Microbial communities of wastewater in the Canadian Arctic

by

Monica Gromala

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Biology

Waterloo, Ontario, Canada, 2019

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Dr. Brendan McConkey and Jared Ellenor assisted with collection of samples in Baker Lake. Members of the Wong lab at the University of Manitoba collected samples in Cambridge Bay.
Rosie Smith and Dr. Heidi Swanson collected samples in Kugluktuk.

Sura Ali assisted with DNA extractions and preparation of samples for sequencing.

Metagenomic sequencing was performed by the Farncombe Metagenomics Facility at McMaster University.

Abstract

Wastewater management in the Canadian Arctic is challenging due to the extreme climate and lack of conventional treatment system infrastructure. To resolve these issues, most Northern communities use waste stabilization ponds (WSPs) as their sole form of wastewater management. In this type of system, wastewater is collected directly from buildings year-round and is transported to the WSP where it is discharged directly into the environment. For approximately nine months of the year, the wastewater remains frozen due to sub-zero temperatures. As the temperature increases in the spring, ice begins to melt, allowing the wastewater to thaw and flow into downstream lakes. Increased sunlight and increased water temperatures stimulate microorganisms to naturally breakdown the various components of the wastewater and reduce concentrations of macronutrients, such as nitrogen and phosphorus. The two to three months during which temperatures remain above zero are collectively referred to as the treatment season as the majority of aerobic and anaerobic degradation occurs during this time.

Currently, the microbial communities of Arctic WSPs are not well understood. Identifying the major microorganisms in the microbial communities of these WSPs is important for understanding whether the wastewater treatment system is effective in the removal of wastewater contaminants from effluent water flowing into downstream lakes. This research aims to characterize the microbial community and identify wastewater indicator organisms of three Arctic WSPs in Baker Lake, Cambridge Bay, and Kugluktuk in Nunavut, Canada, with a focus on the Baker Lake WSP and the downstream lake system. In

Baker Lake, wastewater is discharged into a lagoon, where water is able to flow into downstream lakes before ultimately discharging into Baker Lake. Several sites across this entire system, as well as three reference sites were sampled and analyzed to produce baseline taxonomic profiles. This was achieved through 16S rRNA gene and metagenome sequencing, and subsequent taxonomic assignment to 16S rRNA gene amplicons, metagenome sequences corresponding to the *rpoB* gene, and metagenome-assembled genomes (MAGs).

Both temporal and spatial variation were observed across Baker Lake sampling sites, with clear differences in microbial community composition at samplings sites at the phylum, family, and ASV levels. These trends were observed in each of the 16S rRNA gene, *rpoB* gene, and MAG taxonomic assignments. Wastewater indicator species were identified for the Baker Lake WSP and these indicator microorganisms were not detected in most downstream sites. The three WSPs were also found to have very different microbial communities, with few overlapping ASVs. Despite these differences at the ASV level, the three WSPs contained common groups of microorganisms and microbial community profiles were distinct from Arctic freshwater lake samples. Most of the ASVs detected in all three WSPs belonged to the phylum *Proteobacteria*, which had a relative abundance of >50% in nearly all WSP samples. The proportion of *Proteobacteria* was much lower in lakes found downstream of the Baker Lake WSP. In receiving waters and reference lakes, *Bacteroidetes*, *Actinobacteria*, *Planctomycetes*, and *Verrucomicrobia* were more prevalent.

Wastewater is known to be a hotspot for antibiotic resistance gene (ARG) dissemination and evolution. This thesis also aimed to identify and quantify ARGs within the

WSPs, receiving waters, and reference lakes to determine if ARGs were enriched in samples containing wastewater. Eight gene families related to macrolide resistance, and one gene related to class A betalactamases were detected at higher abundance in wastewater samples than in downstream and reference lakes. Metagenome sequences related to these ARGs were also found to be in bins associated with five MAGs.

This research is an important step in the characterization of microbial communities of wastewater in Arctic WSPs. The taxonomic profiles of the WSPs from Baker Lake, Cambridge Bay, and Kugluktuk provide important baseline data that will assist with monitoring of the wastewater system and gives insights into the efficacy of WSPs as the sole wastewater treatment system for communities in the Arctic. This research will aid in tracking changes in microbial communities as operational and infrastructure upgrades are made to the Baker Lake waste treatment system, to help with developing a safe and effective wastewater treatment system.

Acknowledgements

I would like to thank all of the amazing people that have supported me through this journey.

First, I would like to thank Dr. Brendan McConkey and Dr. Josh Neufeld for their guidance and enthusiasm, and for providing me with the opportunity to work on a such a unique and exciting project. Your constant support fueled my determination and pushed me to work harder than I ever have. I would also like to thank Dr. Laura Hug and Dr. Wayne Parker for being on my committee and providing me with advice and insight.

I would also like to thank the community of Baker Lake, Cambridge Bay, and Kugluktuk for kindly allowing us to sample their waters.

I would also like to thank Sura Ali and Jared Ellenor for their help with lab and field work. Thank you for putting up with me through all of my frustration.

I would also like to thank the many members of the McConkey and Neufeld labs for their continual support and advice. You all kept me smiling throughout this crazy journey, even during the toughest days in the lab. I especially want to thank Katja Engel and Jackson Tsuji for being amazing mentors and assisting me in all of the many times I had to troubleshoot putting up with my constant questions, and for helping me with my scripts. I would like to thank my office mates in the McConkey lab, Janet Lorv, Karsten Rinas, and Mark Lubberts, for always listening to my rants and giving me honest opinions, especially when I was being stubborn.

Finally, I would like to thank my family and friends for supporting me through all of the good times and bad. I could not have accomplished what I did without each and every one of you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AOA	Ammonia-oxidizing archaea
AOB	Ammonia-oxidizing bacteria
ARB	Antibiotic resistant bacteria
ARG	Antibiotic resistance gene
ASV	Amplicon sequence variant
BNR	Biological nitrogen removal
BOD	Biochemical oxygen demand
CARD	Comprehensive Antibiotic Resistance Database
COD	Chemical oxygen demand
dNTP	Deoxynucleotide triphosphate
DO	Dissolved oxygen
EBPR	Enhanced biological phosphorus removal
HMM	Hidden Markov model
HGT	Horizontal gene transfer
IndVal	Indicator value
ISA	Indicator species analysis
MAG	Metagenome-assembled genome
MSA	Multiple sequence alignment
MWWE	Municipal wastewater effluent
NCBI	National Centre for Biotechnology Information

ORF	Open reading frame
OTU	Operational taxonomic unit
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
RPKM	Reads per kilobase of reference sequence per million sample reads
rpoB	Beta subunit of bacterial RNA polymerase
rRNA	Ribosomal RNA
TDS	Total dissolved solids
TSS	Total suspended solids
WSP	Waste stabilization pond
WWTP	Wastewater treatment plant
UPGMA	Unweighted pair group method with arithmetic mean

Chapter 1

Introduction

1.1 Municipal Wastewater Effluent and Eutrophication

More than 80% of Canadians benefit from wastewater treatment system facilities[1]. Effluent from these facilities can deteriorate water quality of receiving waters[2]. Municipal wastewater effluent (MWWE) composition varies depending on region, but in general, MWWE contains solids, dissolved and particulate matter, nutrients, pathogenic and non-pathogenic bacteria, organic chemicals, metals, oils, and plastics[1]. Wastewater treatment plants (WWTPs) aim to reduce as many of these components as possible, focusing on reducing total suspended solids (TSS), chemical and biochemical oxygen demand (COD and BOD, respectively), nitrogen and phosphorus, pathogenic bacteria, and plastics[1]. Chemical oxygen demand is a measure of the amount of oxygen required to chemically oxidize organic matter, whereas BOD is the amount of dissolved oxygen required by microorganisms to oxidize organic matter[3].

When treated improperly, municipal wastewater effluent is a major cause of eutrophication, or nutrient enrichment, of many lakes, ponds, and rivers[4]. MWWE can contain increased concentrations of nitrogen and phosphorus, as well as other micronutrients, that are discharged into receiving waters[4,5]. The amount of nutrients present plays an important role in shaping the biological community composition, and nutrient enrichment can cause both direct and indirect biological changes, such as stimulation of algal growth or changes in fish communities caused by reduced oxygen[4]. Often, the growth of organisms is limited by the amount of nutrients present. In conditions where nutrients are the limiting

factor, eutrophication can stimulate the growth of algal species. This may include species that produce toxins, such as the diatom *Pseudo-nitzschia*, that are released into the water, as well as those that produce large amounts of biomass, such as the green macroalga *Cladophora*, which can cause hypoxia of other aquatic organisms, including microorganisms[4,6]. Fish communities of receiving waters are typically impacted by algal blooms, with increased fish kills in these circumstances[4].

1.1.1 Impact of MWWE on Microbial Communities

Many studies have documented impacts of wastewater effluent on receiving waters. For example, shifts in microbial communities can be detected at the phylum level when comparing wastewater effluent to downstream sites[7], microbial counts increase at sites downstream of wastewater effluent discharge sites[8], and microorganisms from the effluent itself are able to persist at these downstream sites[9]. In addition, microbial community diversity decreases in sediment of downstream sites[10] and these impacted sediment sites are associated with decreased abundances of sulfate reducers, denitrifiers, and methanogens in comparison to sediments of upstream sites[11]. Microplastics from wastewater effluent positively influence the ability of microorganisms to form biofilms and increase the ability of these microorganisms to survive downstream[12]. Overall, microbial community composition is impacted by wastewater effluent discharge.

1.1.2 Role of Microorganisms in Reducing Wastewater Nutrient Contaminants

Microorganisms play an important role in the reduction of nutrients from wastewater and reduce the potential for eutrophication. Specifically, the removal of nitrogen and phosphorus is an important role played by microorganisms in wastewater treatment[13].

Wastewater nitrogen exists primarily as organic nitrogen, ammonia, nitrite, and nitrate[14]. Unionized ammonia is an important contributor to wastewater effluent toxicity, both directly to fish, and also because nitrification decreases the amount of available oxygen for other organisms[15,16]. Because of this, WWTPs commonly employ biological nitrogen removal (BNR) processes in the form of activated sludge to reduce the amount of nitrogen being released into the environment[17]. During BNR, microorganisms in the activated sludge perform both nitrification and denitrification to help reduce inorganic nitrogen[17]. In the process of nitrification, ammonia-oxidizing bacteria (AOB) and ammonia-oxidizing archaea (AOA), use oxygen to convert ammonia to nitrite[16,18]. Nitrite is then oxidized by nitrite-oxidizing bacteria, such as those in the genera *Nitrobacter* and *Nitrospira*, into nitrate[16]. Both nitrite and nitrate can be reduced to dinitrogen gas through the process of denitrification, which then dissipates[16]. The most common denitrifying bacteria are facultative anaerobes, which are able to reduce nitrate and nitrite in both oxic and anoxic conditions[16,19].

Phosphorus is a critical limiting macronutrient for algal growth, and phosphorus removal is also important for reducing eutrophication in receiving waters[20]. The most common form of phosphorus in wastewater is orthophosphate[21]. The biological process to remove phosphorus is referred to as enhanced biological phosphorus removal (EBPR)[22]. Phosphorus removal by EBPR is first done anaerobically, followed by aerobic stages of

treatment[21]. The anaerobic steps cause phosphates to be released from the activated sludge into solution, which is then aerated to stimulate phosphate uptake in higher concentrations than required by the microorganisms[21]. EBPR has been shown to remove more than 90% of soluble phosphorus[21]. Microorganisms involved in this process are primarily associated with *Betaproteobacteria* and *Actinobacteria*[22]. Chemical precipitation is also a commonly used process for the removal of phosphorus, where metal salts are applied to produce metal phosphates, which can then be removed from the water[23,24].

1.1.3 Wastewater and Antibiotic Resistance

Antibiotic resistance has become a growing problem for society due to the frequent misuse of antibiotics. Although antibiotics originate from environmental bacteria, humans have provided strong selective pressures for antibiotic resistance genes (ARGs) since antibiotics were first used to treat and prevent infections[25]. As antibiotics became more prevalent in medical settings, selection of antibiotic resistant bacteria (ARB) and antibiotic resistance genes also increased[26]. These ARGs were thought to have first arisen within antibiotic-producing bacteria, which would allow the microorganisms to protect themselves from the antibiotics[25]. Fecal bacteria also play a role in the selection of ARGs because they may have been exposed to high concentrations of antibiotics[27].

Because ARGs are commonly associated with plasmids that transfer between different bacteria through horizontal gene transfer (HGT), this has led to further dissemination of ARGs, and therefore ARB[28]. Microorganisms can transfer genes by HGT using three different methods: conjugation, transformation, and transduction[29]. During

conjugation, cells interact with one another via a conjugative pilus, which transfers genetic material from a donor to a recipient cell[29,30]. Transformation, in contrast, is limited to naturally competent bacterial cells incorporating free environmental DNA and is far less common mechanism [31]. Transduction occurs when genetic material from a host bacterium is accidentally packaged alongside viral DNA into viral capsids, and transferred to a new recipient by infection via the bacteriophage transducing particle[29]. Within the context of wastewater treatment environments, conjugation is thought to be the primary mechanism for ARG transmission between bacteria through the use of conjugative plasmids[32].

Wastewater has been considered a hotspot for the dissemination and evolution of antibiotic resistance because most antibiotics used by humans appear in and are processed through municipal wastewater[33]. Here, bacteria from a variety of sources are exposed to sublethal concentrations of antibiotics and are able to develop antibiotic resistance. These ARB can pass their ARGs to non-resistant bacteria from different sources through HGT to produce new ARB[34]. This is especially a concern for human pathogens that may be present in wastewater, which also have the potential to acquire ARGs.

With the regular identification of new ARGs, various databases have been created to gather information about ARG classes and further incorporate data about the associated proteins and literature. In particular, the Comprehensive Antibiotic Resistance Database (CARD) is manually curated to provide an updated database of all antibiotic resistance genes and associated metadata[35]. In the CARD, ARGs have been divided into commonly known classes such as aminoglycosides, β-lactams, chloramphenicol, macrolides, tetracyclines, and sulfonamides[35]. This database identifies ARGs from metagenome reads using a tool called

Short, Better Representative Extract Dataset (ShortBRED)[36]. The ShortBRED tool produces protein-family specific marker sequences by first clustering user proteins of interest into families using CD-HIT[36–38]. Multiple sequences alignments (MSA) are then made for each family using MUSCLE[39] and Biopython[40] to produce a consensus sequence for each family. These consensus sequences are then compared against one another as well as against a reference database using BLAST[41], and sequences that are unique, or have minimal overlap from other protein families are identified. The identified marker sequences are then used to determine the relative abundance of the protein families in whole metagenome sequence data using USEARCH[42]. The raw number of hits is converted to a normalized count in reads per kilobase of reference sequence per million sample reads (RPKMs), which are output into a relative abundance table, which can be compared between different samples[36].

1.2 Waste Stabilization Ponds

Waste stabilization ponds (WSPs) are the simplest forms of wastewater management[43]. Most wastewater treatment systems with WSPs require little or no equipment for treatment to occur[43,44]. These ponds are often engineered such that wastewater is stabilized through aerobic and anaerobic processes occurring naturally by microorganisms in the environment. Multiple WSPs can also be combined, each for the removal of different pollutants within the wastewater[45]. For example, facultative and anaerobic ponds are specifically designed for the removal of BOD, whereas aerobic ponds are better for removing pathogens[45]. Although WSPs remove contaminants efficiently, they are optimal for locations with a climate characterized by adequate sunlight exposure and high temperatures[43]. This is due to poor

biological activity in extremely cold conditions, such as those in Arctic environments, where there is limited sunlight and monthly average temperatures below zero for most of the year[46]. Sunlight is also a major factor governing the efficiency of WSPs in the removal of pathogens. Bacteria such as *Escherichia coli*, *Campylobacter jejuni*, and *Salmonella enterica* are sensitive to sunlight, which causes UV-associated DNA damage[47]. Bacteria are also able to repair UV damage in dark conditions, and therefore WSPs that have higher sunlight exposure are expected to be more efficient in the sustained removal of pathogens[47]. There are several types of WSPs, with the most common being facultative ponds, anaerobic ponds, aerated lagoons, maturation ponds, and high-rate algal ponds (HRAPs)[47].

1.2.1 Facultative Waste Stabilization Ponds

The most common type of WSP is facultative WSPs (Figure 1.1). In Arctic settings, these WSPs are often used alone, with effluent discharged directly into the ocean, or with downstream water systems such that multiple stages of treatment can occur. Solid organic material settles to the bottom of the facultative WSP and forms a sludge. The bottom layer of the WSP is anoxic, allowing for anaerobic microorganisms to grow and digest the sludge. Organic carbon is ultimately converted to carbon dioxide and methane through combined anaerobic metabolic processes (e.g., fermentation and methanogenesis) and together, these help deplete organic material from wastewater. In the top layer of the pond, oxygen supports aerobic and facultative microorganisms through the mixing of water by wind and continued discharge of wastewater into the pond[43]. Oxygen also diffuses directly from the air and is produced by photosynthetic organisms, such as cyanobacteria and algae. With increased

sunlight during the treatment season, oxygen levels are highest during this time period, and therefore aerobic respiration decomposes much of the organic material in this layer[45,46,48]. In the most efficient systems, anaerobic and aerobic degradation are able to reduce up to 75% of the BOD[45]. Importantly, Arctic WSP systems are expected to be far less efficient in reducing BOD than temperate WSPs due to the influence of sub-zero temperatures on microbial activity[46].

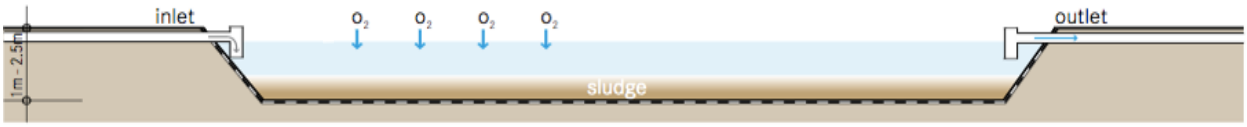


Figure 1.1: Schematic diagram of a facultative waste stabilization pond. Figure adapted from Tilley et al.[45]

1.2.2 Nutrient Removal in WSPs

Given the importance of removing macronutrients from wastewater to prevent eutrophication in receiving waters, important design objectives for WSPs include the capacity for nitrogen and phosphorus removal. Nitrogen is removed in a similar fashion in WSPs as in conventional WWTPs, using a combination of nitrification and denitrification. However, additional physical and biological mechanisms may also be important for WSP treatment. The most important WSP mechanism for nitrogen removal is ammonia stripping[43]. This process works by converting ammonium into its unionized form. Higher concentrations of unionized ammonia are achieved with increased temperature and pH[49]. Unionized ammonia is a gas that can thus be released from the system[49]. Increased pH is achieved in facultative ponds by photosynthesis. Indeed, in highly photosynthetic ponds the pH can increase to

values above 9.0[43]. Photosynthetic algae also assimilate ammonia and nitrate directly[43], which can further deplete wastewater of nitrogen in WSPs.

Phosphorus removal in WSPs occurs by two mechanisms: organic phosphorus is incorporated into algal and bacterial biomass and removed in effluent water, or phosphate precipitation under conditions with high pH[43]. The latter mechanism is more efficient if the pH is above 9.0, and is therefore more efficient in highly photosynthetic ponds, removing up to 80% of the phosphorus in very shallow ponds. However, in facultative ponds, the efficiency can substantially decrease to below 35% phosphorus removed[43].

1.3 Wastewater Treatment in the Canadian Arctic

1.3.1 Factors Affecting Construction of Wastewater Treatment Plants

Wastewater management in northern Canadian communities is limited due to the extreme climate of the Arctic. These communities also have relatively small populations and are isolated from the rest of Canada[50,51]. Because conventional WWTPs are not typically feasible to construct and operate in the Arctic, unique factors must be considered when establishing wastewater management facilities. Waste management solutions for the Arctic seek to minimize costs, prevent deterioration of water quality by eutrophication, and protect community members from potentially pathogenic microorganisms[2,50]. Most northern communities have a majority Inuit population, and many continue to follow traditions and regularly hunt and fish in the area. Many locals also collect their own drinking water directly from rivers which increases potential for exposure to pathogenic microorganisms[48].

To combat these difficulties, many northern Canadian communities use facultative waste stabilization ponds, also referred to as sewage lagoons or wastewater lagoons, for wastewater treatment[46,48]. Wastewater is discharged directly into the environment, and treatment occurs passively through natural biological processes[47]. However, as described previously, Arctic WSPs are far less efficient than temperate WSPs due to sub-zero temperatures and limited sunlight for most of the year. Unfortunately, few studies have explored the efficiency of Arctic WSPs, and these systems are still poorly understood. There is also limited information about the microorganisms found in Arctic WSPs.

1.3.2 Collection and Transport of Wastewater in Nunavut

Many communities in Nunavut use a truck transportation system where sewage is collected directly from buildings. Each building is equipped with two separate storage tanks, one for the delivery of drinking water and another for wastewater collection[52,53]. The trucks remove wastewater from each building several times a week, and discharge it into the designated WSP where it remains frozen for approximately nine months of the year due to extremely cold temperatures of the region[53]. As the temperature increases in the spring, ice begins to thaw and wastewater moves into downstream lakes and streams. The wastewater can either be continuously discharged into downstream bodies of water or retained and discharged once per year. The temperature in the region remains above zero for approximately two to three months. This time period is collectively referred to as the treatment season, during which water temperature increases and longer periods of daylight occur[48]. The warmer conditions and increased light are more favourable for microorganisms to perform aerobic and anaerobic

degradation of organic matter and therefore remove potentially harmful materials from the water system[45].

The Canadian government has established strict guidelines for effluent water quality from wastewater systems. The current national standard for wastewater effluent requires that water not exceed a BOD of 25 mg/L, an average concentration of total suspended solids of 25 mg/L, an average total residual chlorine of 0.02 mg/L, and an unionized ammonia concentration of 1.25 mg/L at 15°C[54]. Due to the factors precluding the construction of WWTPs, the Canadian government has exempted Nunavut from these regulations. Instead, Nunavut must ensure that wastewater does not exceed a maximum TSS of 180 mg/L, and BOD of 120 mg/L[46].

1.3.3 Microorganisms in Arctic Environments and Arctic Wastewater

Numerous studies focusing on the microbial diversity of Arctic soil and permafrost have been done over the past 15 years[55–61]. Many studies agree that *Actinobacteria* is the dominant phylum present within permafrost-affected soil samples. It is thought that *Actinobacteria* are prevalent in these environments due to the ability of *Actinobacteria* species to form cysts, catabolize a wide range of carbon sources, and grow at low temperatures[55,56].

Due to the extreme climate of Arctic environments, microorganisms inhabiting Arctic water bodies require adaptations that allow them to remain viable in the frigid temperatures[62]. Although the diversity of Arctic microbial communities is still not fully understood, interest in climate change has increased focus on the microbiology of Arctic environments. For example, Yergeau and colleagues explored Canadian high-Arctic seawater and found that the

seawater was dominated by operational taxonomic units (OTUs) affiliated with *Nitrosopumilus*, *Flavobacteriales*, and *Oceanospirillaceae*[63]. However, the dominant taxa in these samples made up less than 60% of the predicted microbial community at the sampled sites[63]. Cyanobacteria are also highly abundant in Arctic freshwater and play an important role as primary producers [64]. An Arctic freshwater lake in Alaska was dominated by members of the *Actinobacteria*, *Betaproteobacteria*, *Bacteroidetes*, and *Alphaproteobacteria*, which made up approximately 80% of the detected bacterial community[65].

Microorganisms in Arctic WSPs are not well understood. One study by Gonzalez-Martinez and coworkers identified several dominant OTUs in Arctic influent sewage. The most abundant OTUs were associated with the genera *Trichococcus, Methylorosula, Polaromonas,* and *Arcobacter*, as well as family members of *Leptotrichiaceae, Comamonadaceae, Alcaligenaceae*, and *Holophagaceae*[66]. Another study explored 20 wastewater treatment plants in Denmark and revealed *Tetrasphaera*, *Trichococcus, Candidatus* Microthrix*, Rhodoferax,* and *Rhodobacter* as the most abundant genera in these activated sludge samples[67].

Waterborne diseases are thought to be more prevalent in Arctic Canada in comparison to other parts of Canada[50]. Although WSPs are able to efficiently remove pathogens in warmer climate locations, relatively little is known about the ability of Arctic WSPs to remove pathogenic microorganisms[68]. Outbreaks of waterborne diseases are common in Arctic communities. Pathogens, such as enterohemorrhagic *Escherichia coli* O157:H7, *Salmonella* spp., *Campylobacter* spp., and *Helicobacter pylori*, have been detected in

community water supplies and stool samples[48]. This contamination may have been caused by the inability of current wastewater treatment systems to remove pathogenic microorganisms.

