C1	0.92 (1.7 x 10^6)
C2 and C3	0.94 (3.3 x 10^6)
B1, B2 and B3	2.7 (0.9 x 10^7)

Table S2. SIMPER similarity comparisons between subsamples

Samples	**SIMPER Similarity**	
	Abundance	**Genera with heightened abundance removed**
C2	70.3	70.7
C3	66.0	66.3
C1	71.6	71.8
H1	57.5	58.2
H3	61.2	61.6
H2	53.3	53.9
B2	74.6	75.2
B3	74.1	74.9

B1	76.2	76.6

Table S3. SIMPER similarity comparisons between samples

	SIMPER Dissimilarity	
Samples	**Abundance**	**Genera with heightened abundance removed**
C2 & C3	31.7	31.4
C2 & C1	29.2	28.8
C3 & C1	31.4	31.2
C2 & H1	47.2	47.1
C3 & H1	48.5	48.3
C1 & H1	46.1	45.8
C2 & H3	45.5	44.8
C3 & H3	47.3	46.5
C1 & H3	45.2	44.3
H1 & H3	42.9	41.8
C2 & H2	48.4	48.2
C3 & H2	49.4	49.1
C1& H2	47.3	46.9
H1 & H2	46.5	45.9
H3 & H2	43.4	42.7
C2 & B2	40.6	40.2
C3 & B2	42.0	41.8
C1 & B2	39.1	38.9
H1 & B2	49.8	49.4
H3 & B2	50.0	49.0
H2 & B2	52.3	51.8
C2 & B3	40.5	40.4
C3 & B3	41.8	41.6

C1 & B3	39.2	39.0
H1 & B3	50.1	49.9
H3 & B3	50.9	49.8
H2 & B3	53.1	52.7
B2 & B3	26.6	25.9
C2 & B1	39.1	39.0
C3 & B1	40.5	40.3
C1 & B1	38.3	38.0
H1 & B1	49.1	48.9
H3 & B1	49.9	48.8
H2 & B1	51.5	51.0
B2 & B1	27.2	26.6
B3 & B1	25.5	25.0

Chapter 7

Discussion

OVERVIEW

The research undertaken in this thesis provides new insights into the spatial distributions of microbial communities within marine and freshwater systems. The overall finding within this thesis is that microscale microbial patchiness is higher than previously assumed. This patchiness led to dissimilarity between samples from the same site within large scale studies and high fold changes in microbial abundance in sub-centimetre scale studies of marine and freshwater interface microenvironments. This patchiness agrees with previous findings in early studies by Bainbridge (1957), Bérzins (1958), Birge (1897), Langford (1938), Malone & McQueen (1983), Riley (1963) and Sutcliffe *et al.* (1963) showing patchiness within plankton distributions, as well as recent microscale work by Doubell *et al.* (2006), Seuront *et al.* (2007), Seymour *et al.* (2004, 2005ab, 2006, 2007, 2008) and Waters *et al.* (2003) confirming patchiness in bacterial, viral and phytoplankton communities. However, the research in this thesis builds on this previous work by adding two-dimensional and replicated microscale sampling, which allowed the differentiation of single- and multiple-point maxima as well as gradients and hence revealed patchiness levels that were higher than observed previously. We extended our investigations to look at the taxonomic composition of this patchiness and revealed distinct taxonomic microenvironments within microscale microbial hotspots, coldspots and background regions at the sub-centimetre scale. This final chapter will summarise the major findings of each chapter within the context of the specific aims outlined in Chapter 1.

RESULTS SYNTHESIS

- *Persistence, loss and appearance of river bacterial phylotypes upstream and downstream*

Chapter 2 investigated the bacterial community composition upstream and 3.3 km downstream of a small rural town using the Murray River as a model system. We reported the persistence, loss and appearance of bacterial genotypes upstream and 3.3 km downstream. Persistence was observed as site similarity of dominant archetypal freshwater genotypes, indicating information lengths and colonisation distances equal to or greater than 3.3 km for some bacterial taxa. The loss of microbial genotypes was observed as site exclusivity upstream and downstream and was attributed to mechanisms involved in the kill the winner hypothesis where bacterial genotypes with high growth rates were selected against via viral sweeps and protistan grazing. The appearance of microbial genotypes was observed downstream and was suggestive of immigration from agricultural and domestic land, indicated via the appearance of human- and/or animal-associated microbial genotypes. The persistence of bacterial genotypes has been observed previously in fluvial systems; however focus has primarily been on faecal indicator bacteria, with indigenous bacterial genotypes receiving minimal attention (Eleria & Vogel 2005; Fries *et al.* 2006, 2008; Kreader 1998; Traister & Anisfield 2006). Our results here suggest three fates for bacterial genotypes within river systems, persistence, loss and appearance, with each potentially having a significant impact on system function.

- *Sample dissimilarity in larger scale studies indicating microscale heterogeneity*

In Chapter 2, the investigation of the microbial community composition upstream and 3.3 km downstream of small rural town of Murray Bridge along the Murray River revealed dissimilarity between samples from the same site due to microscale heterogeneous hotspots of discrete Firmicutes and Cyanobacteria species. Local species sorting is suggested as the dominant mechanism for these discrete microenvironments. In addition, high dispersion was observed at the downstream site, hence indicating microbial patchiness exists in freshwater systems. As the primary taxonomic differences between the upstream and downstream sites were due to dissimilarity between samples from the downstream site, this indicates microscale patchiness may impact taxonomic diversity profiles more so than taxonomic patterns over larger distances, here 3.3 km. These results validate the bias of using the “mean field approach” to estimate microbial community dynamics in freshwater systems and suggest the necessity of microscale microbial investigations within this patchy river system.

- *Microscale abundance patchiness*

Following on from the patchiness observed in Chapter 2, in Chapter 3 we performed initial investigations into the microscale abundance variation of virio- and bacterioplankton populations at interface microenvironments within marine systems. Here, via the use of a novel two-dimensional sampler and extensive replication we showed spatial heterogeneity in microbial abundance within the 11.3 cm directly above the sediment-water interface of an estuary and mangrove environment. This heterogeneity was observed as single and multiple point hotspots at Noarlunga

estuary and hotspots and surface gradients at St Kilda mangroves, which resulted in high levels of viral and bacterial abundance near the sediment-water interface. These microbial distribution patterns have been observed previously in marine systems but to a lesser extent. For instance, at St Kilda, bacterial and viral abundance gradients exhibited 45- and 10.5-fold change per centimetre, respectively, compared to the previously reported 1.5-fold change in abundance per centimetre in bacteria (Seymour *et al.* 2007b) and 1.4-fold change per centimetre in viruses (Seymour *et al.* 2006, 2007b). At Noarlunga, bacterial and viral abundance showed a 12.9– and 2584-fold change per centimetre. If lower resolution and one-dimensional sampling was employed these large fold-changes per centimetre would have been missed, therefore highlighting the importance of analysing microbial spatial distribution at the level individual cells experience rather than using a 'mean field' approach. In addition, this high microscale heterogeneity illustrates the necessity and applicability of two-dimensional high resolution sampling, as used throughout this thesis, as this sampler allowed single and multiple point hotspot and/or surface gradient discrimination, which was not possible in past studies as one-dimensional samplers were employed.

Despite hotspots being ubiquitous features of microbial distributions, past studies lacked a universal definition. They qualitatively classified patches as regions of high microbial abundance that 'exceed' or are 'elevated above' background values (Seymour *et al.* 2004; 2005a, 2008). In this thesis we provide a new quantitative method for determining hotspots, coldspots and background values within microscale microbial distributions, which will allow comparative analyses among future microbial microscale distribution studies.

Following on from this initial study was the analysis of microscale abundance distributions at interface microenvironments in freshwater systems in Chapter 4. Here, we report microscale patchiness in freshwater microbial communities with abundance variations of 107 and 80.5 fold per centimetre for virio- and bacterioplankton subpopulations, respectively. This indicates microbial interactions within freshwater systems are likely to differ markedly at the microscale. This patchiness is consistent with observations in marine systems for variation caused by mixing and aggregation, and therefore suggests, as with marine systems, that bulk phase sampling will not provide accurate representation of the dynamics of microbial processes within freshwater systems.

