AQUATIC MICROBIAL DIVERSITY: A SENSITIVE AND ROBUST TOOL FOR ASSESSING ECOSYSTEM HEALTH AND FUNCTIONING

Report to the

WATER RESEARCH COMMISSION

by

GF MATCHER[1*], PW FRONEMAN[2], RA DORRINGTON[1*]

[1] Department of Biochemistry and Microbiology
Rhodes University, Grahamstown, South Africa

[2] Department of Zoology and Entomology
Rhodes University, Grahamstown, South Africa

WRC Report No 2038/1/14

ISBN 978-1-4312-0643-8

March 2015

Obtainable from

Water Research Commission

Private Bag X03

GEZINA, 0031

orders@wrc.org.za or download from www.wrc.org.za

EXECUTIVE SUMMARY

BACKGROUND

Estuaries are ecologically and economically important aquatic systems, functioning as feeding/staging sites for migratory birds, as nurseries for marine fish, and as repositories of high biodiversity. They are also important for the tourism industry and serve as sites for productive fish and invertebrate fisheries and aquaculture. Over the past decade, there has been a dramatic increase in urban, agricultural and industrial development along the southern African coastline, particularly in the vicinity of estuaries, resulting in an escalation in anthropogenic stresses on these delicate ecosystems. Whilst many studies have been carried out on the macro-fauna and macro-flora in South African estuaries, very few studies have been done to assess the microbial foodweb.

Microbes are extremely abundant and diverse in aquatic ecosystems and play critical roles in regulating key biogeochemical cycles such as the carbon, nitrogen and sulphur cycles. Furthermore, due to their small size and rapid proliferation rates, prokaryotes exhibit rapid response rates to changes in nutrient availability and physicochemical changes in the environment, such as those that may be induced by pollution. Characterization of the microbial population within a target system would therefore provide an excellent assessment of ecosystem health and analysis of such data would flag the presence of potential microbial pathogens.

STUDY AIMS

1) To develop protocols for sample collection, template preparation for high throughput Next Generation Sequencing (NGS) and analysis of 16S rRNA sequences in estuarine water and sediment.
2) To use NGS technology to characterise the microbial communities in four distinct Eastern Cape estuarine systems.
3) To establish the links between the physico-chemical characteristics and the diversity and structure of microbial diversity within each estuary.

METHODOLOGY

Four permanently open Eastern Cape estuaries, with diverse hydrodynamics and anthropogenic impacts were selected for this study. The Kariega Estuary is a freshwater-deprived, marine-dominated system with a relatively pristine catchment area. The Kowie Estuary is also marine-dominated, but in contrast to the Kariega, has extensive settlement along the banks and a small harbour at its mouth, with increased urban impacts due to influx of large volumes of sewage and sewerage into the system. The freshwater dominated Sundays Estuary is impacted by commercial agriculture in its catchment area and it is supplemented by the inter-basin transfer scheme from the Gariep dam, which increases the freshwater inflow into this estuary. Finally, the Swartkops Estuary is

also freshwater dominated and was selected because it is highly urbanised and severely impacted by anthropogenic pollution inflicted on this estuary by both urban and industrial activities.

In this study, the water column and the sediment were sampled at three sites along each of these estuaries (*viz*. the estuary mouth, middle reaches and upper reaches). The physico-chemical properties of the water column at each of the sampling sites within the target estuaries were ascertained. For characterisation of the microbial communities, a region within the eubacterial 16S rRNA gene was analysed using 454-pyrosequencing of amplicon libraries constructed from each of the samples. Approximately 179 000 sequence reads were analysed to determine the relative abundances of each of the phylotypes as well as identify those bacteria which occur in low abundances (i.e. rare).

RESULTS

The data revealed that the bacterial communities within South African estuaries differ significantly from estuaries reported in the literature. Principle component analysis showed a close relationship between the Kowie and the Kariega estuaries that likely reflects their strong marine influence. The Sundays Estuary did not cluster with any of the other three estuaries. With respect to the Swartkops Estuary, the upper reaches sampling site clustered more closely with the Sundays Estuary whilst the middle and lower reaches clustered more closely with the Kowie and Kariega estuaries. The analysis revealed different microbial assemblages in the water column versus the sediment. There was a distinct spatial distribution of dominant taxa in the sediment along the length of the estuaries, even in estuaries with relatively small physico-chemical gradients from the mouth to the upper reaches. All the estuaries were found to be net-heterotrophic whilst unicellular algae were responsible for the majority of the phototrophy within these ecosystems. The impact of sewage effluent, and potential subsequent nutrient spikes, into the Kowie Estuary was highlighted by the occurrence of a mono-specific incidence of the cyanobacteria *Synechococcus*, which is indicative of a bloom remnant or early bloom-formation. The dominant phyla represented in the water columns of these estuaries included Bacteroidetes, Gammaproteobacteria and Betaproteobacteria whilst the sediments had elevated levels of Deltaproteobacteria with reduced Betaproteobacterial populations. The species diversity found in the sediment was found to be several-fold higher than that in the water column. This is indicative of the more homogenous character of the water column compared to the more complex environment found within the sediment samples.

CONCLUSIONS

This study has demonstrated the power of NGS technologies in the analysis of aquatic microbial communities. The results have shown that the diversity and structure of microbial communities reflect the physico-chemical characteristics of aquatic systems and provide important insight into the functioning of estuarine ecosystems.

RECOMMENDATIONS FOR FUTURE RESEARCH

1. The advent of new NGS technologies that substantially decrease the cost (up to ten-fold) of generating sequence datasets provide an opportunity to apply the approach taken in this study widely to include other important estuarine systems around the South African coastline. In addition, the reduced cost (less than R 2000 per estuary) could make it feasible to use of this technology for routine monitoring of sensitive estuarine systems.
2. It is now becoming known that numerical abundance of a particular bacterial species does not always directly correlate with metabolic activity or potential growth rate of that species. A more accurate assessment of whether a bacterial taxon is actively metabolising can be achieved by quantifying the rRNA as opposed to the rDNA (i.e. the active component of the ribosome versus the gene encoding for the rRNA). The need for information on which taxa are metabolically active in resident microbial communities is particularly important as a tool for monitoring episodic anthropogenic pollution in urbanized estuaries such as the Swartkops systems.
3. The physico-chemical analysis of water samples is not a very sensitive measure of changes in the type (e.g. inorganic vs organic) nutrient concentration or anthropogenic pollution within the system. We observed changes in the abundance of dominant microbial species in sediment samples that were not observed in the water column along the length of estuaries with apparently small salinity and nutrient gradients. These findings lend strong support for the need to focus on sediment sampling when monitoring estuarine health and aquatic ecosystems in general.

ACKNOWLEDGEMENTS

The authors would like to thank the Reference Group of the WRC Project for the assistance and the constructive discussions over the duration of the project. We would also like to thank those who participated in this project.

Project participant	Contribution	Affiliation
Prof RA Dorrington	Project Leader	Rhodes University
Dr G Matcher	Primary Researcher	Rhodes University
Prof W Froneman	Principle Researcher	Rhodes University
Mrs V Hodgson	Technical and administrative support	Rhodes University
Miss M Isaacs	Technical support	Rhodes University
Miss P Mokgokong	MSc student – method development	Rhodes University
Mr J Dean	Honours' student – method development	Rhodes University
Dr T Henninger	Sample collection	Rhodes University
Mr R Wasserman	PhD student – Sample collection	Rhodes University
Dr T Bornman	Sample collection	SAEON/NMNU

Reference Group	Affiliation
Dr MS Liphadzi	Water Research Commission (Chairman)
Prof CC Bezuidenhout	North West University
Prof B Pletschke	Rhodes University
Prof JB Adams	Nelson Mandela Metropolitan University
Prof Bate	Nelson Mandela Metropolitan University
Dr S Mitchell	Bufo Technologies
Ms MJ Mulubisana	ARC

The project team would like to thank the Water Research Commission for financing the project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
bp	Base pairs
Chl-a	Chlorophyll-a
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
gDNA	Genomic DNA
GPS	Global positioning system
min	Minutes
NCBI	National Centre for Biotechnology Information
NGS	Next generation sequencing
OTU	Operational taxonomic unit
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PSU	Practical salinity units
RDP	Ribosomal Database Project
rDNA	rRNA gene encoding sequence
RFLP	Restriction fragment length polymorphism
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
s	Seconds

1. INTRODUCTION AND OBJECTIVES

1.1 Estuarine ecosystems

Estuaries represent a transition zone between freshwater and marine ecosystems formed by the intersection of freshwater inflow from river systems with ocean tidal waters. As such, estuaries contain strong gradients in both physical and biological aspects. These gradients include salinity, nutrient concentrations, as well as silt and organic loads, which in turn influence fauna and flora distribution and abundance profiles (Flindt et al., 1999; Crump et al., 2004). Furthermore, estuaries are typically relatively shallow and as a consequence, there is often a close relationship between the benthic and pelagic processes (Flindt et al., 1999). Estuaries are well-known as ecosystems with high biodiversity profiles and which play critical roles as feeding and staging sites for migrant birds, as nursery grounds for many aquatic fish species, and typically exhibit high levels of primary and secondary productivity (Nixon et al., 1986; Turpie et al., 2002; Vorwerk and Froneman, 2009).

1.2 Estuarine microbial communities

Bacteria are critical components in key biogeochemical processes including the carbon and nitrogen cycles (Kirchman, 2008). Not only are they involved in nitrogen fixation and photosynthetic primary production, but they are also major players in the recycling and mineralization of organic matter (del Giorgio and Cole, 1998). With regard to aquatic ecosystems, there is a clear delineation between the microbial populations found in marine ecosystems compared to freshwater ecosystems (Bouvier and del Giorgio, 2002; Crump et al., 2004; Wu et al., 2006). Since estuaries represent a convergence of marine and freshwater environments, the bacteria present are derived from both river and oceanic communities advected into the estuary. During high tidal inflow, the estuarine bacterial community is shaped by the marine species whilst during high freshwater input, the estuarine bacterial assemblages more closely resemble those found in the rivers (Fortunato et al., 2012). However, if the residence time of the water column within the estuary is long enough, a unique bacterial assemblage may be established (Crump et al., 2004; Herlemann et al., 2011). The microbiomes present within estuarine ecosystems reflect the dynamic nature of estuaries where steep gradients in several physico-chemical parameters can be observed (Kirchman et al., 2005; Telesh and Khlebovich, 2010; Fortunato et al., 2012). Of these, salinity appears to be a major driving force in affecting microbial diversity profiles in aquatic ecosystems (Bouvier and del Giorgio, 2002; Langenheder et al., 2003; Herlemann et al., 2011; Fortunato et al., 2012).

The best studied estuarine systems are in the Northern Hemisphere, where the microbial communities have been found to be typically net-heterotrophic (Selje and Simon, 2003). The phyla which generally dominate include Bacteroidetes and Proteobacteria (α, β and γ) with Betaproteobacteria being more predominant in lower saline conditions and Alphaproteobacteria more prevalent at higher salinities (Glöckner et al., 1999; Bouvier and del Giorgio, 2002; Wu et al., 2006; Campbell and Kirchmann, 2013). The majority of the studies carried out on the bacterial diversity profiles in estuaries have focused on the water column (Bouvier and del Giorgio, 2002; Langenheder et al., 2003; Herlemann et al., 2011; Fortunato et al., 2012; Campbell and Kirchmann, 2013). Research on bacterial

assemblages within estuarine sediments have typically focussed on one sub-group of bacteria (e.g. ammonia oxidising, sulphur-reducing, etc.) rather than the community as a whole (e.g. Kondo et al., 2004; Bemen and Francis, 2006; Freitag et al., 2006; Sahan and Muyzer, 2008). A study done on sediments of six estuaries in Australia revealed the dominance of Gammaproteobacteria, Deltaproteobacteria and Alphaproteobacteria, representing 23.7%, 21.3% and 8.5% of the bacterial communities respectively (Sun et al., 2013).

1.3 Methods for determining microbial diversity

Historically, the characterisation of bacterial populations within a target ecosystem was done using cultivation techniques followed by subsequent morphological and enzymatic characterization of bacterial isolates. This approach, however, is not only labour-intensive but the most telling disadvantage lies in that the majority (>90%) of the environmental microbial population is extremely difficult to culture or is unculturable (Schloss and Handelsman, 2005). As a result, culture-independent molecular techniques are now typically utilised when assessing the diversity of bacterial assemblages, the mainstay of which is sequence analysis of the 16S rRNA gene (16S rDNA). The 16S rDNA is an excellent target for microbial identification as it is ubiquitous and while the gene is conserved within species there is sufficient evolution-induced variability to allow for inter-species differentiation (Tringe and Hugenholtz, 2008). Structurally, the 16S rDNA gene consists of nine species-specific hypervariable regions (V1-V9) interposed within highly conserved nucleotide regions. The conserved regions allow for the design of PCR primers which are able to hybridize to the 16S rDNA sequence irrespective of the species from which the gene was isolated. Sequence analysis of the hypervariable region subsequently allows for the taxonomic classification of microbial populations in target environments (Wang et al., 2007; Liu et al., 2007; Huse et al., 2008; Wang and Qian, 2009; Claesson et al., 2009).

There are several tools available for the analysis of 16S rDNA sequences from a mixed population of microbes. The most commonly utilised of these include rapid fragment length polymorphism (RFLP), denaturing gradient gel electrophoresis (DGGE) and DNA sequence analysis of the target gene. Techniques such as RFLP and DGGE are relatively inexpensive and provide a good indication of the whether the microbial population between two environments is similar or not. The major drawback of RFLP and DGGE is that these techniques do not provide definitive information as to the identity of the specific taxa within the sample. DNA sequence analysis, while more expensive, not only provides information as to the overlap of microbial assemblages between environments but also allows for the taxonomic identification of the microbes present. At the start of this project in 2011, most studies had utilised traditional Sanger sequencing of an amplicon library, where PCR products derived from a specific environmental sample are cloned into plasmid vectors and the DNA sequence of the individual PCR-derived inserts determined. In general, this approach, in addition to being time-consuming, generates a few hundred sequences which allows for the identification of only the most dominant phylotypes in the target environment. The recent advent of next generation, high throughput next generation sequencing (NGS) technologies allows for the rapid and cost-effective generation of several thousands of sequence reads (Huse et al., 2008) thereby gaining access not

only to the dominant phylotypes but also to the less common/rare phylotypes. Since the advent of NGS analysis, the number of previously unknown bacterial species has increased exponentially and sequence analysis of 16S rDNA has become an indispensable tool in the analysis of aquatic bacterioplankton assemblages (Glöckner et al., 1999; Zinger et al., 2012). Phylogenetic identification of the bacterioplankton assemblages within a given aquatic ecosystem can then be used to infer biological functioning of the bacteria and their contribution to the biogeochemical cycling within the sampled ecosystem.

1.4 Project motivation and aims

Urban, agricultural and industrial development along river systems in South Africa has increased dramatically over the past decade and as a consequence, many of the estuaries are being increasingly negatively impacted by pollutants, human waste products and excessive nutrient loading (Cooper et al., 1994; Orr et al., 2008). These anthropogenically derived pollutants may critically damage estuarine ecosystem functioning and represent potential human health hazards (Turpie et al., 2002). In South Africa, estuarine ecosystem health monitoring to date has focussed on assessing the status of fish assemblages, water quality and aesthetics of the estuary (Cooper et al., 1994; Odume et al., 2012). The major drawback of this approach is that it is time-consuming, labour intensive and requires substantial and knowledgeable manpower in order to accurately identify and quantify the relevant data in the field. Furthermore, the biological health index focuses on the macrofauna without considering the foundation of the food web of which microbes form a critical component.

Microbial populations respond rapidly to changes in the aquatic environment (Atlas et al., 1991; Buckling et al., 2000; Crump et al., 2003) and, due to their small size and quick proliferation rates, microbial communities adjust to environmental changes more rapidly than their multicellular counterparts co-existing in the same habitat. Consequently, physico-chemical alterations in the environment are often characterised by shifts in dominance of one bacterial phylotype over another with a weaker competitive advantage, or in extreme instances, a decrease in overall microbial diversity. Thus the extent of diversity and population structure of bacteria provides a sensitive and robust tool for assessing estuarine ecosystem health and functioning.

Aims:

1) To develop protocols for sample collection, template preparation for high throughput NGS and analysis of 16S rRNA sequences in estuarine water and sediment.
2) To use NGS technology to characterise the microbial communities in four distinct Eastern Cape estuarine systems
3) To establish the links between the physico-chemical characteristics and the diversity and structure of microbial diversity within each estuary

2. METHODOLOGY

2.1. Site selection

Four estuaries were selected for analysis namely the Kariega, Kowie, Sundays and Swartkops estuaries (Figure 1). All four of these estuaries are regarded as permanently open with the Sundays and Swartkops estuaries substantially influenced by freshwater inflow whilst the Kariega and Kowie estuaries demonstrated much higher salinities along their lengths.

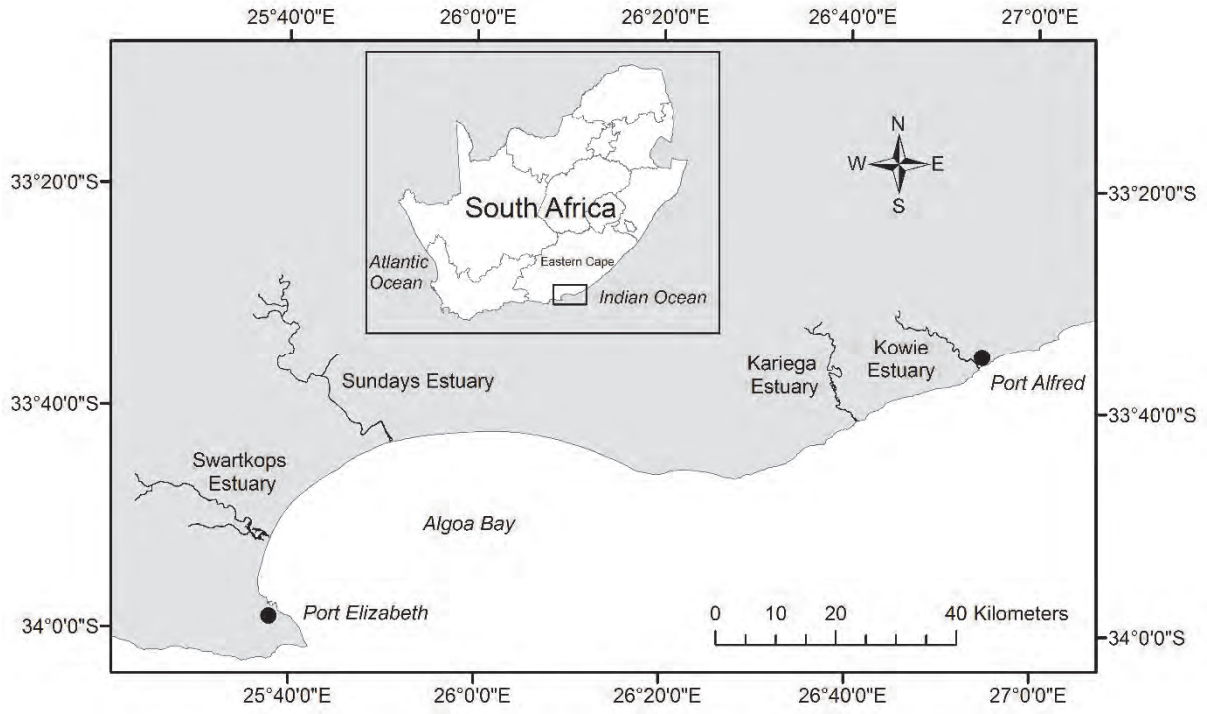


Figure 1. Geographical map of the eastern coastline of South Africa, showing the location of the Kariega, Kowie, Sundays and Swartkops estuaries.

The Kariega Estuary is located approximately 120km east of Port Elizabeth with a water depth ranging from ~3.96 m (lower reaches) to ~1.1m (upper reaches) depending on the tidal state of the estuary. The upper reaches of the Kariega Estuary frequently experience hypersaline (PSU >40) conditions, particularly during the summer and during drought conditions. This is primarily due to a reduced freshwater inflow into the estuary as a combined consequence of a small catchment area (~680 km^2), high evaporation rates and several impoundments along the length of the Kariega River. The impoundments include three major dams and frequently placed small farm weirs (Hodgson, 1987; Grange and Allanson, 1995; Froneman, 2000). Barring the impact of the impoundments, the Kariega Estuary is relatively pristine and its catchment area is predominantly used as wildlife preserves.

The Kowie Estuary flows through the town of Port Alfred, which is located mid-way between East London and Port Elizabeth. The estuary is comprised of a narrow channel (10-15 m in width) and

ranges from 1 to 6 m in depth. The catchment area for the Kowie Estuary is approximately 769 km^2 (Noble and Hemens, 1978). Land usage of the catchment area is extensively allocated for commercial agriculture primarily of pineapples and beef cattle (Cowley and Daniel, 2001). The Estuary itself is also extensively utilised for recreational activities and a small harbour is located at the mouth of the estuary for recreational boats and yachts as well as small commercial fishing boats. The anthropogenic stress on the estuary is further amplified by the location of an informal settlement along the east bank of the lower reaches of the estuary.

The Sundays Estuary discharges into Algoa Bay approximately 30 km northeast of Port Elizabeth (Emmerson, 1989). The catchment area of this estuary is approximately 22 000 km^2 and comprises predominantly of arid Karoo. Freshwater inflow into the Sundays Estuary is supplemented by an inter-basin transfer scheme from the Gariep River system. The primary usage of land surrounding the estuary is for commercial agriculture with a focus on cattle, sheep and citrus farming. As a result of this agricultural activity, the Sundays Estuary is nutrient enriched with a nitrogen:phosphate ratio of 3:1 (Emmerson, 1989; Scharler and Baird, 2003).

The Swartkops Estuary, which also discharges into Algoa Bay, is located ~11 km north of Port Elizabeth and has a catchment area of approximately 1360 km^2 (Baird et al., 1986). This estuary is severely impacted anthropogenically by both urban and industrial pollutants. Industrial activities along the Swartkops Estuary include sewage treatment, salt works, clay mining works as well as wool washers and tanneries (Baird et al., 1986; Binning and Baird, 2001). Furthermore, the lower reaches are heavily urbanised and the Motherwell and Markman stormwater canals drain into the estuary from industrial township and residential areas whilst the Chatty River flows, which flows through informal settlements, also discharges into the Swartkops Estuary (Scharler and Baird, 2003). As a consequence of these activities, the inorganic nutrient load is high in the Swartkops Estuary and there is evidence of heavy metal accumulation within the sediments (Binning and Baird, 2001).

2.2. Sample Collection

Three sites corresponding to the upper, middle and lower reaches of each of the four estuaries were identified (Table 1). At each site, 200 mL of surface water was filtered through a 0.22 μm PES filter (Supor®-200, PALL) whilst sediment samples were collected at a depth of between 0.5 m and 2 m, in sterile 2ml eppendorf tubes. All samples were stored on dry ice immediately after collection in order to minimize degradation of genomic material. Once in the laboratory, the samples were transferred and stored at -20°C until extraction of genetic material was carried out.

Table 1. The date and GPS location of each of the sampling efforts in this study

Estuary	Date of sampling	Tides	GPS coordinates		
			Mouth	Middles reaches	Upper reaches
Kariega	19/04/2011	High	33° 406 Lat, 26° 410 Long	33° 634 Lat, 26° 643 Long	33°560 Lat, 26° 637 Long
Kowie	02/03/2012	High	33° 35.648 Lat, 26° 53.551 Long	33° 35.515 Lat, 26° 52.448 Long	33° 34.225 Lat, 26° 51.605 Long
Sundays	04/10/2012	High	33° 41.88 Lat, 25° 50.20 Long	33° 42.68 Lat, 25° 47.791 Long	33° 41.250 Lat, 25° 46.425 Long
Swartkops	14/10/2013	High	33° 51.348 Lat, 25° 37.176 Long	33° 50.336 Lat, 25° 35.876 Long	33° 48.979 Lat, 25° 32.837 Long

2.3. Physico-chemical analysis

An Aquaread Aquameter (200/ Aquaprobe 800) or Eutech aquameter (CyberScan series 600) was used to measure the salinity (PSU), water temperature (°C) and dissolved oxygen ($mg.L^{-1}$) of the water column at each of the sites. In order to determine the chlorophyll-a (Chl-a) concentration within the water column at each sampling site, 200-500 mL of surface water was filtered through GF/F filters (Whatman). These filters were stored in aluminium foil at -20°C until further processing the laboratory. The Chl-a pigments were extracted from the filters by incubation at 4°C in 6 mL 90% acetone for 24 hours. Modified from Holm-Hansen and Riemann (1978), a Turner 10AU fluorometer was used to determine the Chl-a concentration before and after acidification. Water column samples which had been filtered through 0.22 μm PES filters (Supor®-200, PALL) were sent for analysis of phosphate, nitrite, nitrate and silicate concentrations (R. Roman at the Department of Oceanography, University of Cape Town).

2.4. Amplicon library preparation and multiplex pyrosequencing

The PowerWater DNA and PowerSoil DNA Isolation kits (MoBio Laboratories) were used to extract total genomic DNA (gDNA) from the water column and sediment samples taken. Amplicon libraries of the hypervariable regions 4 and 5 of the bacterial 16S rDNA were generated using the primer pair E517F (5'-CAGCAGCCGCGGTAA-3') and E969-984 (5'-GTAAGGTTCYTCGCGT-3'). Polymerase chain reaction (PCR) amplification was carried out in a two-step manner. Primary PCR used the following parameters: ~10 ng gDNA in a final volume of 25 μL in X1 PCR buffer, 0.3 μM of each primer, 300 μM dNTPs and 0.5 units of high fidelity KAPAHiFi Hotstart DNA Polymerase (KAPA Biosystems). The cycling parameters applied were: 98°C for 5 min (1 cycle); 98°C for 45 s, 45°C for 30 s, 72°C for 45 s (5 cycles); 98°C for 45 s, 50°C for 30 s, 72°C for 45 s (15 cycles); 72°C for 5 min. The PCR products were sized by agarose gel electrophoresis (1% w/v) and the ~450 bp amplification product was purified from the agarose using the Zymo Gel DNA Recovery Kit (Zymo Research). Approximately 2 ng of the purified amplification product was then subjected to a secondary PCR reaction using fusion primers containing Multiplex Identifier Tags in order to barcode the amplicon

libraries from each sample site. The reaction volumes used were as described for the primary PCR. The cycling parameters were: 98°C for 5 min (1 cycle); 98°C for 20 s, 52°C for 45 s, 72°C for 1 min (5 cycles); 98°C for 20 s, 65°C for 45 s, 72°C for 1 min (8 cycles) and a final extension at 72°C for 5 min. The resultant amplicons were agarose gel-purified as described above and pooled in equimolar amounts before pyrosequencing using the GS Junior Titanium Sequencer as per the manufacturer's specification (454 Life Sciences, Roche).