The limited number of studies done in Arctic environments has only just begun to elucidate the microbiomes and the ability of Arctic WSPs to remove wastewater-associated microorganisms. Understanding the microbiota of WSPs and their downstream water bodies is important for characterizing the impact of wastewater discharge on the environment and local community, as well as predicting the impacts of future infrastructure upgrades on Arctic WSP performance.

1.4 Microbial Community Analysis

Studying the microbial communities of wastewater can help with monitoring of wastewater management and determining the effectiveness of treatment systems. Identifying dominant microorganisms in wastewater, as well as determining the presence of potential pathogens, aides with the construction of safer wastewater management systems for Arctic communities. With the improvements in DNA sequencing technologies over the past few decades, studying microbial diversity has become easier[69]. Various methods and pipelines have been developed to simplify the analysis of microbial communities. In particular, amplicon and metagenome sequencing are two very useful approaches for profiling microbial communities[70,71].

1.4.1 16S rRNA Gene Amplicon Sequencing

Amplification and sequencing of the 16S rRNA genes is commonplace for microbial community characterization because of its universal distribution among prokaryotes and relatively slow evolution of the gene sequence[72]. Although much of the 16S rRNA gene is

highly conserved, there are nine variable regions that can be used to distinguish different species of bacteria and archaea[73]. Although the entire 16S rRNA gene can be sequenced, it is also reasonable to amplify a smaller portion of the gene using PCR to produce amplicons that can be subsequently sequenced[70]. This is a targeted approach for profiling microorganisms in communities from many samples, without sequencing entire genomes of constituent members[74].

Post-sequencing analyses of amplicons involves comparing sequences to a database of 16S rRNA gene sequences, such as the SILVA database[75], and assigning taxonomy based on the best database hit. A frequently used analytical pipeline for this task is Quantitative Insights into Microbial Ecology version 2 (QIIME2), combined with a denoising and error correction tool called DADA2[76,77]. Using these tools, it is possible to distinguish 16S rRNA gene amplicons into amplicon sequence variants (ASVs), which are unique sequences that differ by as little as one nucleotide[78]. The relative abundance of microorganisms can then be compared between different samples and locations by multivariate statistical approaches to identify significant differences in community composition.

1.4.2 Metagenome Sequencing

Metagenome sequencing is another common method for studying microbial communities. A metagenome is produced by sequencing all of the genetic material from a single sample[79]. In this approach, metagenomic DNA is sequenced, followed by analysis, and/or assembled and binned into metagenome-assembled genomes (MAGs). Several computational tools have

been developed for metagenome sequence analysis, such as ATLAS[80], which performs the assembly, annotation, and binning of metagenome sequences[80].

The ATLAS pipeline first begins with quality control using various tools in the BBTools suite to remove adapters, filter and trim reads, and perform error correction[81]. In a typical pipeline, *de novo* assembly is then done using either MEGAHIT or metaSPAdes, followed by binning with metabat2, maxbin2, or concoct[82–87]. Binning involves grouping contigs that belong together to make up the genomes of single populations[88]. The DASTool software is also available as a final binner that compares outputs from each of the binning approaches used to produce high quality bins[89]. Once final bins are established, dRep is used to select the best bin for each MAG[90]. Functional annotation is then performed using Prodigal for identifying open reading frames (ORFs), which are translated and mapped to the eggNOG catalogue using DIAMOND to identify genes[91–93]. Finally, taxonomy is inferred using BAT to map genes to GenBank[94,95].

1.5 Thesis Objectives

Currently there is a large gap in knowledge regarding the microbial communities of Arctic wastewater, specifically microorganisms associated with Arctic WSPs. This study aims to identify the major taxa in a representative Arctic WSP systems in Baker Lake, Nunavut to establish a baseline characterization and test the hypothesis that wastewater-associated microorganisms are enriched in wastewater and receiving waters relative to reference lakes, and determine whether these organisms are detected downstream of the WSP. This research

will also identify and relatively quantify antibiotic resistance genes within metagenomes and observe the prevalence of ARGs in the WSPs in comparison to downstream sites.

Chapter 2

Methods

2.1 Study Locations

This study sampled the microbial communities of wastewater lagoons and receiving waters in three traditional indigenous territories in Nunavut (Figure 2.1): Baker Lake, Cambridge Bay, and Kugluktuk. The main focus was on Baker Lake with comparisons to the WSPs in Cambridge Bay and Kugluktuk. The territory of Nunavut has a population of about 37,000 people in 25 communities, with populations ranging from about 150 to 7,500 people per community. Baker Lake, Cambridge Bay and Kugluktuk have populations of approximately 2,000, 1,700, and 1,600 people, respectively[96].

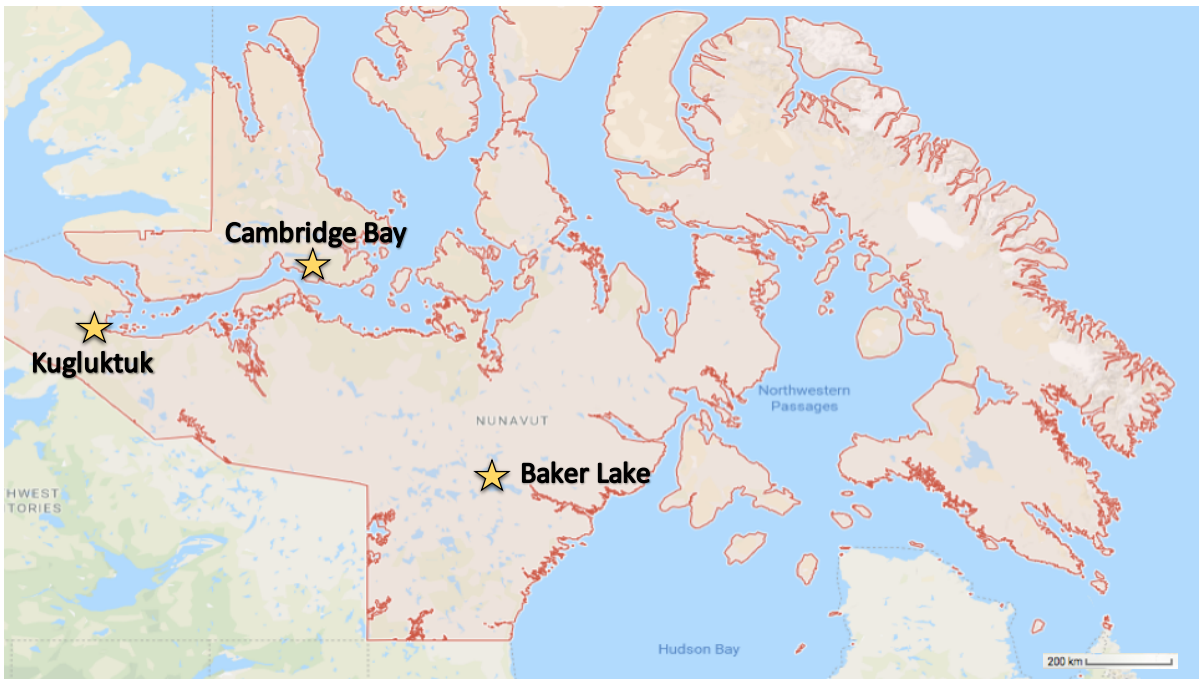


Figure 2.1: Map of study locations in Nunavut, Canada. This map was adapted from Google Maps[97].

2.1.1 Baker Lake, Nunavut

Baker Lake is the only inland community in Nunavut, located west of Hudson's Bay in the Kivalliq region. In Baker Lake, trucks collect wastewater from each building every two to three days and discharge it directly into the wastewater lagoon, located around one kilometre north of the community. The WSP freezes near the end of September and remains frozen until June. Over the nine months when temperatures are sub-zero, the wastewater accumulates in an ice block that spans a field leading down into Lagoon Lake. During the spring thaw in mid-June, water flow is relatively high and first flows into Lagoon Lake. From there, the water flows into Finger Lake, followed by Airplane Lake, and finally discharges into Baker Lake. There are two lakes upstream of Lagoon lake and one lake upstream of Airplane Lake that were used as reference sites that appear to be unaffected by the WSP. Sample collection was carried out at each of the major lakes downstream of the WSP in Baker Lake, Nunavut, as well as three reference lakes (Figure 2.2). Sampling occurred at two times points from July 13th-16th, 2018 and July 22nd-24th, 2018. Samples were labelled according to time points; samples containing the number "1" before site name belong to the group of samples collected July 13th-16, 2018, and samples containing the number "2" before site name correspond to the July 22nd-24th, 2018 sampling time point. Duplicates were indicated by an "A" or "B" following the site name in labels. For example, the first duplicate of the WSP at the first time point is labelled as 1WWLA.

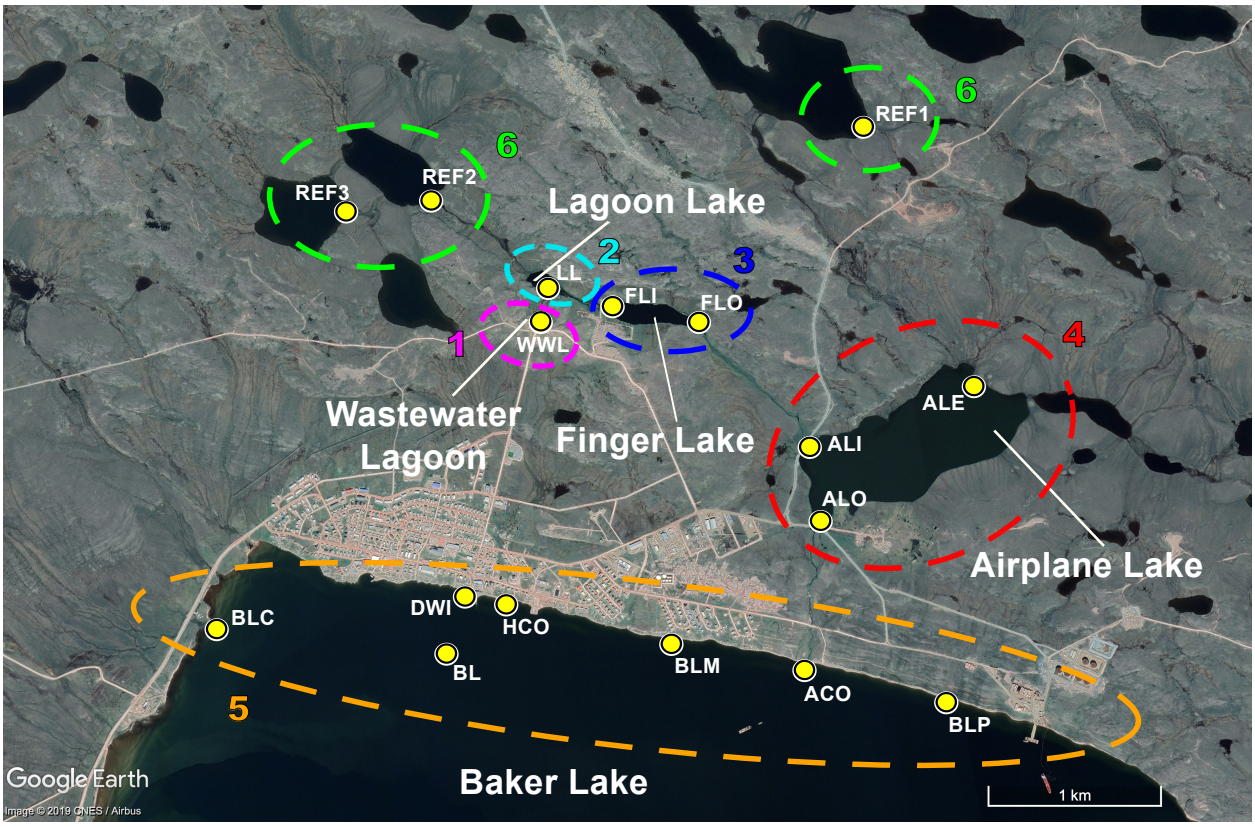


Figure 2.2: Map of Baker Lake, Nunavut. Sampling sites are identified by yellow pins and labeled by sample name. Samples have been grouped by lakes as indicated by the coloured circles. Magenta is the Baker Lake wastewater lagoon, turquoise is Lagoon Lake, blue is Finger Lake, red is Airplane Lake, orange is Baker Lake, and green are the references. This map was adapted from Google Earth[98].

2.1.2 Cambridge Bay, Nunavut

Cambridge Bay is located on Victoria Island in the Kitikmeot region of Nunavut. Throughout the year, wastewater is dumped into a WSP where primary treatment occurs. The wastewater is discharged once a year during the summer, where it passes through a natural tundra wetland that eventually leads into Cambridge Bay. The WSP is located approximately one kilometer north-east of the community. Samples were also taken from the WSP at Cambridge Bay for comparison at two time points on July 4th, 2018 and July 25th, 2018. WSP samples from Cambridge Bay were labelled as CBL for the site name. Similar to the Baker Lake

samples, July 4th, 2018 samples contain a "1" before the site name and July 25th, 2018 samples contain a "2" instead. Duplicates are distinguished by an "A" or "B" following the site name.

2.1.3 Kugluktuk, Nunavut

Kugluktuk is located at the mouth of the Coppermine River in the Kitikmeot region of Nunavut. The community dumps their wastewater into a single-cell, lined sewage lagoon. Similar to Cambridge Bay, the water is discharged once a year through a natural tundra wetland from the northwest corner of the lagoon. After passing through the wetland, the water flows into Coronation Gulf as the receiving water body. Samples were taken from the WSP at Kuglukluk at a single time point on August 16th, 2018. WSP samples from Kugluktuk were labelled as KWWL for the site name. Because only one time point was used, both sample labels contain a "1" before the site, and duplicates are distinguished with an "A" or "B" following the site name.

2.2 Sample Collection and Environmental Data Collection

Water from each sampling site was collected 4-5 m from the shoreline using a telescoping swing sampler and attached bottle to avoid disturbance of sediments. Samples were collected in duplicate from approximately 5-10 cm below the water's surface. The bottle was rinsed multiple times at each sampling site before collecting water. The two sites ALE and BL at Baker Lake were only taken at a single time point due to limited access. For these sites, a boat was taken offshore and anchored. The ALE site samples were taken approximately 20 m offshore, whereas BL was taken approximately 300 m offshore. Water samples for these two

sites were collected using a sample bottle only at about 5-10 cm below the water's surface. Water was filtered through a sterile 0.22 μm Sterivex-GV Pressure Filter Unit (EMD Millipore Corporation, MA, USA) using a 60-ml Luer-Lok Tip Syringe (Becton Dickinson, NJ, USA). Filtration was done until filter pores were plugged or a maximum of 500 ml was filtered. After filtration, residual water was expelled from the filter, leaving minimal amounts of water within the filter unit, then filter units were stored in Ziplock bags surrounded by ice packs. Samples were kept in a cooler and surrounded by frozen ice packs until delivered to the lab. Once received by the lab, samples were stored at -20°C until analyzed.

Environmental data was also collected at each of the sampling sites. Latitude and Longitude were determined using the Gaia GPS app (TrailBehind Inc.). The PC100 Model pH/Conductivity Meter (Cole-Parmer, Montreal, Canada), was used to measure pH, conductivity, total dissolved solids, and salinity. Temperature was also measured on this meter using the pH probe. Dissolved oxygen was measured using the Traceable Dissolved Oxygen Pocket Tester with Calibration (Cole Parmer).

2.3 DNA Extractions

A summary of methods used to process environmental DNA samples in this study is shown in Figure 2.3.

Filter unit housings were opened and filters removed aseptically. Each filter was cut in half, with one half stored at -20°C and the other half used for DNA extraction. Each half used for DNA extraction was cut into smaller pieces to increase extraction efficiency. Although all DNA extractions were done using the PowerSoil DNA Isolation Kit (MO BIO, CA, USA)

following the manufacturer's instructions, PowerBead tubes were heated at 70°C for 10 minutes prior to beadbeating with a FastPrep-24 (MP Biomedicals, CA, USA) at 5.5 m/s for 45 seconds. Extraction kit negative controls were processed alongside samples and were "KC" followed by a number. Genomic DNA was quantified using a NanoDrop 2000 (Thermo Scientific, MA, USA) and Qubit dsDNA High Sensitivity Assay Kit (Thermo Scientific). All genomic DNA samples were stored at -20°C in a sealed 96-well plate until further processing.

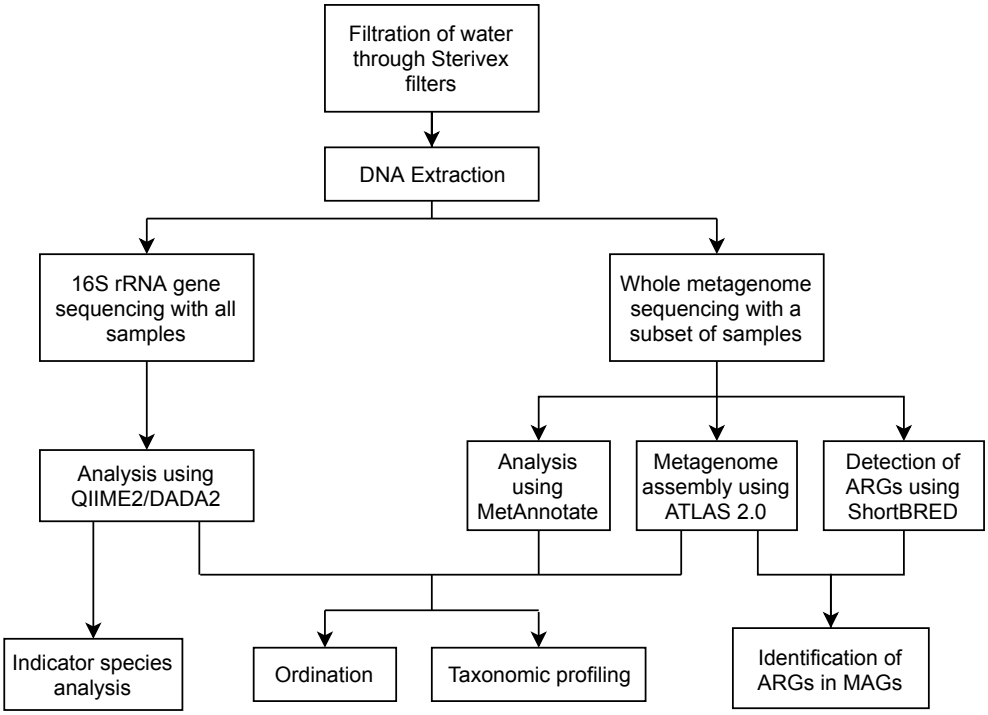


Figure 2.3: Flowchart of DNA processing methods

2.4 16S rRNA Gene Amplicon Library Preparation and Sequencing

Using a modification of a previously published amplification and sequencing method[99], the V4-V5 regions of the 16S rRNA gene were amplified from the genomic DNA extracts using primers 515F-Y/926R[100,101]. Each 25 μl PCR contained 2.5 μl of 10x ThermoPol Reaction Buffer (New England Biolabs, MA, USA), 1.5 μl of 10 mg/ml bovine serum albumin, 0.2 μM 515F primer, 0.2 μM 926R primer, 0.2 mM dNTPs, 0.625 units of *Taq* DNA polymerase (New England Biolabs), 2-50 ng of genomic DNA as template in HyPure Molecular Biology Grade Water (GE Healthcare Life Sciences, UT, USA). Reactions were conducted in triplicate on a T100 Thermal Cycler (BioRad, CA, USA) or DNAEngine Peltier Thermal Cycler (BioRad) and conditions used for PCR were 95°C for 3 minutes, followed by 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, and 68°C for 1 minute, finishing with 68°C for 7 minutes. Samples were randomly assigned barcodes to avoid bias.

Following PCR, triplicate PCR products were run separately on a 1% agarose gel then pooled. Equal amounts of DNA from each sample were combined prior to sequencing. Non-template negative controls were added as 5 μl volumes, and positive controls were added at half the amount of the DNA samples. Negative controls were labelled as "NTC". Positive controls contained equal amounts of *Thermus aquaticus* and *Aliivibrio fischeri* and were labelled as "PC". Pooled samples were then run on another 1% agarose gel and bands corresponding to the 16S rRNA gene amplicons were extracted and purified using the Wizard SV Gel and PCR Clean-Up System (Promega, WI, USA). The eluted DNA library was stored at -20°C until further processing.

The library was prepared for sequencing on a MiSeq (Illumina, CA, USA) using the provided HT1 buffer to dilute the library to about 6 pM. The PhiX library was also added such that it made up 15% of the final library. Paired-end sequencing (2 x 250 bases) was performed on a MiSeq, generating ~18 million paired-end reads. Sequences are available in the Sequence Read Archive (SRA) at NCBI under accession number PRJNA600216 (https://www.ncbi.nlm.nih.gov/sra/PRJNA600216).

2.5 QIIME2 Analysis

Paired-end reads were demultiplexed using the MiSeq Reporter software. Subsequent FASTQ files were imported into QIIME2, and adapter and primer sequences were removed using cutadapt[102]. Reads were then trimmed, denoised, dereplicated, and merged using DADA2 within QIIME2. Reads were trimmed based on a minimum quality score of 25 for the nucleotide position. This produced a feature table containing amplicon sequence variants (ASVs). The April 2018 SILVA release 132, 97% taxonomy classification for the 16S rRNA gene was used to train the Naïve Bayes classifier, and taxonomy was assigned to sequence variants within the feature table. The feature table was collapsed to the phylum level, and the uncollapsed ASV table was rarefied to a sampling depth based on the sample with the lowest number of reads and exported to visualize microbial community diversity in bar plots and bubble plots.

A beta diversity analysis was conducted with scripts contained within the QIIME2 pipeline. Phylogenetic trees were created using the ASV table and beta diversity analysis was performed using the same sampling depth as for rarefaction. Principal coordinates analysis

(PCoA) was performed using the weighted and unweighted UniFrac distance metrics. The PCoA results were visualized using a custom R script.

2.6 Indicator Species Analysis

Rarefied ASV tables exported from QIIME2 were used for an indicator species analysis (ISA)[103], using only Baker Lake WSP samples, Lagoon Lake samples, and reference lake samples. The Baker Lake WSP and Lagoon Lake samples were assigned to one group, and reference lake samples were assigned to another group. Indicator species are species, or in this study, ASVs, that are associated with, or indicative of, an assigned group of samples[104]. Indicator values (IndVals) for each ASV within these samples were calculated using the multipatt function from the indicspecies R package[105]; 1000 permutations were tested to determine statistical significance. The ASVs with an IndVal of at least 0.9 and p value < 0.05 were considered as indicator species and used for subsequent analysis and data visualizations.

2.7 Metagenome Sequencing

Genomic DNA was aliquoted for a subset of samples. Each sample location selected was sequenced in duplicate for all time points available. Sampling sites selected from Baker Lake included both time points and duplicates from WWL, LL, FLO, ALO, DWI, REF1, REF2. All available samples for the WSPs in Cambridge Bay and Kugluktuk were also sequenced. This gave a total of 34 samples that were aliquoted and sent to the Farncombe Metagenomics Facility at McMaster University for library preparation, quality control, and whole metagenome shotgun sequencing. The NEBNext Ultra II DNA Library Prep Kit (New

England Biolabs) was used to prepare the library for sequencing and paired-end sequencing (2 x 250 bases) was performed on two lanes of a HiSeq 1500 (Illumina) generating a total of ~500 million paired-end reads. Sequences are available in the Sequence Read Archive (SRA) at NCBI under accession number PRJNA600216 (https://www.ncbi.nlm.nih.gov/sra/PRJNA600216).

2.8 MetAnnotate Analysis

Raw metagenomic reads were analyzed using a development version of MetAnnotate[106] (available at https://github.com/MetAnnotate/MetAnnotate/tree/develop, commit ID: 6e92c0e). A hidden Markov model (HMM) for the gene for the beta subunit of bacterial RNA polymerase (*rpoB*) was downloaded from the TIGRFAMs database[107] and was used to profile the metagenomes by assigning taxonomy to reads identified as hits to the HMM. Forward and reverse reads were analyzed separately using HMM e values ≤ 1e-3 and RefSeq HMM e values ≤ 1e-6. Default values were used for all other parameters. The *rpoB* output table was then analyzed using a custom R script (available at https://github.com/jmtsuji/metannotate-analysis) to produce taxonomy abundance tables for comparison to 16S rRNA gene taxonomic assignments. Bar plots and bubble plots showing taxonomic profiles were produced with a custom R script.