- *Microscale taxonomic patchiness*

Chapter 5 analysed the taxonomic composition of microscale hotspots, coldspots and background regions within prokaryotic communities found at the sediment-water interface of the Murray River. These results follow on from the abundance patchiness observed in Chapter 4. The results of Chapter 5 showed microscale microbial hotspots and coldspots contain microenvironments that are distinct from the taxonomy of the background community at the sub-centimetre scale. Hotspots were characterised by heightened abundance in discrete genera, with these genera differing between hotspots. Discrete genera showing increased abundance, *Parasporobacterium, Lachnospiraceae incertae sedis, Pseudomonas* and *Bacteroides,* were indicative of animal and/or human inputs indicating a potential anthropogenic effect on microbial patchiness. These findings indicate heterogeneous genus richness and composition exists within a river system, as hotspots and coldspots represent distinct taxonomic microenvironments rather than temporary abundance increases or decreases of the overall community. Therefore, we suggest

small scale fluid parcels persist together long enough to build up genera numbers and create distinctly different proximate taxonomic microenvironments. As it is at the microscale that important microbial interactions occur, the presence of these distinct microenvironments will therefore have a significant impact on microbial biogeochemical processes.

Chapter 6 followed on from the analyses in Chapter 5 to determine whether heterogeneity existed between 1 µl subsamples from hotspot, coldspot and background regions. The results indicated heterogeneity between 1 µl subsamples taken from each region due to increased abundance of specific genera, with this heterogeneity most extreme within hotspot subsamples. Specifically, *Parasporobacterium* and *Lachnospiraceae incertae sedis*, were common amongst all regions whilst *Pseudomonas* and *Bacteroides* were only dominant within hotspot subsamples. Furthermore, *Pseudomonas* and *Bacteroides* caused a shift in dominant genera within hotspot subsamples suggesting exploitation by these genera. Subsamples lacking specific genera increases showed similarities in abundant genera, primarily *Geothrix*, *Nocardioides*, *Flexibacter*, *Kaistia*, *Austwickia* and *Chryseoglobus*, inferring a common taxonomic background and hence indicating genera capable of persistence. Subsample heterogeneity revealed two discrete taxonomic patterns within hotspot subsamples: the dominance of specific individual genera or overall increases in genera common to most subsamples. Shifts observed in the dominant genera of some hotspot subsamples caused distinct taxonomic profiles compared to other subsamples as the whole taxonomic profile changed due to this dominance. Therefore these findings further support that microscale microbial hotspots represent discrete microenvironments, which will have important

implications for nutrient exchange and cellular interactions and hence overall system function.

FUTURE DIRECTIONS

Chapter 2 observed three fates (persistence, loss and appearance) in bacterial genotypes upstream and downstream of Murray Bridge. We suggested these fates are reflected in information lengths and colonisation distances. Future investigations using this data could analyse the effect of method sensitivity on these fates, specifically appearance and loss, as the removal of singletons during quality filtering of taxonomic analysis could have led to specific bacterial genotypes, particularly those with low sequence counts, becoming statistically extinct rather than indicating a true ecological loss. In addition, future directions could involve investigations focussed on genotypes of importance for ecological processes, to enable larger scale temporal and distance analysis of taxa persistence.

Chapters 3 and 4 observed microscale patchiness that was higher than previously reported. This patchiness in virio- and bacterioplankton abundance was due to hotspots, coldspots or surface gradients. Future studies would include nutrient concentration measurements at the microscale to determine whether these patchy regions are correlated with specific nutrient sources. As hotspots in prokaryotic abundance have been attributed to bacterial chemotactic response to high nutrient concentration regions, knowing the concentrations of such food sources, in addition to the taxonomic composition of the microbial communities, would enable a complete understanding of the mechanisms causing these abundance hotspots (Barbara & Mitchell 2003; Blackburn *et al.* 1998; Seymour *et al.* 2005a, 2006, 2007b; Stocker & Seymour 2012). As coldspots are indicative of removal processes, such as viral lysis

or grazing, each which releases organic matter from cell debris, future work measuring nutrients on the microscale would enable a deeper understanding of the nutrient and microbial interactions of such microenvironments (Weinbauer & Höfle 1998). Similarly, as surface gradients are the result of organic matter sinking and its reintegration into the benthos, followed by microbial degradation and transformation leading to increased nutrient concentrations and hence high microbial concentrations directly above the interface (Seymour *et al.* 2007b), knowing the nutrient concentrations within these microscale gradients would enable potential relationships between nutrient point sources and microbial taxonomy to be determined.

Chapter 5 identified the taxonomic composition of microbial hotspots and coldspots. As aforementioned, reduction in microbial communities are attributed to removal processes via viral or grazer action. Identifying the viral taxonomic composition of virioplankton hotspots and coldspots, would enable investigations into whether these microenvironments are also distinct viral taxonomic regions like was observed in the bacterial patches, and hence will indicate whether increased or depleted microscale viral abundance zones are due to specific viral taxa or whether these are overall increases or decreases in all available viral taxa.

CONCLUSION

The results from this thesis provide a detailed investigation on the distributional dynamics of virio- and bacterioplankton within aquatic environments of South Australia. Specifically, flow cytometric enumeration and primer-based 16S sequencing was employed to determine the large scale and microscale abundance and taxonomic distributions of bacterial communities within marine and freshwater

systems, with particular focus given to the Murray River as a model system. To our knowledge, this is the first investigation of the indigenous bacterial community of the Murray River system. Here, the results show that microscale bacterial patchiness is higher than previously assumed in marine and freshwater systems. Microscale patchiness was observed in marine systems as abundance variations of 2584- and 45–fold per centimetre in virio- and bacterioplankton populations, whilst in freshwater, patchiness showed abundance variations of 107-and 80.5-fold per centimetre in virio- and bacterioplankton populations, respectively. This high microscale heterogeneity in abundance confirms that bulk phase sampling does not represent microbial processes in marine or freshwater systems.

In addition, taxonomic analysis of bacterial hotspots and coldspots confirmed that hotspots and coldspots represent distinct taxonomic microenvironments at the sub-centimetre scale, rather than temporary abundance increases or decreases of the overall community, with hotspots characterised by discrete abundance increases in genera. These findings suggest small scale fluid parcels persist together long enough to build up genera numbers and create distinctly different proximate taxonomic microenvironments, which will thereby have implications for microscale biogeochemical processes such as the transformation of carbon, nitrogen and sulphur (Azam 1998; Paerl & Pinckney 1996; Stocker *et al.* 2008). In conclusion, the microscale microbial abundance and taxonomic patchiness found in this thesis may have important implications for nutrient cycling, and hence system function, within aquatic environments.

REFERENCES

Adams HE, Crump BC & Kling GW (2014) Metacommunity dynamics of bacteria in an arctic lake: the impact of species sorting and mass effects on bacterial production and biogeography. *Frontiers in Microbiology* 5 (82): 1-10

Alauzet C, Marchandin H, Courtin P, Mory F, Lemée L, Pons JL, Chapot-Chartier MP, Lozniewski A & Jumas-Bilak E (2014) Multilocus analysis reveals diversity in the genus *Tissierella:* Description of *Tissierella carlieri* sp. nov. in the new class *Tissierellia* classis nov. *Systematic and Applied Microbiology* 37: 23-34

Alonso MC, Jimenez-Gomez F, Rodriguez J, Borrego JJ (2001) Distribution of virus-like particles in an oligotrophic marine environment (Alboran Sea, Western Mediterranean). *Microbial Ecology* 42: 407–415.Anderson DM, Glibert PM & Burkholder JM (2002) Harmful algal blooms and eutrophication: nutrient sources, composition and consequences. *Estuaries* 25 (4b): 704-726

Anderson MJ, Gorley RN & Clarke KR (2008) PERMANOVA+ for PRIMER: Guide to software and statistical methods. PRIMER-E Ltd. Plymouth: Plymouth Marine Laboratory, pp. 87-122

Andersson AF, Lindberg M, Jakobsson H, Bäckhed F, Nyrén P & Engstrand L (2008) Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS One* DOI: 10.1371/journal.pone.0002836

Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, et al. (2006) The marine viromes of four oceanic regions. *PLoS Biology* 4: 2121–2131.