2.5. Computational analyses

2.5.1. Dataset curation and analysis

Quality filtering of the sequence reads generated by pyrosequencing was done using the affiliated GS Junior software provided by 454 Life Sciences and cured of primer and tag sequences. Further curation was done using Mothur software (Schloss et al., 2009) wherein all reads containing ambiguous nucleotides, all reads with homopolymers >7 and all reads shorter than 200 bp were removed from the dataset. The Naïve Bayesian classifier algorithm, hosted by the Ribosomal Database Project (RDP) (Cole et al., 2009, Wang et al., 2007), was used to phylogenetically classify the reads down to the taxon level of genus. Chimeras were detected using UChime (Edgar et al., 2011) and subsequently removed from the dataset using Mothur. The rarefaction calculations were carried out using the rarefaction analysis tool on the RDP Pyrosequencing Pipeline and curves generated for 0.01, 0.03, 0.05 and 0.1 distance values (http://pyro.cme.msu.edu/). Operational taxonomic units (OTUs) were determined at a distance value of 0.03 using Mothur (Schloss et al., 2009) and then classified using Basic Local Alignment Search Tool (BLAST) analysis against the National Centre for Biotechnology Information (NCBI) database (Altschul et al., 1990) or using the Naïve Bayesian classifier algorithm against the Silva 16S rRNA database (Quast et al., 2013). Principal Coordinate analysis (PCoA) was used to determine the degree of similarity between the microbial community in the individual estuaries, the sampling sites and the sediment/water column. OTUs (determined at a distance level of 0.03) which constituted <0.1% of the total sequence reads were removed for each sample and Clearcut software (Evans et al., 2006) used to construct a distance based neighbour-joining phylogenetic tree for use in generating the PCoA plot. The PCoA plot was made using the Fast Unifrac software (http://unifrac.colorado.edu) (Hamady et al., 2010) with weighted (i.e. incorporating abundance data) and normalised Unifrac pairwise distances. Targeted analysis of sub-sets of reads was done using StandAlone BLAST (Comacho et al., 2008) against a self-created database containing Cyanobacterial 16S rRNA gene sequences downloaded from the NCBI database. The sequence reads for each sample sites were deposited in the sequence reads archive (SRA) database of the National Centre of Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/Traces/sra) with the accession number SRP039885, SRP045647, SRP042993 and SRP045660.

2.5.2. Key considerations when analysing data

Sampling depth

In order to determine whether the specific sequence dataset was sufficiently large enough to provide an accurate representation of the bacterial population present at a specific sampling site, rarefaction

statistical analysis is carried out. Rarefaction analyses involve plotting the number of unique phylotypes within the dataset relative to the number of individuals present within a randomly selected sub-set of the sequence dataset. If only a few of the phylotypes present in the target environment or sampling site are represented in the dataset, then the chance of identifying a new phylotype within each sub-set of reads is increased and consequently, the slope of the rarefaction curve will be steep. Rarefaction curves that reach a plateau therefore, indicate that the species diversity within the target environment has been sampled to completion (Hughes and Hellmann, 2005).

It is generally accepted that the sequence divergence present within the 16S rDNA between bacterial species, genera and families/classes occur at 3, 5 and 10% respectively (Stackebrandt and Goebel, 1994; Hugenholtz et al., 1998; Sait et al., 2002). Thus if the percentage of nucleotides which differ between 16S rRNA sequence reads exceeds these values, the sequences in question are not grouped within the same species/genus/family.

When comparing different datasets together, it is important to normalise the number of reads represented in each of the datasets. This is particularly important when investigating operational taxonomic units (OTUs). An OTU refers to a set of reads with similar nucleotide sequences which are grouped together. By specifying the parameters within the algorithm used to determine OTUs, this grouping may be of identical sequences (i.e. each OTU represents an individual strain) or at a specified degree of variation (e.g. reads within a group with <0.03% divergence in their nucleotide sequence reflect a single species). In this study, datasets compared against one another were randomly sub-sampled such that each dataset contained the same number of reads. This was to avoid potentially skewed data in which samples represented by large datasets which, due to the increased number of reads present, appear to contain more OTUs than samples represented by smaller datasets.

Rare versus dominant biosphere

With respect to microbial diversity within a given environment, there are those phylotypes which are numerically dominant and those which occur at extremely low abundances (rare biosphere). It has become evident that the rare biosphere is substantially larger and more diverse than originally supposed (Sogin et al., 2006). The importance of the rare biosphere in contributing the biogeochemical cycling within a given ecosystem should not be underestimated. Studies carried out on the rRNA (representing the metabolically active bacteria) and rDNA (representing the presence of bacteria) on coastal ocean and estuarine samples showed that the rare phyla were often proportionately more active than the abundant phyla and therefore may play an important role in nutrient cycling (Campbell et al., 2011; Campbell and Kirchman, 2013).

One of the major disadvantages of using traditional Sanger sequencing of clone libraries is the limited sampling depth that can be achieved. Since only a few hundred clones can be sequenced, only the dominant, most common phylotypes would be identified whilst the phylotypes which occur in relatively low abundances (rare biosphere) would not be detected. In contrast, pyrosequencing generates

many thousand sequence reads per target environment thereby allowing for the identification of the rare biosphere. An important caveat to keep in mind is the inherent error rate of polymerases, even if they are high fidelity enzymes, used in both the PCR and pyrosequencing processes. This error rate may result in the overestimation of the degree of microbial diversity (Kunin et al., 2010). This potential error can be minimised by stringent quality filtering of the data sets (Huse et al., 2007). In addition, by setting the cut-off threshold for defining OTUs at 0.03% divergence, which is the degree of divergence generally considered as delineating between species (Stackebrandt and Goeble, 1994), will further decrease potential diversity overestimation since sequences which differ by as much as 3% of their nucleotide sequence will still be grouped together a single OTU (Kunin et al., 2010).

An additional caveat to keep in mind when utilising amplicon libraries to ascertain the abundances of dominant phyla relative to the rare phyla is that these libraries are generated via PCR. Under these conditions, the exponential amplification of the numerically dominant microbes would surpass that of the species occurring at very low frequencies. This will not affect the identification of the more abundant bacterial species, but it will exaggerate the relative abundances of the numerically rare species compared to those which are abundant. In order to minimise this occurrence, the number of amplification cycles utilised in the PCR program should be kept to a minimum.

Phylogenetic versus OTU analysis

Classification of the sequence reads within a dataset against a database to generate phylogenetic rankings is an invaluable tool for obtaining information on the type of microbes that are present in the target sampling site as well as the proportions of each taxon relative to one another. There are, however, two major limitations to this type of analysis. Firstly, the scope of the reference database against which the query reads are compared limits the degree to which a read can be assigned a phylogenetic rank. If there is no corresponding reference sequence within the database against which to compare the query read, then classification of the read will be limited to a higher taxon level or will remain unclassified. The second limitation lies in the inability to determine if individual species are unique to a particular sample or if they are found in multiple samples. In order to resolve these limitations, analysis of the operational taxonomic units within a dataset is useful.

Singletons

The term “singletons” refers to instances in which an OTU is represented by a single read. While these should never be discounted, it is important to view singletons with care. This is particularly true if the distance value utilised to determine OTUs is <0.01%. While the rare biosphere within an ecosystem will often be represented by singletons in a sequence dataset, due to the nature of sampling in the open environment, singletons may also represent bacteria that are transient in that they may, for example, represent airborne microbes which have “landed” in the target ecosystem. In addition, larger datasets will have a greater number of singletons that smaller datasets and this results in the potentially incorrect assumption that one sample has increased species diversity compared to the other when analysing OTUs. For the purposes of this study, singletons were

included in the overall phylogenetic classification of sequence reads from each sample, but they were excluded during OTU analyses.

3. COMPARATIVE ANALYSIS OF FOUR EASTERN CAPE ESTUARINE SYSTEMS

3.1. Physicochemical analyses

The physicochemical data collected for the water column at three sites along the length of each of the four target estuaries are presented in Table 2. The salinity measurements for the Kowie and Kariega estuaries remained high along the length of the estuary. This was likely due in part to the reduced freshwater inflow into these estuaries or alternatively it may reflect an increased influence of the ocean tides on this ecosystem. Kariega Estuary is known for its limited freshwater input both as a result of its small catchment area (~680 km^2) as well several upstream impoundments (Hodgson, 1987; Grange and Allanson, 1995; Froneman, 2000). The Kowie estuary also has a relatively small catchment area of 769 km^2 (Noble and Hemens, 1978). In contrast, the Sundays and Swartkops estuaries demonstrate a sharp decrease in salinity from a marine environment at the mouth (PSU 30-35) towards an increasingly freshwater environment upstream (PSU 7-8) (Table 2).

Table 2. Summary of the physicochemical parameters of the water column along the length of the Kariega, Kowie, Sundays and Swartkops estuaries.

Estuary	Site	Temperature (°C)	Salinity (PSU)	Phosphate* (µmol/L)	Nitrite* (µmol/L)	Nitrate* (µmol/L)	Silicate* (µmol/L)	Chlorophyll-a (µg/L)
Kariega	Mouth	19.9	35	-	-	-	-	0.4
	Mid	21.9	30	-	-	-	-	1.95
	Upper	21.7	20	-	-	-	-	1.98
Kowie	Mouth	26	29	-	-	-	-	1.08
	Mid	27.4	19	-	-	-	-	0.27
	Upper	25.5	14	-	-	-	-	0.25
Sundays	Mouth	19.3	35	1.2	1.2	22.8	56.8	0.7
	Mid	20.5	13	1.8	1.3	40.2	107.9	2.1
	Upper	20.4	8	2.4	0.7	50.4	190.6	2.2
Swartkops	Mouth	19.2	30	0.45	0.09	1.18	2.23	0.22
	Mid	18.7	15	7.43	0.43	67.31	7.54	0.98
	Upper	20.2	7	19.64	3.72	16.77	28.68	15.2

* Phosphate, nitrite, nitrate and silicate concentrations were not determined for the Kariega and Kowie estuaries

Analysis of the phosphate, nitrite, nitrate and silicate concentrations was implemented for the Sundays and Swartkops estuaries, both of which were found to be significantly nutrient enriched. Notably, both estuaries contain high concentrations of nitrate and silicate, particularly in the upper reaches (Table 2). The Swartkops Estuary exhibited elevated phosphate and nitrite concentrations compared to those found in the Sundays Estuary. With the exception of the spike in nitrate levels in the middles reaches of the Swartkops Estuary, a distinct increase in measured inorganic nutrients

was observed from the mouths of the estuaries to the upper reaches. The chlorophyll-a (Chl-a) concentrations in the Kariega, Kowie and Sundays estuaries were relatively low (<2.2 μg/L) whilst the upper reaches of the Swartkops Estuary reached Chl-a levels of 15.2 μg/L (Table 2).

3.2. Sequence sampling depth

A total of 179 950 reads were generated from the amplicon libraries prepared during this study. The distribution of these reads to the respective sampling sites and source material (i.e. water column vs. sediment) of the Kariega, Kowie, Swartkops and Sundays estuaries is presented in Table 3.

Table 3. Summary of the sequence reads, generated via pyrosequencing of amplicon libraries, subsequent to sequence curation.

Estuary	Water column			Sediment		
	Mouth	Mid	Upper	Mouth	Mid	Upper
Kariega	8 318	18 994	16 727	6 836	18 580	16 111
Kowie	2 868	3 531	6 560	3 398	4 802	3 070
Sundays	2 244	7 652	5 931	8 616	8 022	6 680
Swartkops	3 518	6 062	6 660	2 408	6 163	6 314

Rarefaction analysis was carried out on all of the datasets outlined in Table 3 in order to determine if, statistically, a good coverage of the overall diversity within the target sampling site had been achieved. This analysis was carried out at distances value of 0.03, 0.05 and 0.1, which correspond to the accepted 16S rDNA sequence divergence between bacterial species, genera, and families, respectively (Stackebrandt and Goebel, 1994; Hugenholtz et al., 1998; Sait et al., 2002). As illustrated in Figure 2, sampling at the taxon level of family was done to completion or near completion for all sites within the Kariega, Kowie, Sundays and Swartkops estuaries with the exception of the Sundays Estuary water column. While the rarefaction curves for the sequence divergence of 5% (genus) and 3% (species) of these four estuaries were not asymptote, the curve had begun to plateau. Rarefaction values resulting in graphs with slight curves suggest that whilst the very rare species may have been missed, the dataset allowed for an accurate representation of the majority of the bacterial population.

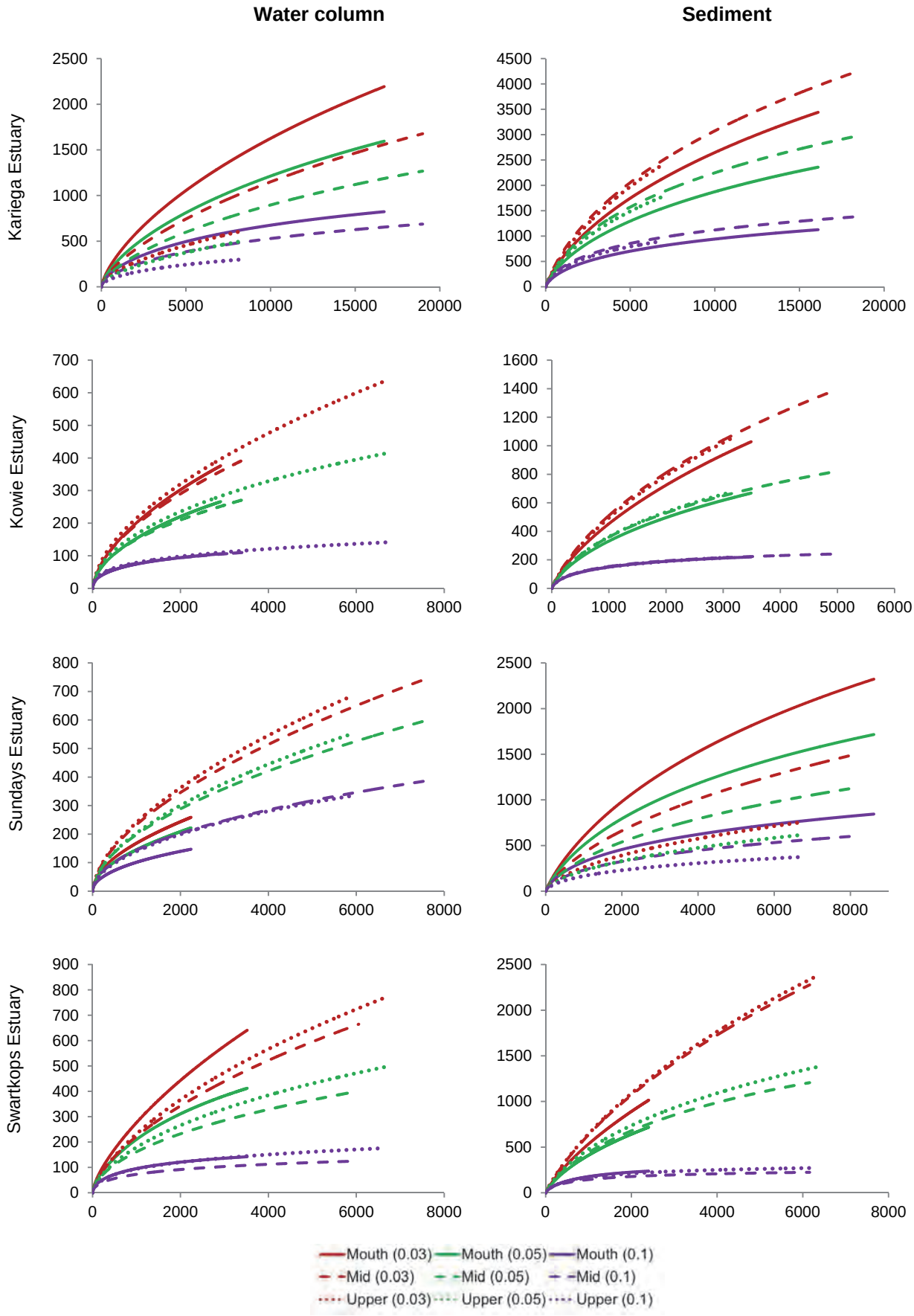


Figure 2. Rarefaction curves generated at distance values of 0.3, 0.5 and 0.1 for the three sites along the length of the Kariega, Kowie, Sundays and Swartkops estuaries. The number of phylotypes and the number of sequence reads are plotted against the x- and y-axis respectively.

3.3 OTU-based comparison of four estuaries

Operational taxonomic units (OTUs) were determined at a sequence distance of 0.03 (i.e. species) and a phylogenetically-based metric (Unifrac) was calculated and used as input for a principal component analysis (PCoA) plot (Figure 3).

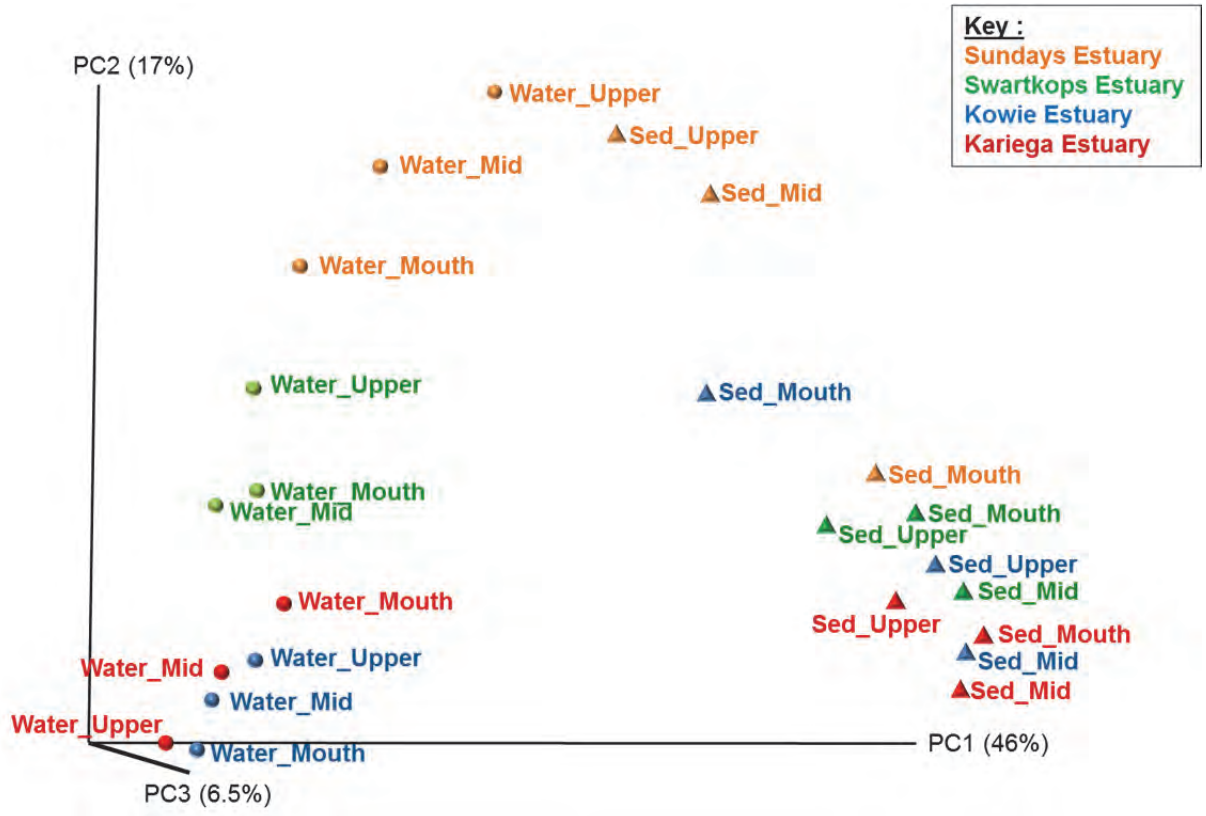


Figure 3. Principle coordinate analysis (PCoA) of the bacterial population occurring within the water column and sediment of the Kariega, Kowie, Sundays and Swartkops estuaries. The PCoA was generated with a Unifrac distance matrix, water and sediment samples for the Kariega (red), Kowie (blue), Sundays (orange) and Swartkops (green) estuaries are shown as circles, while triangles represent sediment samples.

The PCoA plot revealed that in all four estuaries the bacterial communities in the water column was distinctive from those present in the sediment (Figure 3). Additionally, with the exception of the sediment sample collected in the Kowie mouth, a strong clustering was observed for the Kariega and Kowie Estuary water column and sediment samples. The sample representing the Kowie mouth sediment was collected from within the Port Alfred Marina. The Sundays Estuary and the upper reaches of the Swartkops Estuary, however, did not correlate well with the Kowie and Kariega estuaries. In contrast, the lower and middle reaches of the Swartkops Estuary clustered slightly nearer the Kariega and Kowie Estuary samples than to that of the Sundays Estuary. We concluded that this pattern of cluster distribution likely reflected the increased freshwater input into the Sundays Estuary and the upper reaches of the Swartkops Estuary. Consequently, it was decided that the subsequent detailed analysis of the data for these four estuaries would be split into the marine-dominated estuaries (Kowie and Kariega) and the freshwater-dominated estuaries (Sundays and Swartkops).

4. MICROBIAL DIVERSITY IN FRESHWATER-DOMINATED ESTUARIES

4.1. Bacterial taxonomic diversity

Phylogenetic classification of the 16S rDNA sequences generated from the water column and sediment of the Sundays and Swartkops estuaries was done using the 16S rDNA database hosted by the Ribosomal Database Project and the output data are graphically represented at the level of phylum or class (Figure 4). A more exhaustive breakdown of the phylogenetic rankings of the sequences is supplied in Appendix 2.

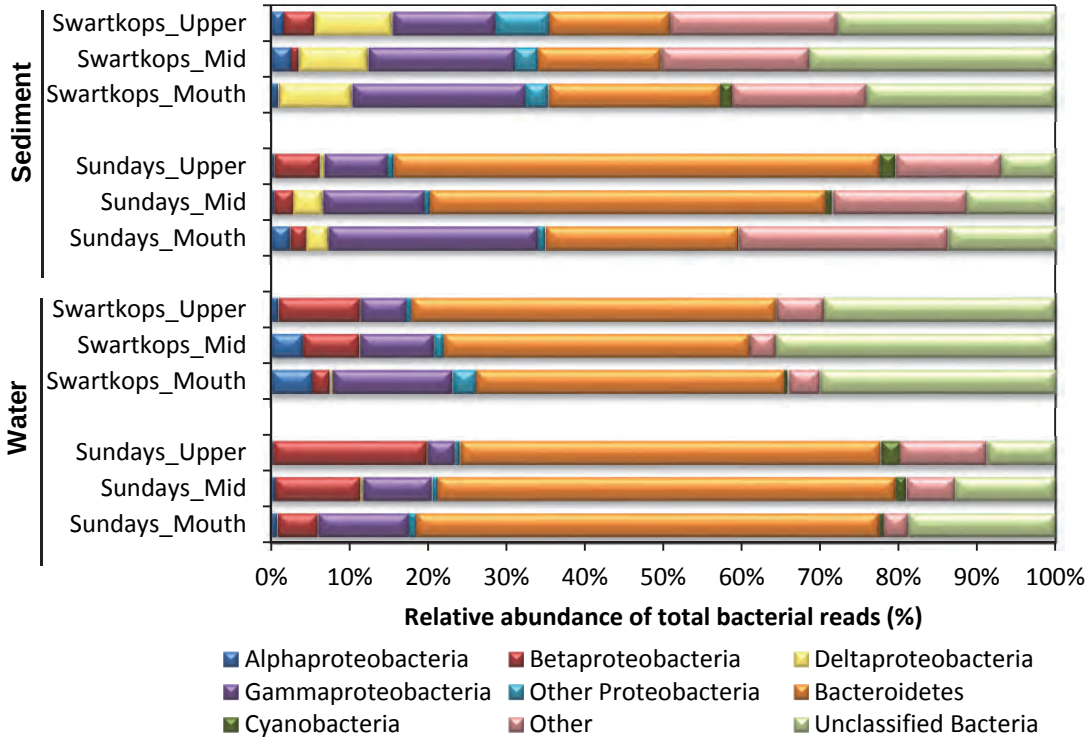


Figure 4. Relative abundances of the dominant phyla within the water column and sediment of the Sundays and Swartkops estuaries. Assignment of reads to taxonomic rankings was done using the Bayesian algorithm against the Ribosomal Database Project. Reads designated as "Other" represent the sum of the reads assigned to less well represented phyla.

The majority of the bacterial 16S rDNA reads within the sediment and water column of the Sundays Estuary belonged to the phylum Bacteroidetes (24.5% to 61.8% of the total reads) and, to a lesser extent, Proteobacteria (15.5% to 34.8% of the total reads) (Figure 4). While this holds true for the water column of the Swartkops Estuary with 39% to 46.5% of the total reads assigned to the phylum Bacteroidetes, the sediment in this estuary was dominated by Proteobacteria (39.9% to 35.4% of the total reads) rather than Bacteroidetes (15.3% to 22% of the total reads). Of the Proteobacteria, Deltaproteobacteria formed a significant proportion of the bacterial community in the sediment of the Swartkops Estuary with a lesser representation within the Sundays Estuary sediment. In both estuaries, the Gammaproteobacteria decreased in abundance whilst the Betaproteobacteria abundances increased as the salinity decreases from the mouth to the upper reaches. The number of

reads that did not exhibit significant homology to any 16S rRNA sequence within the database ("Unclassified bacteria") is much higher in the Swartkops Estuary than in the Sundays Estuary (Figure 4). This may be indicative of the increased inflow into the Swartkops Estuary of complex nutrients, including xenobiotics, from the surrounding urban and industrial activities.

The high proportion of the sum of the reads assigned to the less well represented phyla for the sediment samples in both estuaries (Figure 4, "Other") was expected and likely reflects the more heterogeneous composition of sediments compared to that of the well-mixed water column. A closer examination of the less well represented phyla (Figure 5) substantiate this hypothesis with the water column samples being dominated by Actinobacteria and Verrucomicrobia whilst the sediment samples presented a much wider range of diverse phyla. Representatives of the phylum Chloroflexi were found throughout the length of the Sundays and Swartkops estuary sediments. Members of the phylum Verrucomicrobia were more prevalent in the Sundays Estuary while Planctomycetes occurred more frequently in the Swartkops Estuary.