2.9 Metagenome Assembly and Binning

Metagenomic reads were processed using the ATLAS release 2.0.6, which performs quality control, assembly, annotation, binning, and read mapping[80]. Quality control was performed by the BBtools suite[81] tools to eliminate adapters, remove PCR duplicates, then trim and

filter reads based on quality scores and length of reads. Paired-end reads were assembled using the metaSPAdes[84] assembly tool. Prodigal[91] was used to predict ORFs in contigs and translated gene products were mapped to the eggNOG[92] catalogue using DIAMOND[93]. Taxonomy was assigned using BAT[94], mapping genes to the Genbank protein database[95]. Contigs were then binned using metabat2[85] and maxbin2[86], followed by DASTool[89] as a final binner to dereplicate, aggregate, and score the bins and create high-quality metagenome-assembled genomes (MAGs). Because some of the same genomes were identified in multiple samples, resulting in multiple bins for the same MAG, dRep[90] was also used to select the best bin for each MAG. The MAG tables were then generated using a custom R script (available at https://github.com/jmtsuji/atlas2-helpers), and bar plots and bubble plots were produced with a separate custom R script.

2.10 Mantel Test and Hierarchical Clustering

The ASV tables, *rpoB* tables, and MAG tables were compared to one another with a Mantel test to assess correlations between the three datasets. Bray-Curtis dissimilarity matrices were produced for each of the three tables in R using the vegdist function in the vegan package[108]. The Mantel test was then performed on each pair of matrices using the mantel.rtest function from the ade4 R package[109], with 1000 permutations to verify statistical significance.

Hierarchichal clustering was then performed on each distance matrix using the hclust function in the stats R package[110]. The group average agglomeration method was used, which is equivalent to the unweighted pair-group method (UPGMA) whereby the distance between groups is calculated by averaging all of the distances for all pairs of individuals, assigning

one for each group[111]. Hierarchical clustering results for each dataset were plotted in dendrograms.

2.11 Identification of ARGs using ShortBRED

Forward and reverse reads from each set of paired-end metagenomic read files were analyzed separately to identify and quantify antibiotic resistance genes. This was done using the pre-computed 2017 Antibiotic Resistance Factors marker collection from the ShortBRED[36] documentation pages (available at http://huttenhower.sph.harvard.edu/shortbred). This marker set is based on the CARD. The shortbred-quantify.py function from ShortBRED quantifies the relative abundance of each of the markers in the metagenomic reads. This function also normalizes the counts based on marker length and sequencing depth, giving a normalized count value in reads per kilobase of reference sequence per million sample reads (RPKM). Gene families were grouped manually into larger groups based on CARD information (Appendix 1). The tables of RPKM values were modified in a custom R script to factor in the gene family groupings and plotted in a heatmap with dendrograms produced by hierarchical clustering using the "ward.D2" agglomeration method.

Groupings that appeared to be enriched in wastewater sites were examined in closer detail. The most abundant ARGs in the wastewater samples were selected as genes of interest, and reads from those samples were extracted for further investigation. The extracted reads were mapped to their contigs based on the metagenome assembles done previously. These contigs were then mapped to their bins, and the bins were then mapped to the MAGs to identify which MAGs contained the ARGs that were enriched in wastewater samples.

Chapter 3

Results and Discussion

Arctic wastewater systems are currently not well characterized, and the microbial communities of these systems are not well understood. This study aimed to develop a taxonomic profile of microorganisms within three Arctic WSPs and receiving waters in Baker Lake, Cambridge Bay, and Kugluktuk in the territory of Nunavut, with a focus on Baker Lake. To determine the major taxa present within Arctic wastewaters and verify the presence or absence of these groups in receiving waters, 16S rRNA gene amplicons and metagenomes were sequenced and taxonomy was assigned. Antibiotic resistance genes were also identified and quantified from metagenome sequences. The frequent misuse of antibiotics has created a growing problem of ARB, and this study aims to assess the enrichment of ARGs in wastewater, and determine if these ARGs are also accumulating in receiving waters. Overall, the research presented here will give a better understanding of the efficiency of WSP systems in Arctic conditions, and give insight into future improvements that can be made to the infrastructure of Baker Lake's wastewater management system.

3.1 Wastewater Treatment System of Baker Lake, Nunavut

High-throughput 16S rRNA gene sequencing generated ~18 million paired-end reads, resulting in ~37,000 ASVs from all samples, including those from Cambridge Bay and Kugkluktuk, Nunavut. Read counts varied among samples, with a minimum read count of ~19,000 reads and a maximum read count of ~310,000 (Table 3.1). Almost all negative controls, with the exception of one extraction kit control (KC3), had lower read counts and

possessed ASV profiles that were distinct from samples. Positive controls contained only *Thermus aquaticus* and *Aliivibrio fischeri*, as expected. All ASVs detected were used for ordinations using weighted and unweighted UniFrac distance metrics. Both weighted and unweighted UniFrac distances factor in phylogenetic distance of ASVs present in each sample. Weighted UniFrac distances also consider the abundance of these ASVs instead of just presence or absence. In later analyses, only ASVs with >0.1% abundance in WSP samples were included, excluding sample 2WWLB for filtering steps because it was considered an outlier based on relative abundances of phyla in that samples (Figure 3.2). This reduced the number of ASVs to 71 that were included in the analysis for the Baker Lake WSP. However, other lakes in Baker Lake contained more ASVs at >0.1% abundance that were not found in the WSP, and these ASVs were not studied in detail.

Table 3.1: 16S rRNA gene sequence read counts for all samples

Sample	Read Count	Sample	Read Count	Sample	Read Count
1WWLA	111,491	2WWLA	152,433	1KWWLA	60,033
1WWLB	145,748	2WWLB	162,078	1KWWLB	59,544
1LLA	87,581	2LLA	135,617	KC1	977
1LLB	89,989	2LLB	155,740	KC2	1,048
1FLIA	74,401	2FLIA	116,295	KC3	13,402
1FLIB	110,023	2FLIB	71,642	KC4	491
1FLOA	68,450	2FLOA	37,586	KC5	3,490
1FLOB	190,174	2FLOB	98,601	PCa	112,403
1ALIA	40,696	2ALIA	116,772	PCb	124,341
1ALIB	51,116	2ALIB	104,529	PCc	167,680
1ALOA	60,715	2ALOA	170,630	PCa2	170,665
1ALOB	135,713	2ALOB	117,446	PCb2	168,867
1ALEA	222,454	2ACOA	243,037	PCc2	260,504
1ALEB	108,408	2ACOB	98,806	NTC1a	495
1ACOA	309,507	2BLPA	38,480	NTC1b	254
1ACOB	128,386	2BLPB	176,123	NTC1c	82
1BLPA	95,702	2BLMA	58,054	NTC1a2	2,720
1BLPB	135,337	2BLMB	130,378	NTC1b2	1,094
1BLMA	195,020	2HCOA	33,895	NTC1c2	3,270
1BLMB	239,541	2HCOB	77,445	NTC2a2	793
1HCOA	219,146	2DWIA	36,393	NTC2b2	1,254
1HCOB	84,125	2DWIB	154,192	NTC2c2	3,805
1BLA	81,945	2BLCA	77,511	NTC3	1,020
1BLB	57,651	2BLCB	148,365	NTC3a2	2,288
1DWIA	94,338	2REF1A	108,036	NTC3b2	826
1DWIB	68,186	2REF1B	127,099	NTC3c2	592
1BLCA	19,131	2REF2A	109,562		
1BLCB	47,988	2REF2B	102,131		
1REF1A	72,729	2REF3A	85,693		
1REF1B	126,370	2REF3B	89,265		
1REF2A	100,937	1CBLA	164,233		
1REF2B	176,910	1CBLB	259,568		
1REF3A	47,514	2CBLA	132,595		
1REF3B	146,462	2CBLB	83,738		

3.1.1 Temporal Variability

Microbial community profiling of the wastewater treatment system in Baker Lake, Nunavut, as well as the receiving waters, identified changes in 16S rRNA gene profiles that were influenced by temporal and spatial factors. Specifically, ordinations based on both weighted and unweighted UniFrac metrics (Figure 3.1) demonstrated that samples from the first time point (July 13-16, 2018) separated from samples from the second time point (July 22-24, 2018), with highly similar profiles for replicate samples. Greater variation in the data was explained by the first two ordination axes of the weighted distance metric (50.4%), compared to the unweighted distance metric (24.0%). Samples from Airplane Lake and Baker Lake substantially overlapped in ordination space (Figure 3.1). In general, there are many ASVs in common between these sets of samples. Particularly for the first time point, Airplane Lake and Baker Lake samples have very similar taxonomic profiles (Figure 3.4). This result was expected because incoming water from sources other than the WSP dilute the wastewater such that dominant wastewater microorganisms become less prevalent when mixed with upstream freshwater microbial communities.

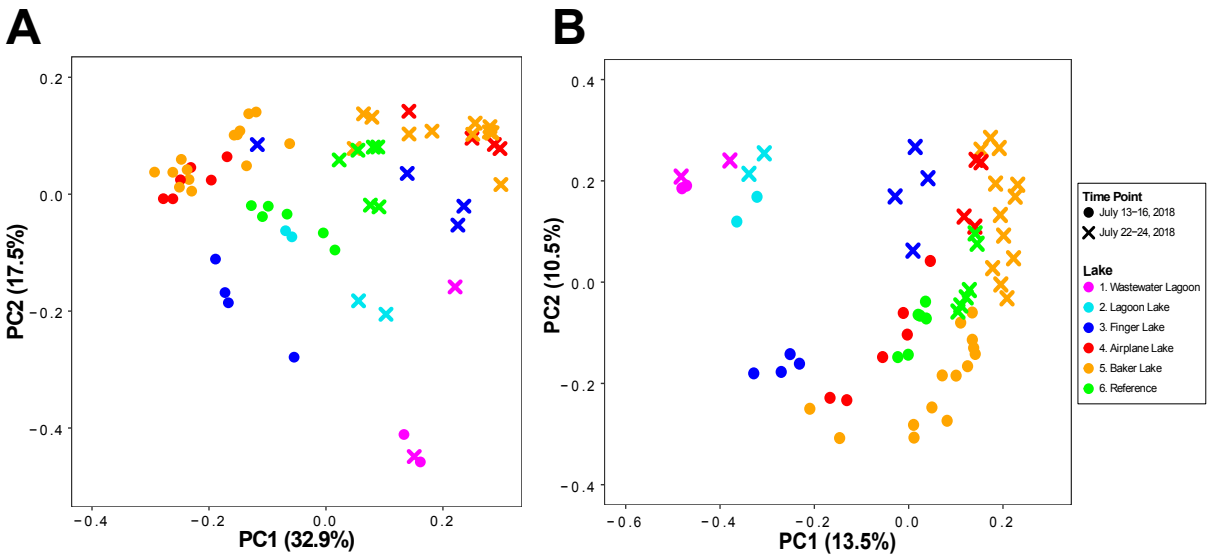


Figure 3.1: Principal coordinates analysis (PCoA) using (A) weighted and (B) unweighted UniFrac distances based on amplicon sequence variants (ASVs) from 16S rRNA gene amplicons for sites in Baker Lake, Nunavut.

A temporal shift in microbial community composition for most samples collected from the Baker Lake site is also reflected at the phylum level. Over the 10-day period between sampling time points, a shift in abundance of each phylum occurred for the majority of sample sites. Most sites shifted from being dominated by *Bacteroidetes* to containing higher levels of *Actinobacteria*, *Cyanobacteria*, and *Verrucomicrobia* (Figure 3.2). The most obvious shift occurred for samples collected from Finger Lake, Airplane Lake, and Baker Lake. In Airplane Lake and Baker Lake samples, members of the *Bacteroidetes* represented ~75% of the detected microbial communities in samples from the first time point, yet decreased to as low as 4% relative abundance at some sampling sites. For samples collected at several of these sites, *Cyanobacteria*, *Patescibacteria*, *Planctomycetes*, and *Verrucomicrobia* were not detected, or were found at very low relative abundance for the first time point. The microbial communities profiled for Finger Lake samples demonstrated a

shift from being almost completely composed of *Bacteroidetes* and *Proteobacteria*, to containing far more *Actinobacteria*. This is particularly noticeable in the samples 2FLIA and 2FLIB, which contained over 50% *Actinobacteria* despite the phylum not being detected at >1% relative abundance at the same site for the first time point (Figure 3.2). The Baker Lake WSP and Lagoon Lake samples were consistent temporally, with the exception of 2WWLB, which contained far more *Firmicutes* than any other Baker Lake samples.

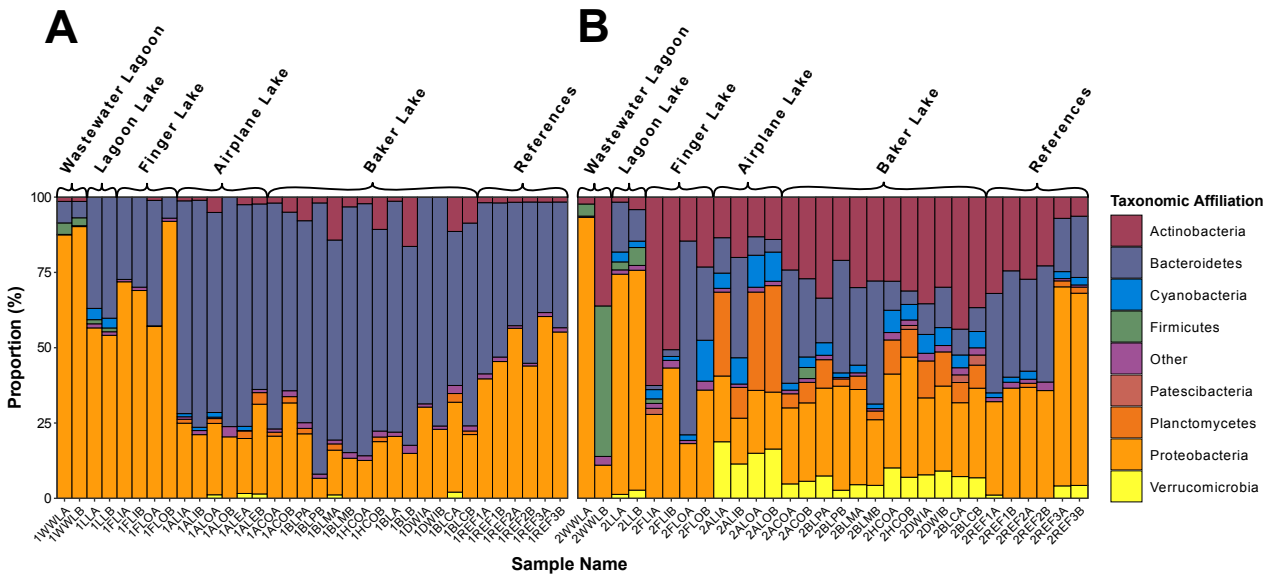
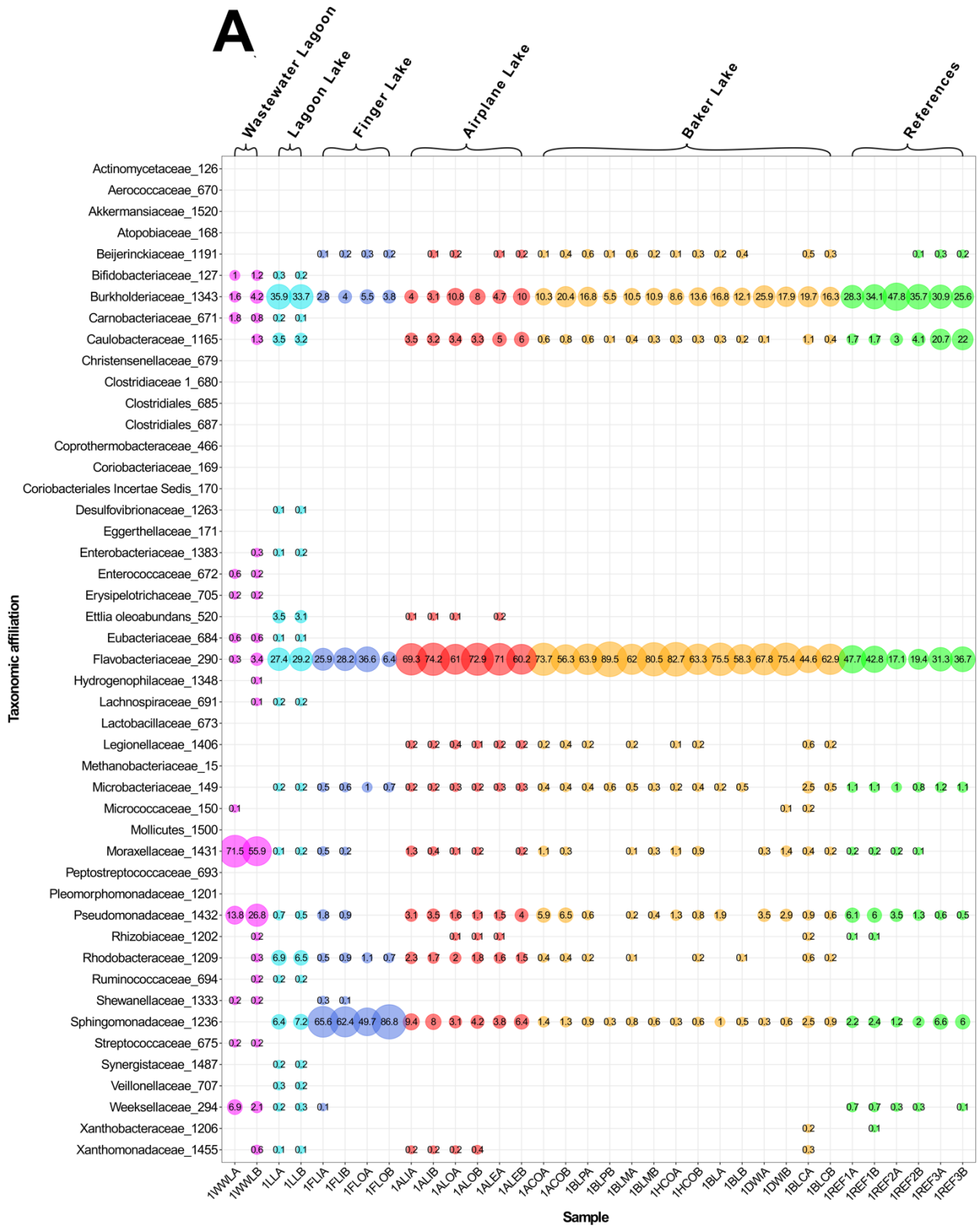


Figure 3.2: Relative abundance of phyla present at each site in Baker Lake, Nunavut. (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. Sample names correspond to sites indicated on the map in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. Corresponding lake names are indicated above bars.

Many of the highly abundant ASVs that affiliated with the *Bacteroidetes* also belonged to the *Flavobacteriaceae* family (Figure 3.3), of which multiple species, particularly those in the genus *Flavobacterium*, have been characterized as being

psychrophilic or psychrotolerant organisms[112,113]. There were two ASVs associated with *Flavobacterium* and *Janthinobacterium* that were more abundant at the first time point than the second time point, particularly for Airplane Lake, Baker Lake, and the reference lake samples. In particular, ASVs affiliated with *Janthinobacterium* were highly abundant in Baker Lake and reference lakes, but decreased in abundance for second time point samples (Figure 3.4). Strains of *Janthinobacterium* (within the Flavobacteriaceae family) have also been identified as psychrotolerant[114,115]. These psychrotolerant bacteria may have been present during the first sampling time point because water temperatures were still relatively cold. However, water temperatures gradually increased over the 10 days between sampling times. The most substantial temperature change occurred for the Baker Lake HCO site, increasing by 5°C over the 10-day period (Table A1). This temperature change may have impacted the microbial community because >60% of the community at the HCO site was dominated by *Flavobacteriaceae*, but this decreased to as low as 2% abundance for samples collected at the second time point (Figure 3.3). The increase in temperature may have induced the growth of other microorganisms, resulting in many low abundance organisms appearing at the second time point that were not detected earlier.


A
Wastewater Lagoon
Lagoon Lake
Finger Lake
Airplane Lake
Baker Lake
References
Taxonomic affiliation
Actinomycetaceae_126
Aerococcaceae_670
Akkermansiaceae_1520
Atopobiaceae_168
Beijerinckiaceae_1191
Bifidobacteriaceae_127
Burkholderiaceae_1343
Carnobacteriaceae_671
Caulobacteraceae_1165
Christensenellaceae_679
Clostridiaceae 1_680
Clostridiales_685
Clostridiales_687
Coprothermobacteraceae_466
Coriobacteriaceae_169
Coriobacteriales Incertae Sedis_170
Desulfovibrionaceae_1263
Eggerthellaceae_171
Enterobacteriaceae_1383
Enterococcaceae_672
Erysipelotrichaceae_705
Ettlia oleoabundans_520
Eubacteriaceae_684
Flavobacteriaceae_290
Hydrogenophilaceae_1348
Lachnospiraceae_691
Lactobacillaceae_673
Legionellaceae_1406
Methanobacteriaceae_15
Microbacteriaceae_149
Micrococcaceae_150
Mollicutes_1500
Moraxellaceae_1431
Peptostreptococcaceae_693
Pleomorphomonadaceae_1201
Pseudomonadaceae_1432
Rhizobiaceae_1202
Rhodobacteraceae_1209
Ruminococcaceae_694
Shewanellaceae_1333
Sphingomonadaceae_1236
Streptococcaceae_675
Synergistaceae_1487
Veillonellaceae_707
Weeksellaceae_294
Xanthobacteraceae_1206
Xanthomonadaceae_1455
1WWLA
1WWLB
1LLA
1LLB
1FLIA
1FLIB
1FLOA
1FLOB
1ALIA
1ALIB
1ALOA
1ALOB
1ALEA
1ALEB
1ACOA
1ACOB
1BLPA
1BLPB
1BLMA
1BLMB
1HCOA
1HCOB
1BLA
1BLB
1DWIA
1DWIB
1BLCA
1BLCB
1REF1A
1REF1B
1REF2A
1REF2B
1REF3A
1REF3B
Sample

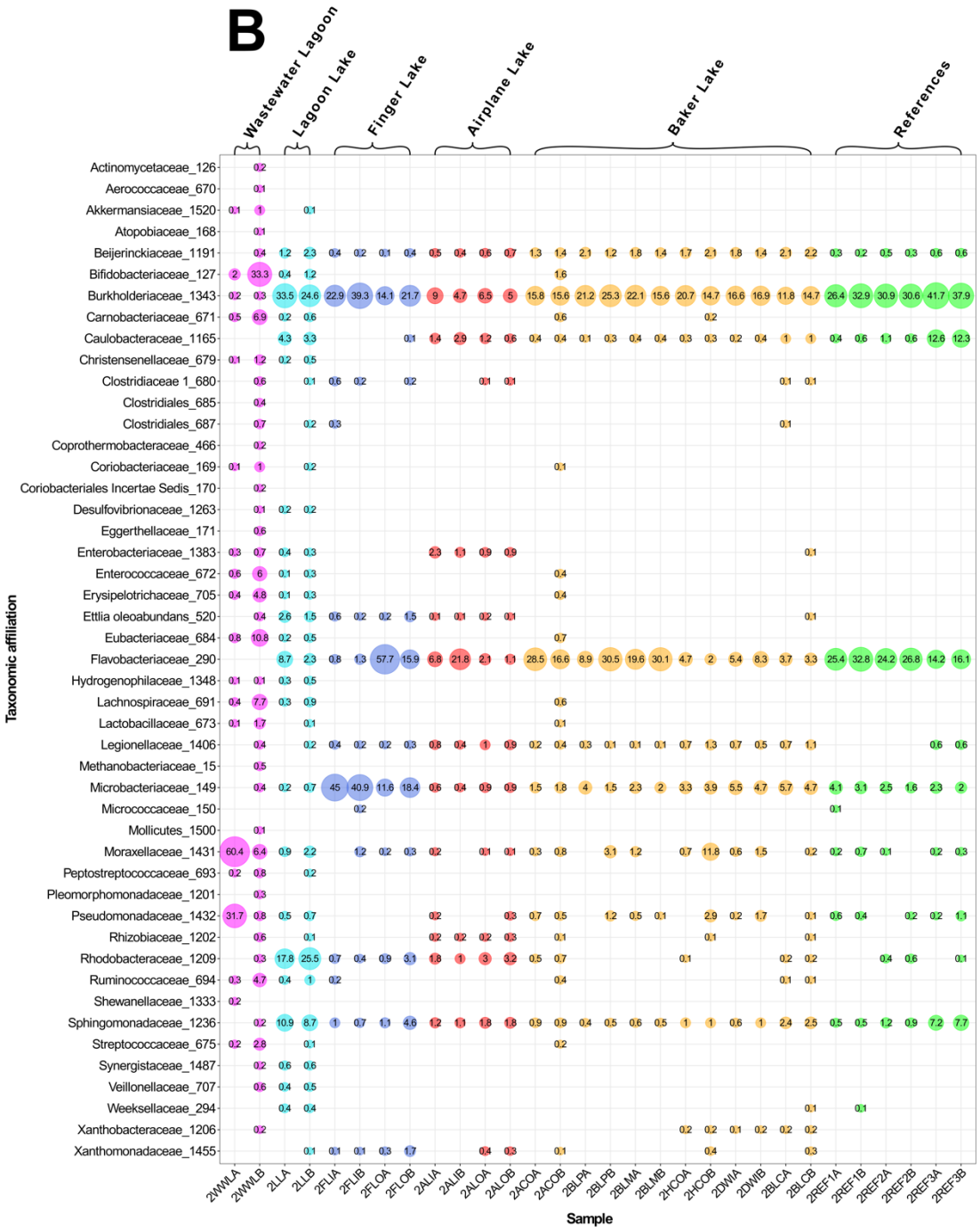


Figure 3.3: Relative abundance of families present at >0.1% abundance at each site in Baker Lake, Nunavut. (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. Sample names correspond to sites indicated on the map in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. Corresponding lake names are indicated above columns.