Azam F (1998) Microbial control of oceanic carbon flux: the plot thickens. *Science* 280: 694–696.

Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nature* 5: 782–791.

Bainbridge R (1957) The size, shape and density of marine phytoplankton concentrations. *Biological Reviews* 32 (1): 91-115

Ballesté E & Blanch AR (2010) Persistence of *Bacteroides* species populations in a river as measured by molecular and culture techniques. *Applied and Environmental Microbiology* 76 (22): 7608-7616

Barbara G, Mitchell JG (1996) Microlaminations of motile bacteria in microbial mats. *Applied and Environmental Microbiology* 62: 3985–3990.

Barbara G, Mitchell JG (2003) Marine bacterial organisation around point-like sources of amino acids. *FEMS Microbiology Ecology* 43: 99–109.

Baric A, Kuspilic G, Matijevic S (2002) Nutrient (N, P, Si) fluxes between marine sediments and water column in coastal and open Adriatic. *Hydrobiologia* 475/ 476: 151–159.

Bassis CM, Tang AL, Young VB & Pynnonen MA (2104) The nasal cavity microbiota of healthy adults. *Microbiome* 2 (27): 1-5

Bergh O, Borsheim KY, Bratbak G, Heldal M (1989) High abundance of viruses found in aquatic environments. *Nature* 340: 467–468.

Bérzins B (1958) Ein planktologisches Querprofil. *Rep. Ist. Freshwat. Res.* Drottningholm 39: 5–22

Bettarel Y, Amblard C, Sime-Ngando T, Carrias JF, Sargos D, Garabétian F & Lavandier P (2003) Viral lysis, flagellate grazing potential and bacterial production in Lake Pavin. *Microbial Ecology* 45: 119-127

Birge EA (1897) Plankton studies on Lake Mendota. II. The crustacea of the plankton in July, 1894 – December 1896. *Trans. Wis. Acad. Sci. Arts Lett.* 11: 274–448.

Blackburn N, Fenchel T, Mitchell JG (1998) Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria. *Science* 282: 2254–2256.

Blackburn N & Fenchel T (1999) Influence of bacteria, diffusion and shear on micro-scale nutrient patches, and implication for bacterial chemotaxis. *Marine Ecology Progress Series* 189: 1-7

Boegman L, Ivey GN (2009) Flow separation and resuspension beneath shoaling nonlinear internal waves. *Journal of Geophysical Research*. 114: 1–15.

Bond NR, Lake PS & Arthington AH (2008) The impacts of drought on freshwater ecosystems: an Australian perspective. *Hydrobiologia* 600: 3-16

Bouvy M, Troussellier M, Got P, Arfi R (2004) Bacterioplankton responses to bottom-up and top-down controls in a West African reservoir (Selingue, Mali). *Aquatic Microbial Ecology* 34: 301–307

Braaten DA, Paw KT, Shaw RH (1990) Particle resuspension in a turbulent boundary layer – observed and modelled. *Journal of Aerosol Science* 21: 613– 628.

Breitbart M (2012) Marine viruses: truth or dare. *Annual Review of Marine Science* 4: 425-448

Brentnall SJ, Richards KJ, Brindley J, Murphy E (2003) Plankton patchiness and its effect on larger-scale productivity. *Journal of Plankton Research* 25: 121–140.

Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, Riddle MJ, Fuhrman JA, Andrews-Pfannkoch C, Hoffman JM, McQuaid JB, Allen A, Rintoul SR & Cavicchioli R (2012) Global biogeography of SAR11 marine bacteria. *Molecular Systems Biology* 8 (595):1-13

Brussaard C, Marie D, Bratbak G (2000) Flow cytometric detection of viruses. *Journal of Virological Methods* 85 (1–2): 175–182.

Brussaard C (2004) Optimisation of procedures for counting viruses by flow cytometry. *Applied and Environmental Microbiology* 70: 1506–1513.

Burns A & Walker KF (2000) Biofilms as food for decapods (Atyidae, Palaemonidae) in the River Murray, South Australia. *Hydrobiologia* 437: 83-90

Caporaso GJ, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Peña A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J & Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7 (5): 335-336

Carrias JF & Sime-Ngando T (2009) 'Bacteria, attached to surfaces' in Likens GE (ed.), *Plankton of Inland Waters*, Elsevier pp. 182-192

Chen F, Lu J, Binder BJ, Liu Y & Hodson RE (2001) Application of digital image analysis and flow cytometry to enumerate marine viruses stained with SYBR Gold. *Applied and Environmental Microbiology* 67: 539-545

Chow CET, Kim DY, Sachdeva R, Caron DA & Fuhrman JA (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *The International Society for Microbial Ecology (ISME) Journal* 8: 816-829

Clarke, KR (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18: 117-143.

Clarke KR & Gorley RN (2006) PRIMER v6: User manual/tutorial. PRIMER-E Ltd. Plymouth: Plymouth Marine Laboratory, pp. 140-142

Clauset A, Shalizi CR & Newman MEJ (2009) Power-law distributions in empirical data. *SIAM Review.* 51 (4): 661-703

Coates JD, Ellis DJ, Gaw CV, Lovley DR (1999) *Geothrix fermentans* gen. nov., sp. nov., a novel Fe(III)-reducing bacterium from a hydrocarbon-contaminated aquifer. *International Journal of Systematic Bacteriology* 49: 1615-1622

Cochlan WP, Wikner J, Steward GF, Smith DC & Azam F (1993) Spatial distribution of viruses, bacteria and chlorophyll *a* in neritic, oceanic and estuarine environments. *Marine Ecology Progress Series* 92: 77-87

Corinaldesi C, Crevatin E, Del Negro P, Marini M, Russo A, et al. (2003) Largescale spatial distribution of virioplankton in the Adriatic Sea: testing the trophic state control hypothesis. *Applied and Environmental Microbiology* 69: 2664–
2673.

Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM & Relman DA (2012) The application of ecological theory towards an understanding of the human microbiome. *Science.* 336: 1255-1262

Cotner JB, Biddanda BA (2002) Small players, large Role: microbial influence on biogeochemical processes in pelagic aquatic ecosystems. Ecosystems 5: 105– 121.

Crump BC, Kling GW, Bahr M & Hobbie JE (2003) Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Applied and Environmental Microbiology* 69 (4): 2253-2268

Curtis TP, Head IM, Lunn M, Woodcock S, Schloss PD, et al. (2006) What is the extent of prokaryotic diversity? *Philosophical Transactions of the Royal Society* 361: 2023–2037.

Dann LM, Mitchell JG, Speck PG, Newton K, Jeffries T & Paterson J (2014) Virio- and bacterioplankton microscale distributions at the sediment-water interface. *PLoS One* 9 (7): 1-14

Del Giorgio PA (2000) Using flow cytometry for counting natural planktonic bacteria

and understanding the structure of planktonic bacterial communities. *Scientia Marina* 64: 197-224

De Menezes AB, Lewis E, O'Donovan M, O'Neill BF, Clipson N & Doyle EM (2011) Microbiome analysis of dairy cows fed pasture or total mixed ration diets. *FEMS Microbial Ecology* 78 (2): 256-265

Department of Environment, Water and Natural Resources (DEWNR) (2014), River Murray Flow Report Public I2 A2, Government of South Australia pp. 1-4

Donlan RM (2002) Biofilms: Microbial Life on Surfaces. *Emerging Infectious Diseases - Perspective.* 8 (9): 881-890

Doubell MJ, Seuront L, Seymour JR, Patten NL & Mitchell JG (2006) High-resolution fluorometer for mapping microscale phytoplankton distributions. *Applied and Environmental Microbiology* 72 (6): 4475-4478

Duarte CM & Vaqué D (1992) Scale dependence of bacterioplankton patchiness. *Marine Ecology Progress Series* 84 (1): 95-100

Durham WM, Kessler JO & Stocker R (2009) Disruption of vertical motility by shear triggers formation of thin phytoplankton layers. *Science* 323: 1067-1070

Durham WM & Stocker R (2012) Thin phytoplankton layers: Characteristics, mechanisms, and consequences. *Annual Reviews of Marine Science* 4: 177-207

Hondzo M & Lyn D (1999) Quantified small-scale turbulence inhibits the growth of a green alga. *Freshwater Biology* 41: 51-61

Drake LA, Choi KH, Haskell AGE, Dobbs FC (1998) Vertical profiles of viruslike particles and bacteria in the water column and sediments of Chesapeake Bay, USA. *Aquatic Microbial Ecology* 16: 17–25.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST, *Bioinformatics* 26 (19), 2460-2461.doi: 10.1093/bioinformatics/btq461

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 15;27(16): 2194-200

Edgar RC (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10 (10): 996-998

Hibbing ME, Fuqua C, Parsek MR & Peterson SB (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology* 8: 15-25

Edwards RA & Rohwer F (2005) Viral metagenomics. *Nature Reviews Microbiology* 3: 504-510

Eleria A & Vogel RM (2005) Prediciting faecal coliform bacteria levels in the Charles River, Massachusetts, USA. *Journal of the American Water Resources Association* 41 (5): 1195-1209

Elser JJ, Stabler BL, Hassett PR (1995) Nutrient limitation of bacterial growth and rates of bacterivory in lakes and oceans: a comparative study. *Aquatic Microbial Ecology* 9: 105–110.