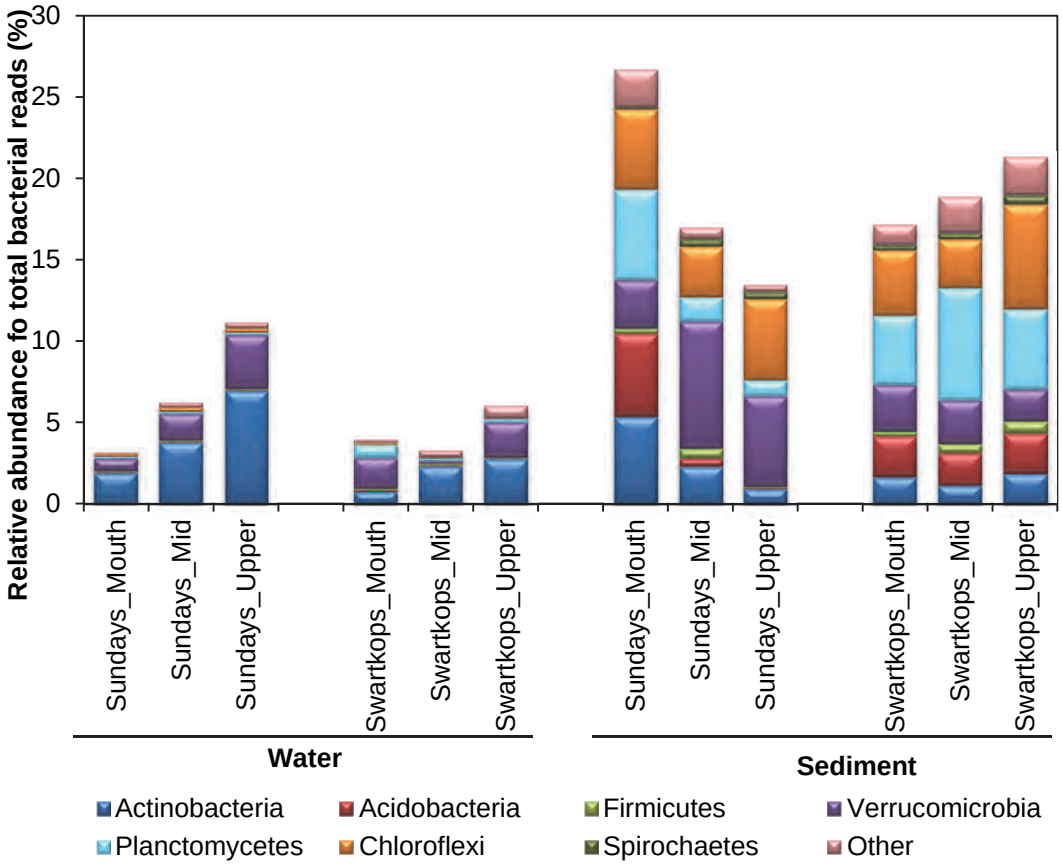
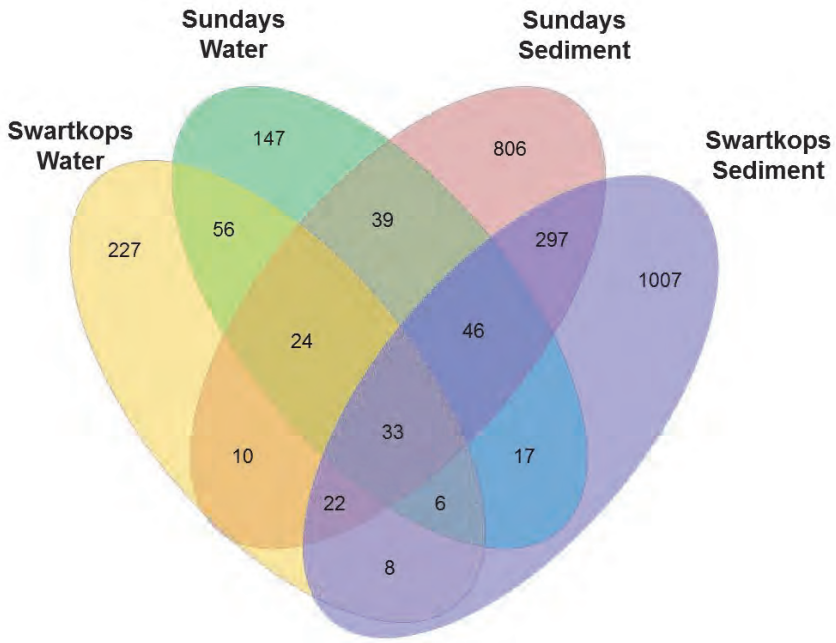


Figure 5. Relative abundances of the less well represented phyla within the dataset generated from water column and sediment samples along the length of the Sundays and Swartkops estuaries. Assignment of reads to taxonomic rankings was done using the Bayesian algorithm against the Ribosomal Database Project.

4.2. Operational taxonomic units (OTU) analysis

OTU analysis was carried out in two ways: firstly by focusing on the identification of OTUs which were found to be common to the different ecosystems and which were unique, and secondly by examining the relative abundance of the reads assigned to the dominant OTUs in each ecosystem. In this analysis, the datasets were randomly sub-sampled to accommodate differences in dataset sizes, OTUs were determined at a distance value of 0.03 (i.e. taxonomic level of species) and singletons were removed (see section 2.5.2).



	Sundays Water	Sundays Sediment	Swartkops Water	Swartkops Sediment
Sundays Water	368	142	119	102
Sundays Sediment	142	1277	89	398
Swartkops Water	119	89	386	69
Swartkops Sediment	102	398	69	1436

Figure 6. Venn diagrams illustrating the distribution of OTUs from the water column and sediment of the Sundays and Swartkops Estuaries.

In both the Sundays and Swartkops estuaries, a greater species diversity was observed in the sediment (1 277 and 1 436 OTUs, respectively) compared to the water column (368 and 386 OTUs, respectively) (Figure 6). The majority of the OTUs identified in each estuary were unique to that system. This was most obvious in the sediment sample datasets in which 63% and 70% of the OTUs from the Sundays and Swartkops sediments respectively, were unique. With respect to the water column, 40% and 59% of the OTUs from the Sundays and Swartkops respectively were unique. The degree of commonality with respect to OTUs between the Sundays and Swartkops water column was observed to be 15.8% of the OTUs, whilst 14.6% of the OTUs found within the sediment datasets were found in both the Sundays and Swartkops estuaries (Figure 6).

Whilst the data illustrated in Figure 6 is valuable, the picture provided is incomplete in that the abundances of the OTUs are not taken into account. For example, a species may be present in two different ecosystems but be prevalent in the one ecosystem and rare in the other. Consequently, a look at the relative abundances of the dominant OTUs were investigated by selecting the top 10 most dominant OTUs for each of the sampling sites and assessing the abundances of the reads assigned to that OTU relative to the total number of reads generated for the corresponding sample site (Figure 7). The taxonomic classification of the dominant OTUs, as determined by BLAST analysis against the NCBI database, is detailed in Table 4.

The most striking characteristic of the distribution and abundance patterns of the dominant OTUs in the Sundays and Swartkops estuaries was the distinct separation of those OTUs found in the water column compared to those from the sediment (Figure 7). This finding correlates with the data illustrated in Figures 4 and 5 and reinforces the apparently distinct ecological niches present in the sediment versus the water column of both estuaries. BLAST analysis of the dominant OTUs from both the Sundays and Swartkops estuary revealed highest relatedness to bacteria isolated from marine environments with only 14 of the 68 OTUs showing high relatedness to microorganisms isolated from freshwater ecosystems (Table 4). Of the 14 sequences extracted from the NCBI database designated as freshwater in origin, the number of reads assigned to the corresponding OTUs for both estuaries was exclusively or predominantly (depending on the OTU) from the upper reaches of the estuaries. This is not surprising as the upper reaches of both estuaries have much lower salinities and are almost completely freshwater environments. Also of interest was that while there is overlap in that the same OTUs are found in both estuaries (e.g. OTUs 28-42), there were also several OTUs present in only one estuary or at low levels in the other (e.g. OTUs 19-27 are associated with the Sundays Estuary, whilst OTUs 45-54 and OTU56-52 are predominantly or exclusively found in the Swartkops Estuary).

Examination of the OTU distribution profiles between the mouth, middle and upper reaches of the estuaries showed that OTUs found at the middle reaches were also found in the upper reaches or mouth (Figure 7). Not surprisingly, less of an overlap is observed when considering the mouth and upper reaches samples. There are dominant OTUs which occur almost exclusively at the mouth (OTUs 11,13,15-17 for the Sundays Estuary and OTUs 50,65-67 in the Swartkops Estuary) and those which are almost exclusively found in the lower salinity upper reaches (OTUs 19-27 in the Sundays Estuary and OTUs 54-57, 30, 33, 44 in the Swartkops Estuary). OTU 29, which occurred in significant numbers in the Swartkops Estuary (representing 17.8% to 26.3% of the total number of reads), showed high homology to the SAR11 clade (Table 4).

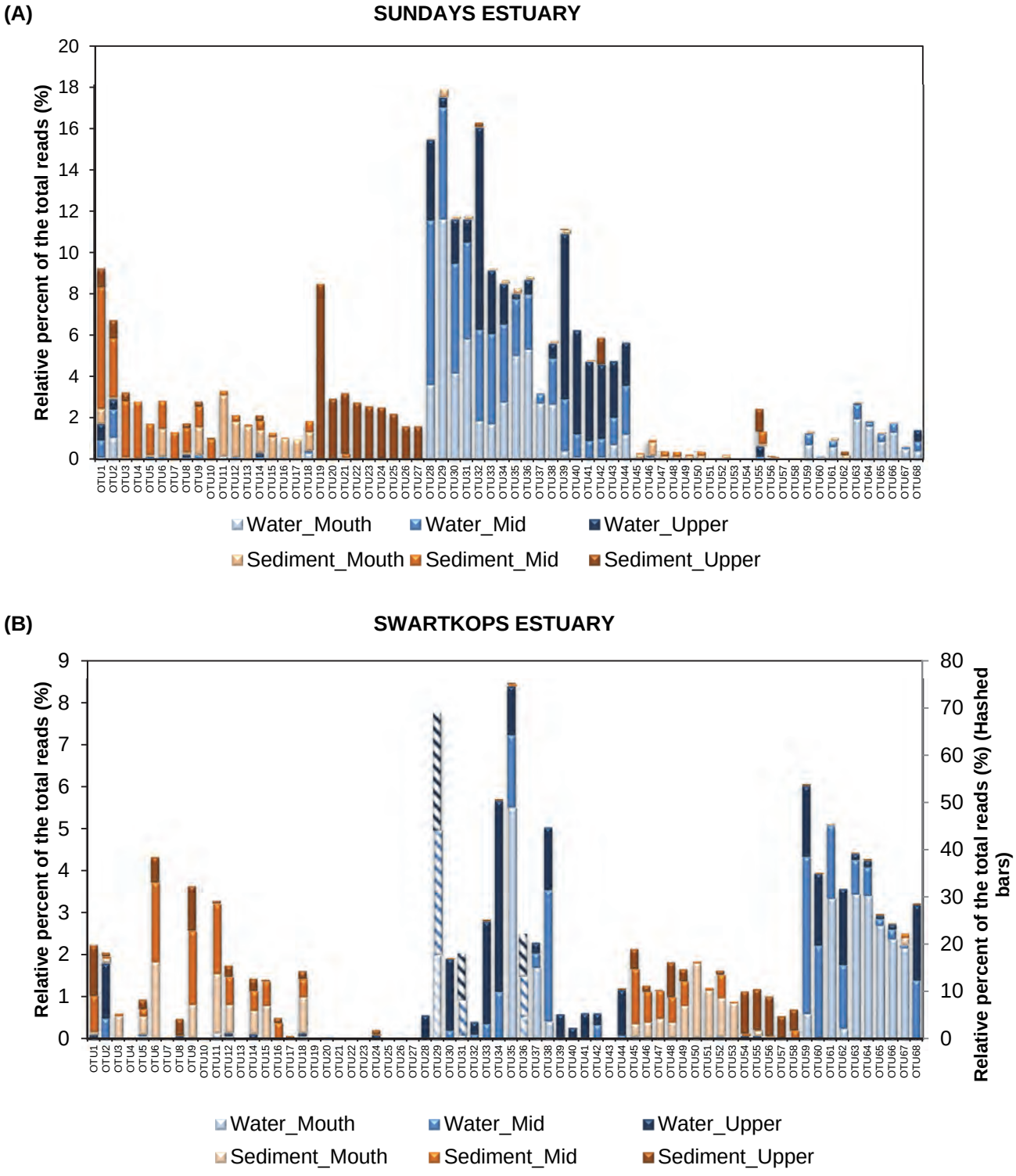


Figure 7. Comparative analysis of the ten most dominant OTUs found within the sediment and water column of the Sundays (A) and Swartkops (B) estuaries. OTUs were determined at a distance value of 0.03.

Table 4. Results of the BLAST analysis of the dominant OTUs from the Sundays and Swartkops estuaries indicating the NCBI database sequence to which the query OTU sequence exhibited highest percentage identity.

OTU	Highest classification of database organism	Phylum	Accession number	Identity (%)	Source
1	Flavobacteriaceae bacterium	Bacteroidetes	JX854356.1	99	North Sea
2	Flavobacteriaceae bacterium	Bacteroidetes	JQ807811.1	99	Intertidal zone (Lingshan Bay)
3	Uncultured bacterium	/	AF159646.1	99	Bacteria associated with sea-grass
4	*Muriicola jejuensis*	Bacteroidetes	EU443206.1	99	Seawater
5	Uncultured bacterium	/	JX391488.1	99	Surface Marine Sediments
6	Uncultured bacterium	/	JF272049.1	99	Marine biofilms
7	Uncultured bacterium	/	HE611103.1	99	Bacteria associated with mud crab
8	Flavobacteriaceae bacterium	Bacteroidetes	JN033800.1	99	Seawater
9	Uncultured bacterium	/	JN839396.1	99	Shallow marine hydrothermal vent
10	Uncultured bacterium	/	EF192878.1	99	Sediments of a temperate artificial lake
11	Flavobacteriaceae bacterium	Bacteroidetes	AM990866.1	100	Coastal Mediterranean ecosystem
12	Uncultured bacterium	/	HQ703828.1	99	Marine sediment
13	*Sediminicola* sp.	Bacteroidetes	JX854358.1	100	North Sea
14	Uncultured bacterium	/	KF185417.1	100	North Adriatic Sea
15	Uncultured γ-proteobacterium	Proteobacteria (γ)	JN672630.1	100	East China sea sediment
17	Uncultured bacterium	/	AY592113.1	99	Mediterranean cold seep
18	Flavobacteriaceae bacterium	Bacteroidetes	DQ993343.1	100	South Korean corals
19	Uncultured Flammovirgaceae	Bacteroidetes	FJ516777.1	99	Biofilm from Central Spanish wetland
20	Uncultured bacterium	/	FJ354921.1	100	Lake Charles
21	*Mariniflexile fucanivorans*	Bacteroidetes	NR_042239.1	99	Mud from the water treatment facility
22	Uncultured Sphingobacteriales	Bacteroidetes	HQ857725.1	98	Hydrocarbon contaminated saline alkaline soil
23	Uncultured Flammovirgaceae	Bacteroidetes	FJ516977.1	92	Upper sediment from wetland
24	*Algoriphagus namhaensis*	Bacteroidetes	NR_109104.1	98	Seawater from the South Sea
25	*Bizionia* sp.	Bacteroidetes	JX844498.1	99	Seawater from Prydz bay, Antarctic
26	Uncultured γ-proteobacterium	Proteobacteria	JN825501.1	99	Alkaline lake Alchichica (Mexico)
27	*Pontibacter xinjiangensis*	Bacteroidetes	AB682652.1	91	Not specified
28	Uncultured Sphingobacteria	Bacteroidetes	AM279196.1	99	Marine plankton
29	Uncultured SAR11 cluster	Proteobacteria (α)	KF786857.1	100	Oil sheen weathering (Deepwater Horizon
30	Uncultured bacterium	/	FJ352183.1	100	Lake Pontchartrain
31	Uncultured Bacteroidetes	Bacteroidetes	HM057748.1	99	Ocean water from the Yellow Sea
32	Uncultured Bacteroidetes	Bacteroidetes	FJ916233.1	99	Temperate lake (Cox Hollow, USA)
33	Uncultured Bacteroidetes	Bacteroidetes	DQ656266.1	100	Surface seawater, Pearl River Estuary
34	Uncultured Comamonadaceae	Proteobacteria (β)	JN591844.1	100	Surface seawater, Puget Sound
35	Uncultured *Flavobacterium* sp.	Bacteroidetes	FJ745102.1	100	Coastal bacterioplankton
36	Uncultured *Owenweeksia* sp.	Bacteroidetes	JX529868.1	100	Southern Ocean
37	Uncultured bacterium	/	KF185669.1	100	North Adriatic Sea
38	Uncultured Bacteroidetes	Bacteroidetes	HM057636.1	100	Yellow Sea water
39	Uncultured *Limnohabitans* sp.	Proteobacteria (β)	HF968603.1	100	Lake water
40	Uncultured *Sediminibacterium*	Bacteroidetes	HM856392.1	100	Yellowstone Lake
41	*Flavobacterium* sp.	Bacteroidetes	KF556685.1	100	Freshwater pond
42	*Hydrogenophaga* sp.	Proteobacteria (β)	AM110076.2	100	Groundwater from a sandstone aquifer
43	Uncultured SAR11 cluster	Proteobacteria (α)	FN665729.1	100	Temperate lakes
44	Uncultured bacterium	/	KF596556.1	100	Gulf of Gdansk
45	Uncultured bacteria	/	JN977359.1	100	Jiaozhou Bay sediments
46	Uncultured γ-proteobacterium	Proteobacteria (γ)	KF463986.1	100	Marine coastal sediment
47	Uncultured δ-proteobacterium	Proteobacteria (δ)	JX241022.1	99	Coastal soil of Gulf of Khambhat
48	Uncultured Bacteroidetes	Bacteroidetes	JQ580152.1	100	Oil-polluted subtidal sediments
49	Uncultured δ-proteobacterium	Proteobacteria	JQ580333.1	100	Oil-polluted subtidal sediments
50	Uncultured γ-proteobacterium	Proteobacteria (γ)	JQ579920.1	100	Oil-polluted subtidal sediments
51	*Phormidium laetevirens*	Cyanobacteria	JF708120.1	99	Intertidal zone (Portuguese coast)
52	Uncultured Chloroflexi	Chloroflexi	AY500104.1	99	Fish farm sediments
53	*Muriicola* sp.	Bacteroidetes	KC839612.1	98	Seawater from the Mariana Trench
54	Helicobacteraceae bacterium	Proteobacteria (ε)	AJ810529.1	99	Central Baltic Sea
55	Uncultured Sinobacteraceae	Proteobacteria (γ)	HQ003540.1	100	Sub-saline Shallow Lake

OTU	Highest classification of database organism	Phylum	Accession number	Identity (%)	Source
56	Uncultured bacterium	/	DQ787711.1	100	Lake sediment
57	Uncultured bacterium	/	JN391906.1	100	Sludge in anoxic tank reactor
58	Uncultured	Proteobacteria (δ)	KC009880.1	99	French Guiana coast
59	Uncultured bacterium	/	KF185805.1	100	North Adriatic Sea
60	Uncultured bacterium	/	KF596579.1	99	Seawater (Gulf Gdansk)
61	Uncultured *Roseobacter* sp.	Proteobacteria (α)	JX530514.1	100	Southern Ocean
62	*Polaribacter* sp.	Bacteroidetes	JX304645.1	100	Seawater (South Korea)
63	Uncultured γ-proteobacterium		JQ515484.1	100	Caribbean reef-building coral
64	Uncultured Bacteroidetes		HM057663.1	99	Yellow Sea water
65	Uncultured SAR11 cluster	Proteobacteria (α)	KF786721.1	100	Oil sheen weathering (Deepwater Horizon
66	Uncultured bacterium	/	JX524984.1	100	Southern Ocean
67	Uncultured Flavobacteria	Bacteroidetes	FN433398.1	99	North Atlantic Ocean
68	Uncultured Actinomycetales	Actinobacteria	FJ745028.1	100	Coastal bacterioplankton

4.3. Chloroplast diversity

Chloroplasts are the organelles responsible for the photosynthetic activity of phototrophic eukaryotes. Due to their similarities in size, morphology and genetic material to that of cyanobacteria, chloroplasts have been hypothesized to be remnants of ancestral endosymbiotic cyanobacterial (Harris et al., 1996; Cavalier-Smith, 2002). Consequently, unicellular algal total genomic DNA also contains chloroplast 16S, 5S and 23S rRNA and, due to the sequence similarity of the chloroplast 16S rRNA to cyanobacterial 16S rRNA, the primer set utilised in this study fortuitously amplified both chloroplast and bacterial 16S rRNA.

The unicellular algae found within the sediment of both the Sundays and Swartkops estuaries were comprised almost exclusively of Bacillariophyta (Figure 8). By contrast, the water column exhibited a more diverse unicellular algal community profile shifting in dominance between the upper reaches and the mouth of the estuaries evident. In both the Sundays and Swartkops estuaries, a greater representation of Chlorophyta was evident near the mouth of the estuary with an increased abundance of Bacillariophyta in the upper reaches (Figure 8). With respect to the Sundays Estuary, the relative abundances of unicellular algae between the middle and upper reaches were very similar. By contrast, each of the sites sampled for the Swartkops Estuary showed differing community profiles with a significant representation of Cryptomonadaceae in the middle reaches and a greater than 12-fold increase in Bacillariophyta between the estuary mouth and upper reaches (Figure 8).

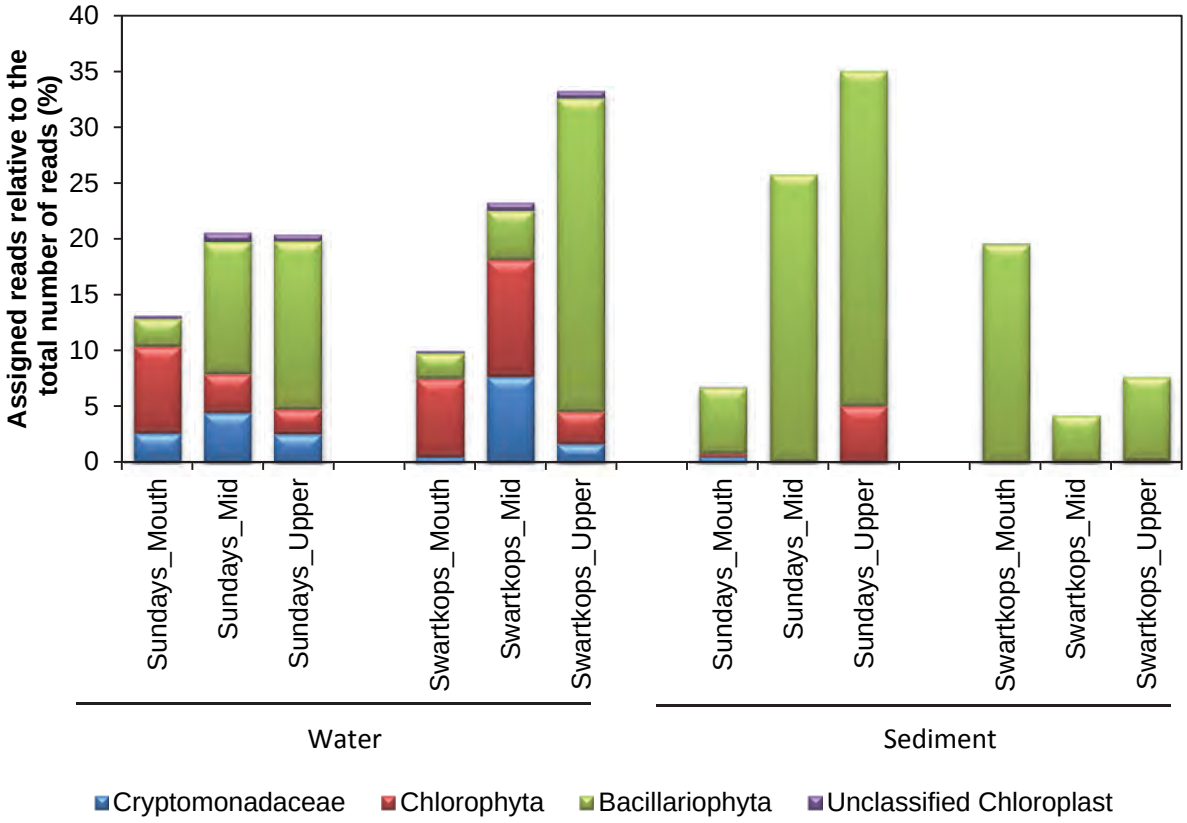


Figure 8. Relative abundance of 16S rDNA reads classified as chloroplast in origin from the water column and sediment of the Sundays and Swartkops estuaries relative to the total number of 16S rDNA reads generated for each site. Mid=middle reaches

4.4. Potential pathogens

Due to the anthropogenic inflow of untreated or poorly treated waste water into many of the estuaries in South Africa, waterborne human pathogens are a major concern. Pathogenic bacteria that are of particular concern include *Enterobacteriaceae, Yersinia enterocolitica, Campylobacter jejuni, Vibrio cholerae, Mycobacterium* sp*., Legionella pneumophila, Pseudomonas aeruginosa, Aeromonas* sp*., Shigella,* and *Salmonella* sp. (Leclerc et al., 2002; Sharma et al., 2003). The 70 270 reads generated for the Sundays and Swartkops estuaries in this study were examined to determine if any 16S rRNA gene sequences with high homology to the above mentioned pathogens could be identified.

The 16S rRNA gene sequences generated in this study were screened for potential pathogens against the RDP database and BLAST analysis against the NCBI database. This screening revealed a complete absence of reads with significant homology to *Enterobacteriaceae, Yersinia enterocolitica, Campylobacter jejuni, Pseudomonas aeruginosa, Aeromonas* spp*., Shigella,* or *Salmonella*. Within the Sundays Estuary dataset, 59 reads were assigned to the genus *Mycobacteria*, however, closer analysis revealed that none of these 16S rRNA sequences belonged to pathogenic Mycobacterium species. With respect to the Swartkops Estuary, the two reads assigned to the genus Mycobacterium showed higher homology to *Mycobacterium vanbaalenii* than to pathogenic species. *M. vanbaalenii*

isolates have been shown to be involved in the degradation of polycyclic aromatic hydrocarbons which are common constituents of fossil fuels (Kim et al., 2005).

When considering the genus *Legionella*, 1 and 13 reads were classified in this genus from the Swartkops and Sundays datasets respectively. However, all of these reads exhibited higher relatedness to species other than the pathogenic *Legionella pneumophila.* Reads from both the Swartkops and Sundays datasets that were classified as belonging to the genus *Vibrio* numbered at 26, of which 23 reads had a higher percentage identity with species other than *Vibrio cholera* (e.g. *V. anquillarum*, *V. azureus, V. breoganni, V. harveyi, V. splendidus*, *V. penctenicida, V. vulnificus*). A single read from the Swartkops Estuary and two from the Sundays Estuary showed 100% sequence homology to both *V. cholera* and *V. alginolyticus*. Whilst *V. alginolyticus* has been shown to infect aquatic animals, it is not a human pathogen (Xie et al., 2005). The absence of any other pathogenic species within these two datasets suggests that these three reads may indeed be *V. alginolyticus* rather than *V. cholera*, however, in order to verify the non-existence of *V. cholera* beyond doubt, a definitive test specific tailored specifically for *V. cholera* should be done.