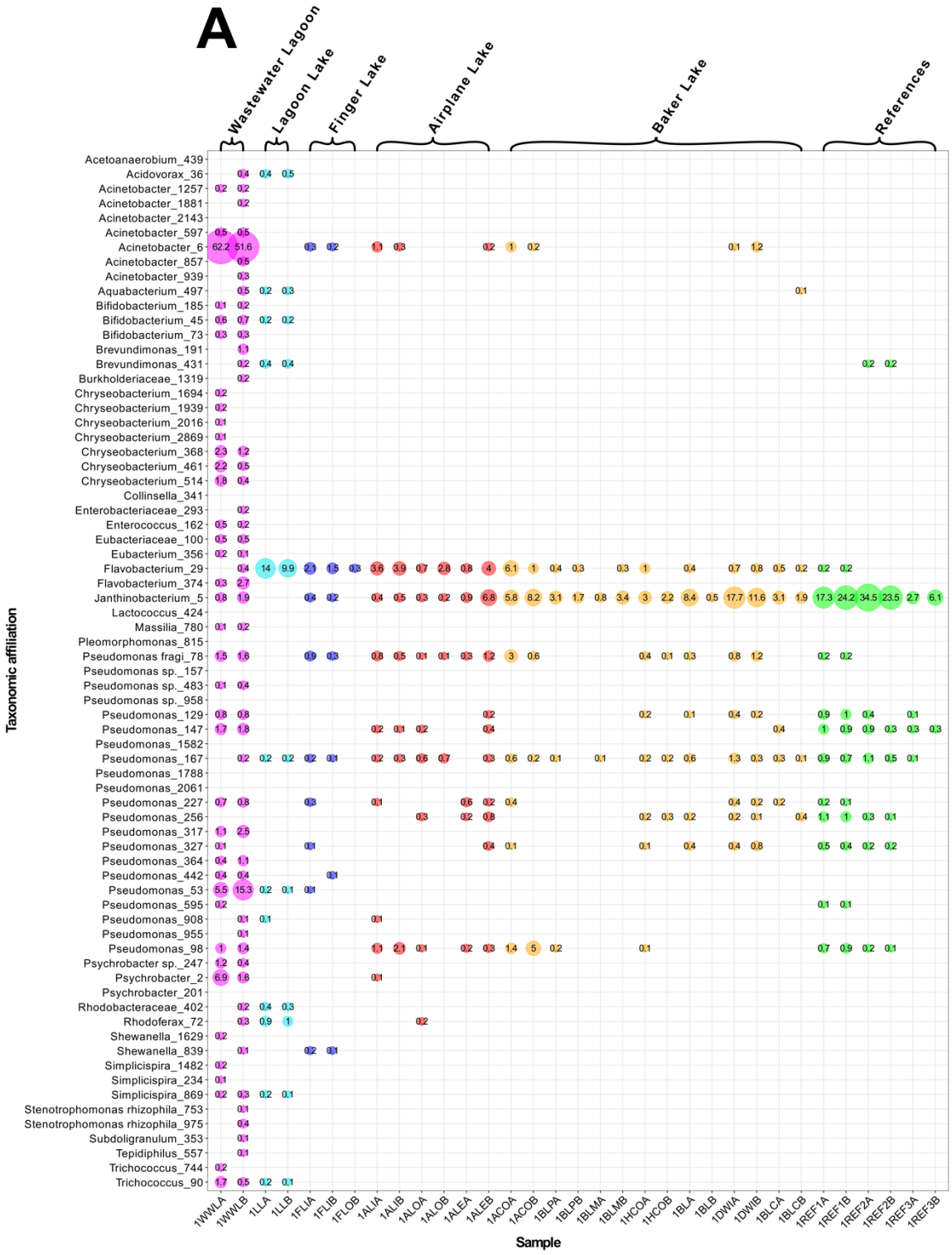

A
Wastewater Lagoon
Lagoon Lake
Finger Lake
Airplane Lake
Baker Lake
References
Taxonomic affiliation
Acetoanaerobium_439
Acidovorax_36
Acinetobacter_1257
Acinetobacter_1881
Acinetobacter_2143
Acinetobacter_597
Acinetobacter_6
Acinetobacter_857
Acinetobacter_939
Aquabacterium_497
Bifidobacterium_185
Bifidobacterium_45
Bifidobacterium_73
Brevundimonas_191
Brevundimonas_431
Burkholderiaceae_1319
Chryseobacterium_1694
Chryseobacterium_1939
Chryseobacterium_2016
Chryseobacterium_2869
Chryseobacterium_368
Chryseobacterium_461
Chryseobacterium_514
Collinsella_341
Enterobacteriaceae_293
Enterococcus_162
Eubacteriaceae_100
Eubacterium_356
Flavobacterium_29
Flavobacterium_374
Janthinobacterium_5
Lactococcus_424
Massilia_780
Pleomorphomonas_815
Pseudomonas fragi_78
Pseudomonas sp._157
Pseudomonas sp._483
Pseudomonas sp._958
Pseudomonas_129
Pseudomonas_147
Pseudomonas_1582
Pseudomonas_167
Pseudomonas_1788
Pseudomonas_2061
Pseudomonas_227
Pseudomonas_256
Pseudomonas_317
Pseudomonas_327
Pseudomonas_364
Pseudomonas_442
Pseudomonas_53
Pseudomonas_595
Pseudomonas_908
Pseudomonas_955
Pseudomonas_98
Psychrobacter sp._247
Psychrobacter_2
Psychrobacter_201
Rhodobacteraceae_402
Rhodoferax_72
Shewanella_1629
Shewanella_839
Simplicispira_1482
Simplicispira_234
Simplicispira_869
Stenotrophomonas rhizophila_753
Stenotrophomonas rhizophila_975
Subdoligranulum_353
Tepidiphilus_557
Trichococcus_744
Trichococcus_90
1WWLA
1WWLB
1LLA
1LLB
1FLIA
1FLIB
1FLOB
1ALIA
1ALIB
1ALOA
1ALOB
1ALEA
1ALEB
1ACOA
1ACOB
1BLPA
1BLPB
1BLMA
1BLMB
1HCOA
1HCOB
1BLA
1BLB
1DWIA
1DWIB
1BLCA
1BLCB
1REF1A
1REF1B
1REF2A
1REF2B
1REF3A
1REF3B
Sample

Figure 3.4: Relative abundance of amplicon sequence variants (ASVs) present at >0.1% abundance at each site in Baker Lake, Nunavut. (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. ASVs with <0.1% abundance in 1WWLA, 1WWLB, and 2WWLA were not included. Samples that did not contain >0.1% abundance of any of these ASVs were not included. Sample names correspond to sites indicated on the map in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. Lake names are indicated above columns.

3.1.2 Spatial Variability

Taxonomic profiling of the WSP lakes system in Baker Lake also revealed spatial variations of the microbial communities in the WSP and downstream lakes. The weighted and unweighted UniFrac ordinations demonstrate that samples are broadly arranged based on their distance from the WSP (Figure 3.1). Sample sites that were geographically closer to the WSP also appeared closer to the WSP within ordination space, with reference sites grouping among the more distant sites, and this was consistent among replicates. Overall, spatial trends appear to be comparable between both ordinations with the exception of the WWL and LL samples, which grouped closer together on the unweighted UniFrac ordination, and separated from the other lakes (Figure 3.1).

Spatial variability was also apparent at the phylum level of community composition. Samples from the WSP had *Proteobacteria* comprising the majority of detected microbial communities, whereas downstream site samples had less *Proteobacteria* and more *Bacteroidetes* for the first time point. However, the second time point samples contained more *Actinobacteria, Verrucomicrobia*, and *Planctomycetes* in sites downstream of the WSP (Figure 3.2). Spatial variability was less apparent when visualized at the ASV and family levels (Figures 3.3 and 3.4). Focusing on microorganisms detected in the Baker Lake WSP at >0.1% abundance, many of the ASVs from the WSP samples were also present in Lagoon Lake samples, albeit at low abundance, but not detected at all in lakes further downstream (Figure 3.4). This may be due to dilution effects caused by inflowing water from the upstream reference lake into Lagoon Lake. At the family level, ~85% of microorganisms in the WSP appear to be related to *Moraxellaceae* and *Pseudomonadaceae* (Figure 3.3). Both of

these groups decreased to below 1% abundance in Lagoon Lake, which may have been caused by differences in both pH and temperature. The lower temperature, and higher pH in Lagoon Lake may have been less favourable to the bacteria affiliated with *Moraxellaceae* and *Pseudomonas*, and more favourable to other microorganisms, causing a shift in microbial community composition as water flowed from the WSP to Lagoon Lake. Lagoon Lake samples also appear to have multiple families in common with both the WSP samples as well as the upstream reference lakes (Figure 3.3). *Flavobacteriaceae* and *Burkholderiacea* were the two most abundant families shared by Lagoon Lake and the reference lakes. In comparison, the WSP contained both of these families in lower abundance than in both Lagoon Lake and the reference lakes. In general, the WSP and Lagoon Lake shared more families that were detected in lower abundance. Although Lagoon Lake receives water from both the WSP and upstream lakes and is expected to contain microbial inputs from both systems, the microorganisms detected in the WSP do not persist in higher abundance than the microorganisms found in the upstream reference lakes. Microorganisms introduced to the environment through the WSP and are likely outcompeted by the microorganisms that are naturally present in the environment.

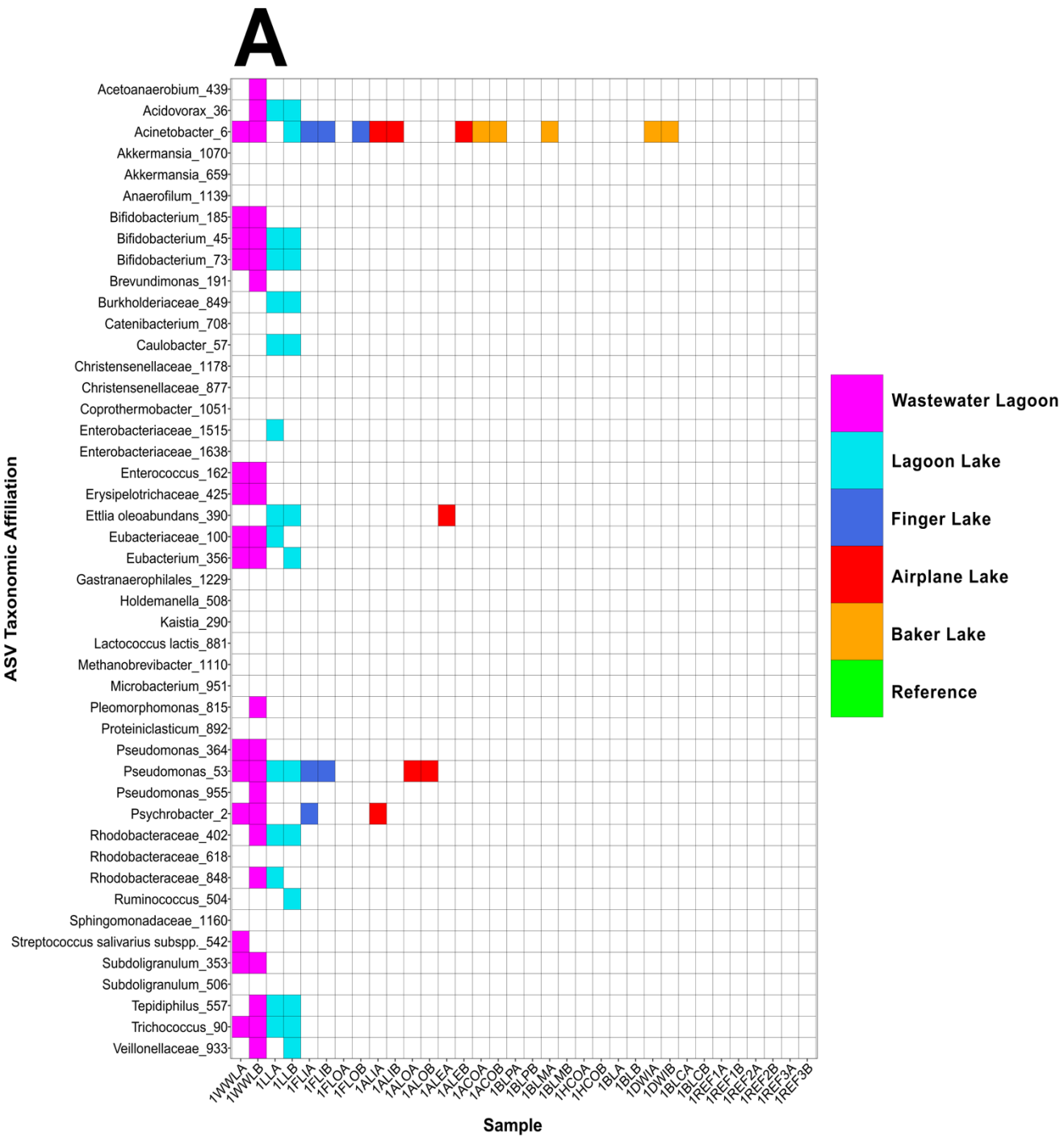
3.1.3 Wastewater Indicators

An indicator species analysis (ISA) identified ASVs specific to WSP-associated samples. An ISA produces indicator values for each ASV found within each group, which are assigned prior to analysis. These indicator values are determined by calculating proportional abundance of an ASV across all groups followed by calculating proportional frequency of the

ASV in each group. These two proportions are combined and ASVs with the highest indicator values are considered to be indicator species for a specific group. Statistical significance is then calculated using a Monte Carlo test to determine whether the observed indicator values are higher than expected by chance[111]. Using indicator values >0.9 and *p* values <0.05, 46 ASVs were identified as indicator species for wastewater (Table 3.2). Several species of the genera associated with these ASVs were identified previously in wastewater, including *Acinetobacter*[116], *Trichococcus*[66], and *Pseudomonas*[117,118]. Other species have been identified in human fecal matter, including *Bifidobacterium*[119], *Ruminococcus*[119], and *Enterococcus*[120]. Although the majority of these ASVs were detected with greater than 10 reads in at least one sample, two ASVs associated with the genera *Kaistia* and *Enterobacteriaceae* had fewer than 10 reads in all samples (Figure 3.5). Out of the 46 indicator ASVs, 17 were identified at sites downstream of Lagoon Lake. However, 13 of these indicator ASVs were found in a single site downstream of Lagoon Lake (i.e. 2ACOB). This sample could be considered an outlier with potential WSP contamination because the replicate (i.e. 2ACOA) did not contain these same wastewater indicator ASVs, except for *Acinetobacter*, which was also detected in other samples.

Table 3.2: Indicator species analysis results for wastewater amplicon sequence variants (ASVs) in Baker Lake, Nunavut. WWL and LL samples were grouped and compared to REFx samples (x indicating lake number). ASVs with an IndVal >0.9 and *p* value <0.05 were included. Asterisks beside ASVs indicate they were found at >0.1% abundance in the waste stabilization pond (WSP).

Taxonomic Affiliation	IndVal	*p* value
Bifidobacterium_45*	1.000	0.001
Pseudomonas_53*	1.000	0.001
Trichococcus_90*	1.000	0.001
Eubacteriaceae_100*	1.000	0.001
Enterococcus_162*	1.000	0.001
Bifidobacterium_185*	1.000	0.001
Subdoligranulum_353*	1.000	0.001
Eubacterium_356*	1.000	0.001
Ettlia oleoabundans_390	1.000	0.001
Erysipelotrichaceae_425	1.000	0.001
Acetoanaerobium_439*	1.000	0.001
Subdoligranulum_506	1.000	0.001
Streptococcus salivarius subspp._542	1.000	0.001
Tepidiphilus_557*	1.000	0.001
Pleomorphomonas_815*	1.000	0.001
Rhodobacteraceae_848	1.000	0.001
Christensenellaceae_877	1.000	0.001
Lactococcus lactis_881	1.000	0.001
Pseudomonas_955*	1.000	0.001
Akkermansia_1070	1.000	0.001
Bifidobacterium_73*	1.000	0.001
Psychrobacter_2*	1.000	0.001
Acinetobacter_6*	0.997	0.004
Acidovorax_36*	0.935	0.001
Caulobacter_57	0.935	0.001
Brevundimonas_191*	0.935	0.001
Kaistia_290	0.935	0.001
Rhodobacteraceae_402*	0.935	0.001
Ruminococcus_504	0.935	0.001
Holdemanella_508	0.935	0.002
Rhodobacteraceae_618	0.935	0.001
Akkermansia_659	0.935	0.001
Catenibacterium_708	0.935	0.002
Burkholderiaceae_849	0.935	0.001
Proteiniclasticum_892	0.935	0.002
Veillonellaceae_933	0.935	0.001
Microbacterium_951	0.935	0.001
Coprothermobacter_1051	0.935	0.001
Methanobrevibacter_1110	0.935	0.001
Anaerofilum_1139	0.935	0.002
Sphingomonadaceae_1160	0.935	0.001
Christensenellaceae_1178	0.935	0.001
Gastranaerophilales_1229	0.935	0.001
Enterobacteriaceae_1515	0.935	0.001
Enterobacteriaceae_1638	0.935	0.001
Pseudomonas_364*	0.933	0.001


A
ASV Taxonomic Affiliation
Acetoanaerobium_439
Acidovorax_36
Acinetobacter_6
Akkermansia_1070
Akkermansia_659
Anaerofilum_1139
Bifidobacterium_185
Bifidobacterium_45
Bifidobacterium_73
Brevundimonas_191
Burkholderiaceae_849
Catenibacterium_708
Caulobacter_57
Christensenellaceae_1178
Christensenellaceae_877
Coprothermobacter_1051
Enterobacteriaceae_1515
Enterobacteriaceae_1638
Enterococcus_162
Erysipelotrichaceae_425
Ettlia oleoabundans_390
Eubacteriaceae_100
Eubacterium_356
Gastranaerophilales_1229
Holdemanella_508
Kaistia_290
Lactococcus lactis_881
Methanobrevibacter_1110
Microbacterium_951
Pleomorphomonas_815
Proteiniclasticum_892
Pseudomonas_364
Pseudomonas_53
Pseudomonas_955
Psychrobacter_2
Rhodobacteraceae_402
Rhodobacteraceae_618
Rhodobacteraceae_848
Ruminococcus_504
Sphingomonadaceae_1160
Streptococcus salivarius subspp._542
Subdoligranulum_353
Subdoligranulum_506
Tepidiphilus_557
Trichococcus_90
Veillonellaceae_933
1WWLA
1WWLB
1LLA
1LLB
1FLIA
1FLIB
1FLOA
1FLOB
1ALIA
1ALIB
1ALOA
1ALOB
1ALEA
1ALEB
1ACOA
1ACOB
1BLPA
1BLPB
1BLMA
1BLMB
1HCOA
1HCOB
1BLA
1BLB
1DWIA
1DWIB
1BLCA
1BLCB
1REF1A
1REF1B
1REF2A
1REF2B
1REF3A
1REF3B
Sample
Wastewater Lagoon
Lagoon Lake
Finger Lake
Airplane Lake
Baker Lake
Reference

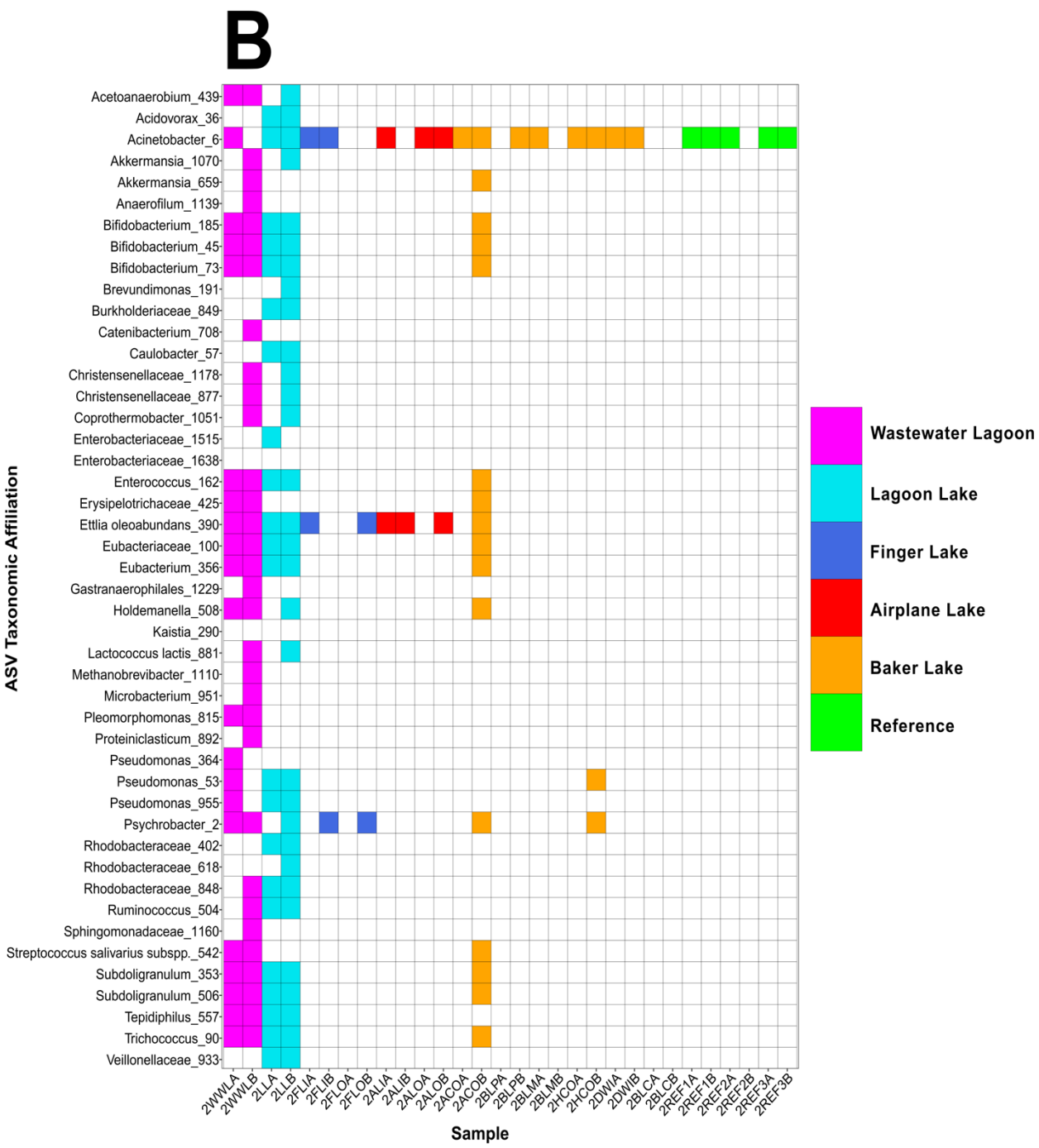


Figure 3.5:Presence/absence heatmap of wastewater indicator amplicon sequence variants (ASVs) at all sites in Baker Lake, Nunavut. (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. Sample names correspond to sites indicated on the map in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. Lake names are indicated above columns. Boxes were filled in if the ASV had at least 10 reads at that site.

Although many of the indicator ASVs were detected at very low relative abundance, 19 out of the 46 ASVs were detected at >0.1% abundance (Table 3.2 and Figure 3.4). Many of these ASVs were detected in the WSP or Lagoon Lake but were not found in downstream sites. This means that wastewater may not be impacting downstream sites as heavily as expected, and that dilution of the wastewater leaving the WSP may be enough to reduce wastewater organisms in downstream sites below detectable levels.