Fenchel T (2001) Eppur si muove: many water column bacteria are motile. *Aquatic Microbial Ecology* 24 (2): 197-201

Finlay BJ, Maberly SC & Cooper JI (1997) Microbial diversity and ecosystem function. *OIKOS* 80: 209-213

Fries J, Characklis G & Noble R (2006) Attachment of faecal indicator bacteria to particles in the Neuse River estuary, N.C. *Journal of Environmental Engineering* 132 (10): 1338-1345

Fries JS, Characklis GW & Noble RT (2008) Sediment-water exchange of *Vibrio* sp. and faecal indicator bacteria: Implications for persistence and transport in the Neuse River estuary, North Carolina, USA. *Water Research* 42 (4-5): 941-950

Fuhrman JA (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* 399: 541-548

Gasol JM, Del Giorgio PA (2000) Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina* 64: 197–224.

Gasol JM, Zweifel UL, Peters F, Fuhrman JA, Hagstrom A (1999) Significance of size and nucleic acid content heterogeneity as measured by flow cytometry in natural planktonic bacteria. *Applied and Environmental Microbiology* 65: 4475–4483.

Gasol JM, Del Giorgio PA (2000) Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina* 64: 197–224.

Geary RC (1954) The contiguity ratio and statistical mapping. *The Incorporated Statistician* 5: 115–145.

Gehrke PC, Brown P, Schiller CB, Moffatt DB & Bruce AM (1995) River regulations and fish communities in the Murray-Darling river system, Australia. *Regulated Rivers: Research & Management* 11: 363-375

Gell P, Tibby J, Little F, Baldwin D & Hancock G (2007) The impact of regulation and salinization on floodplain lakes: the lower River Murray, Australia. *Hydrobiologia* 591: 135-146

Genin A, Greene C, Haury L, Wiebe P, Gal G, et al. (1994) Zooplankton patch dynamics: daily gap formation over abrupt topography. *Deep Sea Research Part I: Oceanographic Research Papers* 41 (5–6): 941–951

Glöckner FO, Zaichikov E, Belkova N, Denissova L, Pernthaler J, Pernthaler A & Amann R (2000) Comparative 16S rRNA analysis of lake bacterioplankton reveals

globally distributed phylogenetic clusters including an abundant group of Actinobacteria. *Applied and Environmental Microbiology* 66(11): 5053-5065

Goddard VJ, Baker AC, Davy JE, Adams DG, De Ville MM, Thackeray SJ, Maberly SC & Wilson WH (2005) Temporal distribution of viruses, bacteria and phytoplankton throughout the water column in a freshwater hypereutrophic lake. *Aquatic Microbial Ecology* 39: 211-223

Goldschmidt P, Ferreira CC, Degorge S, Benallaoua D, Boutboul S, Laroche L, Batellier L & Chaumeil C (2008) Rapid detection and quantification of *Propionibacteriaceae. British Journal of Ophthalmology* 93 (2): 258-262

Goss KF (2003) Environment flows, river salinity and biodiversity conservation: managing trade-offs in the Murray-Darling basin. *Australian Journal of Botany* 51: 619-625

Griffith DA (1987) Spatial autocorrelation: a primer. Resource publications in Geography. Association of American Geographers, Washington, D.C.

Grünbaum D (2012) The logic of ecological patchiness. *Interface Focus* 2: 150– 155.

Haggard BE, Stanley EH & Storm DE (2005) Nutrient retention in a point-source-enriched stream. *Journal of the North American Benthological Society* 24 (1): 29-47

Hahn MW & Höfle MG (1999) Flagellate predation on a bacterial model community: Interplay of size-selective grazing, specific bacterial cell size and bacterial community composition. *Applied and Environmental Microbiology* 65 (11): 4863-4872

Hassett DJ, Cuppoletti J, Trapnell B, Lymar SV, Rowe JJ, Yoon SS, Hilliard GM, Parvatiyar K, Kamani MC, Wozniak DJ, Hwang SH, McDermott TR & Ochsner UA (2002) Anaerobic metabolism and quorum sensing by *Pseudomonas aeruginosa*

biofilms in chronically infected cystic fibrosis airways: rethinking antibiotic treatment strategies and drug targets. *Advanced Drug Delivery Reviews* 54 (11): 1425-1443

Hillmer I, Reenen PV, Imberger J, Zohary T (2008) Phytoplankton patchiness and their role in the modelled productivity of a large, seasonally stratified lake. *Ecological Modelling* 218: 49–59.

Hollibaugh J (2001) Dynamic bacterial and viral response to an algal bloom at subzero temperatures. *Limnology and Oceanography* 46: 790-801

Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6: 65–70.

Hurwitz BL, Deng L, Poulos BT & Sullivan MB (2013) Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. *Environmental Microbiology* 15 (5): 1428-1440

Huws SA, Kim EJ, Lee MRF, Scott MB, Tweed JKS, Pinloche E, Wallace RJ & Scollan ND (2011) As yet uncultured bacteria phylogenetically classified as *Prevotella, Lachnospiraceae incertae sedis,* and unclassified *Bacteroidales, Clostridiales* and *Ruminococcaceae* may play a predominant role in ruminal biohydrogenation. *Environmental Microbiology* doi:10.1111/j.1462-2920.2011.02452.x

Jacquet S & Bratbak G (2004) Spring phytoplankton bloom dynamics in Norwegian coastal waters: Microbial community succession and diversity. *Limnology and Oceanography* 49: 180-190

Jacquet S, Domaizon I, Personnic S, Ram ASP, Hedal M, Duhamel S & Sime-Ngando (2005) Estimates of protozoan- and viral-mediated mortality of bacterioplankton in Lake Bourget (France). *Freshwater Biology* 50: 627-645

Jacquet S, Miki T, Noble R, Peduzzi P & Wilhelm S (2010) Viruses in aquatic

ecosystems: important advancements of the last 20 years and prospects for the future in the field of microbial oceanography and limnology. *Advances in Oceanography and Limnology* 1: 97-141

Jezbera J, Sharma AK, Brandt U, Doolittle WF & Hahn MW (2009) '*Candidatus* Planktophila limnetica', an actinobacterium representing one of the most numerically important taxa in freshwater bacterioplankton. *International Journal of Systematic and Evolutionary Microbiology* 59: 2864-2869

Jezbera J, Jezberová J, Koll U, Horňák K, Šimek K & Hahn MW (2012) Constrasting trends in distribution of four major planktonic betaproteobacterial groups along a pH gradient of epilimnia of 72 freshwater habitats. *FEMS Microbiology Ecology* 81: 467-479

John SG, Mendez CB, Deng L, Poulos B, Kauffman AKM, Kern S, Brum J, Polz MF, Boyle EA & Sullivan MB (2011) A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environmental Microbiology Reports* 3(2): 195-202

Jolly ID, Williamson DR, Gilfedder M, Walker GR, Morton R, Robinson G, Jones H, Zhang L, Dowling TI, Dyce P, Nathan RJ, Nandakumar N, Clarke r & McNeill V (2001) Historical stream salinity trends and catchment salt balances in the Murray-Darling Basin, Australia. *Marine and Freshwater Research* 52: 53-63

Jones SE, Newton RJ & McMahon (2008) Potential for atmospheric deposition of bacteria to influence bacterioplankton communities. *FEMS Microbiol Ecol* 64: 388-394

Joux F (2001) Does the high nucleic acid content of individual bacterial cells allow us to discriminate between active cells and inactive cells in aquatic systems. *Applied and Environmental Microbiology* 67: 1775–1782.