5. MICROBIAL DIVERSITY IN MARINE-DOMINATED ESTUARIES

5.1. Bacterial taxonomic diversity

Classification of the 16S rRNA sequences generated from the water column and sediment of the Kariega and Kowie estuaries was done using the Ribosomal Database Project (RDP) database and the data are graphically represented at the level of phylum or class.

With respect to the water column in the Kariega Estuary, the majority of the reads did not show significant similarity to any sequence in the RDP database (Figure 9, “Unclassified”). The combined contribution of bacteria belonging to the Proteobacteria and Bacteroidetes phyla ranged from 37% to 60% of the total number of reads with the relative contribution increasing from the upper reaches towards the mouth of the estuary. With respect to the water column in the Kowie Estuary, the most striking difference is the high prevalence of Cyanobacteria (17% to 22% of the total reads). A closer examination of the reads assigned to this phylum revealed that >95% of these reads belong to the genus *Synechococcus*. With respect to the Proteobacteria, Gammaproteobacteria dominated the Kariega Estuary (13% to 17% of the reads) whilst the Kowie Estuary had increasing numbers of Betaproteobacteria in the upper reaches (7% of the total reads in the Kowie Estuary versus 2% in the Kariega Estuary). With regards to the phylum Bacteroidetes, an increased prevalence of bacteria belonging to this phylum was observed in the Kariega Estuary (16% to 31% of the total reads) compared to that found in the Kowie Estuary (13% to 16% of the total reads).

Shifting the focus to the bacterial community profiles within the sediment of these two estuaries (Figure 9), the most striking difference when comparing the sediment to the water column was the increase in Proteobacteria. This was particularly evident in the Kowie Estuary where a greater than 2-

fold increase was observed in the middle and upper reaches. In both the Kariega and Kowie estuaries, the most dominant Proteobacteria belonged to the Gamma- and Deltaproteobacteria whereas in the water column, the Gamma- and Betaproteobacteria dominated. Also of interest in the sediments of both estuaries was the high prevalence of reads assigned to less dominant phyla (Figure 9, "Other"), with between 18% and 20% of the total reads falling within this category. A breakdown of the phyla represented in this group ("Other") is outlined in Figure 10.

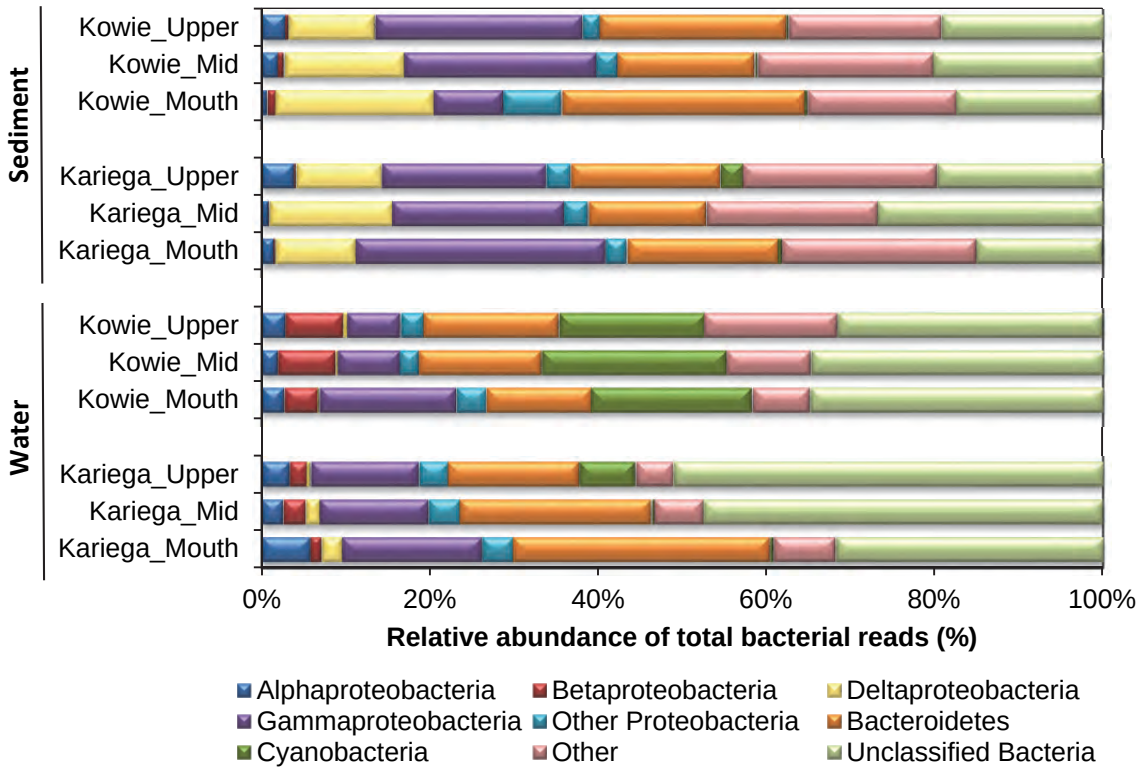


Figure 9. Relative abundances of the dominant phyla within the water column and sediment of the Kariega and Kowie estuaries. Assignment of reads to taxonomic rankings was done using the Bayesian algorithm against the Ribosomal Database Project. Reads designated as "Other" represent the sum of the reads assigned to less well represented phyla.

A more exhaustive analysis of the reads to phylogenetic rankings below that of phylum is supplied in Appendix 2. Of interest is that within the phylum Bacteroidetes, the Flavobacteria, Sphingobacteria and unclassified Bacteroidetes contributed to the majority of the reads for both the sediment and water samples for the Kariega Estuary. Within the Kowie Estuary, the Flavobacteria were well represented in the water column whilst the Bacteroidetes found within the sediment consisted of Sphingobacteria in addition to Flavobacteria.

While the dominant phyla are central to any given ecosystem, the occurrence of numerically less well represented bacterial taxa should not be discounted. Not only do these bacteria represent the potential to become dominant should the physico-chemical conditions in the ecosystem shift to favour their growth over that of the currently dominant bacteria, but these less abundant bacterial

representatives may still contribute significantly to the complex biotic cycling within their given ecological niche (Sogin et al., 2006; Huse et al., 2008; Jones and Lennon, 2010).

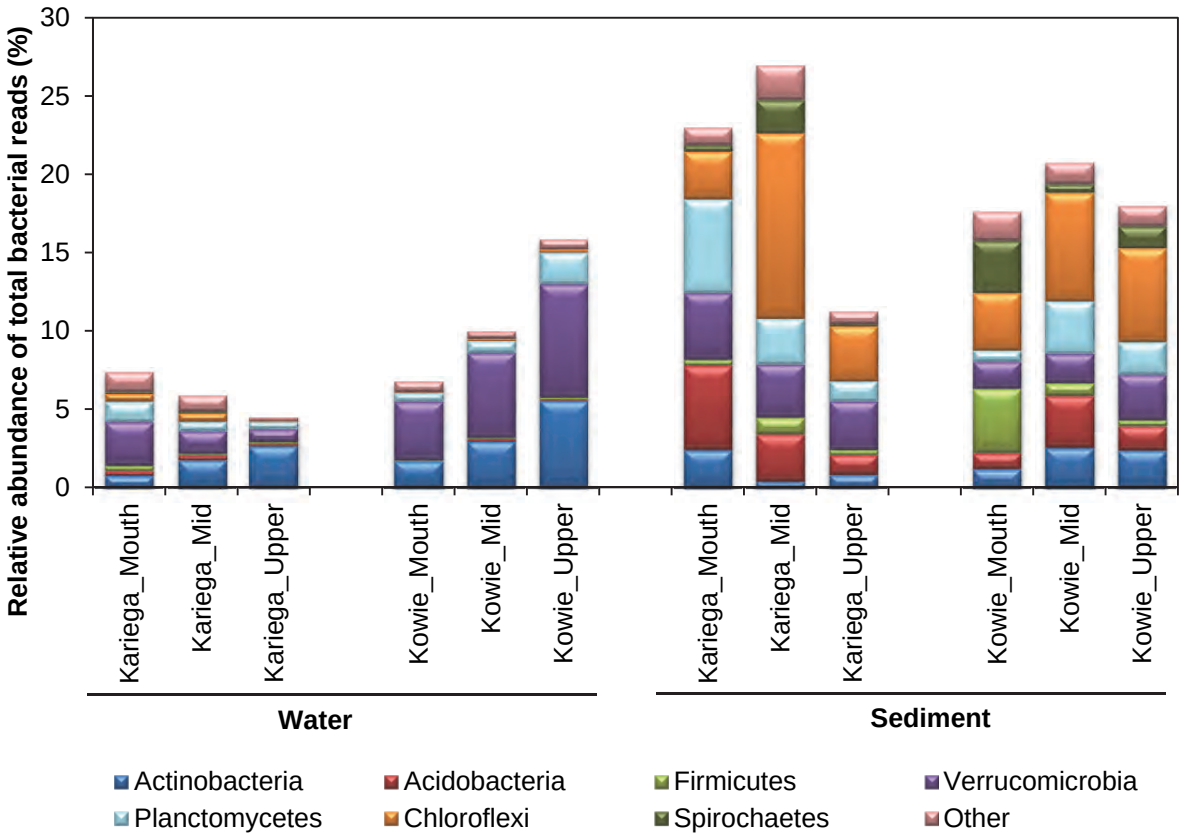
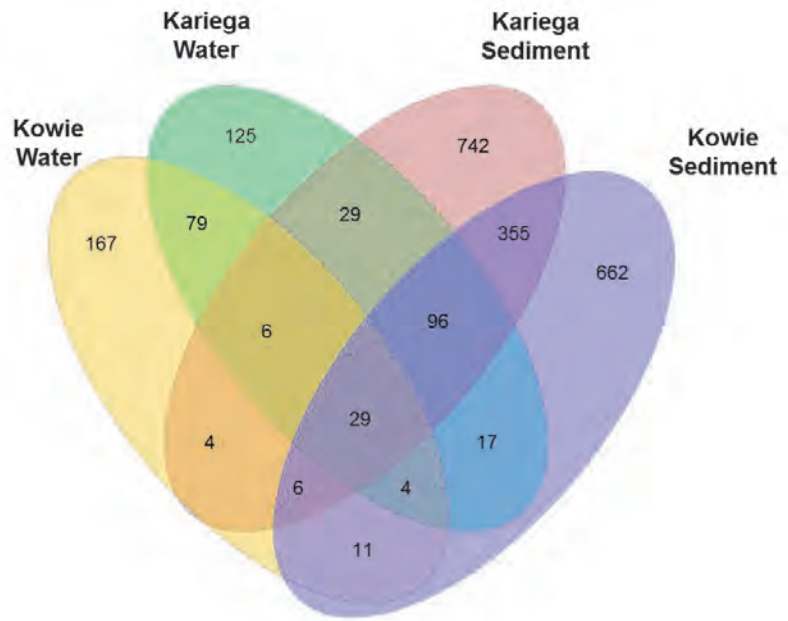


Figure 10. Relative abundances of the less well represented phyla within the dataset generated from water column and sediment samples along the length of the Kariega and Kowie estuaries. Assignment of reads to taxonomic rankings was done using the Bayesian algorithm against the Ribosomal Database Project.

An investigation of the less well represented phyla in the Kariega and Kowie estuaries (Figure 10), revealed that Actinobacteria form a significant proportion of the reads within the water column with increasing levels of abundance from the mouth to the upper reaches of both estuaries. With respect to the water column in the Kowie Estuary, Verrucomicrobia were also prevalent. An obvious increase in the diversity of phyla represented in the sediment samples compared to that of the water column was observed (Figure 10). This is indicative of the more complex and heterogeneous composition of the sediment, requiring a greater diversity of metabolic capabilities, compared to the well-mixed and more homogenous composition of the water column. With respect to the Kariega Estuary, the relative abundances of Planctomycetes and Acidobacteria decreased further upstream and a spike in bacteria belonging to the phylum Chloroflexi was observed in the middle reaches. The Kowie Estuary showed increased abundances of Firmicutes and Spirochaetes at the mouth and increased representation of Chloroflexi, Planctomycetes and Actinobacteria in the middle and upper reaches.

5.2. Operational Taxonomic Units (OTU) analysis

In light of the results of the phylogenetic classification (Figure 9, 10) it was not surprising that the number of OTUs found within the sediment samples of both estuaries exceeded that of the water column by more than 3-fold (Figure 11). As discussed above, this may reflect a more complex environment in the sediment. Alternatively, the dormant bacteria have a greater residence time in the sediment than in the water column where they are more often than not diluted by seawater during high tide and then discharged into the sea. The percentage of OTUs common to both the sediment and water column for the Kariega and Kowie estuaries was 9.7% and 3.4% respectively (Figure 11) which is similar to the results obtained from the Sundays (8.6%) and Swartkops (3.8%) estuaries (Figure 6).



	Kariega Water	Kariega Sediment	Kowie Water	Kowie Sediment
Kariega Water	382	160	118	143
Kariega Sediment	160	1267	45	486
Kowie Water	118	45	306	50
Kowie Sediment	143	486	50	1177

Figure 11. Venn diagrams illustrating the relative distribution of OTUs from the water column and sediment of the Kariega and Kowie estuaries. (Singletons have been removed from the dataset for this analysis)

Over half of the OTUs identified in the Kowie estuary were unique to either the sediment or water column, and more than half of the Kowie Estuary OTUs did not occur in the Kariega Estuary (54.6% and 56.2%, of the Kowie water and sediment OTUs respectively, are unique). This occurrence of unique OTUs was less evident in the Kariega water column where only 32.7% of the OTUs were unique, whilst the sediment samples in the Kariega Estuary followed a similar trend to that of the

Kowie Estuary with 58.6% of the OTUs being unique (Figure 11). When comparing sediment samples from both estuaries to one another, a 19.9% overlap in OTUs was observed between the Kowie and Kariega estuaries. With respect to the water column, a 17.2% of the OTUs were found in both of the estuaries.

With respect to the numerically dominant OTUs (Figure 12) there is a high degree of overlap between the Kowie and Kariega estuaries. The exceptions being OTUs 47, 49, 55, 59 and 60 which are poorly represented or absent in the Kariega Estuary and OTUs 17, 27 and 39 for the Kowie Estuary. Distinct spatial distribution of the representative reads for each of the OTUs between the mouth, middle reaches and upper reaches of both estuaries was evident in the sediment, particularly in the case of the Kowie Estuary where several OTUs were found almost exclusively at the mouth of the estuary (e.g. OTUs 43-49). In contrast, the OTUs in the water column of the Kowie Estuary were more evenly distributed between the mouth, middle reaches and upper reaches indicating that there is a high degree of mixing between the different sampling sites. While most of the OTUs in the Kariega Estuary water column were found throughout the sampling sites, there was some evidence that some OTUs were more abundant in the upper regions than in the mouth and visa versa (e.g. OTU31 is almost exclusively found at the mouth whilst OTUs 37 and 40 are almost exclusively found in the upper reaches) (Figure 12). This might indicate less mixing within the Kariega Estuary as compared to the Kowie Estuary.

Subsequent to BLAST analysis, OTU59 was the only dominant OTU associated with a freshwater organism, whilst the remainder of the OTUs were closely related to microorganisms from marine ecosystems (Table 5). This reflects the impact of the marine environment on these estuaries, in particular the Kowie Estuary. Evidence of this is provided by the fact that the most prominent OTU in the Kariega Estuary was undisputedly OTU23, classified within the SAR11 cluster of the Alphaproteobacteria, representing 25% to 38% of the total reads for each of the water column sampling sites (Figure 12A, Table 5). This OTU was also very highly abundant in the Kowie Estuary (Figure 12B) with 19% to 24% of the total reads assigned to this OTU. Further sequence analysis of this OTU indicated a 99% identity with *Pelagibacter ubique*, a bacterial species abundant in the ocean, but also found in freshwater ecosystems (Rappe et al., 2002). Other OTUs that were significantly dominant included OTU40, accounting for 17% to 21% of the total reads within the Kowie Estuary. OTU40 was identified as *Synechococcus* sp. (Table 5), a genus that represents >95% of the cyanobacterial reads found in the Kowie Estuary (Figure 9). Within the sediments of the Kariega and Kowie estuaries, the role of prokaryotes in the cycling of sulphur was emphasized with dominant OTUs including Desulfobulbaceae, Desulfobacteraceae, Desulfuromonadaceae, *Sulfurovum* sp. and *Sulfurimonas* sp. (OTUs 3, 4, 10, 19, 45, 46, 48) (Figure 12, Table 5, Supplementary Table 2).

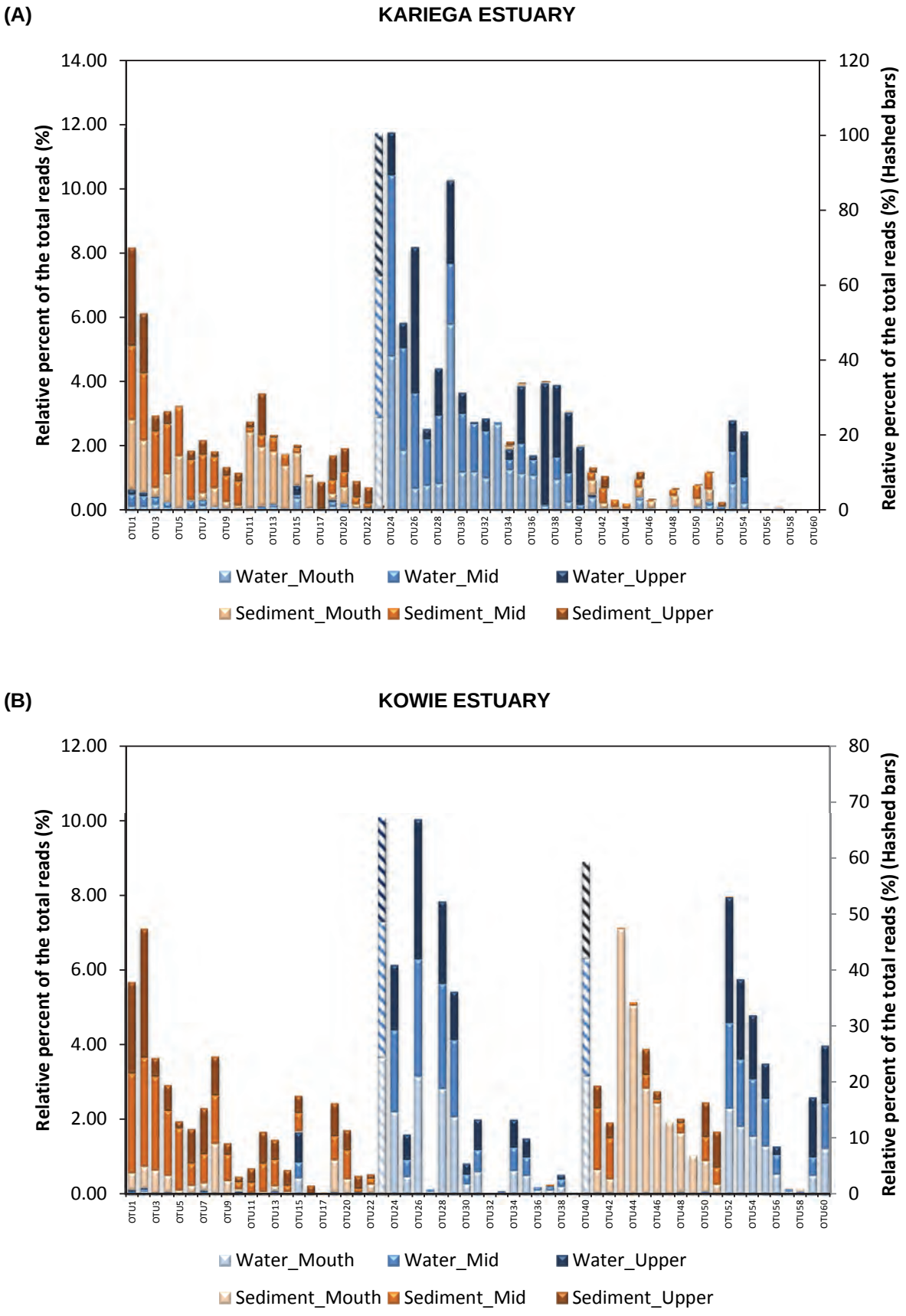


Figure 12. Comparative analysis of the ten most dominant OTUs found within the sediment and water column of the Kariega (A) and Kowie (B) estuaries. OTUs were determined at a distance value of 0.03. Mid=middle reaches

Table 5. Results of the BLAST analysis of the dominant OTUs from the Kariega and Kowie estuaries indicating the NCBI database sequence to which the query OTU sequence exhibited highest percentage identity.

OTU	Highest classification of database organism	Phylum	Accession number	Identity (%)	Source
1	Uncultured bacterium	/	JN839396.1	99	Shallow marine hydrothermal vent
2	Uncultured bacterium	/	JF272049.1	99	Marine biofilms
3	Uncultured bacterium	/	JN977359.1	100	Jiaozhou Bay sediments
4	Uncultured δ-proteobacteria	Proteobacteria	JQ580333.1	100	Oil-polluted subtidal sediments
5	Uncultured γ-proteobacteria	Proteobacteria (γ)	JQ579920.1	100	Oil-polluted subtidal sediments
6	Uncultured γ-proteobacteria	Proteobacteria (γ)	KF463986.1	100	Marine coastal sediment
7	Uncultured bacterium	/	HQ703828.1	99	Marine sediment
8	Uncultured Bacteroidetes	Bacteroidetes	JQ580152.1	100	Oil-polluted subtidal sediments
9	Uncultured Chloroflexi	Chloroflexi	FJ949414.1	99	Artificial oil-spill in coastal marine sediment
10	Uncultured δ-proteobacteria	Proteobacteria	KC009880.1	99	French Guiana coast
11	Uncultured γ-proteobacteria	Proteobacteria	JF344164.1	99	Oil-polluted subtidal sediments
12	Uncultured γ-proteobacteria	Proteobacteria	JN672630.1	100	East China sea sediment
13	Uncultured bacterium	/	HQ191089.1	100	Intertidal sediment
14	Uncultured bacterium	/	HQ191056.1	100	Coastal, intertidal sediment
15	Uncultured bacterium	/	FJ626909.1	99	Sublittoral Gulf of Mexico Sands
16	*Muriicola* sp.	Bacteroidetes	KC839612.1	98	Seawater from the Mariana Trench
17	Uncultured bacterium	/	JN534891.1	99	hypersaline evaporation pond
19	Uncultured δ-proteobacteria	Proteobacteria	JX241022.1	99	Coastal soil of Gulf of Khambhat
20	Uncultured bacterium	/	JX391641.1	99	Surface Marine Sediments
21	Uncultured bacterium	/	JX391810.1	99	Surface Marine Sediments
22	Uncultured bacterium	/	HE574860.1	100	Nidamental Gland of the Squid
23	Uncultured SAR11 cluster	Proteobacteria (α)	KF786857.1	100	Oil sheen weathering
24	Uncultured *Flavobacterium* sp.	Bacteroidetes	FJ745102.1	100	Coastal bacterioplankton
25	Uncultured proteobacterium	Proteobacteria	EU852566.1	100	Marine sponge
26	Uncultured α-proteobacteria	Proteobacteria	GU230260.1	100	South Atlantic Ocean
27	Uncultured SAR11 cluster	Proteobacteria (α)	KF786721.1	100	Oil sheen weathering
28	Uncultured bacterium	/	KF185805.1	100	North Adriatic Sea
29	Uncultured *Owenweeksia* sp.	Bacteroidetes	JX529868.1	100	Southern Ocean
30	*Formosa* sp.	Bacteroidetes	KF023501.1	100	Marine heterotrophic bacteria
31	Uncultured Flavobacteriaceae	Bacteroidetes	JQ515455	99	Caribbean reef
32	Uncultured γ-proteobacteria	Proteobacteria	KF786921.1	100	Oil sheen weathering
33	Uncultured *Roseobacter* sp.	Proteobacteria (α)	JX530514.1	100	Southern Ocean
34	Uncultured Verrucomicrobia	Verrucomicrobia	DQ778251.1	100	Mediterranean coastal waters
35	Uncultured *Marinovum* sp.	Proteobacteria (α)	FJ745019.1	100	Coastal bacterioplankton
36	Uncultured bacterium	/	KF185669.1	100	North Adriatic Sea
37	*Synechococcus* sp.	Cyanobacteria	CP006882.1	100	Sargasso Sea
38	Uncultured γ-proteobacteria	Proteobacteria	JQ515484.1	100	Caribbean reef-building coral
39	Uncultured Acidimicrobiales	Actinobacteria	KF786730.1	99	Oil sheen weathering
40	Uncultured *Synechococcus*	Cyanobacteria	FJ999601.1	100	Sponge-associated bacteria
41	Flavobacteriaceae bacterium	Bacteroidetes	AM990866.1	100	Coastal Mediterranean ecosystem
42	Uncultured Bacteroidetes	Bacteroidetes	JQ580134.1	100	Oil-polluted subtidal sediments
43	*Cytophaga fermentans*	Bacteroidetes	NR_044696.1	100	Not specified
44	Uncultured bacterium	/	JX391257.1	99	Surface Marine Sediments
45	Uncultured bacterium	/	KF440287.1	100	Mediterranean Sea mud volcano
46	Uncultured bacterium	/	KC471295.1	99	Marine sediments
47	Uncultured bacterium	/	KC527450.1	98	Bacteria associated with coral
48	Uncultured *Sulfurimonas* sp.	Proteobacteria (ε)	KF464251.1	100	Alang-Sosiya ship breaking yard
49	Uncultured bacterium	/	KC527459.1	98	Bacteria associated with coral
50	Uncultured bacterium	/	JX391488.1	99	Surface Marine Sediments
51	Uncultured bacterium	/	AY897351.1	100	Fish farm sediments
52	Uncultured bacterium	/	EU182121.1	100	Coastal waters of the South china sea
53	Uncultured Bacteroidetes	Bacteroidetes	HM057636.1	100	Yellow Sea water

OTU	Highest classification of database organism	Phylum	Accession number	Identity (%)	Source
54	Uncultured γ-proteobacteria	Proteobacteria	FJ666204.1	100	Ballast tank of a commercial ship
55	Uncultured α-proteobacteria	Proteobacteria	HQ692004.1	100	Lagoon in the North Pacific
56	Uncultured γ-proteobacteria	Proteobacteria	FJ745203.1	100	Coastal bacterioplankton
57	Uncultured bacterium	/	JQ269281.1	99	Estuarine bacteria (China)
58	*Thiomicrospira frisia*	Proteobacteria (γ)	NR_028679.1	99	Culture collection sulfur-oxidizing strain
59	Uncultured SAR11 cluster	Proteobacteria (α)	FN665729.1	100	Temperate lakes
60	Uncultured α-proteobacteria	Proteobacteria	GU230190.1	100	Coastal South Atlantic Ocean

5.3. Chloroplast diversity

As was seen in the Swartkops and Sundays estuaries, the sediment samples in the Kariega and Kowie estuaries was almost exclusively dominated by Bacillariophyta (Figure 13). With respect to the water column in the Kariega Estuary, the middle reaches and mouth showed similar diversity profiles with Bacillariophyta dominating and a sizeable representation by algae belonging to the Chlorophyta. Once in the upper reaches, the relative abundances of Chlorophyta and Cryptomonadaceae increase approximately four-fold whilst Bacillariophyta abundances decrease. A less dramatic shift in relative abundances of chloroplasts was observed in the water column of the Kowie Estuary with a decrease in Chlorophyta from 4.6% to 1.7% and an increase in Bacillariophyta from 3.3% to 4.7% of the total reads generated from the mouth and upper reaches respectively (Figure 13).