A single wastewater indicator ASV was detected in reference lakes as well as in the WSP. This ASV was associated with the genus *Acinetobacter* and was detected in the WSP at ~60% relative abundance. Species within this genus are relatively ubiquitous[121], and therefore it is not surprising that this ASV was detected in every lake examined, albeit at below 1% abundance for samples taken from most sites. This could be a potential health concern because a number of species within this genus are human pathogens and several of which have been associated with multi-drug resistant infections[121,122]. However, the source of this ASV cannot be directly associated with wastewaters because it was also found in samples from reference sites, which should have no wastewater input. Further testing would need to be done to identify the origin of this ASV. Species-level classification was not identified for the majority of ASVs, and therefore it is not possible to know whether this ASV was a potential pathogen, and further tests would also need to be performed to better characterize the *Acinetobacter*-affiliated taxa detected in the wastewater.

3.2 Comparison of WSPs from Baker Lake, Cambridge Bay, and Kugluktuk

Ordinations prepared with WSP samples from Baker Lake, Cambridge Bay, and Kugluktuk demonstrated that samples grouped distinctly based on WSP location on both axes (Figure 3.6), with the exception being that first time point Cambridge Bay WSP samples grouped with the Baker Lake WSP samples. The weighted UniFrac ordination grouped these samples closer together than the unweighted UniFrac ordination. In total, ~75.5% of the variance was explained on the first two axes of the weighted UniFrac ordination, compared to only ~46.7% of the variance explained on the first two axes of the unweighted UniFrac ordination. The weighted UniFrac ordination also separated the Baker Lake sample, 2WWLB, from the other Baker Lake samples. Sample 2WWLB was previously established as an outlier, so this trend was expected. The Kugluktuk samples also did not group with any other samples from different sampling sites (Figure 3.6).

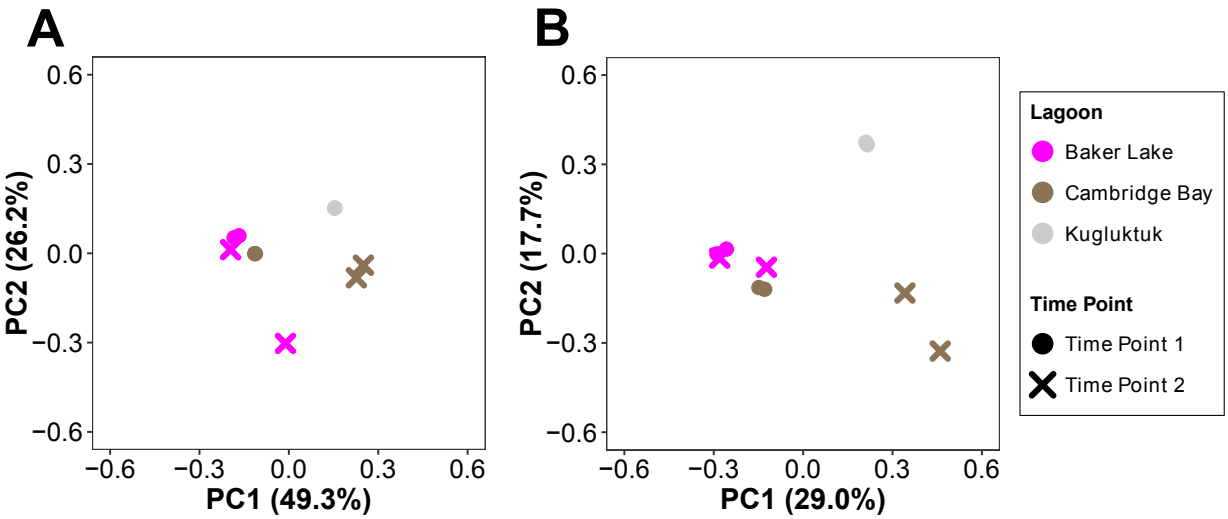


Figure 3.6: Principal coordinates analysis (PCoA) ordination plots using (A) weighted and (B) unweighted UniFrac distances based on amplicon sequence variants (ASVs) from 16S rRNA gene amplicons for waste stabilization pond (WSP) sites in Baker Lake, Cambridge Bay, and Kugluktuk.

Despite site-specific grouping within ordination space, samples from the three WSPs had similar taxonomic profiles, with *Proteobacteria* representing >75% of the microbial communities detected for all samples, except 2WWLB and the samples from the second time point of Cambridge Bay (Figure 3.7). As expected based on the ordinations, microbial community ASV profiles detected for Kugluktuk samples were distinct (Figure 3.8), with only two ASVs at >1% abundance in common for the other two locations. These two ASVs were associated with the genus *Pseudorhodobacter,* which was in common with the Cambridge Bay WSP and the genus *Brevundimonas*, which was shared with the Baker Lake WSP. The most abundant ASVs in the Kugluktuk WSP were affiliated with *Acidovorax* and *Rhodoferax* genera. Multiple strains of *Acidovorax* were reported for activated sludge samples from municipal WWTPs, and are known to be denitrifiers[123,124]. Strains of *Rhodoferax* have also been identified as denitrifiers involved in the removal of nitrogen in WWTPs[125], and several strains have also been described as psychrotolerant[126,127]. It is likely that these samples contained a mixture of wastewater and sludge, indicating disturbance of the sediments within the WSP caused by the continuous dumping of new wastewater into the WSP.

The first time point Cambridge Bay samples grouped more closely to samples from Baker Lake than to the second time point Cambridge Bay samples (Figures 3.6). The second time point Cambridge Bay samples contained more phyla than identified in the first time point, including *Cyanobacteria*, *Planctomycetes*, and *Verrucomicrobia* (Figure 3.7). *Actinobacteria* ASVs were also present in higher proportions for second time point Cambridge Bay samples than in the rest of the samples, apart from 2WWLB. Differences in

microbial community profiles for Cambridge Bay samples were likely linked to temporal factors. The first set of WSP samples from Cambridge Bay were collected prior to wastewater discharge and the second set of samples were collected after the initial release of wastewater had occurred. The lagoon became shallower as a result of the release, presumably allowing for a rapid water temperature increase. Previous research showed that *Cyanobacteria* increased in shallow lakes with higher temperature[128], and thus may have proliferated in the Cambridge Bay WSP as water levels decreased. This same growth induction may have occurred for other microorganisms as temperatures increased. The dominant ASV for the first time point was associated with *Psychrobacter*, at a relative abundance of >80% (Figure 3.8), yet, this ASV was not detected in the second time point. Instead, ASVs associated with other genera became more prevalent such that none of the ASVs identified during the second time point were identified in the first at an abundance of >1%. *Psychrobacter* species have been identified as psychrophiles[112] and a shift from colder to warmer temperatures is highly likely associated with these abundance changes. Unfortunately, temperature measurements were not taken for either time point due to sampling logistics constraints and therefore this cannot be verified.

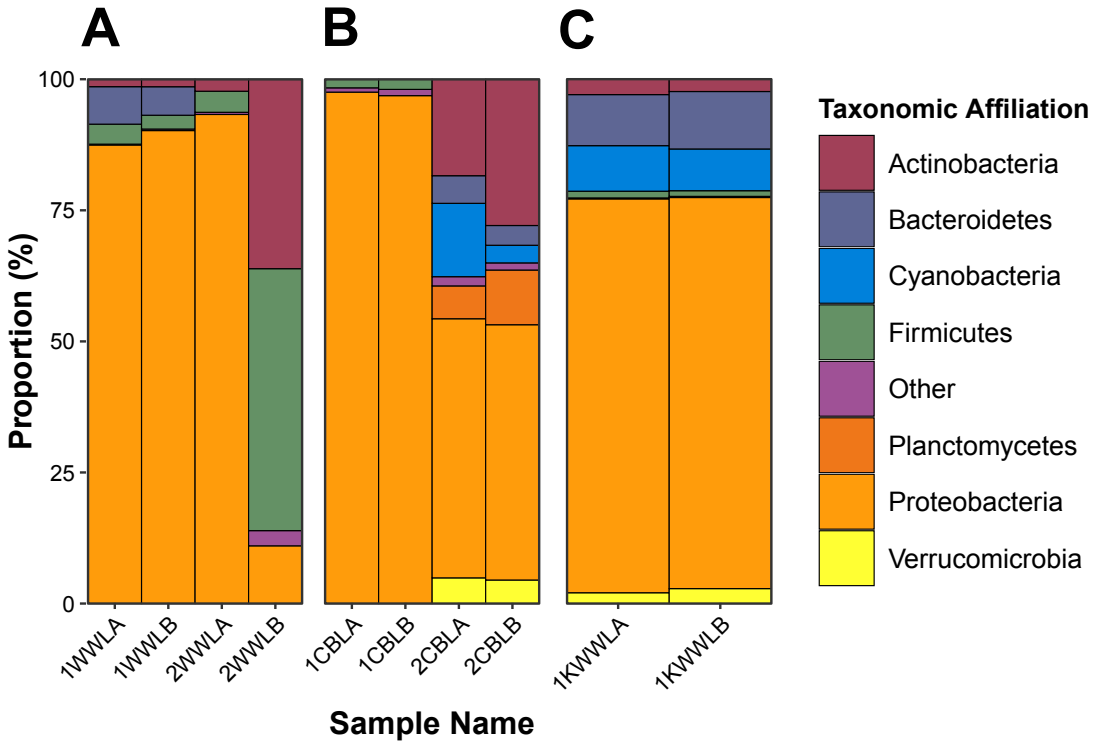


Figure 3.7: Relative abundance of phyla present at waste stabilization pond (WSP) sites in Baker Lake, Cambridge Bay, and Kugluktuk. (A) Samples collected from the WSP at Baker Lake. (B) Samples collected from the WSP at Cambridge Bay. (C) Samples collected from the WSP at Kugluktuk. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point.

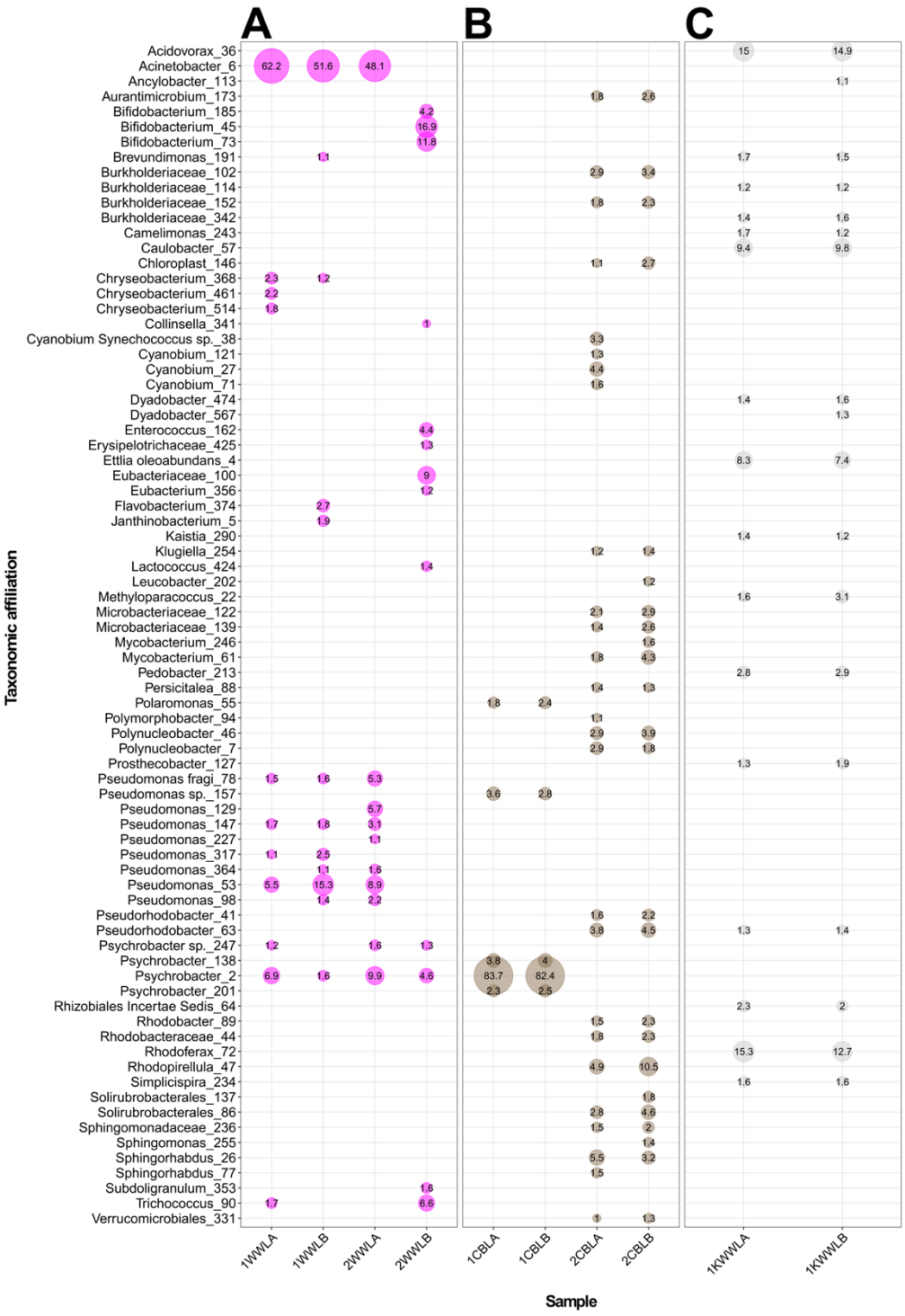


Figure 3.8: Relative abundance of amplicon sequence variants (ASVs) present at >1% abundance at waste stabilization pond (WSP) sites in Baker Lake, Cambridge Bay, and Kugluktuk. (A) Samples collected from the WSP at Baker Lake. (B) Samples collected from the WSP at Cambridge Bay. (C) Samples collected from the WSP at Kugluktuk. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point.

3.3 Comparison of 16S rRNA Genes and Metagenomes

A subset of the samples that were analyzed to compare 16S rRNA gene profiles were selected for metagenomic sequencing, including a single sample from each of the lakes downstream of the Baker Lake WSP, as well as from the WSP itself. The sampling sites closest to the outflow of those lakes were selected. The DWI site in Baker Lake was selected because it is located near the water intake pipe for the community. Two out of the three reference sites were included for comparison. The Cambridge Bay and Kugluktuk WSP samples were also sequenced. For each of the sites selected from Baker Lake and Cambridge Bay, duplicates at both time points were sequenced. Because Kugluktuk only had a single time point sampled, there were only duplicates sequenced.

Metagenome sequencing generated ~600 million paired-end reads. Raw reads were analyzed using MetAnnotate[106], which identified all sequences that matched with the *rpoB* gene HMM. This produced a microbial community profile that could be compared to the 16S rRNA gene profile. All reads were also assembled using the ATLAS 2.0 pipeline[80], and MAGs were generated. The ATLAS 2.0 pipeline also assigns taxonomy to the MAGs based on taxonomic assignments to all binned genes. Although these profiles were also compared to the 16S rRNA gene profiles, only a portion of the reads were assigned to bins. This portion of unassigned reads varied between samples, ranging between 20-80% of reads. These reads were labelled as “unclassified” (Figures 3.10C-12C). Therefore, the MAG taxonomic profiles reflect a subset of DNA sequences obtained.

3.3.1 Consistent Taxonomic Profiles

Taxonomic profiles based on 16S rRNA gene amplicons, taxonomic assignments to sequences associated with *rpoB,* and taxonomic assignments to MAGs showed similar groupings of samples (Figure 3.9). Each of the ordinations display spatial variability with regards to geographic distance of sampling sites from one another. The WSP samples were more similar to sample sites located geographically closer to Baker Lake's WSP samples than Airplane Lake, Baker Lake, and reference lake samples. This gradient was expected because wastewater entering the downstream lake systems is diluted, and therefore microbial communities associated with the wastewater are also diluted and outcompeted by other microorganisms. The Baker Lake WSP grouped closely with the first time point of the Cambridge Bay WSP, whereas the Kugluktuk WSP and the second time point of the Cambridge Bay WSP grouped more closely with Lagoon Lake and Finger Lake samples. Based on the phylum-level taxonomic profiles (Figures 3.10-12), this grouping was expected because similar proportions of each phyla were observed for samples that grouped together.

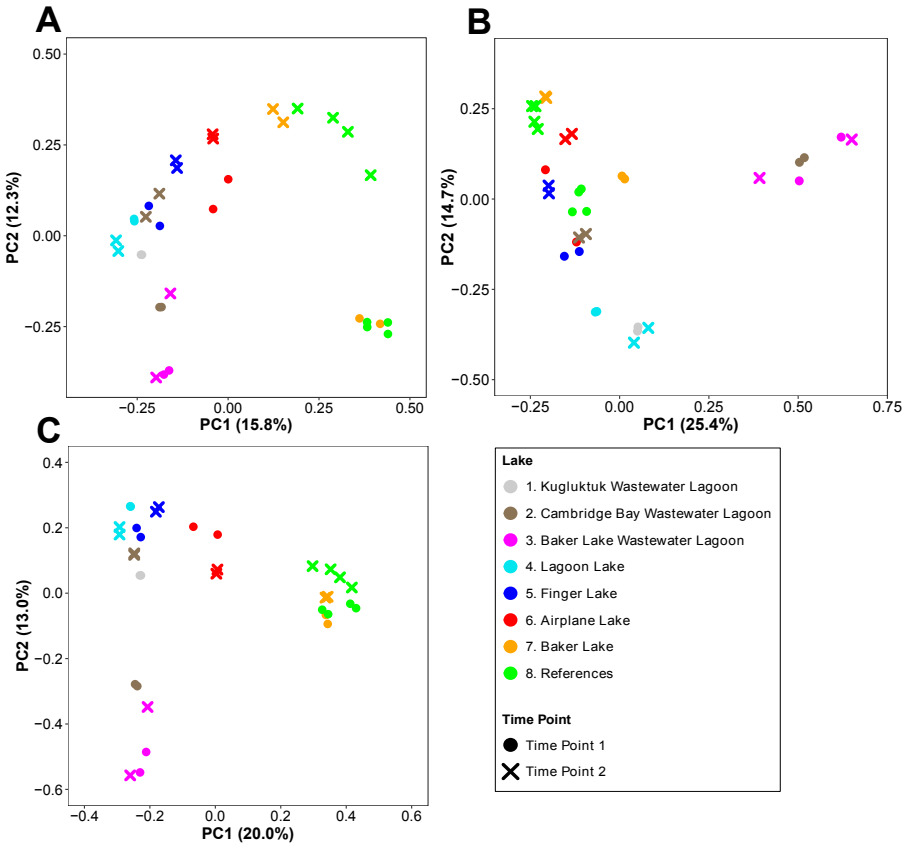


Figure 3.9: Principal coordinates analysis (PCoA) ordination plots using Bray-Curtis distances based on (A) amplicon sequence variants (ASVs) from 16S rRNA gene amplicons, (B) taxonomic assignments to metagenome sequences corresponding to the rpoB hidden Markov model (HMM), and (C) taxonomic assignments to metagenome-assembled genomes (MAGs).

In general, similar proportions of phyla for the same samples occurred when comparing the 16S rRNA gene taxonomic assignments and *rpoB* gene taxonomic assignments (Figures 3.10-12). However, this pattern was not as clear when observing MAG taxonomic assignments. Only a small proportion of all metagenomic reads were assigned to bins, and therefore many reads were not classified. However, similar ratios of the phyla that were detected in each sample were still observed when comparing MAG taxonomic assignments to the other two datasets. *Proteobacteria* was the dominant phylum in WSP samples within all three datasets. *Bacteroidetes* and *Actinobacteria* are also highly abundant

phyla in the 16S rRNA gene taxonomic assignments as well as in the *rpoB* gene taxonomic assignments. However, these groups are underrepresented in the MAG taxonomic assignments, likely due to the low proportion of mapped reads. The only difference with the *rpoB* taxonomic classifications is that the *Actinobacteria* phylum was more prominent than displayed in the 16S rRNA gene taxonomic assignments. This is most noticeable in the first Baker Lake time point where relative abundance of the phylum is only as high as ~5% in the sample 1ALOA based on the 16S rRNA genes, but reaches up to ~30% for the same sample based on the *rpoB* gene (Figure 3.10). Forward and reverse metagenomic reads were used separately to identify taxa present based on the *rpoB* HMM. Both sets of reads also had consistent taxonomic profiles.

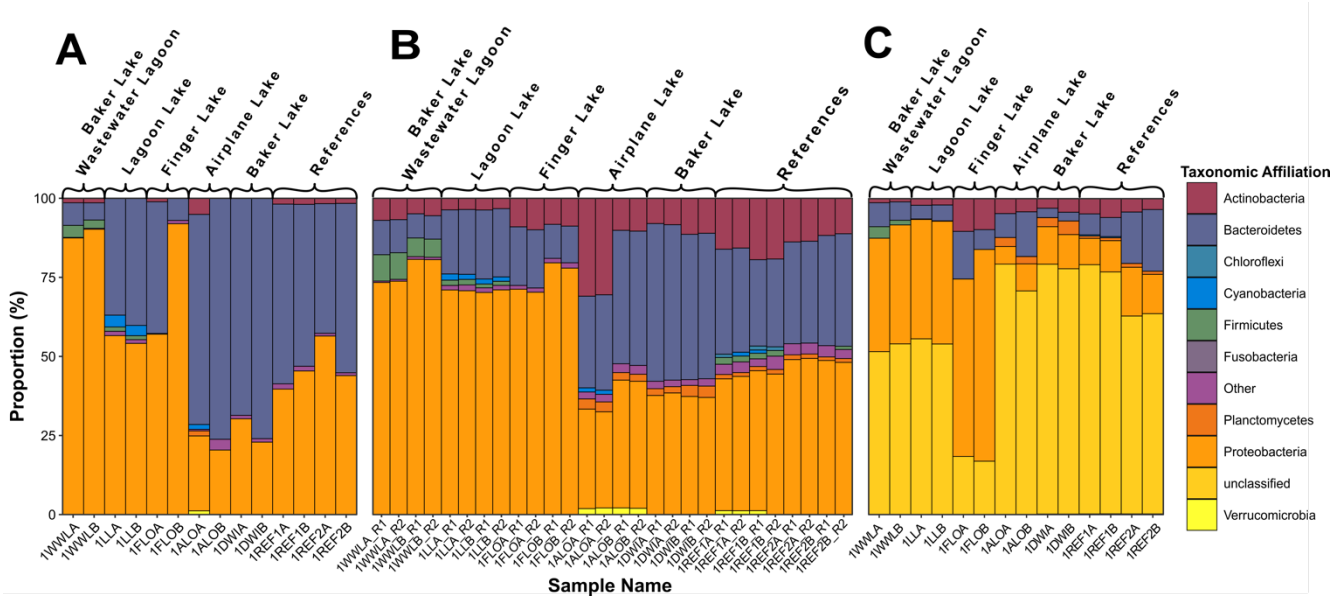


Figure 3.10: Relative abundance of phyla present at metagenome sequenced sites from Baker Lake samples collected between July 13-16, 2018. (A) Taxonomic assignments to 16S rRNA gene amplicons. (B) Taxonomic assignments to sequences corresponding to the *rpoB* hidden Markov model (HMM). (C) Taxonomic assignments to metagenome-assembled genomes (MAGs). Sample names correspond to sites indicated in Figure 2.2. Letters A and B at the end of sample names refer to replicates at the same site and time point. R1 and R2 refer to forward and reverse reads, respectively. Lake names are indicated above bars.