Jin L, Kim KK, Lee HG, Ahn CY & Oh HM (2012) *Kaistia defluvii* sp. nov., isolated from river sediment. *International Journal of Systematic and Evolutionary Microbiology* 62: 2878-2882

Karl DM, Dore JE (2001) Microbial ecology at sea: sampling, subsampling and incubation considerations. *Methods in Microbiology* 30: 13–39.

Karr JR & Chu EW (2000) Sustaining living rivers. *Hydrobiologia* 422/423: 1-14

Keller-Costa T, Jousset A, van Overbeek L, van Elsas JD & Costa R (2014) The freshwater sponge *Ephydatia fluviatilis* harbours diverse *Pseudomonas* species (*Gammaproteobacteria, Pseudomonadales)* with broad-spectrum antimicrobial activity. *PLOS One* DOI: 10.1371/journal.pone.0088429

Kester W (2009) What the Nyquist criterion means to your sampled data system design. *Analog Devices* 1–12.

Kiem AS & Verdon-Kidd DC (2011) Steps toward “useful” hydroclimatic scenarios for water resource management in the Murray-Darling Basin. *Water Resources Research* 47: 1-14

Kolmogorov AN (1991) Dissipation of energy in the locally isotropic turbulence. *Proceedings: Mathematical and Physical Sciences* 434: 15-17

Konopka A (2009) What is microbial community ecology? *The International Society for Microbial Ecology (ISME) Journal* 3: 1223-1230

Kreader CA (1998) Persistence of PCR-detectable *Bacteroides distasonis* from human faeces in river water. *Applied* Lebaron P, Servais P, Agogue H, Courties C,

Lamendella R, Santo Domingo JW, Oerther DB, Vogel JR & Stoeckel DM (2007) Assessment of fecal pollution sources in a small northern-plains watershed using PCR and phylogenetic analyses of *Bacteroidetes* 16s rRNA gene. *Federation of European Microbiological Societies* 59: 651-660

Langford RR (1938) Diurnal and seasonal changes in the distribution of limnetic crustacea in Lake Nipissing, Ontario*. Univ. Toronto Stud. Biol. Ser.* 45: 1–142.

Larsen A, Fonnes Flaten GA, Sandaa RA, Castberg T, Thyrhaug R, Erga SR, Jacquet S & Bratbak G (2004) Spring phytoplankton bloom dynamics in Norwegian coastal waters: Microbial community succession and diversity. *Limnology and Oceanography* 49: 180-190

Lear G, Bellamy J, Case BS, Lee JE & Buckley HL (2014) Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *The International Society for Microbial Ecology (ISME) Journal* 8: 1715-1726

Lebaron P, Servais P, Baudoux A, Bourrain M, Courties C, et al. (2002) Variations of bacterial-specific activity with cell size and nucleic acid content assessed by flow cytometry. *Aquatic Microbial Ecology* 28: 131–140.

Leff LG, McArthur JV & Shimkets LJ (1992) Information spiralling: movement of bacteria and their genes in streams. *Microbial Ecology* 24: 11-24

Levin S, Whitfield M (1994) Patchiness in marine and terrestrial systems: from individuals to populations. *Philosophical Transactions of the Royal Society* 343 (1303): 99–103.

Liu H, Imai K, Suzuki K, Nojiri Y, Tsurushima N, et al. (2002) Seasonal variability of picophytoplankton and bacteria in the western subarctic Pacific Ocean at station KNOT. *Deep Sea Research Part II: Topical Studies in Oceanography* 49: 24–25.

Liu Z, Huang S, Sun G, Xu Z & Xu M (2011) Phylogenetic diversity, composition and distribution of bacterioplankton community in the Dongjiang River, China. *FEMS Microbiology Ecology* 80: 30-44

Liu AC, Chou CY, Chen LL & Kuo CH (2015) Bacterial community dynamics in a swine wastewater anaerobic reactor revealed by 16s rDNA sequence analysis. *Journal of Biotechnology* 194: 124-131

Lomans BP, Leijdekkers P, Wesselink JJ, Bakkes P, Pol A, van der Drift C, Op den Camp HJM (2001) Obligate sulphide-dependent degradation of methoxylated aromatic compounds and formation of methanethiol and dimethyl sulphide by a freshwater sediment isolate, *Parasporobacterium paucivorans* gen. nov., sp. nov. *Applied Environmental Microbiology* 67 (9): 4017-4023

Long RA, Azam F (2001) Microscale patchiness of bacterioplankton assemblage richness in seawater. *Aquatic Microbial Ecology* 26: 103–113.

Maheshwari BL, Walker KF & McMahon (1995) Effects of regulation on the flow regime of the River Murray, Australia. *Regulated Rivers: Researach & Management* 10: 15-38

Maier HR, Burch MD & Bormans M (2001) Flow management strategies to control blooms of the cyanobacterium, *Anabaena circinalis,* in the River Murray at Morgan, South Australia. *Regulated Rivers: Research & Management* 17: 637-650

Malone BJ & McQueen DJ (1983) Horizontal patchiness in zooplankton populations in two Ontario kettle lakes. *Hydrobiologia* 99: 101-124

Maranger R & Bird DF (1995) Viral abundance in aquatic systems – a comparison between marine and freshwaters. *Marine Ecology Progress Series* 121: 217-226

Mariat D, Firmesse O, Levenez F, Guimarães VD, Sokol H, Doré J, Corthier G & Furet JP (2009) The *Firmicutes/Bacteroidetes* ratio of the human microbiota changes with age. *BMC Microbiology* 9 (123): 1-6

Marie D, Partensky F, Jacquet S, Vaulot D (1997) Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the

nucleic acid stain SYBR Green I. *Applied and Environmental Microbiology* 63: 186–193.

Marie D, Brussaard C, Thyrhaug R, Bratbak G, Vauolt D (1999) Enumeration of marine viruses in culture and natural samples by flow cytometry. *Applied and Environmental Microbiology* 65: 45–52.

Matz C & Jürgens K (2005) High motility reduces grazing mortality of planktonic bacteria. *Applied and Environmental Microbiology* 71 (2): 921-929

McArthur JV, Leff LG & Smith MH (1992) Genetic diversity of bacteria along a stream continuum. *Journal of the North American Benthological Society* 11 (3): 269-277

McArthur JV & Tuckfield RC (1997) Information length: spatal and temporal parameters among stream bacterial assemblages. *Journal of the North American Benthological Society* 16 (2): 347-357

McBride MJ & Zhu Y (2013) Gliding motility and por secretion system genes are widespread among members of the phylum *Bacteroidetes Journal of Bacteriology* 195 (2): 270-278

Meckenstock RU, von Netzer F, Stumpp C, Lueders T, Himmelberg AM, Hertkorn N, Schmitt-Kopplin P, Harir M, Hosein R, Haque S & Schulze-Makuch D (2014) Water droplets in oil are microhabitats for microbial life. *Science* 345 (6197): 673-676

Meehan CJ & Beiko RG (2014) A phylogenomic view of ecological specialisation in the *Lachnospiraceae*, a family of digestive tract-associated bacteria. *Genome Biology and Evolution* 6 (3): 703-713

Meyer JL (1994) The microbial loop in flowing waters. *Microbial Ecology* 28: 195-199

Meyer KA, O'Neil JM, Hitchcock GL & Heil CA (2014) Microbial production along the West Florida Shelf: Responses of bacteria and viruses to the presence and phase of

Karenia brevis blooms. *Harmful Algae* DOI: 10.1016/j.hal.2014.04.015

Mitchell JG, Okubo A, Fuhrman JA (1985) Microzones surrounding phytoplankton form the basis for a stratified marine microbial ecosystem. *Nature* 316: 58–59.

Mitchell JG & Fuhrman JA (1989) Centimetre scale vertical heterogeneity in bacteria and chlorophyll a. *Marine Ecology Progress Series* 54: 141-148

Mitchell JG (2002) The energetic and scaling of search strategies in bacteria. *American Naturalist* 160: 727-740.