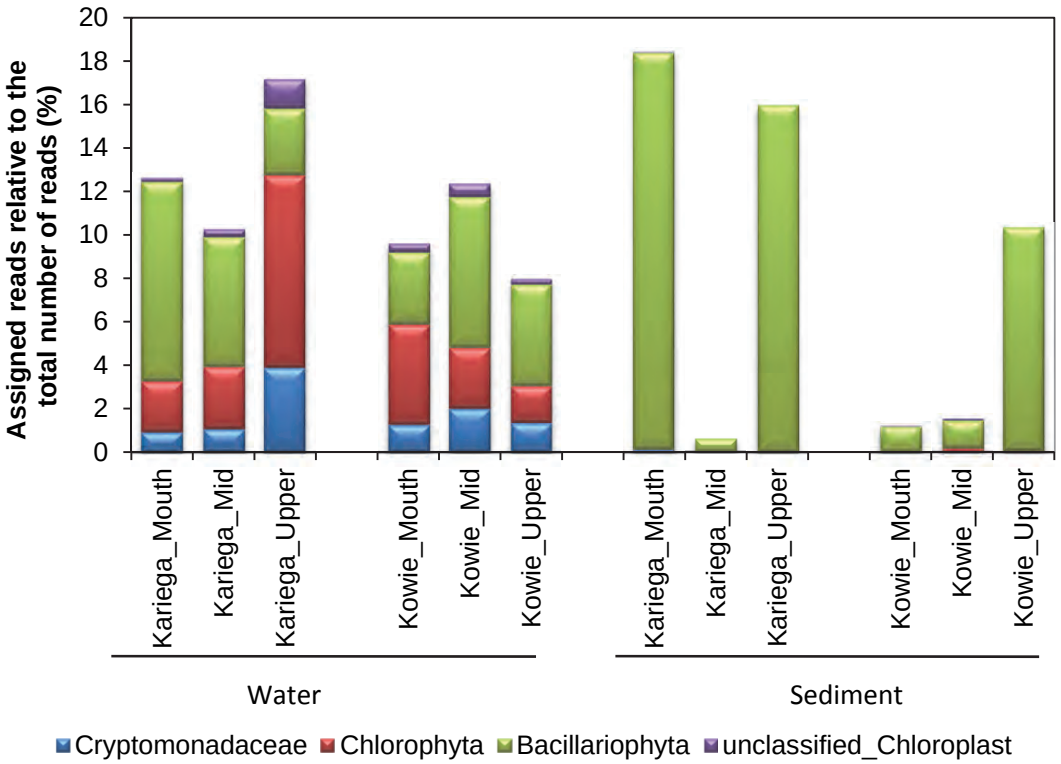


Figure 13. Relative abundance of 16S rDNA reads classified as chloroplast in origin from the water column and sediment of the Kowie and Kariega estuaries. Mid=middle reaches

5.4. Potential pathogens

Analysis of the 16S rRNA reads generated for the Kowie and Kariega estuarine systems were scrutinised for the presence of potential pathogenic strains. The waterborne pathogens targeted in this analysis included *Enterobacteriaceae, Campylobacter jejuni, Yersinia enterocolita, Vibrio cholera*, *Mycobacterium* sp., *Pseudomonas aeruginosa, Aeromonas* sp., *Legionella pneumophila, Salmonella* sp., and *Shigella* (LeClerc et al., 2002; Sharma et al., 2003). With regard to the Kowie Estuary, 57 reads from a total of 109 795 reads were classified into the genera representing potential water-borne pathogens. However, closer inspection using BLAST analysis revealed that these reads exhibited higher relatedness to non-pathogenic species within these genera than to the pathogenic species. In the Kariega Estuary, a total of 65 reads (which equates to 0.15% of the total number of reads) were assigned to the genus *Vibrio*. Species belonging to the genus *Vibrio* include many heterotrophic marine bacteria (Thompson et al., 2004) and BLAST analysis of the 65 reads for the Kariega estuary resulted in higher sequence identity with *Vibrio* species other than the pathogen *Vibrio cholera*. A further two reads from the sediment samples of the Kariega estuary were assigned to the genus *Legionella*, however, these reads also showed higher identity to non-pathogenic *Legionella* species. Only a single read (which equates to <<0.01% of the total reads) fell within the *Enterobacteriaceae* taxonomic group.

In addition to pathogenic bacteria, the presence of toxin-producing cyanobacteria is also of concern particularly in aquatic systems. Several cyanobacterial species have been shown to produce a variety of toxins harmful to humans and fish including hepatotoxins, neurotoxins and dermotoxins (Briand et al., 2003). BLAST analysis of the cyanobacterial reads in the Kariega Estuary indicated that these reads presented closer phylogenetic relationships with benign cyanobacteria rather than toxin-producing cyanobacteria. With respect to the Kowie Estuary, the majority (>95%) of the cyanobacterial reads were assigned to the genus *Synechococcus* which are widely distributed in marine ecosystems (Partensky et al., 1999). While some strains of *Synechococcus* have been shown to produce neurotoxins and hepatotoxins (Martins et al., 2005), compared to other genera of Cyanobacteria, incidences of toxin-producing *Synechococcus* strains is limited.

6. DISCUSSION AND CONCLUSIONS

The composition and functioning of the biotic population within any given ecosystem is dependent on the complex interfaces between several factors including the physical, chemical and nutritional elements within that ecosystem (Dolan, 2005). The aim of this study was to investigate how these factors influenced the functioning of estuarine ecosystems by focussing on the bacterial community structures in four distinct Eastern Cape estuaries. Two estuaries, namely the Sundays and Swartkops, are considered fresh-water dominated, where a significant inflow of freshwater resulted in lower salinities in the upper reaches. Both of these systems are considered to be heavily impacted by human activity resulting in high nutrient loads. However they are different in that the Sundays is agriculturally impacted while the Swartkops is an urbanize estuary. On the other hand, the Kowie and Kariega estuaries represent marine-dominated systems, where a reduced salinity gradient occurs between the estuary mouth and the upper reaches. These two estuarine systems also differ in that the Kariega is a fresh-water deprived system considered to be relatively pristine, while the Kowie is

impacted by human settlement along the reaches of the estuary and agricultural activities in the catchment area.

In general, it is widely accepted that nutrient composition and availability is likely a major influence on the type of organisms that may thrive in an ecosystem. Of the physical factors, salinity has been identified as one of the most important factors in determining prokaryote community composition in aquatic ecosystems (Crump et al., 1999; Bouvier and del Giorgio, 2002; Kirchman et al., 2005; Lupozone and Knight, 2007). Therefore one would have expected the distinct spatial distribution of bacterial communities in the Sundays and Swartkops estuaries, with their steep salinity gradients (7-8 PSU in the upper reaches). By contrast, the Kariega and Kowie have relatively small salinity gradients from the mouth to the upper reaches. It was therefore quite surprising to discover shifts in bacterial community profiles between the estuary mouth and the upper reaches for all four of the estuaries suggesting that while salinity may have played a role in determining the community structure, it is not the sole driver of prokaryotic population dynamics in these estuaries. This is supported by the observation that the Swartkops and Sundays estuaries support distinct communities in spite of similar salinity gradients (Figure 2). We conclude that the most important drivers of microbial community composition in these estuaries are nutrients and in particular, the concentration of nitrates and phosphates.

Autotrophy and heterotrophy are the most basic components of biotic factors contributing to ecosystem functioning and represent the entry of energy and nutrients into the ecosystem as well as cycling of those nutrients within the ecosystem. Autotrophy in estuarine ecosystems is typically carried out by algae as well as photosynthetic bacteria such as Cyanobacteria and various genera from within the Chlorobi, Firmicutes, Chloroflexi, Betaproteobacteria and Alphaproteobacteria (Garrity et al., 2001). With the exception of the Kowie Estuary, primary productivity in the estuaries examined in this study is predominately represented by unicellular algae with a reduced input from photosynthetic bacteria. With respect to the Kowie Estuary, the high incidence of Cyanobacteria observed can be almost entirely (>95% of the cyanobacterial reads) attributed to isolates belonging to a single genus, namely *Synechococcus*. Typically the occurrence of a single cyanobacterial genus in an aquatic ecosystem is indicative of a bloom event. The most common culprits of cyanobacterial bloom formation include the genera *Anabaena*, *Aphanizomenon*, *Nodularia, Cylindrospermopsis*, *Oscillatoria* and *Microcystis* (Paerl et al., 2001). Blooms formed by *Synechococcus*, whilst they do occur (Cook and Holland, 2012), are not as common. Gippsland Lakes (Australia) has been affected by *Nodularia* blooms sporadically since 1965 but in 2007-2008, a *Synechococcus* bloom was recorded in this lake system for the first time. The *Synechococcus* bloom observed in Gippsland Lakes was found to coincide with a dramatic increase in chlorophyll-a (Chl-a) concentrations (~5 μg/L to 25 μg/L) (Cook and Holland, 2012). In contrast, the Chl-a concentrations observed in the Kowie Estuary were <1.08 μg/L indicating that while *Synechococcus* is dominant, it is not a full-scale bloom. Although the Kowie Estuary is not yet eutrophic as a result of a bloom, the occurrence of mono-specific cyanobacterial representation is of concern and the impact of anthropogenic influx of nutrients into this ecosystem should be closely monitored.

The dominant phyla in the four estuaries discussed in this study are the Proteobacteria and Bacteroidetes. The Bacteroidetes are considered to be chemoorganotrophic bacteria whose role in nutrient cycling involves the uptake and degradation of complex organic substrates. Other abundant heterotrophic bacterial taxa found in all four estuaries in this study include Alteromonadales (Gammaproteoabacteria), Pseudomonadales (Gammaproteobacteria), many of the genera in Xanthomonadaceae (Gammaproteobacteria), *Cycloclasticus* (Gammaproteobacteria), Burkholderiales (Betaproteobacteria), Actinobacteria, Verrucomicrobia and Planctomycetes (Bowman and McMeekin, 2001; Nishihara, 2001; Reichenbach, 1992; Pérez-Pantoja et al., 2012; Cottrell and Kirchmann 2000; Bull et al., 2005; Gade et al., 2004; DeLong et al., 1993; Cottrell and Kirchman, 2000; Kirchman, 2002; Eilers et al., 2001; Webster et al., 2001; Abell and Bowman, 2005; Grossart et al., 2005; Thomas et al., 2011). Based on these taxa, the relative percentage of heterotrophic bacterial populations in the water column was generally high: Kariega (21-39%), Kowie (21-26%), Sundays (69-81%) and Swartkops (43-59%). The dominance of heterotrophic bacteria is likely to be even higher since there are likely to be additional heterotrophic bacteria in these estuaries. This is because a large proportion of the reads obtained in this study remain unclassified, particularly within the Gammaproteobacteria subphylum where between 50% and 91% of the reads assigned to this class could not be classified further. Another factor is that the phenotypes of many of the bacterial taxons for which 16S rRNA sequence is available, have yet to be determined. Despite this, the large proportion of heterotrophic bacteria in these estuaries clearly highlights the importance of the bacterial-driven detrital foodweb in these estuarine ecosystems. Overall, all four of the estuaries appear to be net-heterotrophic which is expected in estuarine ecosystems with a high particulate and dissolved organic load (Selje and Simon, 2003).

The substantial contribution of Deltaproteobacteria to the bacterial community in the sediments in all four estuaries is interesting. The Deltaproteobacteria that were identified were dominated by representatives from the orders Desulforomonadales and Desulfobacterales which contain bacteria known to reduce sulphate (Kuever et al., 2005; Sylvan et al., 2012). These bacteria thus contribute to the cycling of sulphur within ecosystems and by extension, play an important role in the mineralization of organic matter within benthic sediments (Miletto et al., 2007).

While significant volumes of research has been carried out on the microbial assemblages found in the oceans and seas, relatively few studies on the microbial populations of estuaries have been done (Zinger et al., 2012). Of the microbial community profiles within estuaries that are reported in the literature (Kirchman et al., 2005; Feng et al., 2009; Campbell and Kirchman, 2013) it is interesting to note that the described abundances of Bacteroidetes is fairly low (4.6% to ~16%) whilst in the estuaries examined in this study, Bacteroidetes formed a major component of the bacterial assemblage within the water column where, with the exception of the Kowie Estuary, the relative abundance of Bacteroidetes ranged from 16% to 59% of the total bacterial community in the water column. Other studies have demonstrated that with respect to the occurrence of Proteobacteria along salinity gradients, Alphaproteobacteria are generally associated with high salinities whilst Betaproteobacteria are reported to occur in higher abundances in freshwater (Bouvier and del Giorgio, 2002; Feng et al., 2009; Herlemann et al, 2011; Campbell and Kirchman, 2013). This is supported by reports of the prevalence of Alphaproteobacteria and Gammaproteobacteria in marine ecosystems (Pommier et al., 2007; Biers et al., 2009) and the incidence of Betaproteobacteria in

freshwater habitats (Glöckner et al., 1999; Zwart et al., 2002; Mueller-Spitz et al., 2009). In the Kariega and Kowie estuaries, the relative abundance of Betaproteobacteria and Alphaproteobacteria along the length of the estuaries did not change very much which is most likely due to the marine-dominated nature of these two estuaries. In contrast, the relative abundances of these two classes in the Swartkops Estuary, between the mouth and the upper reaches, increased or decreased by more than five-fold for Betaproteobacteria and Alphaproteobacteria respectively. The relative abundance of Alphaproteobacteria was very low in the Sundays Estuary but the Betaproteobacteria followed a similar trend to that found in the Swartkops Estuary.

Benthic bacteria are central players in bio-geochemical cycling within aquatic ecosystems in addition to which, they present an important potential for bioremediation of anthropogenic pollution. This is important as chemical pollutants readily bind to fine particles and, as a result, the sediment in estuaries represents a sink of these pollutants. This is of particular importance in the Kowie and Swartkops estuaries, which are heavily impacted by urban and industrial activities respectively. Acute influx of pollutants is known to negatively impact on the species diversity in an ecosystem. If the source of the pollutant is constant, however, bacterial communities can exhibit their resilience and proliferation of tolerant species can restore the species diversity levels albeit with an altered diversity profile (Gillan et al., 2005). The number of OTUs described in this study for both impacted estuaries was high and compared well with the results reported for the Port Jackson, Botany Bay and Port Kembla estuaries (Sun et al., 2013). These authors reported microbial diversity profiles with a dominance of Gammaproteobacteria (23.7% of the reads) and Deltaproteobacteria (21.3%). The contribution of Alphaproteobacteria, Bacteroidetes and Chloroflexi to the community composition was reported as 8.5%, 6.9%and 6.5% respectively (Sun et al., 2013). When compared to the community profiles observed in the Swartkops and Kowie estuaries, the relative abundances of Gammaproteobacteria are slightly low (13.1%-22.1% in the Swartkops Estuary and 8.2%-24.4% in the Kowie Estuary) whilst an even lower abundance of Deltaproteobacteria was observed (8.7%-9.9% in the Swartkops Estuary and 10.1%-18.8% for the Kowie Estuary). A major difference between the Eastern Cape estuaries and those from New South Wales lies with the Bacteroidetes which are far more prevalent in the Eastern Cape estuaries (13.8% to 61.8%). This may be indicative of a greater availability of complex organic matter in the Eastern Cape estuaries compared to those of New South Wales.

In conclusion, each of the four estuaries examined in this study showed some broad overlaps in terms of their population profiles at a high taxonomic ranking. However, there was a significant proportion of the bacterial species, which were unique to each of the estuaries. This is not surprising as two of the estuaries are influenced by high freshwater inflow, one by agricultural input, a second by industrial pollutants and two others by urbanisation. Any of these factors could result in a bacterial population profile uniquely adapted to the physical and nutritional pressures in each ecosystem. Together these four study sites offer important insight into the complex networks of microbial communities in estuarine systems. This study has laid the groundwork for future research into the role of the microbiota in the functioning of estuarine systems and importantly, how resident communities respond to changes in the physico-chemical conditions in the water column brought about by increased urbanization and agricultural activities in the catchment areas.

7. RECOMMENDATIONS FOR FUTURE RESEARCH

7.1. The use of next generation sequencing (NGS) technologies is a sensitive and affordable tool for continuous monitoring of aquatic ecosystem functioning and health

The aim of this research was to determine whether the diversity of microbial communities could be used to assess the functioning and health of estuarine ecosystems. The approach was to use NGS technologies (particularly 454 pyrosequencing) to characterise microbial diversity by analysing rRNA gene sequences in water column and sediment samples. The results of the study clearly demonstrate the correlation between microbial populations and key, dominant taxa and the physico-chemical conditions along the length of the estuary. Also, we have been able to show that different types of estuarine systems harbour distinct microbial assemblages that have adapted to physico-chemical conditions and anthropogenic pollution.

During the course of this research project, the cost of 454 pyrosequencing a set of six estuary samples (mouth, middle and upper reaches, water column and sediment) was approximately R 2 000 per sample (R0.50 per read). New NGS technologies, with increased efficiency and more affordable reagents have now been developed where the cost per sample could decrease as much as ten-fold. This provides the opportunity to apply the approach taken in this study widely to include other important estuarine systems around the South African coastline. In addition, the reduced cost could make it feasible to use of this technology for routine monitoring of sensitive estuarine systems.

7.2. The need to assess microbial metabolic activity in addition to numerical abundance

It is known that numerical abundance of a particular bacterial species does not necessarily directly correlate with metabolic activity or potential growth rate of that species. Microbial species that occur in lower numbers may be more metabolically active than their more abundant counterparts. A more accurate assessment of whether a bacterial taxon is actively metabolising can be achieved by quantifying the rRNA as opposed to the rDNA (i.e. the active component of the ribosome versus the gene encoding for the rRNA). This was demonstrated in a recent study in the Delaware Bay, where the numerical abundance of microbial taxa did not change, but there were changes in the metabolically active taxa in response to physical, biological or chemical gradients along the (Campbell and Kirchman, 2013). The need for information on which taxa are metabolically active in resident microbial communities is particularly important as a tool for monitoring episodic anthropogenic pollution in urbanized estuaries such as the Swartkops systems.

7.3. Water column versus sediment sampling

Traditionally, aquatic systems are monitored by the analysis of water samples, which are not very sensitive measures of changes in type (e.g. inorganic vs organic) nutrient concentration or anthropogenic pollution within the system. We observed changes in the abundance of dominant microbial species in sediment samples that were not observed in the water column along the length of estuaries with apparently small salinity and nutrient gradients. Also, it appears that microbes in the sediment or associated with particulate matter in the water column are metabolically far more active than free-living bacterial cells in the water column (Campbell and Kirchman, 2013). Together, these

findings lend strong support for the need to focus on sediment sampling when monitoring estuarine health and aquatic ecosystems in general.

8. REFERENCES

ABELL GCJ and BOWMAN JP (2005) Colonization and community dynamics of class *Flavobacteria* on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *FEMS Microbiol. Ecol.* **53** 379-391

ALTSCHUL SF, GISH W, MILLER W, MYERS EW and LIPMAN DJ (1990) Basic local alignment search tool. *J. Mol. Biol.* **215:**403-410

ATLAS RM, HOROWITZ A, KRICHEVSKY M, and BEJ AK (1991) Response of microbial populations to environmental disturbance. Microbial Ecology **22**(3) 249-256

BAIRD D, HANEKOM NM, GRINDLEY JR (1986) Estuaries of the Cape. Part II. Synopses of available information on individual systems (pp. 82). Report No. 23. Swartkops (CSE3). CSIR Research Report 422

BEMEN JM, FRANCIS CA (2006) Diversity of Ammonia-Oxidizing Archaea and Bacteria in the Sediments of a Hypernutrified Subtropical Estuary: Bahía del Tóbari, Mexico*. Appl. Environ. Microbiol.* **72**:7767-7777

BIERS EJ, SUN S and HOWARD EC (2009) Prokaryotic Genomes and Diversity in Surface Ocean Waters: Interrogating the Global Ocean Sampling Metagenome. *Appl. Environ. Microbiol.* 75:2221-2229

BINNING K and BAIRD D (2001). Survey of heavy metals in the sediments of the Swartkops River Estuary, Port Elizabeth, South Africa. *Water SA* **27**:61-466

BOUVIER TC and DEL GIORGIO PA (2002) Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnol. Oceanogr.* 47(2) 453-470

BOWMAN JP and McMEEKIN TA (2001) in Bergey's manual of systematic bacteriology. Volume 2B. 2nd edition. Springer-Verlag, New York. p443

BRIAND J, JACQUET S, BERNARD C and HUMBERT J (2003) Health hazards for terrestrial vertebrates from toxic cyanobacteria in surface water ecosystems. *Vet. Res.* **34**:361-377.

BUCKLING A, KASSEN R, BELL G, and RAINEY PB (2000) Disturbance and diversity in experimental microcosms. Nature, 408 961-964.

BULL AT, STACH JEM, WARD AC, and GOODFELLOW M (2005) Marine actinobacteria: perspectives, challenges, future directions. *Antonie van Leeuwenhoek* **87**:65-79

CAMACHO C, COULOURIS G, AVAGYAN V, MA N, PAPADOPOULOS J, BEALER K, and MADDEN TL (2008) BLAST+: architecture and applications. *BMC Bioinformatics* **10:**421

CAMPBELL BJ, YU L, HEIDELBERG JF, KIRCHMAN DL (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci* **108**:12776-12781

CAMPBELL BJ and KIRCHMAN DL (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J*. **7**:210-220

CAVALIER-SMITH T (2002) Chloroplast evolution: Secondary symbiogenesis and multiple losses. *Curr. Biol.* **12** R62-R64

CLAESSON MJ, O'SULLIVAN O, WANG Q, NIKKILA J, MARCHESI JR, SMIDT H, DE VOS WM, ROSS RP and O'TOOLE PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* **4** (8) e6669. doi: 10.1371/journal.pone.0006669.

COLE JR, WANG Q, CARDENAS E, FISH J, CHAI B, FARRIS RJ, KULAM-SYED-MOHIDEEN AS, MCGARRELL DM, MARSH T, GARRITY GM and TIEDJE JM (2009) The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *NAR*. **37** (Database Issue): D141-D145.

COOK PLM and HOLLAND DP (2012) Long term nutrient loads and chlorophyll dynamics in a large temperate Australian lagoon system affected by recurring blooms of cyanobacteria. *Biogeochem.* **107** 261-274.

COOPER JAG, RAMM AEL, HARRISON TD (1994) The Estuarine Health Index: A new approach to scientific information transfer. *Ocean Coastal Management* **25**:103-141

COTTRELL MT and KIRCHMAN DL (2000) Natural assemblages of marine Proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* **66:**1692-1697

COWLEY PD and DANIEL C. (2001). Estuaries of the Ndlambe Municipality. Institute of Natural Resources Investigational Report, No. 229. University of KwaZulu-Natal, Pietermaritzburg, South Africa

CRUMP BC, ARMBRUST EV and BAROSS JA (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl. Environ. Microbiol.* **65** 3192-3204

CRUMP BC, KLING GW, BAHR M, and 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

CRUMP BC, HOPKINSON CS, SOGIN ML, HOBBIE JE (2004) Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Appl Environ Microbiol* **70**:1494-1505

DEL GIORGIO PA and COLE JC (1998) Bacterial growth efficiency in natural aquatic ecosystems. *Annu Rev Ecol Sys* **29** 503-541

DELONG EF, FRANKS DG, and ALLDREDGE AL (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* **38**:924-934.

DOLAN JR (2005) An introduction to the biogeography of aquatic microbes. *Aquat. Microb. Ecol.* **41** 39-48.