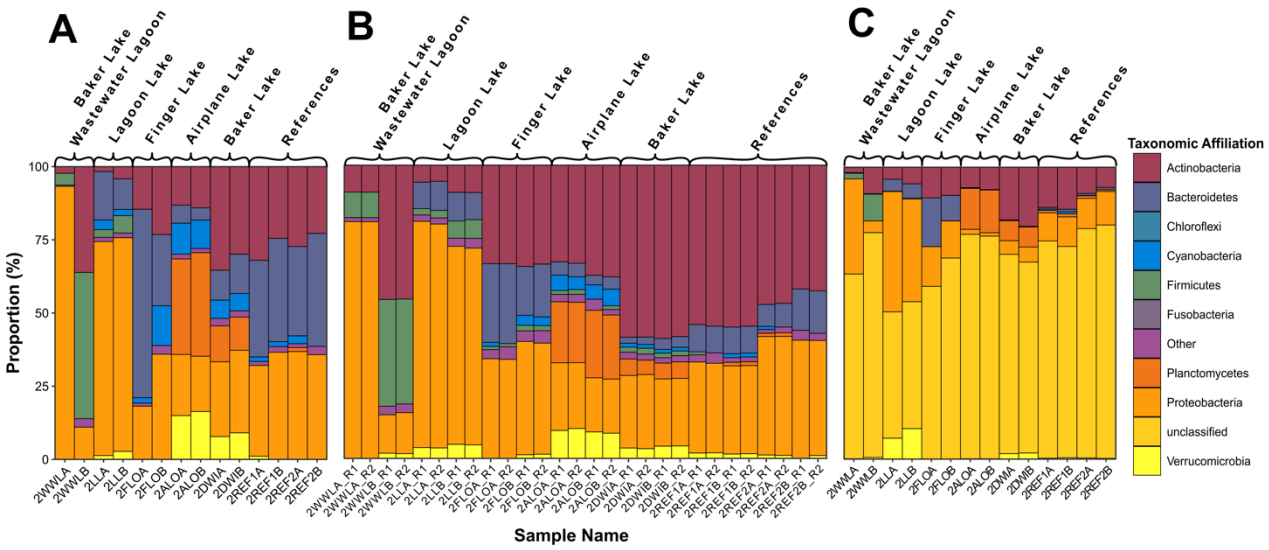


Figure 3.11: Relative abundance of phyla present at metagenome sequenced sites from Baker Lake samples collected between July 22-24, 2018. (A) Taxonomic assignments to 16S rRNA gene amplicons. (B) Taxonomic assignments to sequences corresponding to the *rpoB* hidden Markov model. (C) Taxonomic assignments to metagenome-assembled genomes (MAGs). Sample names correspond to sites indicated in Figure 2.2. Letters A and B at the end of sample names refer to replicates at the same site and time point. R1 and R2 forward and reverse reads, respectively. Lake names are indicated above bars.

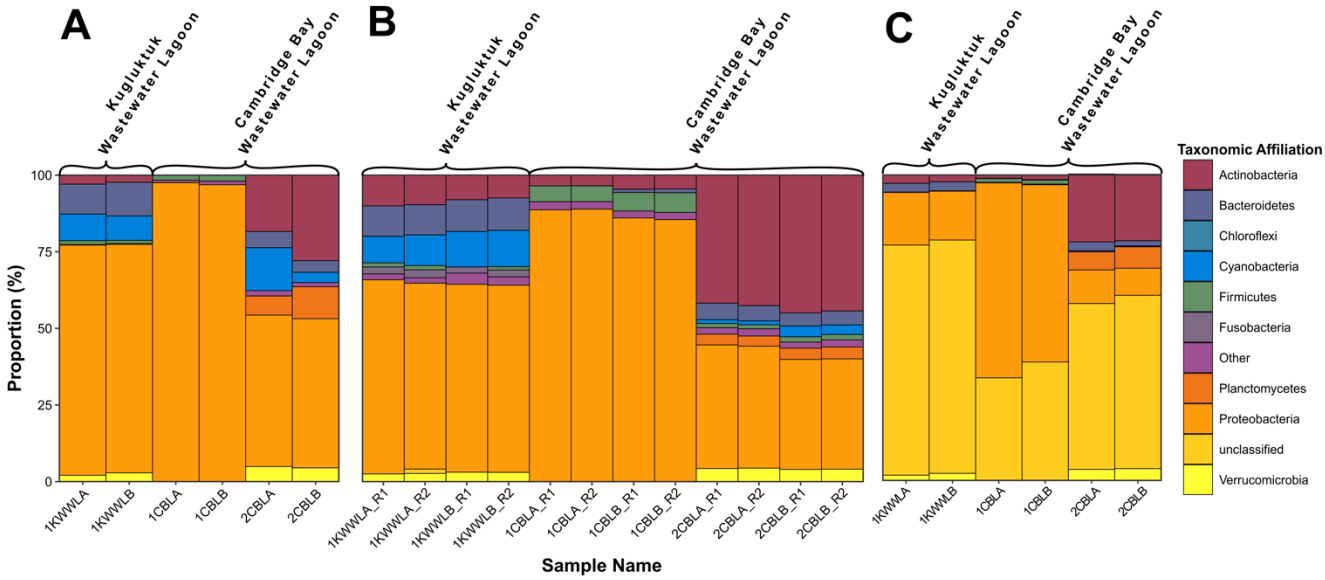


Figure 3.12: Relative abundance of phyla present at metagenome sequenced sites from the WSPs in Cambridge Bay, and Kugluktuk. (A) Taxonomic assignments to 16S rRNA gene amplicons. (B) Taxonomic assignments to sequences corresponding to the *rpoB* hidden Markov model. (C) Taxonomic assignments to metagenome-assembled genomes (MAGs). Sample names correspond to sites indicated in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. R1 and R2 refer to forward and reverse reads, respectively. Lake names are indicated above bars.

Comparing these 16S rRNA gene, *rpoB* gene, and MAG datasets at the family level, the *rpoB* gene taxonomic assignment identified the most families (Figures 3.13-15). The *rpoB* gene taxonomic assignment identified 82 families, while only 56 were identified using the 16S rRNA gene, and 28 were identified using MAG taxonomy. Studies have shown that assigning taxonomy using *rpoB* gene amplicons allows for the identification of more species[129,130]. In this research, *rpoB* genes were identified using HMMs; however, the larger number of detected taxa may be attributed to *rpoB* providing better resolution among species, although this does not explain why bacteria from families such as *Comamonadaceae* were detected at relatively high abundance in the *rpoB* and MAG taxonomic assignment, but were not detected at all in the 16S rRNA gene analysis. However, the 16S rRNA gene analysis may have not detected the *Comamonadaceae* family due to primer bias against the 16S rRNA genes of the family.

Overall, seven of the most abundant taxa were identified in all three datasets, which includes *Bifidobacteriaceae, Carnobacteriaceae, Caulobacteriaceae, Flavobacteriaceae, Moraxellaceae, Pseudomonadaceae, Rhodobacteraceae,* and *Sphingomonadaceae*. *Moraxellaceae* was the dominant family in the Cambridge Bay WSP at the first time point as well as in both time points in the Baker Lake WSP. The *Sphingomonadaceae* family was highly abundant in Finger Lake samples, and *Flavobacteriaceae* was identified at high proportions in nearly all samples from Baker Lake. In contrast, *Burkholderiaceae* was identified in the 16S rRNA gene and *rpoB* gene analysis, but was not in the MAG taxonomy assignment. This may be due to the large number of reads that were not mapped to bins and thus many reads that were not assigned to a MAG may correspond to this family. This may

also have been caused by the inability to resolve taxonomic assignments to lower levels, and therefore some of the MAGs unresolved to the family level might also be associated with *Burkholderiaceae*, specifically those within the *Burkholderiales* order (Table 3.3).

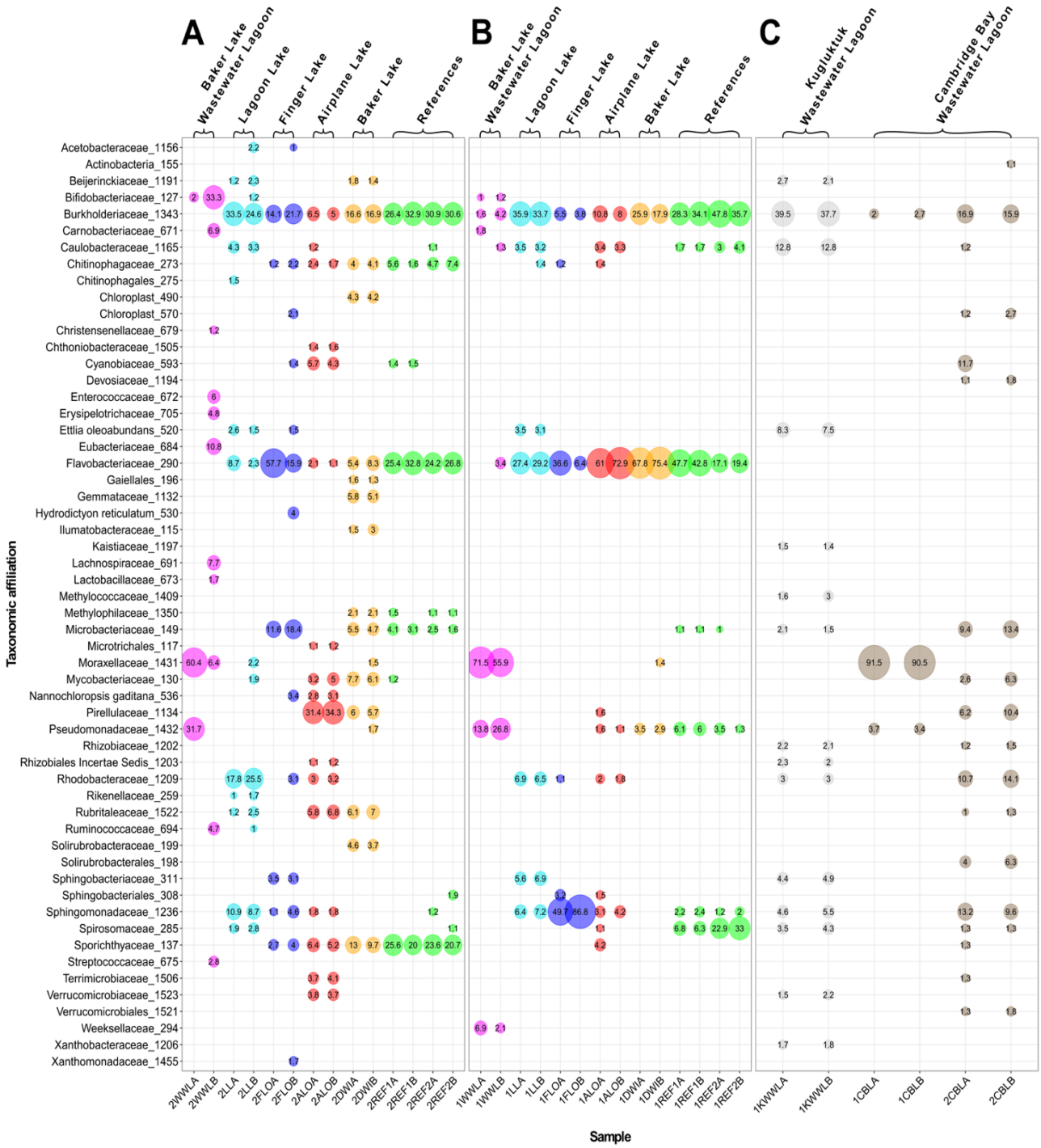
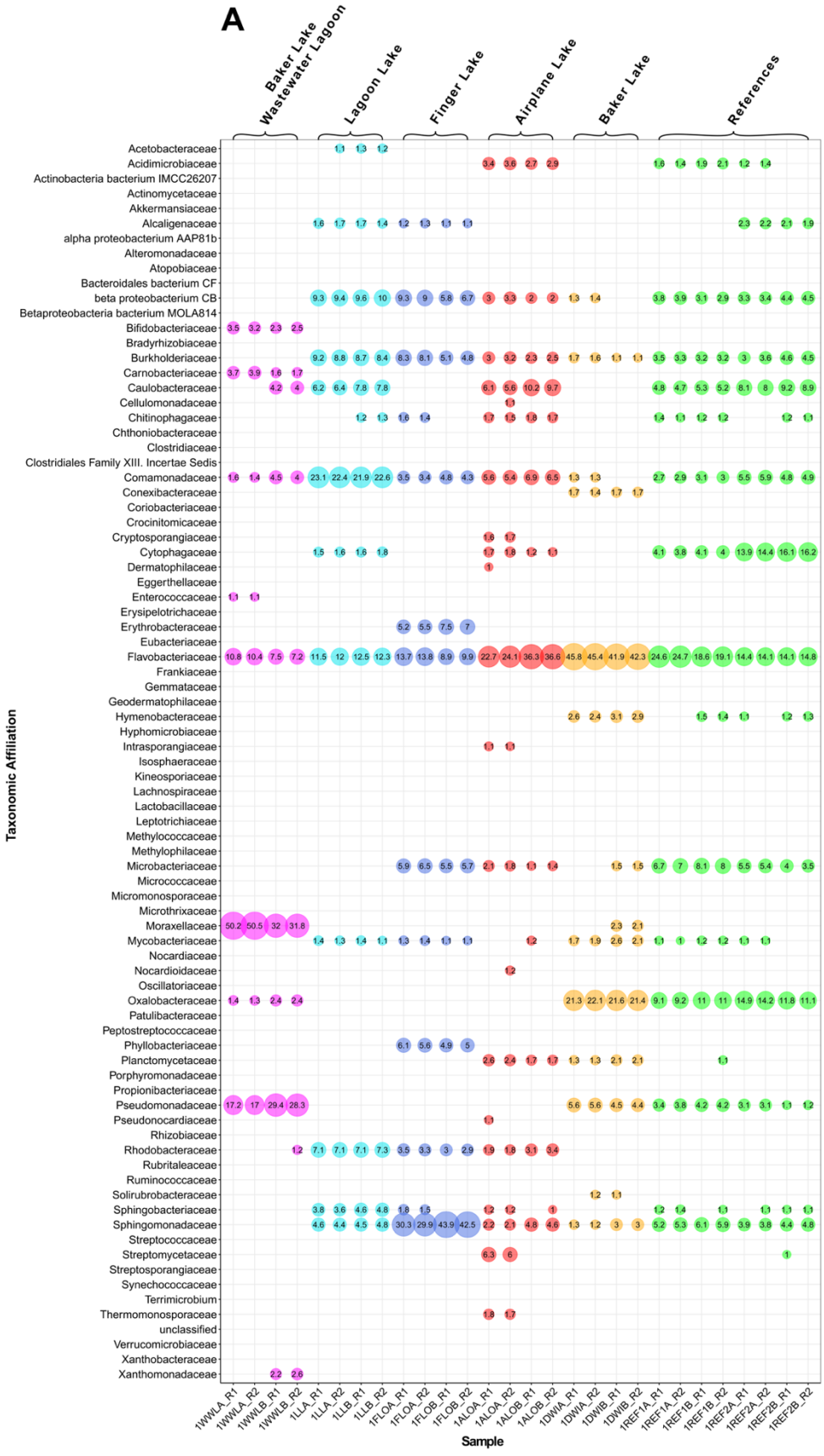


Figure 3.13: Relative abundance of families present at >1% abundance in metagenome sequenced sites from Baker Lake, Cambridge Bay, and Kugluktuk, based on 16S rRNA gene amplicons. (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. (C) Samples collected from the Cambridge Bay and Kugluktuk waste stabilization ponds (WSPs). Baker Lake sample names correspond to sites indicated in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and those beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. Lake names are indicated above columns.


A
Baker Lake Wastewater Lagoon
Lagoon Lake
Finger Lake
Airplane Lake
Baker Lake
References
Taxonomic Affiliation
Acetobacteraceae
Acidimicrobiaceae
Actinobacteria bacterium IMCC26207
Actinomycetaceae
Akkermansiaceae
Alcaligenaceae
alpha proteobacterium AAP81b
Alteromonadaceae
Atopobiaceae
Bacteroidales bacterium CF
beta proteobacterium CB
Betaproteobacteria bacterium MOLA814
Bifidobacteriaceae
Bradyrhizobiaceae
Burkholderiaceae
Carnobacteriaceae
Caulobacteraceae
Cellulomonadaceae
Chitinophagaceae
Chthoniobacteraceae
Clostridiaceae
Clostridiales Family XIII. Incertae Sedis
Comamonadaceae
Conexibacteraceae
Coriobacteriaceae
Crocinitomicaceae
Cryptosporangiaceae
Cytophagaceae
Dermatophilaceae
Eggerthellaceae
Enterococcaceae
Erysipelotrichaceae
Erythrobacteraceae
Eubacteriaceae
Flavobacteriaceae
Frankiaceae
Gemmataceae
Geodermatophilaceae
Hymenobacteraceae
Hyphomicrobiaceae
Intrasporangiaceae
Isosphaeraceae
Kineosporiaceae
Lachnospiraceae
Lactobacillaceae
Leptotrichiaceae
Methylococcaceae
Methylophilaceae
Microbacteriaceae
Micrococcaceae
Micromonosporaceae
Microthrixaceae
Moraxellaceae
Mycobacteriaceae
Nocardiaceae
Nocardioidaceae
Oscillatoriaceae
Oxalobacteraceae
Patulibacteraceae
Peptostreptococcaceae
Phyllobacteriaceae
Planctomycetaceae
Porphyromonadaceae
Propionibacteriaceae
Pseudomonadaceae
Pseudonocardiaceae
Rhizobiaceae
Rhodobacteraceae
Rubritaleaceae
Ruminococcaceae
Solirubrobacteraceae
Sphingobacteriaceae
Sphingomonadaceae
Streptococcaceae
Streptomycetaceae
Streptosporangiaceae
Synechococcaceae
Terrimicrobium
Thermomonosporaceae
unclassified
Verrucomicrobiaceae
Xanthobacteraceae
Xanthomonadaceae
1WWLA_R1
1WWLA_R2
1WWLB_R1
1WWLB_R2
1LLA_R1
1LLA_R2
1LLB_R1
1LLB_R2
1FLOA_R1
1FLOA_R2
1FLOB_R1
1FLOB_R2
1ALOA_R1
1ALOA_R2
1ALOB_R1
1ALOB_R2
1DWIA_R1
1DWIA_R2
1DWIB_R1
1DWIB_R2
1REF1A_R1
1REF1A_R2
1REF1B_R1
1REF1B_R2
1REF2A_R1
1REF2A_R2
1REF2B_R1
1REF2B_R2
Sample

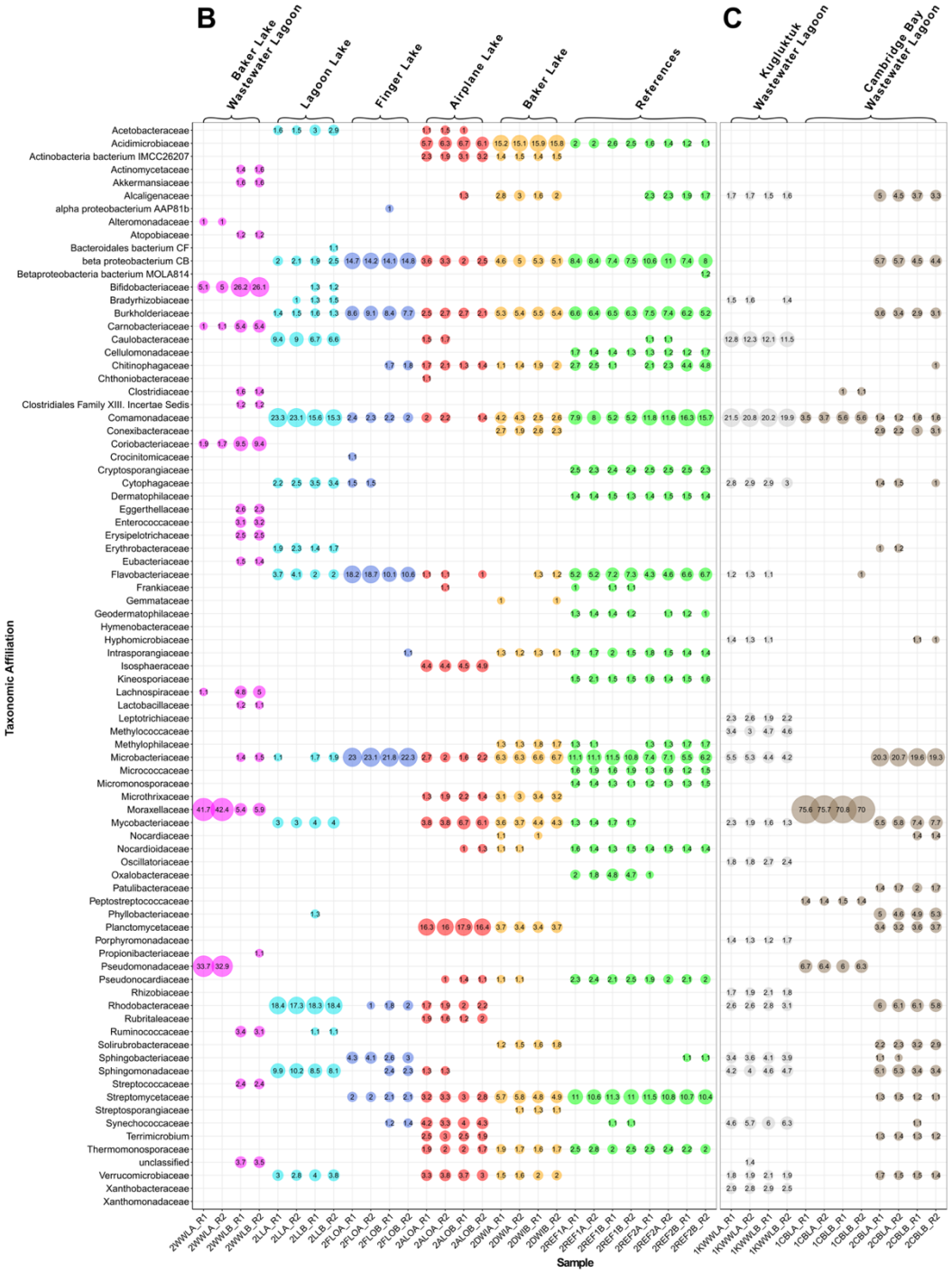


Figure 3.14: Relative abundance of families present at >1% abundance in metagenome sequenced sites from Baker Lake, Cambridge Bay, and Kugluktuk, based on taxonomic assignments to reads corresponding to the rpoB hidden Markov model (HMM). (A) Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. (C) Samples collected from the Cambridge Bay and Kugluktuk waste stabilization ponds (WSPs). Baker Lake sample names correspond to sites indicated in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and those beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point. R1 and R2 refer to forward and reverse reads, respectively. Lake names are indicated above columns.

Figure 3.15: Relative abundance of families present at metagenome sequenced sites from Baker Lake, Cambridge Bay, and Kugluktuk, based on taxonomic assignments to metagenome-assembled genomes (MAGs). Samples collected between July 13-16, 2018. (B) Samples collected between July 22-24, 2018. (C) Samples collected from the Cambridge Bay and Kugluktuk waste stabilization ponds (WSPs). MAGs with unresolved taxonomic assignment at the family level were grouped into "Unresolved_MAGs". Baker Lake sample names correspond to sites indicated in Figure 2.2. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. A and B at the end of sample names refer to replicates at the same site and time point. Lake names are indicated above columns.

Table 3.3: Metagenome assembled genomes (MAGs) with unresolved family-level taxonomy. The entry with an asterisk indicates insufficient data for classification for the Bin Annotation Tool (BAT) during assembly.

Number of MAGs	Taxonomic Affiliation	MAG IDs
22	Actinobacteria	MAG001;MAG004;MAG024;MAG027;MAG037; MAG038;MAG049;MAG059;MAG062;MAG078; MAG079;MAG085;MAG087;MAG088;MAG089; MAG098;MAG100;MAG108;MAG130;MAG132; MAG135;MAG137
2	Bacteroidetes	MAG094;MAG109
5	Betaproteobacteria	MAG011;MAG033;MAG083;MAG091;MAG103;
10	Burkholderiales	MAG006;MAG016;MAG044;MAG077;MAG096; MAG102;MAG107;MAG116;MAG121;MAG124
1	Candidatus Parcubacteria	MAG084
2	Clostridiales	MAG018;MAG090
2	Micrococcales	MAG069;MAG133
8	Planctomycetes	MAG021;MAG046;MAG050;MAG053;MAG092; MAG106;MAG139;MAG140
1	Rhizobiales	MAG020
6	Solirubrobacterales*	MAG010;MAG026;MAG057;MAG113;MAG119; MAG142
1	Sphingobacteriales	MAG048
3	Verrucomicrobia	MAG047;MAG095;MAG123

To determine how similar these three datasets are to one another, hierarchical clustering and the Mantel tests were conducted. The Mantel test determined, based on distance matrices for all samples from each dataset, that the *rpoB* dataset and 16S rRNA gene dataset were most similar, with a statistically significant observed correlation (Table 3.4). This is a strong positive correlation, indicating that samples within each dataset had similar distances to one another. The MAG dataset had a weaker correlation with the *rpoB* and 16S rRNA gene datasets (Table 3.4), but these were still considered significantly positive

correlations. Again, this is likely due to the underrepresentation of taxa in the MAG dataset as a result of fewer reads mapping to the MAGs.

Table 3.4: Mantel test results comparing distance matrices for each pair based on 16S rRNA gene amplicons, taxonomic assignment to sequences corresponding to the rpoB hidden Markov model (HMM), and taxonomic assignment to metagenome-assembled genomes (MAGs).