Mitchell JG (2004) Rank-size analysis and vertical phytoplankton distribution patterns. In: Handbook of Scaling Methods in Aquatic Ecology: Measurement, Analysis, Simulation (L. Seuront, P.G. Strutton eds) CRC, Boca Raton, 257– 278.

Mitchell JG, Yamazaki H, Seuront L, Wolk F & Li H (2008) Phytoplankton patch patterns: Seascape anatomy in a turbulent ocean. *Journal of Marine Systems* 69 (3-4): 247-253

Moran PAP (1950) Notes on continuous stochastic phenomena. *Biometrika* 37: 17–23.

Moore GF, Dunsmore BC, Jones SM, Smejkal CW, Jass J, Stoodley P, & Lappin-Scott HM (2000), 'Microbial detachment from biofilms', in *SYMPOSIA-SOCIETY*

Mosley LM, Zammit B, Leyden E, Heneker TM, Hipsey MR, Skinner D & Aldridge KT (2012) The impact of extreme low flows on water quality of the lower Murray Riber and lakes (South Australia). *Water Resource Management* 26: 3923-3946

Nava GM, Stappenbeck TS (2011) Diversity of the autochthonous colonic microbiota. *Gut Microbes* 2 (2): 99-104

Nevin KP & Lovley DR (2002) Mechanisms for accessing insoluble Fe(III) oxide during dissimilatory Fe(III) reduction by *Geothrix fermentans. Applied Environmental Microbiology* 68 (5): 2294-2299

Newbold JD, Elwood JW, O'Neil RV & Van Winkle W (1981) Measuring nutrient spiralling in streams. *Canadian Journal of Fisheries and Aquatic Sciences* 38: 860-863

Newbold JD, Elwood JW, O'Neill RV & Sheldon AL (1983) Phosphorus dynamics in a woodland stream ecosystem: a study of nutrient spiralling. *Ecology* 64 (5): 1249-1265

Newton RJ, Kent AD, Triplett EW & McMahon KD (2006) Microbial community dynamics in a humic lake: differential persistence of common freshwater phylotypes. *Environmental Microbiology* 8(6): 956-970

Newton RJ, Jones SE, Helmus MR & McMahon KD (2007) Phylogenetic ecology of the freshwater *Actinobacteria* acI lineage. *Applied and Environmental Microbiology* 73(22): 7169-7176

Newton RJ, Jones SE, Eiler A, McMahon KD & Bertilsson S (2011) A guide to natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews* 75 (1): 14-49

Nicholson KW (1988) A review of particle resuspension. *Atmospheric Environment* 22: 2639–2651.

Ochman H, Worobey M, Kuo CH, Ndjango JBN, Peeters M, Hahn BH & Hugenholtz P (2010) Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biology* 8 (11): 1-8

Okabe S, Okayama N, Savichtcheva O & Ito T (2007) Quantification of host-specific *Bacteroides-Prevotella* 16s rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiological Biotechnology* 74: 890-901

Oliver RL & Ganf GG (2002) 'Freshwater blooms' in BA Whitton & M Potts (ed.) *The Ecology of Cyanobacteria,* Springer Netherlands, pp 149-194

Ortmann AC, Lawrence JE & Suttle CA (2002) Lysogeny and lytic viral production during a bloom of cyanobacterium *Synechococcus* spp. *Microbial Ecology* 43: 225-231

Pace N (1997) A molecular view of microbial diversity and the biosphere. *Science* 276: 734–740.

Paerl HW & Pinckney JL (1996) A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. *Microbial Ecology* 31: 225-247

Palijan G (2012) Abundance and biomass responses of microbial food web components to hydrology and environmental gradients within a floodplain of the River Danube. *Microbial Ecology* 64: 39-53

Palleroni NJ (1981) "Introduction to the family *Pseudomonadaceae"* in Starr *et al.* (eds.), *The Prokaryotes,* Springer, Berlin Heidelberg, pp. 655-665

Pan LA, Zhang LH, Zhang J, Gasol JM, Chao M (2005) On-board flow cytometric observation of picoplankton community structure in the East China Sea during the fall of different years. *FEMS Microbiology Ecology* 52: 243–253.

Pan LA, Zhang J, Zhang LH (2007) Picophytoplankton, nanophytoplankton, heterotrophic bacteria and viruses in the Changjiang Estuary and adjacent coastal waters. *Journal of Plankton Research* 29: 187–197.

Park J, Lappe M & Teichmann SA (2001) Mapping protein family interactions: intramolecular and intermolecular protein family interaction repertoires in the PDB and yeast. *Journal of Molecular Biology* 307: 929-938

Parvathi A, Zhong X & Jacquet S (2012) Dynamics of various viral groups infecting

autotrophic plankton in Lake Geneva. *Advances in Oceanography and Limnology* 3 (2): 171-191

Paterson JS, Nayar S, Mitchell JG & Seuront L (2012) A local upwelling controls viral and microbial community structure in South Australian continental shelf waters. *Estuarine, Coastal and Shelf Science* 96: 197-208

Patten N, Bourne DG & Mitchell JG (2005b) Spatial dynamics of virus-like particles and heterotrophic bacteria within a shallow coral reef. *Marine Ecology Progress Series* 288: 1-8

Patten NL, Seymour JR & Mitchell JG (2006) Flow cytometric analysis of virus-like particles and heterotrophic bacteria within coral-associated reef water. *Journal of the Marine Biological Association of the United Kingdom* 86: 563-566

Payet JP & Suttle CA (2008) Physcial and biological correlates of virus dynamics in the southern Beaufort Sea and Amundsen Gulf. *Journal of Marine Systems* 74: 933-945

Peakall R, Ruibal M & Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. *Evolution* 57: 1182-1195

Peduzzi P & Schiemer F (2004) Bacteria and viruses in the water column of tropical freshwater reservoirs *Environmental Microbiology* 6(7): 707-715

Pickett S, Cadenasso M (1995) Landscape ecology: Spatial heterogeneity in ecological systems. *Science* 269 (5222): 331–334

Pollard PC & Ducklow H (2011) Ultrahigh bacterial production in a eutrophic subtropical Australian river: Does viral lysis short-circuit the microbial loop? *Limnology and Oceanography* 56 (3): 1115-1129

Ponce-Terashima R, Koskey AM, Reis MG, McLellan SL & Blanton RE (2014) Sources and distribution of surface water fecal contamination and prevalence of Schistosomiasis in a Brazilian village. *PLoS – Neglected Tropical Diseases.* 8 (10): 1-10

Pontes D, Lima-Bittencourt C, Chartone-Souza E, Nascimento A (2007) Molecular approaches: advantages and artifacts in assessing bacterial diversity. *Journal of Industrial Microbiology & Biotechnology* 34: 463–473.

Pylro VS, Roesch LFW, Morais DK, Clark IM, Hirsch PR & Tótola MR (2014) Data analysis for 16s microbial profiling from different benchtop sequencing platforms. *Journal of Microbiological Methods* 107: 30-37

Riley GA (1963) Zooplankton and the food cycle. *AIBS Bulletin* 13 (5): 60-61+69

Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, et al. (2010), Viral and microbial community dynamics in four aquatic environments. *International Society for Microbial Ecology (ISME)*, 4: 739–751.

Rossi RE, Mulla DJ, Journel AG, Franz EH (1992) Geostatistical tools for modelling and interpreting ecological spatial dependence. *Ecological Monographs* 62: 277–314.

Rutgers van der Loeff MM, Meyer R, Rudels B, Rachor E (2002) Resuspension and particle transport in the benthic nepheloid layer in and near Fram Strait in relation to faunal abundances and Th depletion. *Deep-Sea Research* 49: 1941– 1958.

Sandilands RG, Mudroch A (1983) Nepheloid layer in Lake Ontario. *Journal of Great Lakes Research* 9: 190–200.