EDGAR RC, HAAS BJ, CLEMENTE JC, QUINCE C, KNIGHT R (2011) UCHIME improves sensitivity and speed of chimera detection, *Bioinformatics* **27**: 2194-2200

EILERS H, PERNTHALER J, PEPLIES J, GLÖCKNER FO, GERDTS G, and AMANN R (2001) Isolation of novel pelagic bacteria from the German Bight and their seasonal contribution to surface picoplankton. *Appl. Environ. Microbiol.* **67**:5134-5142

EMMERSON WD (1989) The nutrient status of the Sundays River estuary. *Water Research* **23** 1059-1067

EVANS J, SHENEMAN L, FOSTER JA (2006) Relaxed Neighbor-Joining: A fast distance-based phylogenetic tree construction method. *J Mol Evol* **62**:785-792

FENG B, LI X, WNAG J, HU Z, MENG H, XIANG L and QUAN Z (2009) Bacterial diversity of water and sediment in the Changjiang estuary and coastal area of the East China Sea. *FEMS Microbiol. Ecol.* **70**:236-248

FLINDT MR, PARDAL MA, LILLEBØ AI, MARTINS I, MARQUES JC (1999) Nutrient cycling and plant dynamics in estuaries: A brief review. *Acta Oecologica* **20**:237-248

FORTUNATO CS, HERFORT L, ZUBER P, BAPTISTA AM, CRUMP BC (2012) Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. *ISME J* **6**:554-563

FREITAG TE, CHANG L, PROSSER JI (2006) Changes in the community structure and activity of betaproteobacterial ammonia-oxidizing sediment bacteria along a freshwater-marine gradient. *Environ Microbiol* **8**: 684-696

FRONEMAN PW (2000) Feeding studies of selected zooplankton in a temperate estuary, South Africa. *Estuarine, Coastal and Shelf Science* **51** 543-552

GADE D, SCHLESNER H, GLÖCKNER FO, AMANN R, PFEIFFER S, and THOMM M (2004) Identification of *Planctomycetes* with order-, genus-, and strain-specific 16S rRNA-targeted probes. *Microb. Ecol*. **47**:243-251

GARRITY GM, BELL JA, LILBURN T (Eds) (2001) in Bergey's Manual of systematic bacteriology. 2nd edition. Spinger-Verlang New York. pg 1, 63-116, 443

GILLAN DC, DANIS B, PERNET P, JOLY G and DUBOIS P (2005) Structure of sediment-associated microbial communities along a heavy-metal contamination gradient in the marine environment. *Appl Environ Microbiol* **71:** 679-690

GLÖCKNER FO, FUCHS BM and AMANN R (1999) Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Appl. Environ. Microbiol*. 65:3721-3726

GRANGE N and ALLANSON BR (1995) The influence of freshwater inflow on the nature, amount and distribution of seston in estuaries of the Eastern Cape, South Africa. *Estuarine, Coastal Shelf Sci.* 40 403-420

GROSSART HP, LEVOLD F, ALLGAIER M, SIMON M and BRINKHOFF T (2005) Marine diatom species harbour distinct bacterial communities. *Environ. Microbiol.* **7** 860-873

HAMADY M, LOZUPONE C, KNIGHT R (2010) Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* **4**:17-27

HARRIS EH, BOYNTON JE and GILLHAM NW (1994) Chloroplast ribosomes and protein synthesis. *Microb. Rev.* **58**(4) 700-754

HERLEMANN DPR, LABRENZ M, JÜRGENS K, BERTILSSON S, WANIEK JJ, ANDERSSON AF (2011) Transitions in bacterial communities along the 2000km salinity gradient of the Baltic Sea. *ISME J* **5**:1571-1579

HODGSON AN (1987) Distribution and abundance of macrobenthic fauna of the Kariega Estuary. *S. Afr. J. Zool.* **22** 153-162

HOLM-HANSEN O and RIEMAN B (1978) Chlorophyll a determination: improvements in methodology. *Oikos* **30**:438-447

HUGENHOLTZ P, GOEBEL BM and PACE NR (1998) Impact of culture independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol*. **180** 4765-4774

HUGHES JB and HELLMANN JJ (2005) The application of rarefaction techniques to molecular inventories of microbial diversity. In: *Methods in Enzymology* (Leadbetter JE, eds.) Elsevier, **397** 292-308

HUSE SM, HUBER JA, MORRISON HG, SOGIN ML and WELCH DM (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* **8**(R143) doi: 10.1186/gb-2007-8-7-r143

HUSE SM, DETHLEFSEN L, HUBER JA, WELCH DM, RELMAN DA and SOGIN ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* **4** (11) e1000255

HUSON DH, AUCH AF, QI J and SCHUSTER SC (2007) MEGAN Analysis of Metagenomics Data. *Genome Research*. **17:**377-386.

JONES SE and LENNON JT (2010) Dormancy contributes to the maintenance of microbial diversity. *PNAS* **107**(13) 5881-5886.

KIM Y, FREEMAN JP, MOODY JD, ENGESSER K, CERNIGLIA CE (2005) Effects of pH on the degradation of phenanthrene and pyrene by *Mycobacterium vanbaalenii* PYR-1. *Appl Microbiol Biotechnol* **67**:275-285.

KIRCHMAN DL (2002) The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* **39** 91-100

KIRCHMAN DL, DITTEL AI, MALMSTROM RR and COTTRELL MT (2005) Biogeography of major bacterial groups in the Delaware Estuary. *Limnol. Oceanogr.* **50**(5) 1697-1706

KIRCHMAN DL (2008) New light on an important microbe in the ocean. Proceedings of the National Academy of Science 105(25) 8487-8488

KONDO R, NEDWELL DB, PURDY KJ, SILVA SQ (2004) Detection and enumeration of sulphate-reducing bacteria in estuarine sediments by competitive PCR. *Geomicrobiol J* **21** 145-157

KUEVER J, RAINEY FA and WIDDEL F (2005) Order III. Desulfobacterales ord.nov. In Bergey's Manual of Systematic Bacteriology. 2nd Ed. Eds BRENNER DJ, KRIEG NR and STALEY JT, Springer, USA **2**C 959-991

KUNIN V, ENGELBREKTSON A, OCHMAN H and HUGENHOLTZ P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol*. **12**(1) 118-123

LANGENHEDER S, KISAND V, WIKNER J, TRANVIK LJ (2003) Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. *FEMS Microbiol Ecol* **45**:189-202

LECLERC H, SCHWARTZBROD L and DEI-CAS E (2002) Microbial agents associated with waterborne diseases. *Crit. Rev. Microbiol.* **28**(4) 371-409

LIU Z, LOZUPONE C, HAMADY M, BUSHMAN FD and KNIGHT R (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. *NAR* **35** (18) e120.

LOZUPONE CA and KNIGHT R (2007) Global patterns in bacterial diversity. *PNAS* **104**(27) 11436-11440

MARTINS R, PEREIRA P, WELKER M, FASTNER J and VASCONCELOS MV (2005) Toxicity of culturable cyanobacteria strains isolated from the Portuguese coast. *Toxicon* **46** 454-464.

MILETTO M, BODELIER PLE, LAANBROEK HJ (2007) Improved PCR-DGGE for high resolution diversity screening of complex sulfate-reducing prokaryotic communities in soils and sediments. *J. Microbiol. Methods* **70**:103-111.

MUELLER-SPITZ SR, GOETZ GW and MCLELLAN SL (2009) Seasonal and spatial variability in nearshore bacterioplankton communities of Lake Michigan. *FEMS Microbiol Ecol* **67**:511-522

NISHIHARA H (2001) in Bergey's manual of systematic bacteriology. Volume 2B. 2nd edition. Springer-Verlag, New York. pg 188

NIXON SW, OVIATT CA, FRITHSEN J, Sullivan b (1986) Nutrients and the productivity of estuarine and costal marine ecosystems. *J. Limn Soc SA* **12**:43-71

NOBLE R and HEMENS J (1978) Inland water ecosystems in South Africa – A Review of Research Needs. South African National Science Programmes. Report No. **34**.

ODUME ON, MULLER WJ, PALMER CG and ARIMORO FO (2012) Mentum deformities in *Chironomidae* communities as indicators of anthropogenic impacts in Swartkops River. *Physics and Chemistry of the Earth* **50-52** 140-148

ORR KK, BURGESS JE, FRONEMAN PW (2008). The effects of mouth-phase and rainfall on the concentration of heavy metals in the sediment and water of select Eastern Cape estuaries, South Africa. *Water SA* **34** 39-52

PAERL HW, FULTON RS, MOISANDER PH and DYBLE J (2001) Harmful freshwater algal blooms, with an emphasis on Cyanobacteria. *Scientific World* **1** 76-113

PARTENSKY F, BLANCHOT J and VAULOT D (1999) Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. *Bull. Inst. Océanogr. Monaco Special no.**19*** 457-475.

PÉREZ-PANTOJA D, DONOSO R, AGULLÓ L, CÓRDOVA M, SEEGER M, PIEPER DH, GONZÁLEZ B (2012) "Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales" Environmental Microbiology, **14** 1091-1117

POMMIER T, CANBÄCK B, RIEMANN L, BOSTRÖM KH, SIMU K, LUNDBERG P, TUNLID A and HAGSTRÖM A (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology* **16**:867-880

QUAST C, PRUESSE E, YILMAZ P, GERKEN J, SCHWEER T, YARZA P, PEPLIES J, GLÖCKNER FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res.* **41**:D590-596

REICHENBACH H (1992) The Order Cytophagales. *The Prokaryotes Vol. 4* (2nd edn.) (Balows A, Trüper HG, Dworkin M, Harder W and Schleifer KH eds) pp. 3631-3687. Springer, Berlin

SAHAN E, MUYZER G (2008) Diversity and spatio-temporal distribution of ammonia-oxidizing Archaea and Bacteria in sediments of the Westerschelde estuary. *FEMS Microbiol Ecol* **64**:175-186

SAIT M, HUGENHOLTZ P, and JANSSEN PH (2002) Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environ. Microbiol*. **4** 654-666

SCHARLER U and BAIRD D (2003) The influence of catchment management on salinity, nutrient stoichiometry and phytoplankton biomass of Eastern Cape estuaries, South Africa. *Estuarine Coastal Shelf Sci* **56**:735-748

SCHLOSS PD and HANDELSMAN J (2005) Metagenomics for studying unculturable microorganisms: Cutting the Gordian knot. *Genome Biol.* **6** (8) 229

SCHLOSS PD, WESTCOTT SL, RYABIN T, HALL JR, HARTMANN M, HOLLISTER EB, LESNIEWSKI RA, OAKLEY B, PARKS D, ROBINSON C, SAHL J, STRES B, THALLINGER G, VAN HORN D and WEBER C (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. **75**(23):7537-41

SELJE N and SIMON M (2003) Composition and dynamics of particle-associated and free-living bacterial communities in the Weser estuary, Germany. *Aquat Microb Ecol* **30**: 221-237

SHARMA S, SACHDEVA P and VIRDI JS (2003) Emerging water-borne pathogens. *Appl. Microbiol. Biotechonol.* **61** 424-428

SOGIN ML, MORRISON HG, HUBER JA, WELCH DM, HUSE SM, NEAL PR, ARRIETA JM and HERNDL GJ (2006) Microbial diversity in the deep sea and the unexplored 'rare biosphere'. *Proc. Natl. Acad. Sci. U.S.A.* **103** 12115-12120

STACKEBRANDT E and GOEBEL BM (1994) A place for DNA-DNA reassociation and 16S rRNA sequence-analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol*. **44** 846-849

SUN MY, DAFFORN KA, JOHNSTON EL, BROWN MV (2013) Core sediment bacteria rive community response to anthropogenic contamination over multiple environmental gradients. *Environ Microbiol* **15**:2517-2531

SYLVAN JB, TONER BM, EDWARDS KJ (2012) Life and death of deep-sea vents:Bacterial diversity and ecosystem succession on inactive hydrothermal sulfides. *mBio* **3** doi:10.1128/mBio.00279-11

TELESH IV, KHLEBOVICH VV (2010) Principal processes within the estuarine salinity gradient: a review. *Mar Pollut Bull* **61**:149-155

THOMAS F, HEHEMANN J, REBUFFET E, CZJZEK M and MICHEL G (2011) Environmental and gut *Bacteroidetes*: the food connection. *Frontiers in Microbiol.* **2** 1-15

THOMPSON JR, RANDA MA, MARCELINO LA, TOMITA-MITCHELL A, LIM E and POLZ MF (2004) Diversity and dynamics of North Atlantic coastal Vibrio community. *Appl. Environ. Microbiol.* **70**(7) 4103-4110.

TRINGE SG and HUGENHOLTZ P (2008) A renaissance for the pioneering 16S rRNA gene. *Curr. Opin. Microbiol.* 11 442-446.

TURPIE JK, , ADAMS JB, JOUBERT, A, HARRISON TD, COLLOTY BM, MAREE RC, WHITFIELD AK, WOOLDRIDGE TH, LAMBERTH SJ, TALJAARD S and VAN NIEKERK L (2002) Assessment of the conservation priority status of South African estuaries for use in management and water allocation. *WaterSA* **28**(2) 191-206

VORWERK PD and FRONEMAN PW (2009) The importance of estuarine derived carbon for the near shore marine environment: studies on contrasting southern African estuaries. *Afr. J. of Aqua. Sci.* **34**: 67-77

WANG Q, GARRITY GM, TIEDJE JM, and COLE JR (2007) Naïve Bayesian Classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**(16) 5261-7

WANG Y and QIAN P (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* **4** (10) e7401

WEBSTER NS, WILSON KJ, BLACKALL LL and HILL RT (2001) Phylogenetic diversity of bacteria associated with the marine sponge *Rhopaloeides odorabile. Appl. Environ. Microbiol.* **67** 434-444

WU QL, ZWART G, SCHAUER M, KAMST-VAN AGTERVELD MP, HAHN MW (2006) Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan plateau, China. *Appl Envrion Microbiol* **72** 5478-5485

XIE ZY, HU CQ, CHEN C, ZHANG LP, REN CH (2005) Investigation of seven Vibrio virulence genes among *Vibrio alginolyticus* and *Vibrio parahaemolyticus* strains from the coastal mariculture systems in Guangdong, China. *Letters in Applied Microbiology* **41:**202-207.

ZINGER L, GOBET A, POMMIER T (2012) Two decades of describing the unseen majority of aquatic microbial diversity. *Mol Ecol* **21**:1878-1896

ZWART G, CRUMP BC, AGTERVELD M, HAGEN F and HAN SK (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* **28**:141-155

Appendix 1

Supplementary Table 1. Taxonomic classification of the dominant OTUs from the Sundays and Swartkops estuaries using a Naïve Bayesian classification algorithm against the Ribosomal Database Project trainset9.

OTU#	Classification to lowest taxonomic rank (percentage confidence levels in brackets)
OTU1	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100)
OTU2	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100)
OTU3	Verrucomicrobia(100);Opitutae(100);Puniceicoccales(100);Puniceicoccaceae(100);*Coraliomargarita*(100)
OTU4	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Muriicola*(98)
OTU5	Bacteroidetes(100);unclassified
OTU6	Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU7	Bacteroidetes(100);unclassified
OTU8	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU9	Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU10	Proteobacteria(100);Betaproteobacteria(100);Rhodocyclales(99);Rhodocyclaceae(99);unclassified
OTU11	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Lutimonas*(100)
OTU12	Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU13	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU14	Proteobacteria(100);Gammaproteobacteria(100);Alteromonadales(100);Alteromonadaceae(100);*Haliea*(100)
OTU15	Proteobacteria(100);Gammaproteobacteria(99);unclassified
OTU16	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU17	Acidobacteria(100);Acidobacteria_Gp10(100);Acidobacteria_Gp10_family_incertae_sedis(100);*Gp10*(100)
OTU18	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU19	Bacteroidetes(100);Sphingobacteria(100);Sphingobacteriales(100);Flammeovirgaceae(100)
OTU20	Proteobacteria(100);Betaproteobacteria(100);Burkholderiales(100);unclassified
OTU21	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Mariniflexile*(96)
OTU22	Bacteroidetes(100);Sphingobacteria(100);Sphingobacteriales(100);Flammeovirgaceae(93);unclassified
OTU23	Bacteroidetes(100);unclassified
OTU24	Bacteroidetes(100);Sphingobacteria(100);Sphingobacteriales(100);Cyclobacteriaceae(100);*Algoriphagus*(99);
OTU25	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU26	Proteobacteria(100);Gammaproteobacteria(100);Xanthomonadales(98);Xanthomonadaceae(98);unclassified
OTU27	Bacteroidetes(100); unclassified
OTU28	Bacteroidetes(100); unclassified
OTU29	Bacteria(99);unclassified
OTU30	Bacteroidetes(100);unclassified
OTU31	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU32	Bacteroidetes(100);Sphingobacteria(100);Sphingobacteriales(100);Cyclobacteriaceae(100);*Algoriphagus*(100)
OTU33	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Maribacter*(93)
OTU34	Proteobacteria(100);Betaproteobacteria(100);Burkholderiales(99);unclassified;
OTU35	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU36	Bacteroidetes(100);unclassified
OTU37	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(99);unclassified
OTU38	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU39	Proteobacteria(100);Betaproteobacteria(100);Burkholderiales(100);Comamonadaceae(100);unclassified
OTU40	Bacteroidetes(100);Sphingobacteria(100); Sphingobacteriales(100);Chitinophagaceae(100);*Sediminibacterium*(92);
OTU41	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Flavobacterium*(100);
OTU42	Proteobacteria(100);Betaproteobacteria(100);Burkholderiales(100);Comamonadaceae(100);unclassified
OTU43	Bacteria(100);unclassified
OTU44	Bacteroidetes(100);unclassified
OTU45	Proteobacteria (100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobulbaceae(100);unclassified
OTU46	Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU47	Proteobacteria(100);Deltaproteobacteria(100);Desulfuromonadales(100);Desulfuromonadaceae(100);unclassified
OTU48	Bacteroidetes(100);unclassified
OTU49	Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobacteraceae(100);unclassified
OTU50	Proteobacteria(100);Gammaproteobacteria(97);unclassified
OTU51	Cyanobacteria (100); Cyanobacteria_order_incertae_sedis(100);Family_XIII(100);*GpXIII*(100)
OTU52	Bacteria(100);Chloroflex(82);unclassified
OTU53	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU54	Proteobacteria(100);Epsilonproteobacteria(100);Campylobacterales(100);Helicobacteraceae(100);*Sulfurimonas*(99)

OTU#	Classification to lowest taxonomic rank (percentage confidence levels in brackets)
OTU55	Proteobacteria(100);Gammaproteobacteria(100);Xanthomonadales(97);Sinobacteraceae(95);*Steroidobacter*(95)
OTU56	Proteobacteria(100);Betaproteobacteria(100);Hydrogenophilales(100);Hydrogenophilaceae(100);*Thiobacillus*(100)
OTU57	Bacteroidetes(100);unclassified
OTU58	Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobacteraceae(100);unclassified
OTU59	Proteobacteria(100);Betaproteobacteria(100);unclassified
OTU60	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU61	Proteobacteria(100);Alphaproteobacteria(100);Rhodobacterales(99);Rhodobacteraceae(99);unclassified
OTU62	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Polaribacter*(83)
OTU63	Proteobacteria(99);Gammaproteobacteria(93);unclassified
OTU64	Bacteroidetes(99);Sphingobacteria(94);Sphingobacteriales(94);unclassified
OTU65	Proteobacteria(80);unclassified
OTU66	Proteobacteria(100);Gammaproteobacteria(98);unclassified
OTU67	Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU68	Actinobacteria(100);Actinobacteria(100);Actinomycetales(99);unclassified

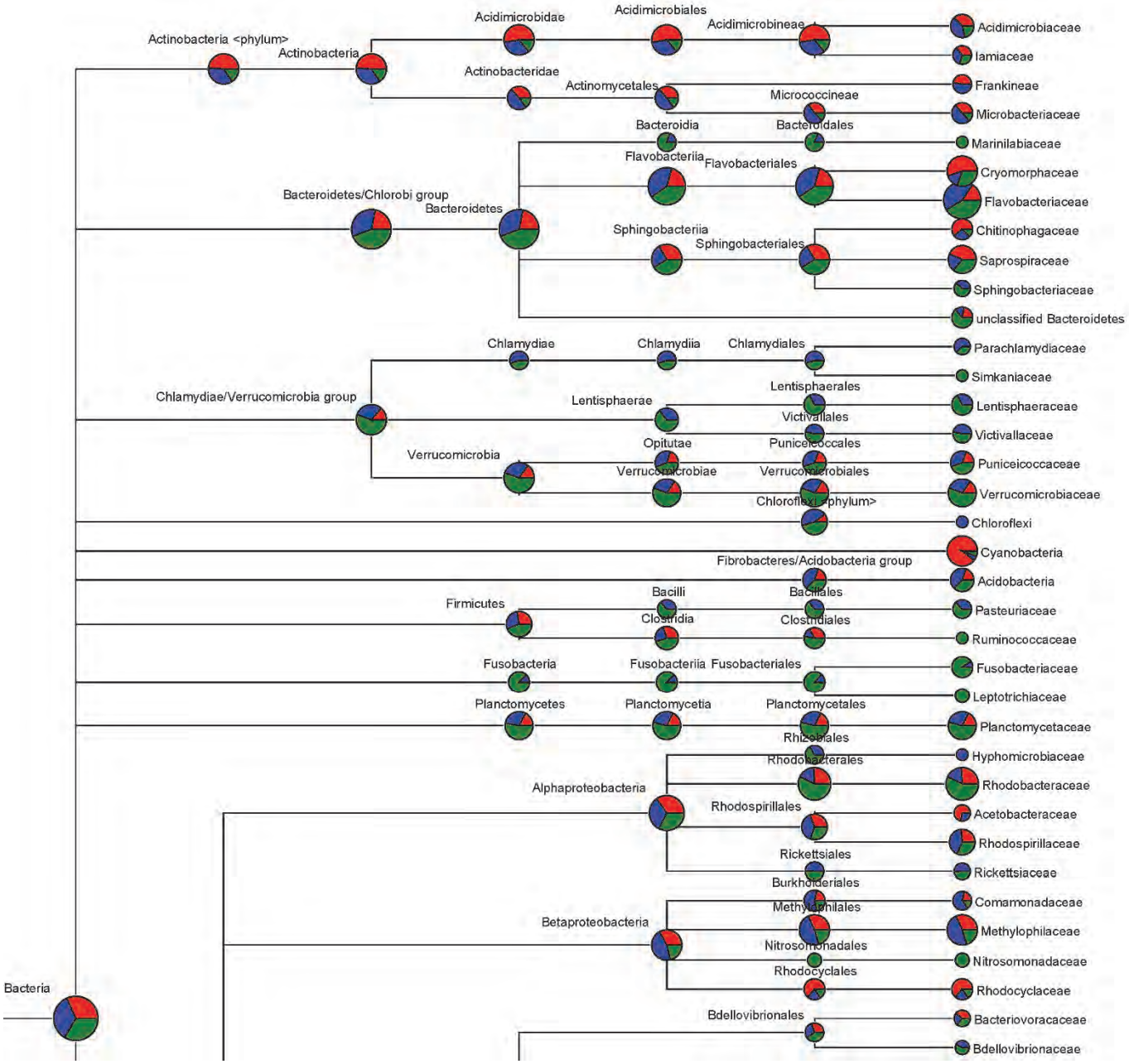
Supplementary Table 2. Taxonomic classification of the dominant OTUs from the Kariega and Kowie estuaries using a Naïve Bayesian classification algorithm against the Ribosomal Database Project trainset9.

OTU#	Classification to lowest taxonomic rank (percentage confidence levels in brackets)
OTU1	Bacteria(100); Proteobacteria(99);Gammaproteobacteria(99);unclassified
OTU2	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100); unclassified
OTU3	Bacteria(100);Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobulbaceae(100); unclassified
OTU4	Bacteria(100);Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobacteraceae(100); unclassified
OTU5	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(99);unclassified
OTU6	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU7	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU8	Bacteria(100);Bacteroidetes(100);unclassified
OTU9	Bacteria(100);Chloroflexi(99);Anaerolineae(99);Anaerolineales(99);Anaerolineaceae(99);unclassified
OTU10	Bacteria(100);Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobacteraceae(100);unclassified
OTU11	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU12	Bacteria(100);"Proteobacteria"(100);Gammaproteobacteria(99);Gammaproteobacteria_incertae_sedis(81); unclassified
OTU13	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);Gammaproteobacteria_incertae_sedis(98); unclassified
OTU14	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU15	Bacteria(100);Planctomycetes(100);Planctomycetacia(100);Planctomycetales(100);Planctomycetaceae(100); *Rhodopirellula*(100)
OTU16	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU17	Bacteria(100);Verrucomicrobia(100);Opitutae(100);Puniceicoccales(100);Puniceicoccaceae(100);*Coraliomargarita*(97)
OTU19	Bacteria(100);Proteobacteria(100);Deltaproteobacteria(100);Desulfuromonadales(100);Desulfuromonadaceae(100); unclassified
OTU20	Bacteria(100);Bacteroidetes(100);Sphingobacteria(95);Sphingobacteriales(95); unclassified
OTU21	Bacteria(100);unclassified
OTU22	Bacteria(100);Verrucomicrobia(98);Opitutae(98);unclassified
OTU23	Bacteria(100);unclassified
OTU24	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU25	Bacteria(100);unclassified
OTU26	Bacteria(100);unclassified
OTU27	Bacteria(100);Proteobacteria(83);unclassified
OTU28	Bacteria(100);Proteobacteria(100);Betaproteobacteria(100); unclassified
OTU29	Bacteria(100);Bacteroidetes(100);unclassified
OTU30	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);*Formosa*(100)
OTU31	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU32	Bacteria(100);Proteobacteria(97);Gammaproteobacteria(90);unclassified
OTU33	Bacteria(100);Proteobacteria(100);Alphaproteobacteria(100);Rhodobacterales(100);Rhodobacteraceae(100);unclassified
OTU34	Bacteria(100);Verrucomicrobia(100);Verrucomicrobiae(100);Verrucomicrobiales(100);Verrucomicrobiaceae(91);unclassified