	16S rRNA gene vs. *rpoB* gene	**16S rRNA gene vs. MAGs**	***rpoB* gene vs. MAGs**
Observed Correlation	0.83	0.63	0.52
***p* value**	0.001	0.001	0.001

The average agglomeration method was used for hierarchical clustering of samples. Clustering performed using the single and complete agglomeration methods produced similar dendrograms. Although hierarchical clustering produced similar clusters of samples among different datasets, there were discrepancies in the clusters generated from the *rpoB* dataset in comparison to the other two datasets. The dendrograms all consistently clustered Baker Lake WSP samples with samples from the first time point from the Cambridge Bay WSP (Figure 3.16). This trend was consistent across all analyses done, indicating that those samples were more similar to each other than to other samples, despite having few ASVs in common. Inconsistencies occur with Lagoon Lake samples that clustered with Baker Lake and reference lakes samples in the *rpoB* dendrogram, but clustered more closely with Kugluktuk WSP and Cambridge Bay WSP second time point samples in the other dendrograms. The family-level profile shows that Lagoon Lake samples had many families in common with the microbial communities of Airplane Lake, Baker Lake, and reference lakes (Figure 3.14).

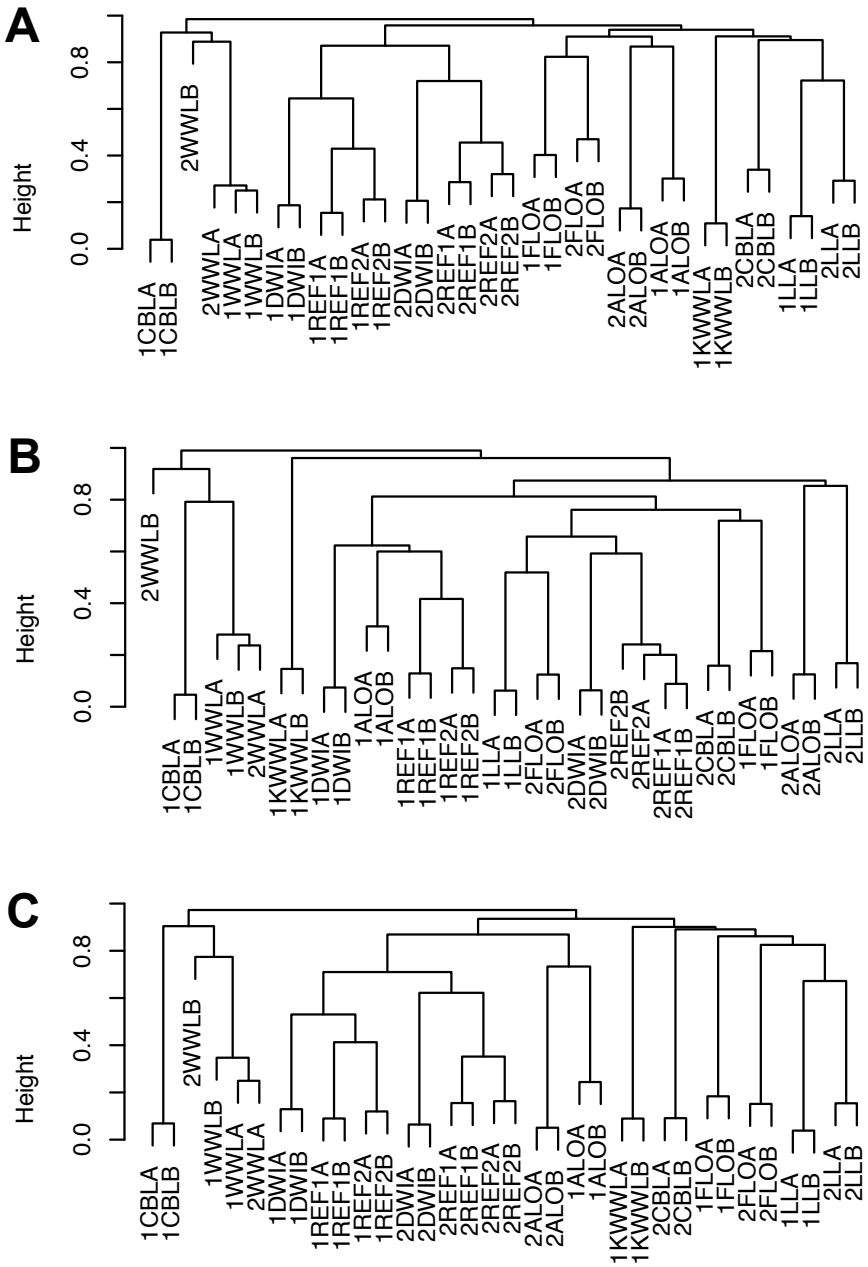


Figure 3.16: Dendrograms based on hierarchical clustering done using the average agglomeration method and Bray-Curtis distances for (A) 16S rRNA gene amplicons, (B) taxonomic assignment to sequences corresponding to the rpoB hidden Markov model (HMM), and (C) taxonomic assignment to metagenome-assembled genomes (MAGs). Baker Lake sample names correspond to sites indicated in Figure 2.2. CBL refers to the Cambridge Bay waste stabilization pond (WSP). KWWL refers to the Kugluktuk WSP. Sample names beginning with 1 correspond to samples collected during the first time point, and samples names beginning with 2 correspond to samples collected during the second time point. Letters A and B at the end of sample names refer to replicates at the same site and time point.

3.4 Antibiotic Resistance Genes in Metagenomes

Antibiotic resistance genes were identified and quantified from metagenomic reads using ShortBRED[36]. The ShortBRED tool was used with pre-computed markers for ARGs based on the CARD to search through forward and reverse metagenomic reads separately and quantify relative abundances of ARG families within the reads. These relative abundances were normalized to RPKM and visualized (Figure 3.17). The RPKM values of forward and reverse reads for each sample were summed to determine which genes to use in downstream analyses.

Detection and quantification of antibiotic resistance genes from raw metagenomic reads revealed an enrichment of ARGs in WSP samples (Figure 3.17). The pre-computed ARG markers used were manually grouped into larger classes (Table A2). Although many sequences associated with the class "genes modulating resistance" were identified, the three WSPs and Lagoon Lake appeared to have an increased abundance of genes related to macrolide resistance, class A beta-lactamases, ABC-F ribosomal protection, and rRNA methyltransferases. Within these four ARG classes, nine genes were selected for further analysis because they were present at a level of >10 RPKMs in at least two sample sites (Table 3.5). Background abundances were calculated using all samples that did not cluster with the WSPs and Lagoon Lake, including both duplicates of the second time point from the Cambridge Bay WSP (Figure 3.17). The reads containing these nine gene families were mapped to bins associated with MAGs, and five MAGs were identified. Even though none of the bins that were associated with the MAGs that were selected as the best bin during dereplication, it is still possible that the representative best bins contain these genes. Further

investigation needs to be done to better connect the resistome to microbial community members.

Four of the five ARG-containing MAGs were associated with genera that had been identified in the indicator species analysis (Table 3.2). These included *Psychrobacter*, *Trichococcus*, *Acetoanaerobium*, and *Enterococcus* (Table 3.6). The MAG associated with *Psychrobacter* (MAG030) was highly abundant in the Cambridge Bay WSP, whereas the other three MAGs associated with indicator species were found at much lower abundances. Although it has not been confirmed whether these bacteria contained the highly abundant ARGs, there is some evidence that wastewater indicator species encode ARGs. However, the wastewater indicators, as previously described, appear to be contained within the Baker Lake WSP and Lagoon Lake, and are only found at very low abundance at further downstream sites. Therefore, the Baker Lake WSP system may limit the spread of potentially pathogenic microorganisms containing ARGs into bodies of water that are regularly used for recreational purposes.

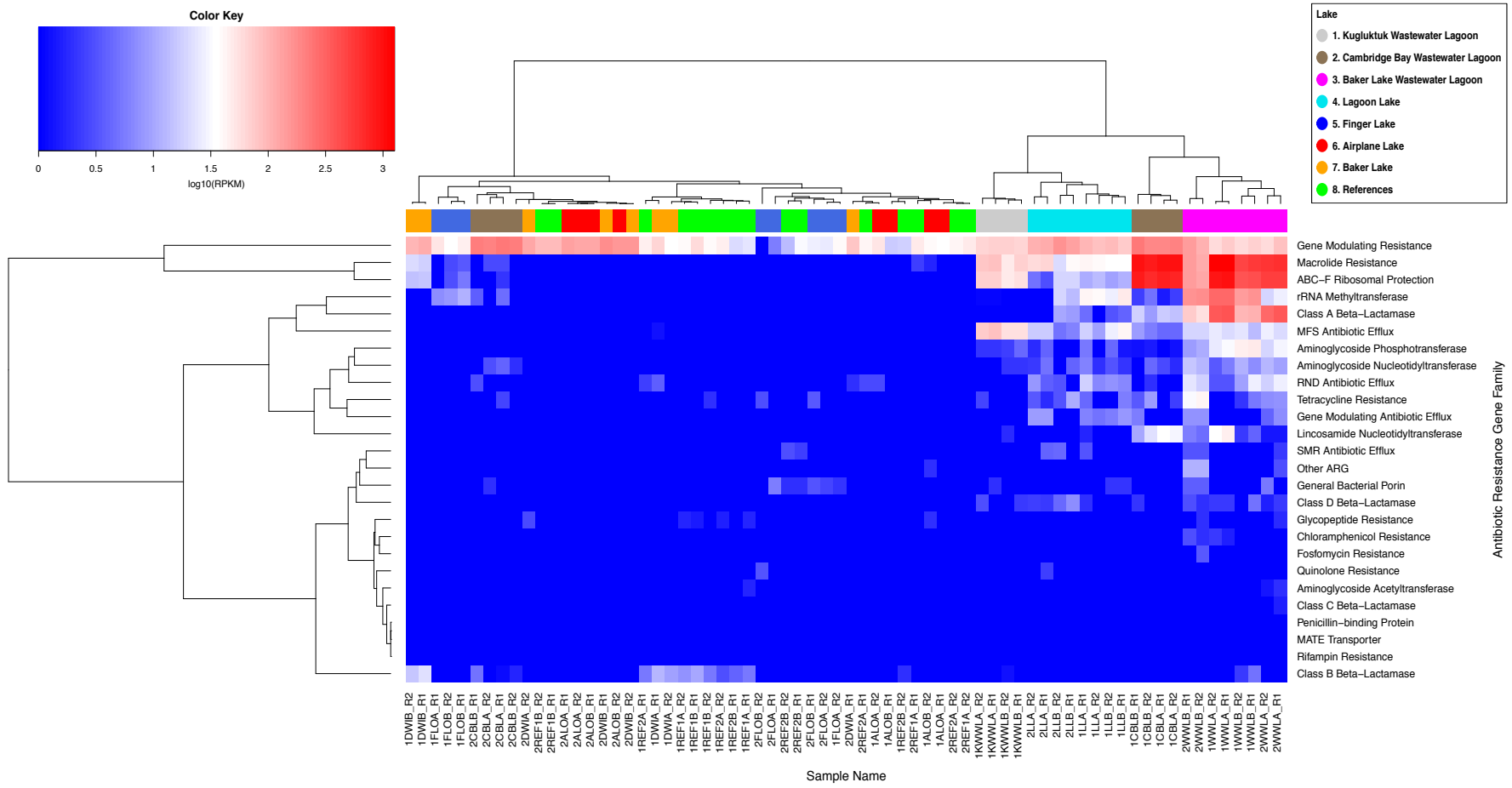


Figure 3.17: Heatmap of log10-transformed RPKM values of antibiotic resistance gene (ARG) families based on raw metagenomic reads. Baker Lake sample names correspond to sites indicated in Figure 2.2. CBL refers to the Cambridge Bay waste stabilization pond (WSP). KWWL refers to the Kugluktuk WSP. Sample names beginning with the number 1 correspond to samples collected during the first time point, and those beginning with the number 2 correspond to samples collected during the second time point. The letters A and B at the end of sample names refer to replicates at the same site and time point. Dendrograms for samples and ARG families have also been included

Table 3.5: Abundant antibiotic resistance gene (ARG) families selected for downstream analysis. The RPKM values for forward and reverse reads from the same sampling site were summed. The RPKM ranges for duplicates were included for waste stabilization pond (WSP) and Lagoon Lake sampling sites that clustered together (Figure 3.17). Background abundances were also determined by calculating the average and standard deviation using all other samples. Bolded entries refer to those that contained >10 RPKMs in at least one of the duplicates.

Resistance Class	CARD gene family	RPKMs						
		Background abundance	Time Point 1				Time Point 2	
			Cambridge Bay WSP	Kugluktuk WSP	Lagoon Lake	Baker Lake WSP	Lagoon Lake	Baker Lake WSP
Class A Beta-Lactamase	CARB-16	0	**25.03-29.69**	0	0	**189.08-727.71**	0-13.82	**124.43-630.88**
rRNA Methyltransferase	*ermB*	0	1.87-2.94	0	1.26-1.66	**97.94-264.40**	0-4.73	**28.52-236.33**
rRNA Methyltransferase	*ermF*	1.38 ± 4.02	0-2.50	0	**71.78-77.04**	**204.94-307.13**	**0-26.09**	3.43-5.03
ABC-F Ribosomal Protection	*mefA*	0	2.53-2.65	4.06-4.39	0	**12.05-16.19**	0	**11.58-53.21**
ABC-F Ribosomal Protection	*mel*	0	**0-11.50**	0	0	0-6.15	0	**1.93-15.11**
Macrolide Resistance	*mphA*	0	0-1.82	0-1.78	**34.63-34.76**	**5.34-58.25**	**16.98-75.90**	4.70-5.63
Macrolide Resistance	*mphE*	1.99 ± 7.93	**1883.07-1918.60**	**115.00-164.04**	**21.80-25.77**	**964.03-2437.73**	**7.03-31.61**	**249.08-1213.86**
Macrolide Resistance	*mrx*	0	0-1.59	0	**11.89-19.68**	**4.74-35.69**	**10.24-40.80**	5.22-6.96
ABC-F Ribosomal Protection	*msrE*	1.61 ± 6.08	**1508.78-1524.03**	**100.46-132.75**	**18.00-23.17**	**801.25-1892.30**	6.13-36.00	**184.63-1009.19**

Table 3.6: Metagenome-assembled genomes (MAGs) identified as being associated with bins that contain reads corresponding to antibiotic resistance genes (ARGs). Asterisks indicate taxonomic classifications that were previously identified as wastewater indicator organisms.

MAG ID	Associated Taxonomy
MAG030	*Psychrobacter**
MAG040	*Trichococcus**
MAG058	*Acetoanaerobium sticklandii**
MAG063	*Mycobacteriaceae*
MAG080	*Enterococcus**

Of the nine most abundant ARG families (Table 3.5), all but CARB-16 confer resistance to macrolide antibiotics. These ARG families were found in higher abundances in Baker Lake WSP samples than in other WSP and Lagoon Lake samples (Table 3.5). Macrolide antibiotics have been used to treat clinical infections for several decades[131] and are most commonly used to treat infections of the upper respiratory tract and skin infections[132]. In 2016 there was an outbreak of whooping cough among community members in Baker Lake[133]. Currently, common treatment for whooping cough includes taking antibiotics such as azithromycin, clarithromycin, or erythromycin, all of which are macrolide antibiotics[134]. Although speculative, it is possible that Baker Lake residents were prescribed macrolide antibiotics to treat various illnesses, which may have included whooping cough, thus releasing low concentrations of antibiotics into the environment that accumulated over time. This may have caused selection of bacteria containing ARGs and transfer of these genes to other bacteria through HGT, thus creating an abundance of ARGs concentrated in the WSP.

The ARBs may also have the opportunity to transfer resistance genes to pathogenic bacteria, specifically human pathogens, which coexist with other ARB and ARGs in wastewater[135].

3.4.1 CARB-16

CARB-16, also referred to as bla_{RTG-6}, belongs to a group of Class A beta-lactamases that are able to hydrolyze carbenicillin[136]. This gene has been identified on the pKLH80 plasmid, which is able to replicate in strains of *Acinetobacter* and *Psychrobacter*[137]. The plasmid itself was isolated from a psychrotolerant strain of *Psychrobacter maritimus* from permafrost[137]. Bins associated with MAG030, but not picked as the representative bin, contained the most abundant ARGs (Table 3.5). This MAG was found to be associated with the genus *Psychrobacter* and was detected within the Baker Lake WSP but was not detected further downstream. Large proportions of *Acinetobacter* were also identified within the microbial community of the Baker Lake WSP. This is concerning as species of *Acinetobacter* have been identified as human pathogens[121] and if *Psychrobacter* species containing resistance genes for CARB-16 are able to pass these genes onto *Acinetobacter* species, this could create an antibiotic resistant strain of a potential human pathogen. However, further investigation must be done in order to confirm the presence of a CARB-16 gene, and possibly the pKLH80 plasmid itself, within the genome of a *Psychrobacter* species.

3.4.2 *ermB* and *ermF*

Erm antibiotic resistance genes are a family of 23S methyltransferases that methylate the 23S rRNA gene, reducing the affinity of macrolide antibiotics to bacterial ribosomes[138].

The *ermB* gene is known to be induced by the presence of erythromycin[139]. Both *ermB* and *ermF* have been detected in various *Enterococcus* species, some of which are important human pathogens[140,141]. MAG080, corresponding to the genus *Enterococcus*, was associated with bins containing the most abundant ARGs (Table 3.6), and was also detected in the Baker Lake WSP. An ASV corresponding to *Enterococcus* was also identified as a wastewater indicator species (Figure 3.6). However, MAG080 and the indicator ASV corresponding to *Enterococcus* were not detected at sites downstream of the WSP, and therefore the WSP system may be removing potentially pathogenic species of *Enterococcus* from the lake system. Alternatively, dilution effects may be decreasing the abundance of this ASV to below detection limits.

3.4.3 *mefA* and *mel*

The *mefA* and *mel* genes are a part of the same operon as *mefE*[142]. The most current version of the CARD classifies these genes under the same accession number, possibly due to their similarities in function and sequence. These genes encode ATP-binding cassette (ABC) ribosomal protection proteins that catalyze macrolide efflux[142]. However, the expression of *mel* is insufficient to confer resistance to macrolides and also requires *mefE* expression for efflux[142]. The *mel* gene was found in lower abundance than the other highly abundant ARGs (Table 3.5), which may be due to the lack of *mefE*, and therefore there is likely no resistance conferred to organisms within the metagenomes containing the *mel* gene. The *mefA* gene was also found in relatively low abundance compared to the other highly abundant ARGs. However, the *mefE* gene has also been described as a subclass of the *mefA*

gene[143], suggesting that *mefA* and *mefE* may be the same gene. If this is the case, *mel* combined with *mefA* would be sufficient to confer resistance to macrolides. Further investigation must be done in order to determine if these genes were found within the same bacteria, whether they are able to cause resistance to macrolides together.

3.4.4 *mphE* and *msrE*

The *mphE* and *msrE* genes have been found in close proximity on a plasmid in an *Acinetobacter baumannii* isolate[144], as well as on plasmids found in *Klebsiella pneumoniae*, *Escherichia coli*, and *Citrobacter freundii*[145]. The *mphE* gene encodes a macrolide kinase that phosphorylates the 2'-hydroxyl group of macrolide antibiotics[146]. The *msrE* gene encodes an ABC ribosomal protection protein that protects ribosomes by binding to macrolides and reducing affinity of the antibiotics to the ribosomes[147]. These genes were found at similar abundance in the metagenomes of Baker Lake's WSP, the first time point of Cambridge Bay's WSP, Kugluktuk's WSP and Lagoon Lake, which could be indicative of co-occurrence within the same bacteria, possibly on the same plasmids.

3.4.5 *mphA* and *mrx*

The *mphA* gene encodes a macrolide 2'-phosphotransferase that inactivates macrolide antibiotics[148]. However, the protein encoded by the *mrx* gene is an unidentified hydrophobic protein[148]. Although the *mrx* gene no longer exists in the CARD, it has been found to occur on the same operon as *mphA*, and both genes are thought to be required for

macrolide resistance[148,149]. These two genes were also found at similar abundance to one another within metagenomes corresponding to the Baker Lake WSP and Lagoon Lake.

Chapter 4

Conclusions and Future Directions

The unique and frigid climate of Arctic environments provides the need for specialized wastewater treatment solutions. In general, the microbial communities of Arctic WSPs are not well understood. This research has provided important insight into the composition of microbial communities in Arctic WSPs through 16S rRNA gene amplicon and metagenome analysis. The dominant phylum identified within the WSPs was *Proteobacteria*, which was consistent across all three WSPs observed in this study, with high abundances of ASVs associated with the family *Moraxellaceae* in both Baker Lake and Cambridge Bay samples. These results were consistent across 16S rRNA gene amplicons, taxonomic assignments to sequences identified by the *rpoB* HMM and taxonomic assignments to MAGs. Despite these similarities across the different WSPs, at the ASV level, the microbial communities were found to be quite different, with very few overlapping ASVs. The dominant ASV in the Baker Lake WSP was *Acinetobacter*, which was found at >50% relative abundance, whereas *Psychrobacter* was the dominant ASV in the Cambridge Bay WSP, at >80% relative abundance. Although the Kugluktuk WSP did not have a single, dominant ASV at the family level, *Burkholderiaceae* and *Caulobacteraceae* were the most prominent families, according to 16S rRNA gene amplicons. This research demonstrated that, despite the many differences in community composition at the ASV level, wastewater from different locations contain common groups of bacteria that are distinct from other Arctic freshwater sites. Future research involving the three WSPs observed should aim to compare the downstream lake or

wetland systems of the WSPs in detail to determine the efficiency of each WSP in the removal of potentially harmful wastewater microorganisms.

This research characterized the microbial communities of Arctic wastewater, with a focus on the Baker Lake WSP system. Over the next five years, the Baker Lake WSP system will likely undergo infrastructure upgrades to improve the performance of the system in the removal of wastewater contaminants from the environment, in an attempt to meet government guidelines with respect to wastewater effluent standards. The research presented here reveals both spatial and temporal changes in microbial communities across the Baker Lake WSP and downstream lake systems (Figures 3.1-3.4). The microbial communities of sites located further downstream from the WSP are less similar to sites located closer to the WSP, as well as the WSP itself. Temporally, shifts in relative abundance were observed at the phylum level where ten days between sampling times produced very different taxonomic profiles. In lakes downstream of the WSP, microbial communities shifted from being dominated by microorganisms associated with *Bacteroidetes*, to containing higher levels of organisms associated with *Actinobacteria*, *Cyanobacteria*, *Patescibacteria*, and *Verrucomicrobia*. These patterns were consistent across the three different datasets, which was further confirmed with Mantel tests providing significant, positive correlations. This research has also shown that many wastewater indicator ASVs were not detected in downstream lakes, demonstrating that the WSP effluent dilutes effectively within downstream receiving waters. Future research should ideally test whether these results are consistent across a multi-year timeframe and in multiple seasons.

Another important objective of this research was to identify and quantify antibiotic resistance genes within the three WSPs, as well as sites downstream of the Baker Lake WSP. Wastewater is known as a hotspot for ARG dissemination and evolution. This was further confirmed through this research, because an enrichment of ARGs was detected within WSP samples. This was consistent across all three WSPs. In particular, genes related to macrolide resistance were highly abundant in all three WSPs (Table 3.5). Although this may be related to the use of antibiotics within the local community, no information about the administration of antibiotics in Arctic communities was found. This research also displayed evidence of ARGs associated with specific MAGs within the metagenomes. Five MAGs were identified as being associated with bins that contained reads identified as ARG sequences. However, more evidence is required to better connect the genes identified with their respective hosts. This could be accomplished by sequencing plasmids and associating them with their host by observing DNA methylation patterns using methods proposed by Beaulaurier and colleagues[150]. Plasmids could also be assembled using the metaplasmidSPAdes tool[151], which could potentially be incorporated into the ATLAS pipeline used in this research.

Overall, this research has demonstrated the composition of microbial communities in the Canadian Arctic, specifically focusing on the unique microbial assemblages in Arctic WSPs and associated freshwater lakes. The 16S rRNA gene amplicon and metagenome sequencing approaches allowed for in-depth profiling of the microbial communities. Although further research is required to be able to fully understand the microorganisms in Arctic wastewater, and the efficiency of WSPs in the removal of wastewater contaminants,

this study has provided a baseline characterization of microbial communities in Arctic WSPs which will aid the future development of a safe and effective wastewater treatment system for the Baker Lake community.

Bibliography

1. Chambers, P. A. *et al.* Impacts of municipal wastewater effluents on Canadian waters: a review. *Water Qual. Res. J. Canada* **32**, 659–713 (1997).

2. Gunnarsdóttir, R., Jenssen, P. D., Erland Jensen, P., Villumsen, A. & Kallenborn, R. A review of wastewater handling in the Arctic with special reference to pharmaceuticals and personal care products (PPCPs) and microbial pollution. *Ecol. Eng.* **50**, 76–85 (2013).