Ritchie RJ (2008) Universal chlorophyll equations for estimating chlorophylls *a*, *b*, *c* and *d* and total chlorophylls in natural assemblages of photosynthetic organisms using acetone, methanol, or ethanol solvents. *Photosynthetica* 46(1): 115-126

Robin A, Ravet V, Personnic S, Theil S, Colombet J & Sime-Ngando T (2012) Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* 7 (3): doi: 10.1371/journal.pone.0033641

Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, Buchanan J, Desnues C, Dinsdale E, Edwards R, Felts B, Haynes M, Liu H, Lipson D, Mahaffy J, Martin-Cuadrado AB, Mira A, Nulton J, Pašić L, Rayhawk S, Rodriguez-Mueller J, Rodriguez-Valera F, Salamon P, Srinagesh S, Thingstad TF, Tran T, Vega Thurber R , Willner D, Youle M & Rohwer F (2010) Viral and microbial community dynamics in four aquatic environments. *The International Society for Microbial Ecology (ISME) Journal* 4: 739-751

Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B, Pašić L, Thingstad TF, Rohwer F & Mira A (2009) Explaining microbial population genomics through phage predation. *Nature Reviews Microbiology* 7: 828-836

Roudnew B, Seymour JR, Jeffries TC, Lavery TJ, Smith RJ & Mitchell JG (2012) Bacterial and virus-like particle abundances in purged and unpurged groundwater depth profiles. *Groundwater Monitoring and Remediation* 32 (4): 72-77

Roudnew B, Lavery TJ, Seymour JR, Jeffries TC & Mitchell JG (2014) Variability in bacteria and virus-like particle abundances during purging of unconfined aquifers. *Groundwater* 52 (1): 118-124

Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hastings A, Holt RD, Mayfield MM, O'Connor MI & Rice WR (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution* 22 (9): 465-471

Schapira M, Buscot MJ, Leterme SC, Pollet T, Chapperon C, et al. (2009) Distribution of heterotrophic bacteria and virus-like particles along a salinity gradient in a hypersaline coastal lagoon. *Aquatic Microbial Ecology* 54: 171– 183.

Schmieder R, Lim YW & Edwards R (2012): Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* 2012, 28:433-435.

Seuront L, Lacheze C, Doubell MJ, Seymour JR, Van Dongen-Vogels V, et al. (2007a) The influence of Phaeocystisglobosa on microscale spatial patterns of chlorophyll a and bulk-phase seawater viscosity *Biogeochemistry* 83: 173–188.

Seuront L, Lacheze C, Doubell MJ, Seymour JR, Dongen-Vogels VV, Newton K, Alderkamp AC & Mitchell JG (2007b) The influence of *Phaeocystis globosa* on microscale spatial patterns of chlorophyll *a* and bulk-phase seawater viscosity. *Biogeochemistry.* 83: 173-188

Seymour JR, Mitchell JG, Pearson L, Waters R (2000) Heterogeneity in bacterioplankton abundance from 4.5 millimetre resolution sampling. *Aquatic Microbial Ecology* 22: 143–153.

Seymour JR, Mitchell JG, Seuront L (2004) Microscale heterogeneity in the activity of coastal bacterioplankton communities. *Aquatic Microbial Ecology* 35: 1–16.

Seymour JR, Seuront L, Mitchell JG (2005a) Microscale and small-scale temporal dynamics of a coastal planktonic microbial community. *Marine Ecology Progress Series* 300: 21–37.

Seymour JR, Patten N, Bourne DG, Mitchell JG (2005b) Spatial dynamics of virus-like particles and heterotrophic bacteria within a shallow coral reef system. *Marine Ecology Progress Series* 288: 1–8.

Seymour JR, Seuront L, Doubell M, Waters RL, Mitchell JG (2006) Microscale patchiness of virioplankton. *Journal of the Marine Biological Association of the United Kingdom* 86: 551–561.

Seymour JR, Seuront L, Mitchell JG (2007) Microscale gradients of planktonic microbial communities above the sediment surface in a mangrove estuary. *Estuarine, Coastal and Shelf Sciences* 73: 651–666.

Seymour JR, Seuront L, Doubell M, Mitchell JG (2008) Mesoscale and microscale spatial variability of bacteria and viruses during a Phaeocystisglobosa bloom in the Eastern English Channel. *Estuarine, Coastal and Shelf Science* 80: 589–597.

Seymour JR, Marcos & Stocker R (2009) Resource patch formation and exploitation throughout the marine microbial food web. *The American Naturalist.* 173 (1): 15-29

Seymour JR, Ahmed T, Durham WM, Stocker R (2010) Chemotactic response of marine bacteria to the extracellular products of Synechococcus and Prochlorococcus. *Aquatic Microbial Ecology* 59: 161–168.

Shade A, Jones SE & McMahon KD (2008) The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. *Environmental Microbiology* 10 (4): 1057-1067

Shrout JD, Chopp DL, Just CL, Hentzer M, Givskov M & Parsek MR (2006) The impact of quorum sensing and swarming motility on *Pseudomonas aeruginosa* biofilm formation is nutritionally conditional. *Molecular Microbiology* 62 (5): 1264-1277

Simberloff D (1978) 'Use of rarefaction and related methods in ecology' in KL Dickson *et al.* (ed.) *Biological data in water pollution assessment: quantitative and statistical analyses,* American Society for Testing and Materials, pp. 150-165

Smith SC & Bruce CW (2008) *CrimeStat®III User Workbook,* The National Institute of Justice, viewed on: 8th August 2012 <http://www.icpsr.umich.edu/CrimeStat/workbook/CrimeStat_Workbook.pdf >

Smith RJ, Paterson JS, Sibley CA, Hutson JL & Mitchell JG (2015) Putative effect of aquifer recharge on the abundance and taxonomic composition of endemic microbial communities. *PLoS One* 10 (6): e0129004. doi:10.1371/journal.pone.0129004

Sokal RR (1978) Spatial autocorrelation in biology: 1. Methodology. Biological Journal of the Linnean Society 10: 199–228.

Sonthiphand P, Hall MW & Neufeld JD (2014) Biogeography of anaerobic ammonia-oxidising (anammox) bacteria. *Frontiers in Microbiology.* 5 (399): 1-13.

Stackebrandt E, Cummins CS & Johnson JL (2006) 'Family *Propionibacteriaeceae*: The genus *Propionibacterium'* in M. Dworkin *et al.* (ed.) *The Prokaryotes,* Springer New York, pp. 400-418

Stocker R, Seymour JR, Samadani A, Hunt DE & Polz MF (2008) Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 105: 4209-4214

Stocker R & Seymour JR (2012) Ecology and physics of bacterial chemotaxis in the ocean. *Microbiology and Molecular Biology Reviews* 76: 792-812

Stocker R (2012) Marine microbes see a sea of gradients. *Science* 338: 628-633

Strutton PJ, Mitchell JG, Parslow JS, Greene RM (1997) Phytoplankton patchiness: quantifying the biological contribution using Fast Repetition Rate Fluorometry. *Journal of Plankton Research* 19: 1265–1274.

Sutcliffe Jr WH, Baylor ER & Menzel DW (1963) Sea surface chemistry and Langmuir circulation. *Deep Sea Research and Oceanographic Abstracts* 10 (3): 233, IN7, 236-237, IN8, 243.

Suttle CA (2005) Viruses in the sea. *Nature* 437: 356-361Wilhelm SW & Matteson AR (2008) Freshwater and marine virioplankton: a brief overview of commonalities and differences. *Freshwater Biology* 53: 1076-1089

Suttle CA (2007) Marine viruses – major players in the global ecosystem. *Nature Reviews* 5: 801–812.

Taylor JR, Stocker R (2012) Trade-offs of chemotactic foraging in turbulent water. *Science* 338: 675–679.

Telgmann U, Horn H & Morgenroth E (2004) Influence of growth history on sloughing and erosion from biofilms. *Water Research* 38 (17): 3671-3684

Teneva I, Dzhambazov B, Koleva L, Mladenov R & Schirmer K (2005) Toxic potential of five freshwater Phormidium species (Cyanoprokaryota) *Toxicon* 45 (6): 711-725

Thar R, Ku¨hl M (2003) Bacteria are not too small for spatial sensing of chemical gradients: An experimental evidence. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 100: 5748–5753.