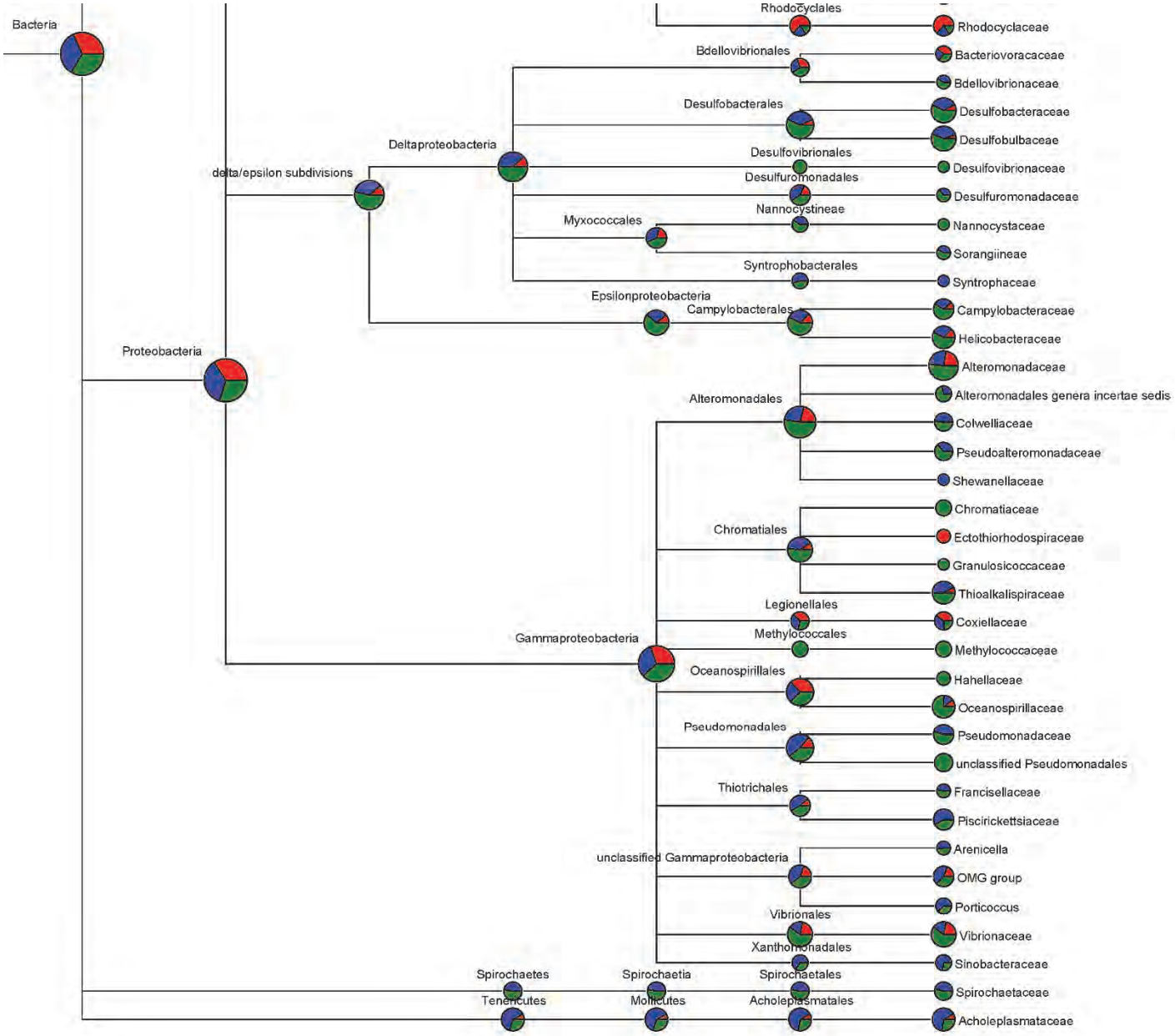
OTU#	Classification to lowest taxonomic rank (percentage confidence levels in brackets)
OTU35	Bacteria(100);Proteobacteria(100);Alphaproteobacteria(100);Rhodobacterales(100);Rhodobacteraceae(100);*Roseovarius*(96)
OTU36	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU37	Bacteria(100);Cyanobacteria(100);Cyanobacteria(100);Family_II(100);GpIIa(100);unclassified
OTU38	Bacteria(100);Proteobacteria(98);Gammaproteobacteria(93);unclassified
OTU39	Bacteria(100);Actinobacteria(100);Actinobacteria(100); unclassified
OTU40	Bacteria(100);Cyanobacteria(100);Cyanobacteria(100);Family_II(100);GpIIa(100);unclassified
OTU41	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);"Flavobacteriales"(100);Flavobacteriaceae(100);*Lutimonas*(100)
OTU42	Bacteria(100);Bacteroidetes(100);Sphingobacteria(96);Sphingobacteriales(96);unclassified
OTU43	Bacteria(100);Bacteroidetes(100);Sphingobacteria(100);Sphingobacteriales(100);Cytophagaceae(100);*Cytophaga*(100)
OTU44	Bacteria(100);Bacteroidetes(100);unclassified
OTU45	Bacteria(100);Proteobacteria(100);Epsilonproteobacteria(100);Campylobacterales(100);Helicobacteraceae(100);*Sulfurovum*(100)
OTU46	Bacteria(100);Proteobacteria(100);Deltaproteobacteria(100);Desulfobacterales(100);Desulfobulbaceae(100);unclassified
OTU47	Bacteria(100);Spirochaetes(100);Spirochaetes(100);Spirochaetales(100);Spirochaetaceae(100);*Spirochaeta*(100)
OTU48	Bacteria(100);Proteobacteria(100);Epsilonproteobacteria(100);Campylobacterales(100);Helicobacteraceae(100);unclassified
OTU49	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100); unclassified
OTU50	Bacteria(100);Bacteroidetes(100);unclassified
OTU51	Bacteria(100);Proteobacteria(99);Gammaproteobacteria(98);unclassified
OTU52	Bacteria(100);Verrucomicrobia(100);Verrucomicrobiae(100);Verrucomicrobiales(100);Verrucomicrobiaceae(99);unclassified
OTU53	Bacteria(100);Bacteroidetes(100);Flavobacteria(100);Flavobacteriales(100);Flavobacteriaceae(100);unclassified
OTU54	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(91);unclassified
OTU55	Bacteria(100);unclassified
OTU56	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(99);unclassified
OTU57	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);unclassified
OTU58	Bacteria(100);Proteobacteria(100);Gammaproteobacteria(100);Thiotrichales(100);Piscirickettsiaceae(100);*Thiomicrospira*(100)
OTU59	Bacteria(100);unclassified
OTU60	Bacteria(100);unclassified

Appendix 2

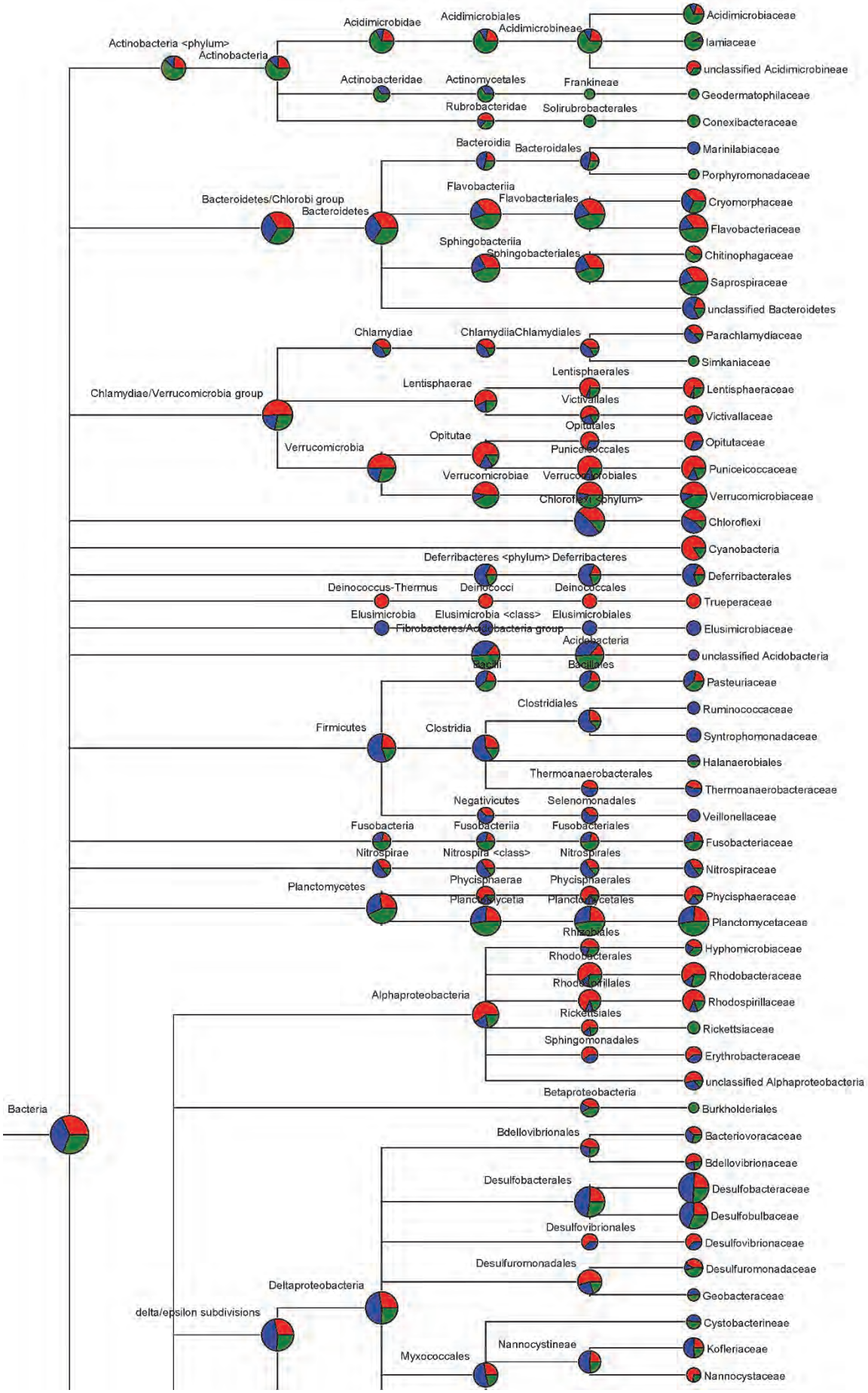
In all of the following figures, the lowest taxonomic ranking is that of family. Taxonomic classification was done using the Ribosomal Database Project and the phylogram generated using MEGAN software version 5.5.3 using square Root Normalization for optimal visualisation of taxonomic nodes (Huson et al., 2007).



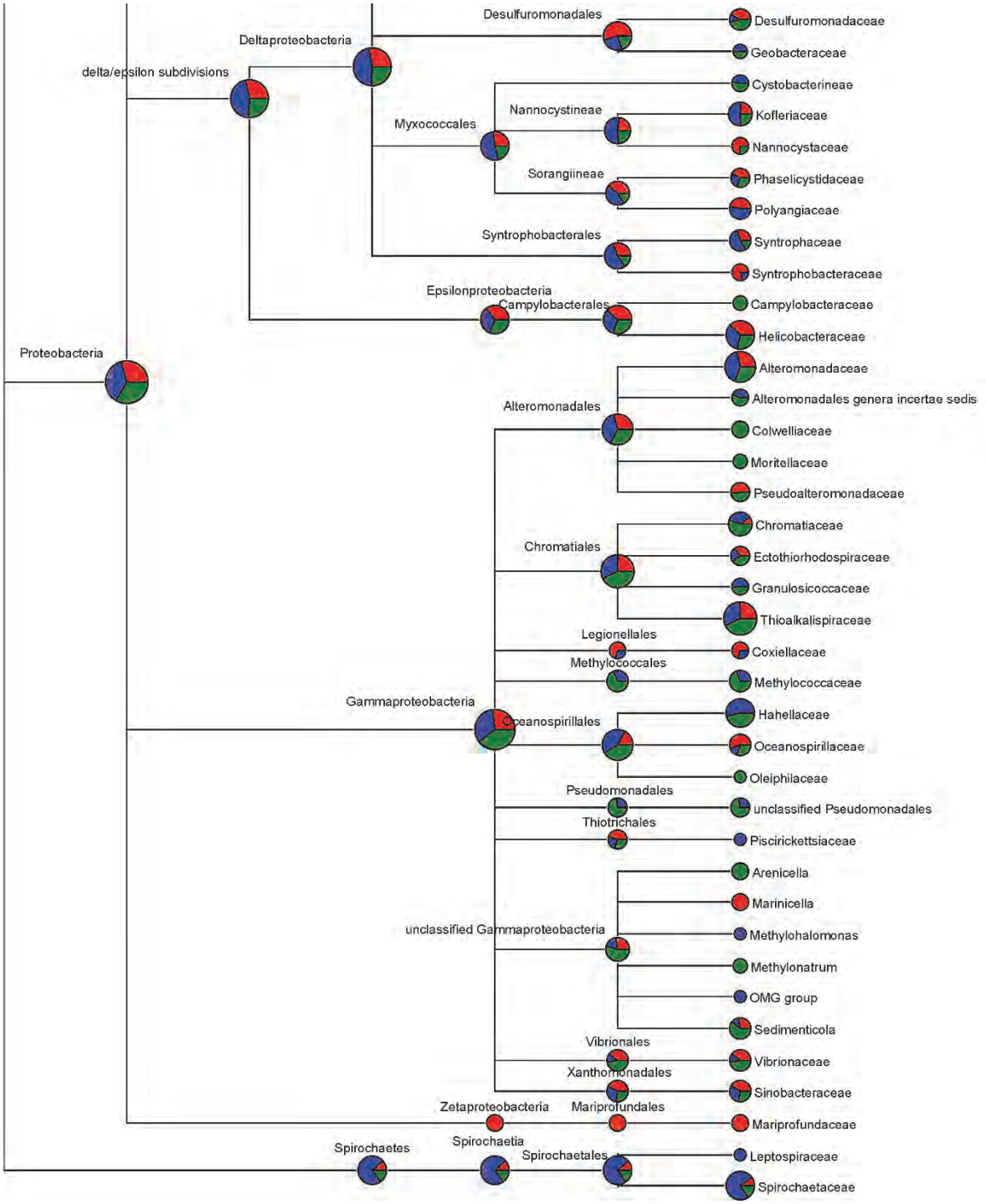
Continued on next page….



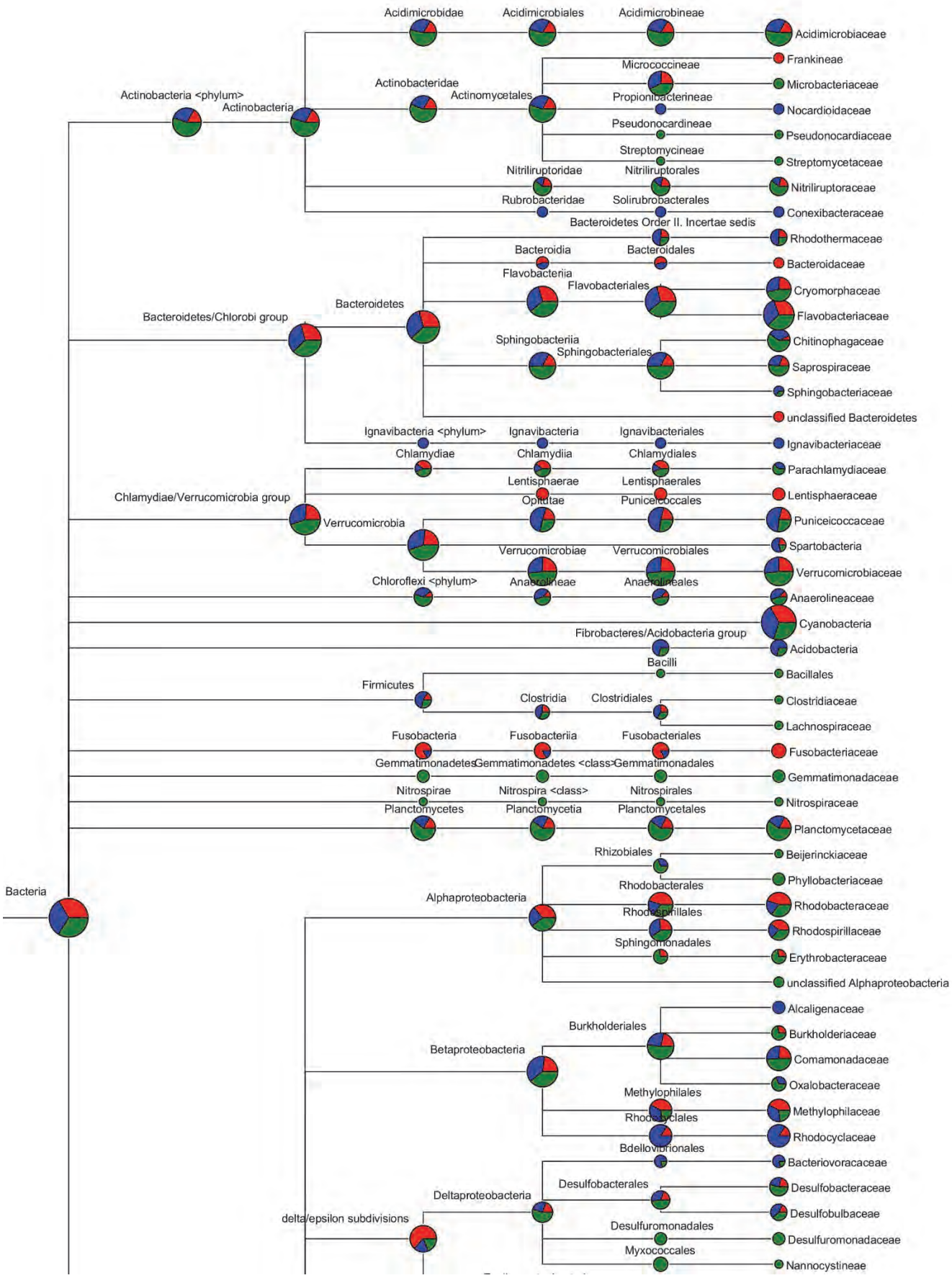
Supplementary Figure 1. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the water column of the Kariega Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



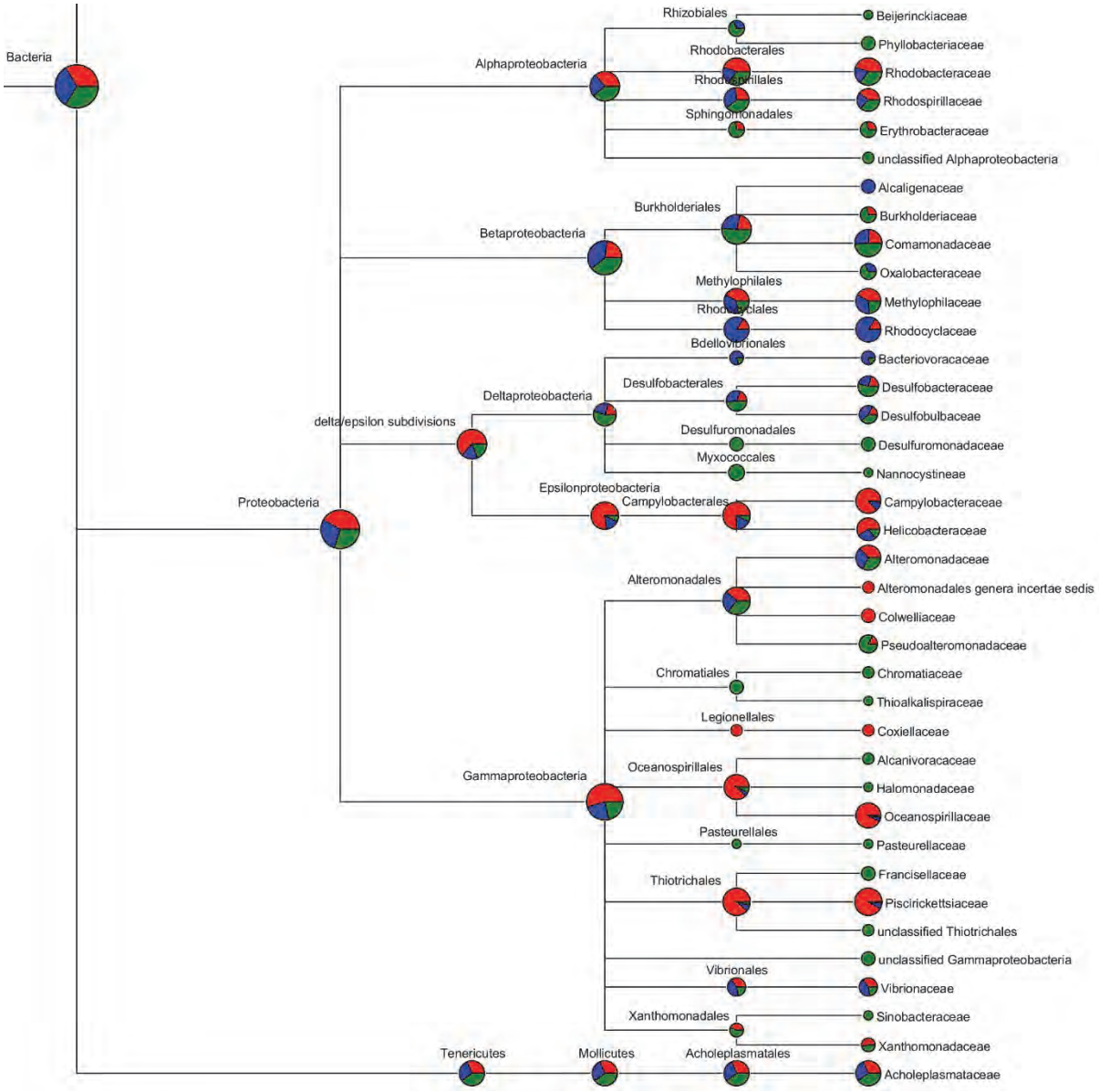
Continued on next page….



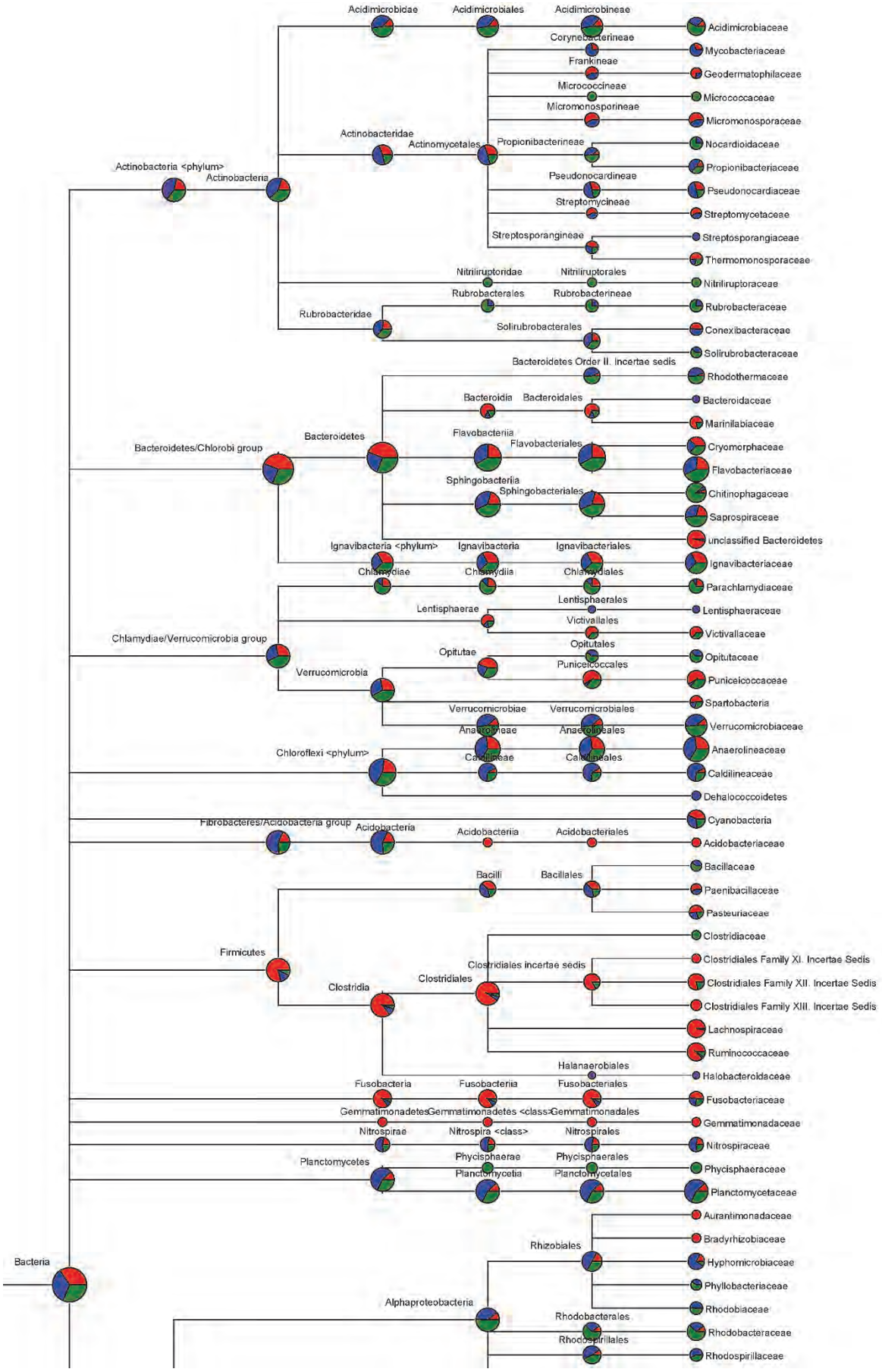
Supplementary Figure 2. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the sediment of the Kariega Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



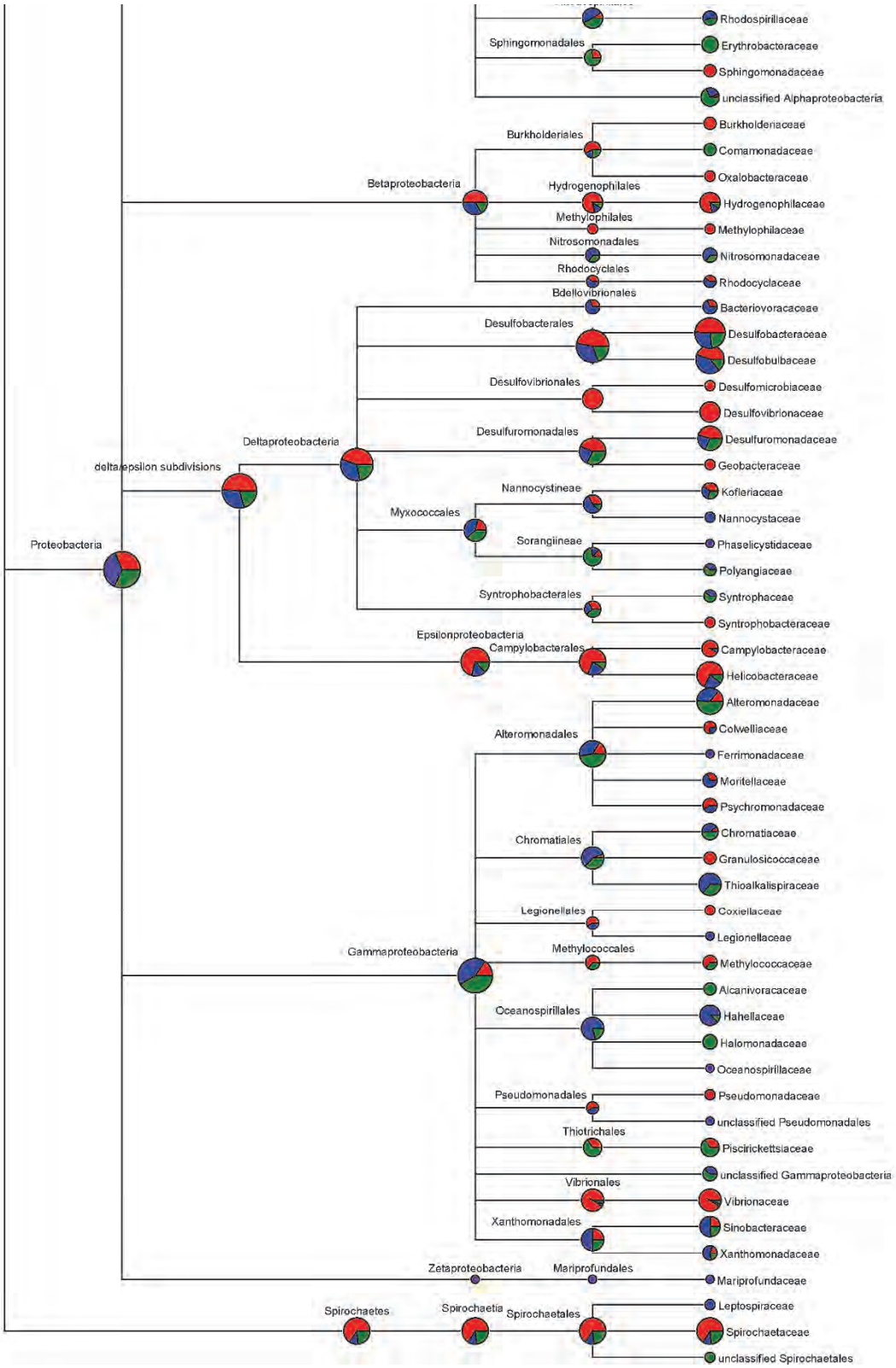
Continued on next page….