Thingstad FT & Lignell R (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* 13: 19-27

Thingstad FT (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography* 45 (6): 1320-1328

Thomas R, Berdjeb L, Sime-Ngando T & Jacquet S (2011) Viral abundance, production, decay rates and life strategies (lysogeny versus lysis) in Lake Bourget (France). *Environmental Microbiology* 13: 616-630

Thomson PG, Davidson AT, Enden R, Pearce I, Seuront L, et al. (2010) Distribution and abundance of marine microbes in the Southern Ocean between 30 and 80°E. *Deep-Sea Research II* 57: 815–827.

Tonon LAC, Moreira APB & Thompson F (2014) 'The family *Erythrobacteraceae'* in E. Rosenberg *et al.* (ed.) *The prokaryotes – Alphaproteobacteria and Betaproteobacteria* Springer-Verlag Berlin Heidelberg pp. 213-235

Topp E, Mulbry WM, Zhu H, Nour SM & Cuppels D (2000) Characterisation of *S*-triazine herbicide metabolism by a *Nocardioides* sp. isolated from agricultural soils. *Applied and Environmental Microbiology* 66 (8): 3134-3141

Traister E & Anisfeld SC (2006) Variability of indicator bacteria at different time scales in the upper Hoosic River watershed. *Envionmental Science and Technology* 40 (16): 4990-4995

Tsai AY, Gong GC, Sanders RW & Huang JK (2013) Contribution of viral lysis and nanoflagellate grazing to bacterial mortality in the inner and outer regions of the Chingjiang River plume during Summer. *Journal of Plankton Research* doi: 10.1093/plankt/fbt074

Tsubouchi T, Shimane Y, Usui K, Shimamura S, Mori K, Hiraki T, Tame A, Uematsu K, Maruyama T & Hatada Y (2013) *Brevundimonas abyssalis* sp. nov., a dimorphic prosthecate bacterium isolated from deep-subsea floor sediment. *International Journal of Systematic and Evolutionary Microbiology* 63: 1987-1994

Ueda HR, Hayashi S, Matsuyama S, Yomo T, Hashimoto S, Kay SA, Hogenesch JB, Iino M (2004) Universality and flexibility in gene expression from bacteria to human. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 101 (11): 3765-3769

Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S, Jeppesen E, JM, Schwenk K, Zwart G, Degans H, Vyverman W & 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 of the United States of America (PNAS)* 104 (51): 20404-20409

Van Dijk AIJM, Hairsine PB, Arancibia JP & Dowling TI (2007) Reforestation, water availability and stream salinity: a multi-scale analysis in the Murray-Darling Basin, Australia. *Forest Ecology and Management* 251: 94-109

Van Hannen EJ, Zwart G, Van Agterveld MP, Gons HJ, Ebert J & Laanbroek HJ (1999) Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. *Applied and Environmental Microbiology* 65 (2): 795-801

Venter J, Remington K, Heidelberg J, Halpern A, Rusch D, et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66–74.

Vrba ES & Gould SJ (1986) The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology* 12 (2): 217-228

Wagner M & Horn M (2006) The *Plantomycetes, Verrucomicrobia, Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Current Opinions in Biotechnology* 17 (3): 241-249

Walker KF & Thoms MC (1993) Environmental effects of flow regulation on the lower River Murray, Australia. *Regulated Rivers: Research & Management* 8: 103-119

Wang Q, Garrity GM, Tiedje JM & Cole JR (2007) Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied Environmental Microbiology* 73 (16): 5261-7

Wang Y, Hammes F, Boon N, Chami M, Egli T (2009) Isolation and characterisation of low nucleic acid (LNA)-content bacteria. *The International Society for Microbial Ecology (ISME) Journal* 3: 889–902.

Wang Y, Sheng HF, He Y, Wu JY, Jiang YX, Tam NFY & Zhou HW (2012) Comparison of the levels of bacterial diversity in freshwater, intertidal wetland and marine sediments by using millions of Illumina tags. *Applied and Environmental Microbiology* 78 (23): 8264-8271

Warnecke F, Amann R & Pernthaler J (2004) Actinobacterial 16S rRNA genes from freshwater habitats cluster in four distinct lineages. *Environmental Microbiology* 6(3): 242-253

WaterConnect Government of South Australia, Department of Environment, Water and Natural Resources 2012, *River Murray Data,* viewed on: 18th June 2012, <https://apps.waterconnect.sa.gov.au/TelemetryViewer/TimeSeries.aspx?App=RMW&Period=Hrly&Site=A4261162&SiteName=Long+Island&Params=100%2c12>

Waters RL, Mitchell JG (2002) Centimetre-scale spatial structure of estuarine in vivo fluorescence profiles. *Marine Ecology Progress Series* 237: 51–63.

Waters RL, Mitchell JG, Seymour J (2003) Geostatistical characterisation of centimetre-scale spatial structure of in vivo fluorescence. *Marine Ecology Progress Series* 251: 49–58.

Webster JR (1975) Analysis of potassium and calcium dynamics in stream ecosystems on three southern Appalachian watersheds of contrasting vegetation.

Webster JR & Patten BC (1979) Effects of watershed perturbation on stream potassium and calcium dynamics. *Ecological Monographs* 19: 51-72

Weinbauer MG & Höfle MG (1998) Significance of viral lysis and flagellate grazing as factors controlling bacterioplankton production in a eutrophic lake. *Applied and Environmental Microbiology* 64 (2): 431-438

Weinbauer MG, Suttle CA (1999) Lysogeny and prophage induction in coastal and offshore bacterial communities. *Aquatic Microbial Ecology* 18: 217–225.

Weinbauer MG (2004) Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* 28: 127-181

Weinbauer MG, Hornák K, Jezbera J, Nedoma J, Dolan JR & Šimek K (2006) Synergistic and antagonistic effects of viral lysis and protistan grazing on bacterial biomass, production and diversity. *Environmental Microbiology* doi:10.1111/j.1462-2929.2006.01200.x

Wiebe PH (1970) Small-scale spatial distribution in oceanic zooplankton. *American Society of Limnology and Oceanography* 15: 205–217. doi: 10.4319/lo.1970.15.2.0205

Wiese J, Thiel V, Gärtner A, Schmaljohann R & Imhoff JF (2009) *Kiloniella laminariae* gen. nov., sp. nov., an alphaproteobacterium from the marine macroalga *Laminaria saccharina*. *International Journal of Systematic and Evolutionary Microbiology* 59: 350-356

Winter C, Bouvier T, Weinbauer MG & Thingstad TF (2010) Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiology and Molecular Biology Reviews* 74 (1): 42-57

Wise MG, Shimkets LJ & McArthur JV (1995) Genetic structure of a lotic population of Burkholderia (Pseudomonas) cepacia. *Applied and Environmental Microbiology* 61 (5): 1791-1798

Wommack KE & Colwell RR (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiology and Molecular Biology Reviews* 64: 69-114

Wu X, Xi W, Ye W & Yang H (2007) Bacterial community composition of a shallow hypertrophic freshwater lake in China, revealed by 16S rRNA gene sequences. *FEMS Microbiology Ecology* 61(1): 85-96

Yager P, Connelly T, Mortazavi B, Wommack K, Bano N, Bauer J, Opsahl S & Hennes KP & Simon M (1995) Significance of bacteriophage for controlling bacterioplankton growth in a mesotrophic lake. *Applied and Environmental Microbiology* 61: 333-340

Zhang J, Yang Y, Zhao L, Xie S & Liu Y (2014) Distribution of sediment bacterial and archaeal communities in plateau freshwater lakes. *Applied Microbiological Biotechnology* DOI 10.1007/s00253-014-6262-x

Zheng G, Yampara-Iquise H, Jones JE & Carson AC (2009) Development of *Faecalibacterium* 16s rRNA gene marker for identification of human faeces. *Journal of Applied Microbiology* 106 (2): 634-641

Zhinger L, Gobet A & Pommier T (2012) Two decades of describing the unseen

majority of aquatic microbial diversity. *Molecular Ecology* 21 (8): 1878-1896

Zhong X, Pradeep Ram AS, Colombet J & Jacquet S (2014) Variations in abundance, genome size, morphology, and functional role of the virioplankton in lakes Annecy and Bourget over a 1-year period. *Microbial Ecology* 67 (1): 66-82

Zwart G, Crump BC, Agterveld MPK, Hagen F & Han SK (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquatic Microbial Ecology* 28: 141-155

APPENDICES

Published manuscripts arising from this thesis