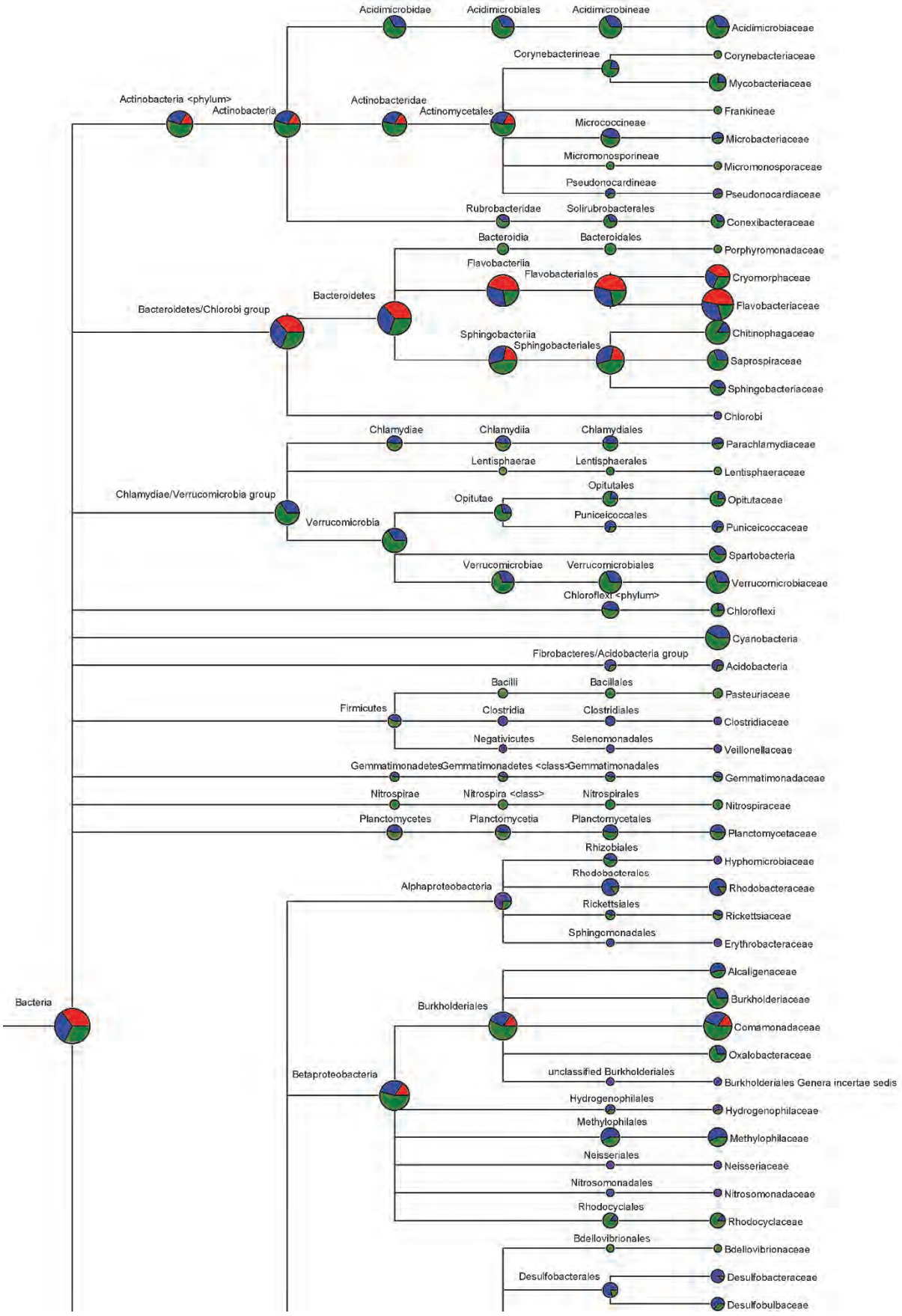
Supplementary Figure 3. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the water column of the Kowie Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



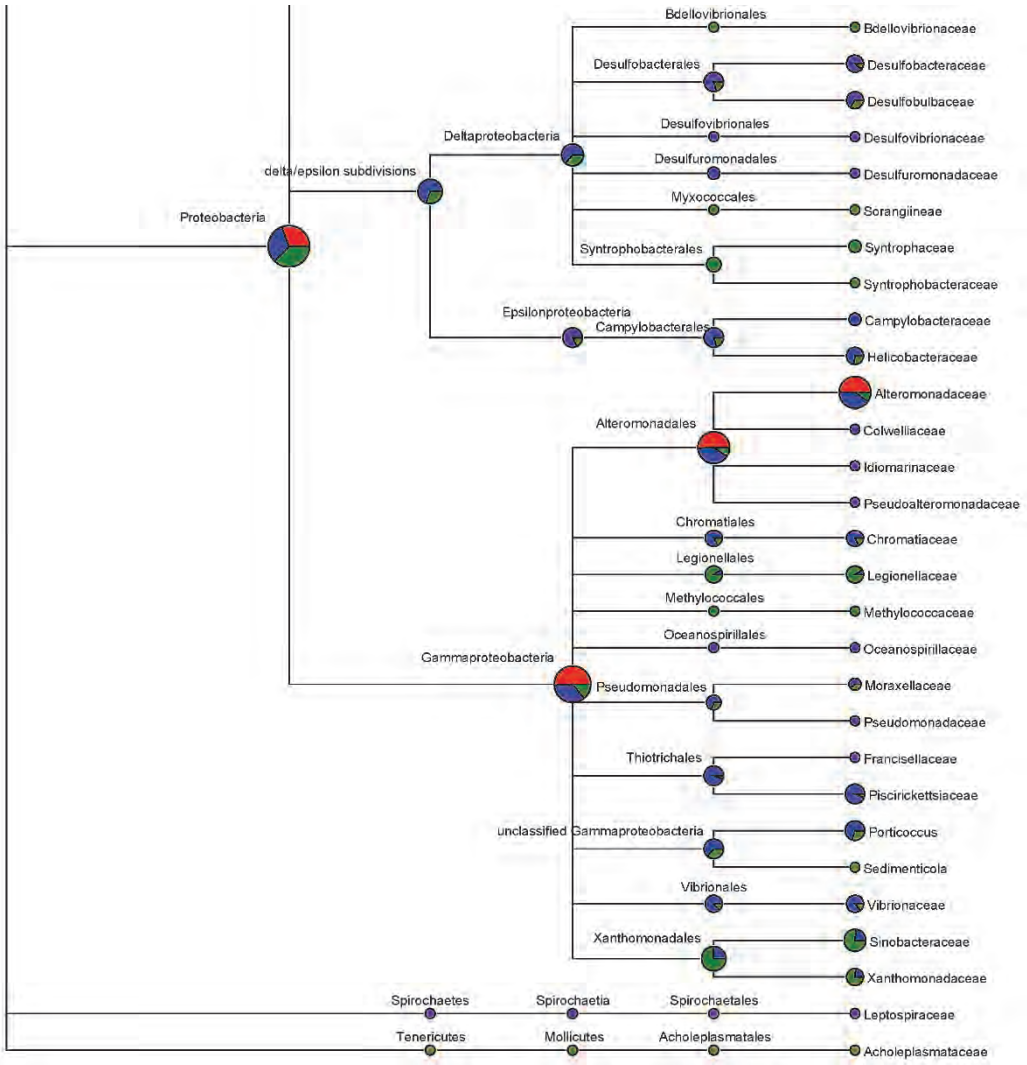
Continued on next page….



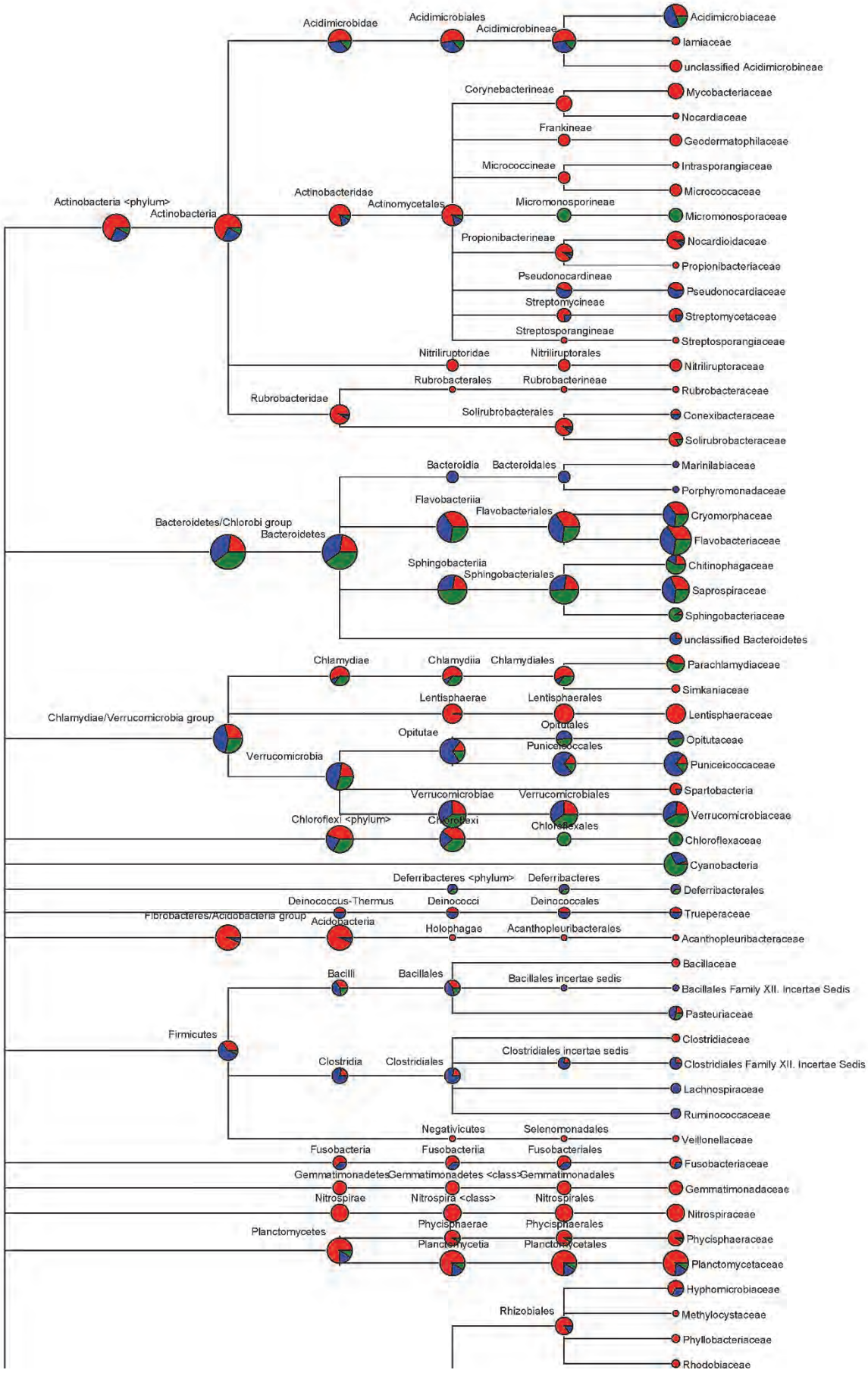
Supplementary Figure 4. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the sediment of the Kowie Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



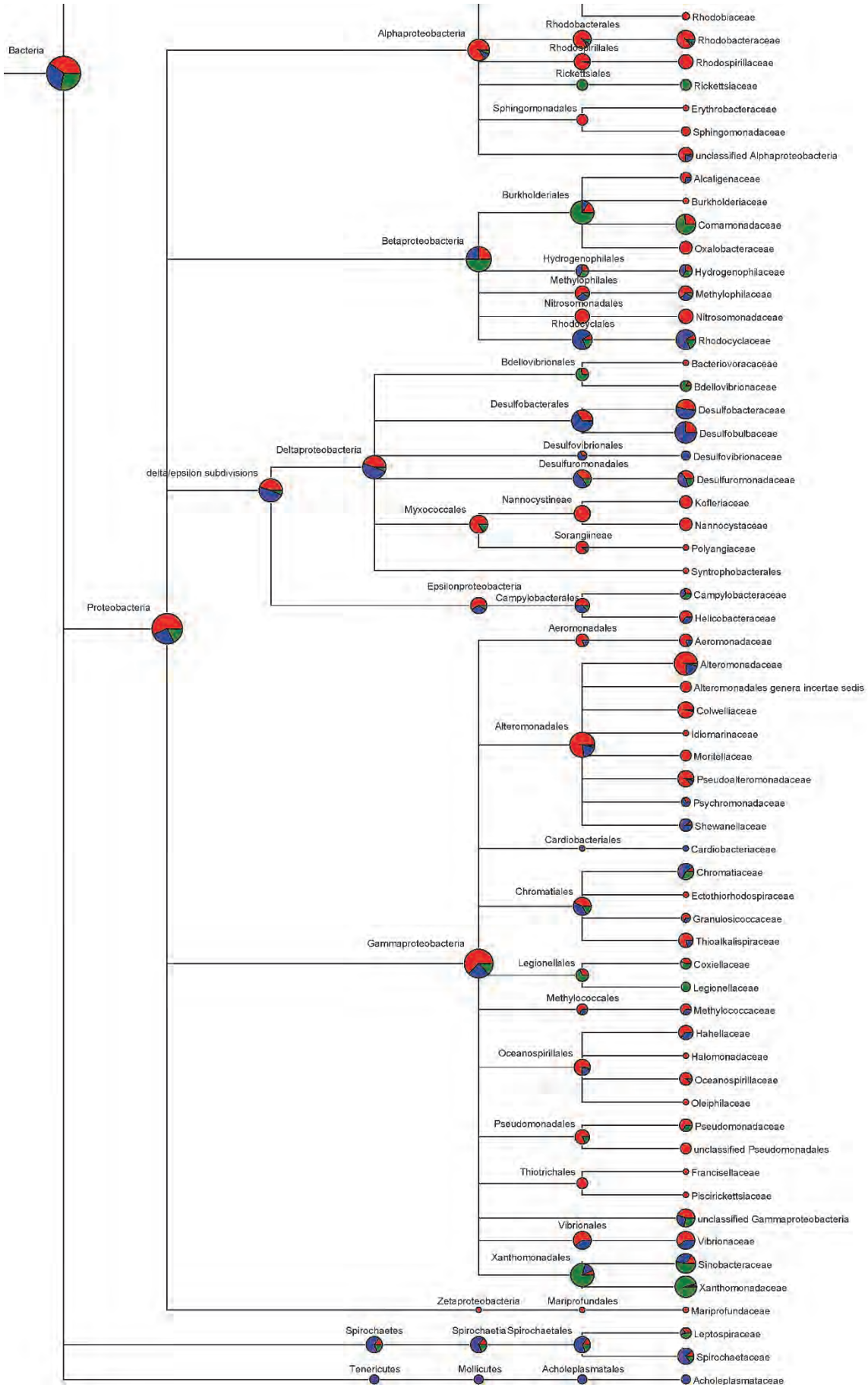
Continued on next page….



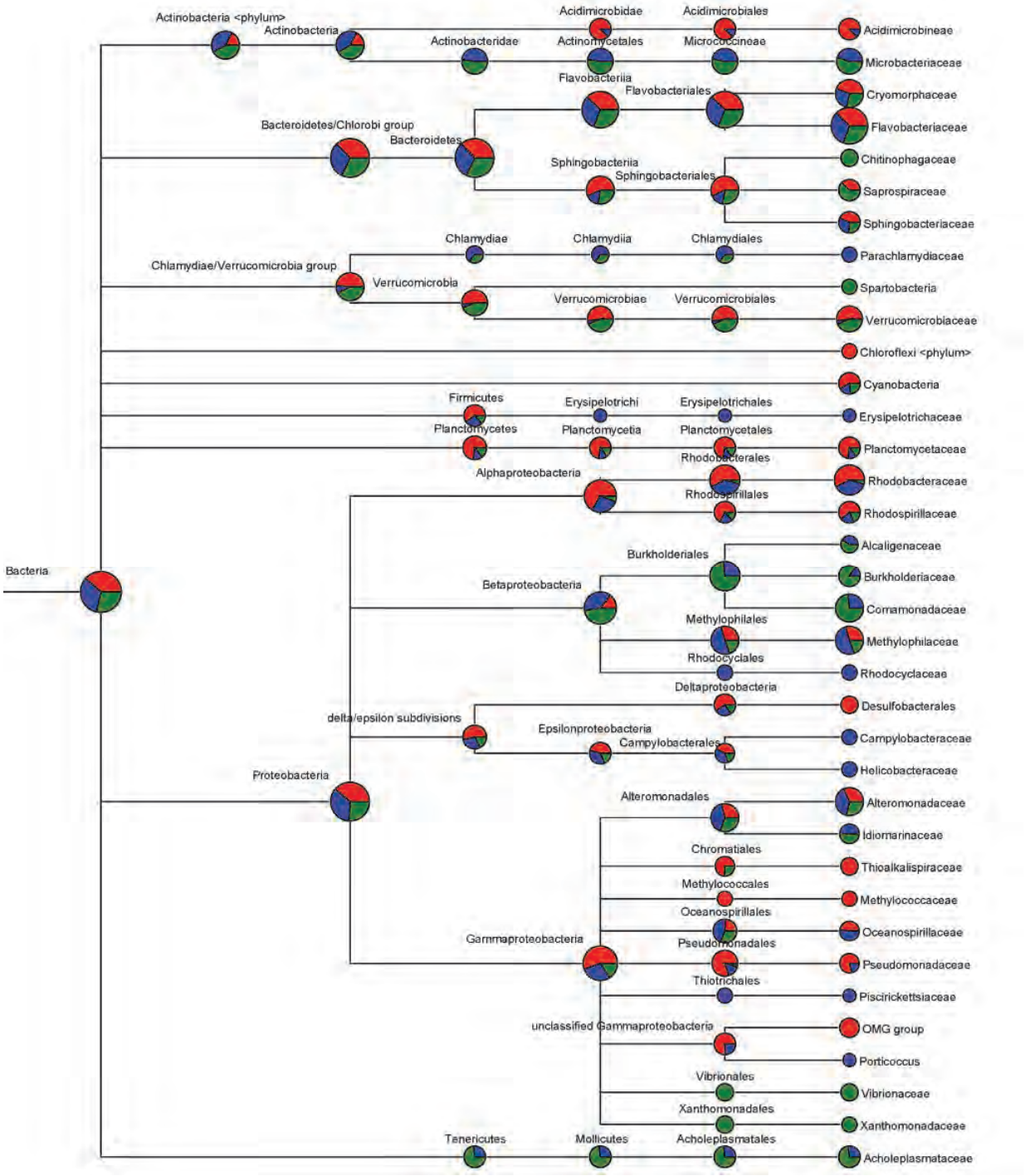
Supplementary Figure 5. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the water column of the Sundays Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



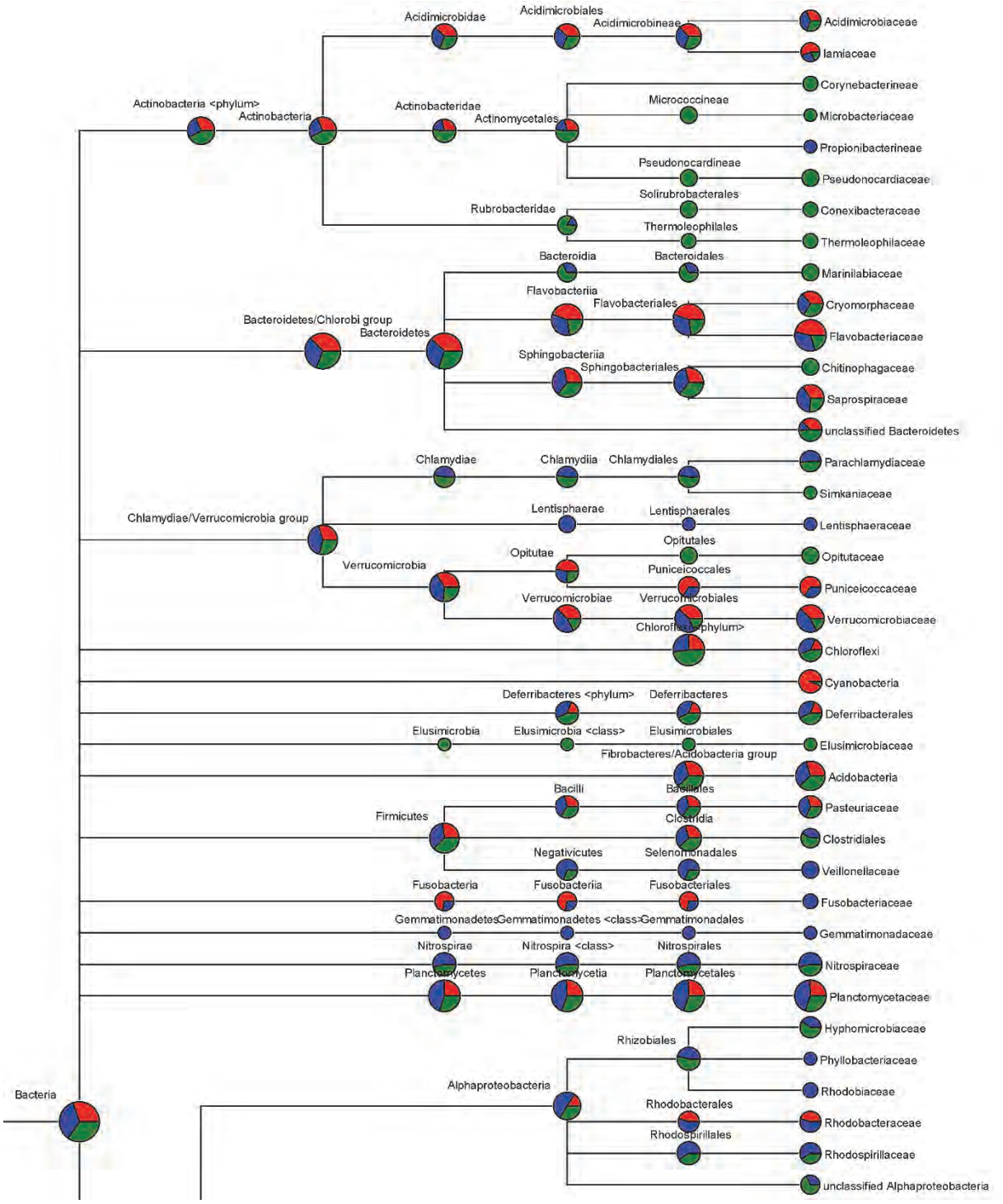
Continued on next page….



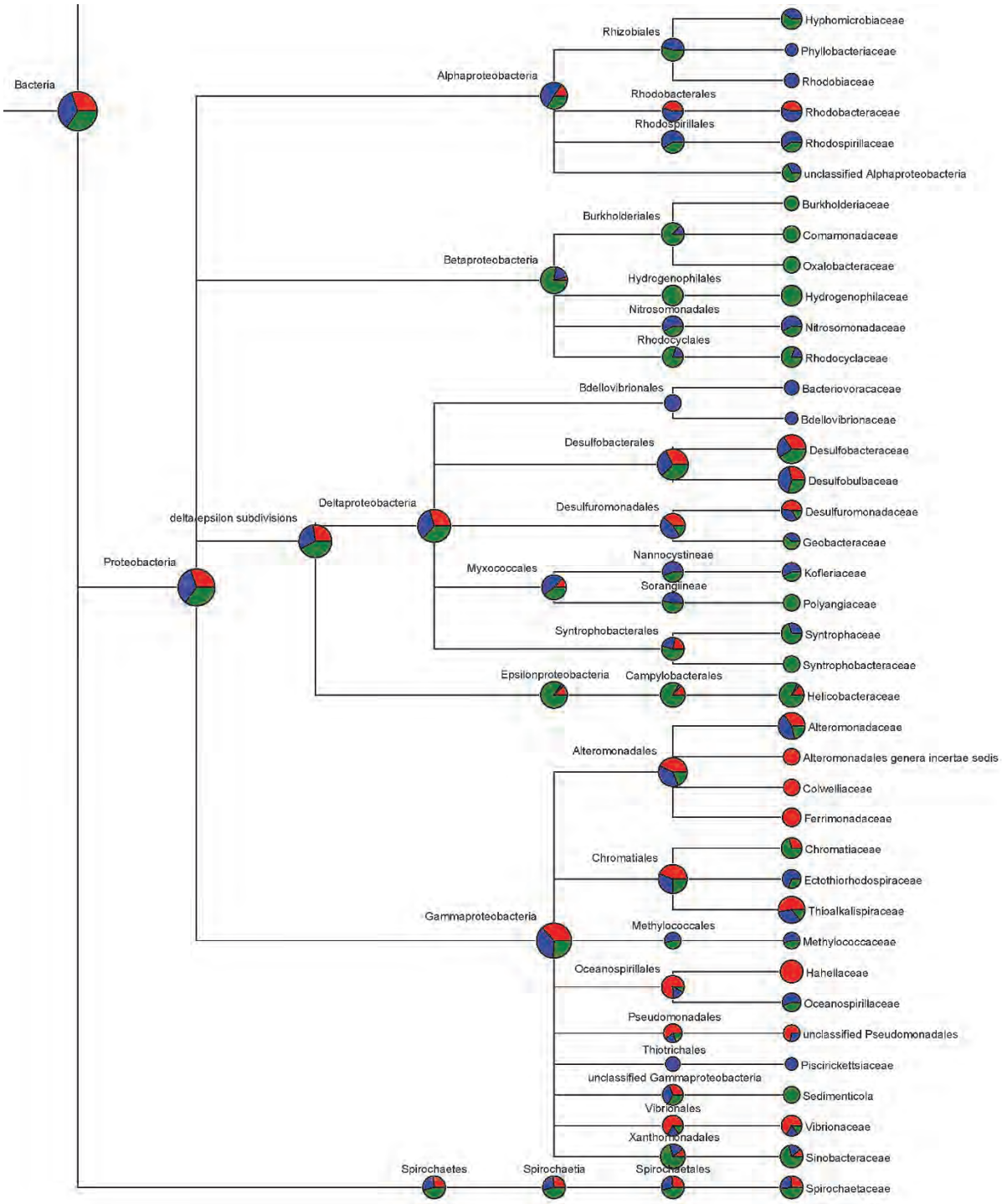
Supplementary Figure 6. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the sediment of the Sundays Estuary.



Supplementary Figure 7. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the water column of the Swartkops Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)



Continued on next page….



Supplementary Figure 8. Graphical representation of a phylogram comparison of the 16S rRNA amplicon libraries generated from the sediment of the Swartkops Estuary. (Red: Mouth; Blue: Middle reaches; Green: Upper reaches)