3. Samudro, G. & Mangkoedihardjo, S. Review on BOD, COD and BOD/COD ratio: a triangle zone for toxic, biodegradable and stable levels. *Int. J. Acad. Res.* **2**, 235–239 (2010).

4. Harper, D. What is eutrophication? in *Eutrophication of Freshwaters* 1–28 (Springer Netherlands, 1992).

5. Hu, Z., Houweling, D. & Dold, P. Biological nutrient removal in municipal wastewater treatment: new directions in sustainability. *J. Environ. Eng.* **138**, 307–317 (2012).

6. Heisler, J. *et al.* Eutrophication and harmful algal blooms: A scientific consensus. *Harmful Algae* **8**, 3–13 (2008).

7. Wakelin, S. A., Colloff, M. J. & Kookana, R. S. Effect of wastewater treatment plant effluent on microbial function and community structure in the sediment of a freshwater stream with variable seasonal flow. *Appl. Environ. Microbiol.* **74**, 2659–68 (2008).

8. Cébron, A., Coci, M., Garnier, J. & Laanbroek, H. J. Denaturing gradient gel electrophoretic analysis of ammonia-oxidizing bacterial community structure in the lower Seine River: impact of Paris wastewater effluents. *Appl. Environ. Microbiol.* **70**, 6726–37 (2004).

9. Goni-Urriza, M., Capdepuy, M., Raymond, N., Quentin, C. & Caumette, P. Impact of an urban effluent on the bacterial community structure in the Arga River (Spain), with special references to culturable Gram-negative rods. *Can. J. Microbiol.* **45**, 826–832 (1999).

10. Drury, B., Rosi-Marshall, E. & Kelly, J. J. Wastewater treatment effluent reduces the abundance and diversity of benthic bacterial communities in urban and suburban rivers. *Appl. Environ. Microbiol.* **79**, 1897–905 (2013).

11. Atashgahi, S. *et al.* Impact of a wastewater treatment plant on microbial community composition and function in a hyporheic zone of a eutrophic river. *Sci. Rep.* **5**, 17284 (2015).

12. Eckert, E. M. *et al.* Microplastics increase impact of treated wastewater on freshwater microbial community. *Environ. Pollut.* **234**, 495–502 (2018).

13. Henze, M., van Loosdrecht, M. C. M., Ekama, G. A. & Brdjanovic, D. *Biological wastewater treatment: principles, modelling and design*. (IWA Publishing, 2008).

14. Akpor, O. B. Wastewater effluent discharge: effects and treatment processes. in *Proceedings of 2011 3rd International Conference on Chemical, Biological and Environmental Engineering (ICBEE 2011)* 85–91 (2011).

15. Holeton, C., Chambers, P. A. & Grace, L. Wastewater release and its impacts on Canadian waters. *Can. J. Fish. Aquat. Sci.* **68**, 1836–1859 (2011).

16. Wiesmann, U. Biological nitrogen removal from wastewater. in 113–154 (Springer, Berlin, Heidelberg, 1994).

17. Czerwionka, K., Makinia, J., Pagilla, K. R. & Stensel, H. D. Characteristics and fate of organic nitrogen in municipal biological nutrient removal wastewater treatment plants. *Water Res.* **46**, 2057–2066 (2012).

18. Yin, Z., Bi, X. & Xu, C. Ammonia-Oxidizing Archaea (AOA) play with Ammonia-Oxidizing Bacteria (AOB) in nitrogen removal from Wastewater. *Archaea* **2018**, 1–9

(2018).

19. Lu, H., Chandran, K. & Stensel, D. Microbial ecology of denitrification in biological wastewater treatment. *Water Res.* **64**, 237–254 (2014).

20. Schindler, D. W. Eutrophication and recovery in experimental lakes: implications for lake management. *Science (80-.).* **184**, 897–899 (1974).

21. Yeoman, S., Stephenson, T., Lester, J. N. & Perry, R. The removal of phosphorus during wastewater treatment: a review. *Environ. Pollut.* **49**, 183–233 (1988).

22. Bond, P. L., Erhart, R., Wagner, M., Keller, J. & Blackall, L. L. Identification of some of the major groups of bacteria in efficient and nonefficient biological phosphorus removal activated sludge systems. *Appl. Environ. Microbiol.* **65**, 4077–84 (1999).

23. Sun, W. *et al.* Effective treatment of high phosphorus pharmaceutical wastewater by chemical precipitation. *Can. J. Chem. Eng.* **95**, 1585–1593 (2017).

24. Oleszkiewicz, J. *Options for improved nutrient removale and recovery from municipal wastewater in the Canadian context*. (2015).

25. Alonso, A., Sanchez, P. & Martinez, J. L. Environmental selection of antibiotic resistance genes. *Environ. Microbiol.* **3**, 1–9 (2001).

26. Finley, R. L. *et al.* The scourge of antibiotic resistance: the important role of the environment. *Clin. Infect. Dis.* **57**, 704–710 (2013).

27. Garcia-Armisen, T. *et al.* Antimicrobial resistance of heterotrophic bacteria in sewage-contaminated rivers. *Water Res.* **45**, 788–796 (2011).

28. Karkman, A., Do, T. T., Walsh, F. & Virta, M. P. J. Antibiotic-resistance genes in waste water. *Trends Microbiol.* **26**, 220–228 (2018).

29. Huddleston, J. R. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infect. Drug Resist.* **7**, 167–76 (2014).

30. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: building the web

of life. *Nat. Rev. Genet.* **16**, 472–482 (2015).

31. Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3**, 711–721 (2005).

32. Jacquiod, S. *et al.* Deciphering conjugative plasmid permissiveness in wastewater microbiomes. *Mol. Ecol.* **26**, 3556–3571 (2017).

33. Li, A.-D. *et al.* Effects of sample preservation and DNA extraction on enumeration of antibiotic resistance genes in wastewater. *FEMS Microbiol. Ecol.* (2017).

34. Bueno, I. *et al.* Impact of point sources on antibiotic resistance genes in the natural environment: a systematic review of the evidence. *Anim. Heal. Res. Rev.* 1–16 (2017).

35. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–57 (2013).

36. Kaminski, J. *et al.* High-specificity targeted functional profiling in microbial communities with ShortBRED. *PLOS Comput. Biol.* **11**, e1004557 (2015).

37. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).

38. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).

39. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

40. Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).

41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

42. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

43. von Sperling, M. *Biological wastewater wreatment series: volume three - waste stabilisation ponds*. (IWA Publishing, 2007).

44. Tartakovsky, B., Kleiner, Y. & Manuel, M.-F. Bioelectrochemical anaerobic sewage treatment technology for Arctic communities. *Environ. Sci. Pollut. Res.* **25**, 32844–32850 (2018).

45. Tilley, E., Ulrich, L., Lüthi, C., Reymond, P. & Zurbrügg, C. *Compendium of sanitation systems and technologies*. (2014).

46. Ragush, C. M. *et al.* Performance of municipal waste stabilization ponds in the Canadian Arctic. *Ecol. Eng.* **83**, 413–421 (2015).

47. Verbyla, M., von Sperling, M. & Maiga, Y. Waste Stabilization Ponds. in *Global Water Pathogen Project* 3–18 (2017).

48. Huang, Y., Truelstrup Hansen, L., Ragush, C. M. & Jamieson, R. C. Disinfection and removal of human pathogenic bacteria in Arctic waste stabilization ponds. *Environ. Sci. Pollut. Res. Int.* (2017).

49. Zhang, L., Lee, Y.-W. & Jahng, D. Ammonia stripping for enhanced biomethanization of piggery wastewater. *J. Hazard. Mater.* **199–200**, 36–42 (2012).

50. Daley, K., Jamieson, R., Rainham, D. & Truelstrup Hansen, L. Wastewater treatment and public health in Nunavut: a microbial risk assessment framework for the Canadian Arctic. *Environ. Sci. Pollut. Res. Int.* **25**, 32860–32872 (2018).

51. Krkosek, W. H. *et al.* Treatment performance of wastewater stabilization ponds in Canada's far North. in *Cold Regions Engineering 2012* (2012).

52. Hayward, J. L., Jackson, A. J., Yost, C. K., Truelstrup Hansen, L. & Jamieson, R. C. Fate of antibiotic resistance genes in two Arctic tundra wetlands impacted by municipal wastewater. *Sci. Total Environ.* **642**, 1415–1428 (2018).

53. Neudorf, K. D. *et al.* Antibiotic resistance genes in municipal wastewater treatment systems and receiving waters in Arctic Canada. *Sci. Total Environ.* **598**, 1085–1094

(2017).

54. Department of Fisheries and Oceans. Fisheries Act: wastewater systems effluent regulations. *Canada Gaz. Part II* **146**, (2012).

55. Frank-Fahle, B. A., Yergeau, É., Greer, C. W., Lantuit, H. & Wagner, D. Microbial functional potential and community composition in permafrost-affected soils of the NW Canadian Arctic. *PLoS One* **9**, e84761 (2014).

56. Hansen, A. A. *et al.* Viability, diversity and composition of the bacterial community in a high Arctic permafrost soil from Spitsbergen, Northern Norway. *Environ. Microbiol.* **9**, 2870–2884 (2007).

57. Larose, C., Dommergue, A. & Vogel, T. M. The dynamic Arctic snow pack: an unexplored environment for microbial diversity and activity. *Biol. 2013, Vol. 2, Pages 317-330* **2**, 317–330 (2013).

58. Nemergut, D. R. *et al.* Structure and function of alpine and arctic soil microbial communities. *Res. Microbiol.* **156**, 775–784 (2005).

59. Wilhelm, R. C., Niederberger, T. D., Greer, C. & Whyte, L. G. Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. *Can. J. Microbiol.* **57**, 303–315 (2011).

60. Steven, B., Pollard, W. H., Greer, C. W. & Whyte, L. G. Microbial diversity and activity through a permafrost/ground ice core profile from the Canadian high Arctic. *Environ. Microbiol.* **10**, 3388–3403 (2008).

61. Steven, B. *et al.* Characterization of the microbial diversity in a permafrost sample from the Canadian high Arctic using culture-dependent and culture-independent methods. *FEMS Microbiol. Ecol.* **59**, 513–523 (2007).

62. Methé, B. A. *et al.* The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10913–10918 (2005).

63. Yergeau, E. *et al.* Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Sci. Rep.* **7**, 42242 (2017).

64. Chrismas, N. A. M., Barker, G., Anesio, A. M. & Sánchez-Baracaldo, P. Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium *Phormidesmis priestleyi* BC1401. *BMC Genomics* **17**, 533 (2016).

65. Crump, B. C., Amaral-Zettler, L. A. & Kling, G. W. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J.* **6**, 1629–1639 (2012).

66. Gonzalez-Martinez, A. *et al.* Microbial ecology of full-scale wastewater treatment systems in the Polar Arctic Circle: archaea, bacteria and fungi. *Sci. Rep.* **8**, 2208 (2018).

67. McIlroy, S. J. *et al.* MiDAS: the field guide to the microbes of activated sludge. *Database (Oxford)*. **2015**, bav062 (2015).

68. Huang, Y. *et al.* Removal of human pathogens in wastewater stabilization ponds in Nunavut. in *CSCE 2014 13th International Environmental Specialty Conference* (2014).

69. Reuter, J. A., Spacek, D. V & Snyder, M. P. High-throughput sequencing technologies. *Mol. Cell* **58**, 586–97 (2015).

70. Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D. & Dangl, J. L. Practical innovations for high-throughput amplicon sequencing. *Nat. Methods* **10**, 999–1002 (2013).

71. Chen, K. & Pachter, L. Bioinformatics for whole-genome shotgun sequencing of microbial communities. *PLoS Comput. Biol.* **1**, e24 (2005).

72. Ward, D. M., Weller, R. & Bateson, M. M. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**, 63–65 (1990).

73. Gray, M. W., Sankoff, D. & Cedergren, R. J. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* **12**, 5837–52 (1984).

74. Zaheer, R. *et al.* Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci. Rep.* **8**, 5890 (2018).

75. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

76. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

77. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

78. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).

79. Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. & Goodman, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **5**, R245-9 (1998).

80. Kieser, S., Brown, J., Zdobnov, E. M., Trajkovski, M. & McCue, L. A. ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. *bioRxiv* 737528 (2019).

81. Bushnell, B. BBTools. (2019). Available at: sourceforge.net/projects/bbmap/.

82. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

83. Li, D. *et al.* MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).

84. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).

85. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).

86. Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2016).

87. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).

88. Strous, M., Kraft, B., Bisdorf, R. & Tegetmeyer, H. E. The binning of metagenomic contigs for microbial physiology of mixed cultures. *Front. Microbiol.* **3**, 410 (2012).

89. Sieber, C. M. K. *et al.* Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836–843 (2018).

90. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).

91. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

92. Huerta-Cepas, J. *et al.* eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286–D293 (2016).

93. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).

94. von Meijenfeldt, F. A. B. von, Arkhipova, K., Cambuy, D. D., Coutinho, F. H. & Dutilh, B. E. Robust taxonomic classification of uncharted microbial sequences and

bins with CAT and BAT. *bioRxiv* 530188 (2019).

95. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **44**, D7–D19 (2016).

96. Nunavut Bureau of Statistics. Population estimates. **2017**, (2016).

97. Google Maps. Nunavut. (2018). Available at: https://www.google.ca/maps/place/Nunavut/@67.3334876,-103.3705765,4.7z/data=!4m5!3m4!1s0x4d0823492a24a5ed:0xffd371054ff1a9b6!8m2!3d70.2997711!4d-83.107577. (Accessed: 15th February 2018)

98. Google Earth 7.3.2. Baker Lake, Nunavut, Canada, 64°19’7.40"N, 96° 0’37.94"W. (2019).

99. Bartram, A. K., Lynch, M. D. J., Stearns, J. C., Moreno-Hagelsieb, G. & Neufeld, J. D. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl. Environ. Microbiol.* **77**, 3846–52 (2011).

100. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).

101. Quince, C., Lanzen, A., Davenport, R. J. & Turnbaugh, P. J. Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**, 38 (2011).

102. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

103. Dufrêne, M. & Legendre, P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* **67**, 345–366 (1997).

104. Bakker, J. D. Increasing the utility of Indicator Species Analysis. *J. Appl. Ecol.* **45**, 1829–1835 (2008).

105. De Cáceres, M. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3574 (2009).

106. Petrenko, P., Lobb, B., Kurtz, D. A., Neufeld, J. D. & Doxey, A. C. MetAnnotate: function-specific taxonomic profiling and comparison of metagenomes. *BMC Biol.* **13**, 92 (2015).

107. Haft, D. H., Selengut, J. D. & White, O. The TIGRFAMs database of protein families. *Nucleic Acids Res.* **31**, 371–3 (2003).

108. Oksanen, J. *et al.* vegan: community ecology package. R package version 2.5-6. (2019).

109. Bougeard, S. & Dray, S. Supervised multiblock analysis in R with the ade4 package. *J. Stat. Softw.* **86**, 1–17 (2018).

110. R Core Team. R: A language and environment for statistical computing. (2019).

111. McCune, B. & Grace, J. B. *Analysis of ecological communities*. (MjM Software Design, 2002).

112. Moyer, C. L. & Morita, R. Y. Psychrophiles and psychrotrophs. in *Encyclopedia of Life Sciences* (John Wiley & Sons, Ltd, 2007).

113. Bernardet, J.-F. & Nakagawa, Y. An introduction to the family *Flavobacteriaceae*. in *The Prokaryotes* 455–480 (Springer New York, 2006).

114. Gong, X. *et al.* High quality draft genome sequence of *Janthinobacterium psychrotolerans* sp. nov., isolated from a frozen freshwater pond. *Stand. Genomic Sci.* **12**, 8 (2017).

115. Schloss, P. D. *et al.* Psychrotrophic strain of *Janthinobacterium lividum* from a cold Alaskan soil produces prodigiosin. *DNA Cell Biol.* **29**, 533–541 (2010).

116. Wiedmann-al-Ahmad, M., Tichy, H. V & Schön, G. Characterization of *Acinetobacter* type strains and isolates obtained from wastewater treatment plants by PCR

fingerprinting. *Appl. Environ. Microbiol.* **60**, 4066–71 (1994).

117. Slekovec, C. *et al.* Tracking down antibiotic-resistant *Pseudomonas aeruginosa* isolates in a wastewater network. *PLoS One* **7**, e49300 (2012).

118. Igbinosa, I. H., Nwodo, U. U., Sosa, A., Tom, M. & Okoh, A. I. Commensal *Pseudomonas* species isolated from wastewater and freshwater milieus in the Eastern Cape Province, South Africa, as reservoir of antibiotic resistant determinants. *Int. J. Environ. Res. Public Health* **9**, 2537–49 (2012).

119. Ng, C. *et al.* Characterization of metagenomes in urban aquatic compartments reveals high prevalence of clinically relevant antibiotic resistance genes in wastewaters. *Front. Microbiol.* **8**, 2200 (2017).

120. Beukers, A. G. *et al.* Comparative genomics of *Enterococcus* spp. isolated from bovine feces. *BMC Microbiol.* **17**, 52 (2017).

121. Al Atrouni, A., Joly-Guillou, M.-L., Hamze, M. & Kempf, M. Reservoirs of Non-*baumannii Acinetobacter* species. *Front. Microbiol.* **7**, 49 (2016).

122. Higgins, P. G., Hrenovic, J., Seifert, H. & Dekic, S. Characterization of *Acinetobacter baumannii* from water and sludge line of secondary wastewater treatment plant. *Water Res.* **140**, 261–267 (2018).

123. Heylen, K., Lebbe, L. & De Vos, P. *Acidovorax caeni* sp. nov., a denitrifying species with genetically diverse isolates from activated sludge. *Int. J. Syst. Evol. Microbiol.* **58**, 73–77 (2008).

124. Schulze, R. *et al.* Genotypic diversity of *Acidovorax* strains isolated from activated sludge and description of *Acidovorax defluvii* sp. nov. *Syst. Appl. Microbiol.* **22**, 205–214 (1999).

125. McIlroy, S. J. *et al.* Identification of active denitrifiers in full-scale nutrient removal wastewater treatment systems. *Environ. Microbiol.* **18**, 50–64 (2016).

126. Kaden, R., Sproer, C., Beyer, D. & Krolla-Sidenstein, P. Rhodoferax saidenbachensis

sp. nov., a psychrotolerant very slowly growing bacterium within the family *Comamonadaceae*, proposal of appropriate taxonomic position of *Albidiferax ferrireducens* strain T118T in the genus *Rhodoferax*. *Int. J. Syst. Evol. Microbiol.* **64**, 1186–1193 (2014).

127. Finneran, K. T., Johnsen, C. V & Lovley, D. R. *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *Int. J. Syst. Evol. Microbiol.* **53**, 669–673 (2003).

128. Kosten, S. *et al.* Warmer climates boost cyanobacterial dominance in shallow lakes. *Glob. Chang. Biol.* **18**, 118–126 (2012).

129. Vos, M., Quince, C., Pijl, A. S., de Hollander, M. & Kowalchuk, G. A. A comparison of *rpoB* and 16S rRNA as markers in pyrosequencing studies of bacterial diversity. *PLoS One* **7**, e30600 (2012).

130. Khamis, A., Raoult, D. & La Scola, B. Comparison between *rpoB* and 16S rRNA gene sequencing for molecular identification of 168 clinical isolates of *Corynebacterium*. *J. Clin. Microbiol.* **43**, 1934–6 (2005).

131. Omura, S. *Macrolide antibiotics : chemistry, biology, and practice.* (Academic, 2002).

132. Sturgill, M. G. & Rapp, R. P. Clarithromycin: review of a new macrolide antibiotic with improved microbiologic spectrum and favorable pharmacokinetic and adverse effect profiles. *Ann. Pharmacother.* **26**, 1099–1108 (1992).

133. The Government of Nunavut Department of Health. *Confirmed cases of whooping cough in Baker Lake and Rankin Inlet*. (2016).

134. Gopal, D. P., Barber, J. & Toeg, D. Pertussis (whooping cough). *BMJ* **364**, l401 (2019).

135. Nnadozie, C. F., Kumari, S. & Bux, F. Status of pathogens, antibiotic resistance genes and antibiotic residues in wastewater treatment systems. *Reviews in Environmental Science and Biotechnology* **16**, 491–515 (2017).

136. Labia, R., Guionie, M., Barthélémy, M. & Philippon, A. Properties of three carbenicillin-hydrolysing β-lactamases (CARB) from *Pseudomonas aeruginosa* : identification of a new enzyme. *J. Antimicrob. Chemother.* **7**, 49–56 (1981).

137. Petrova, M., Kurakov, A., Shcherbatova, N. & Mindlin, S. Genetic structure and biological properties of the first ancient multiresistance plasmid pKLH80 isolated from a permafrost bacterium. *Microbiology* **160**, 2253–2263 (2014).

138. Dzyubak, E. & Yap, M. N. The expression of antibiotic resistance methyltransferase correlates with mRNA stability independently of ribosome stalling. *Antimicrob. Agents Chemother.* **60**, 7178–7188 (2016).

139. Min, Y.-H., Kwon, A.-R., Yoon, E.-J., Shim, M.-J. & Choi, E.-C. Translational attenuation and mRNA stabilization as mechanisms of *erm(B)* induction by erythromycin. *Antimicrob. Agents Chemother.* **52**, 1782–9 (2008).

140. Portillo, A. *et al.* Macrolide resistance genes in *Enterococcus* spp. *Antimicrob. Agents Chemother.* **44**, 967–71 (2000).

141. Nowakiewicz, A. *et al.* Characterization of multidrug resistant *E. faecalis* strains from pigs of local origin by ADSRRS-fingerprinting and MALDI-TOF MS; evaluation of the compatibility of methods employed for multidrug resistance analysis. *PLoS One* **12**, e0171160 (2017).

142. Ambrose, K. D., Nisbet, R. & Stephens, D. S. Macrolide efflux in *Streptococcus pneumoniae* is mediated by a dual efflux pump (*mel* and *mef*) and is erythromycin inducible. *Antimicrob. Agents Chemother.* **49**, 4203–9 (2005).

143. Ardanuy, C. *et al.* Distribution of subclasses *mefA* and *mefE* of the *mefA* gene among clinical isolates of macrolide-resistant (M-phenotype) *Streptococcus pneumoniae*, viridans group streptococci, and *Streptococcus pyogenes*. *Antimicrob. Agents Chemother.* **49**, 827–9 (2005).

144. Blackwell, G. A. & Hall, R. M. The tet39 determinant and the msrE-mphE genes in

Acinetobacter plasmids are each part of discrete modules flanked by inversely oriented *pdif* (XerC-XerD) Sites. *Antimicrob. Agents Chemother.* **61**, (2017).

145. Bonnin, R. A., Nordmann, P., Carattoli, A. & Poirel, L. Comparative genomics of IncL/M-type plasmids: evolution by acquisition of resistance genes and insertion sequences. *Antimicrob. Agents Chemother.* **57**, 674–6 (2013).

146. Pawlowski, A. C. *et al.* The evolution of substrate discrimination in macrolide antibiotic resistance enzymes. *Nat. Commun.* **9**, 112 (2018).

147. Dinos, G. P. The macrolide antibiotic renaissance. *Br. J. Pharmacol.* **174**, 2967–2983 (2017).

148. Noguchi, N. *et al.* Nucleotide sequence and characterization of erythromycin resistance determinant that encodes macrolide 2'-phosphotransferase I in Escherichia coli. *Antimicrob. Agents Chemother.* **39**, 2359 (1995).

149. Papagiannitsis, C. C., Miriagou, V., Giakkoupi, P., Tzouvelekis, L. S. & Vatopoulos, A. C. Characterization of pKP1780, a novel IncR plasmid from the emerging *Klebsiella pneumoniae* ST147, encoding the VIM-1 metallo-β-lactamase. *J. Antimicrob. Chemother.* **68**, 2259–2262 (2013).

150. Beaulaurier, J. *et al.* Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat. Biotechnol.* **36**, 61–69 (2018).

151. Antipov, D., Raiko, M., Lapidus, A. & Pevzner, P. A. Plasmid detection and assembly in genomic and metagenomic data sets. *Genome Res.* **29**, 961–968 (2019).

Appendix A1

Table A1: Environmental data from sites in Baker Lake, Nunavut. Site names correspond to those on the map in Figure 2.2. Data from both sampling time points has been included.

Site Name	Latitude	Longitude	Temperature (°C)		pH		Conductivity (µS/cm)		TDS (mg/L)		Salinity (ppt)		DO (mg/L)	
			July 13-16	July 22-24	July 13-16	July 22-24	July 13-16	July 22-24	July 13-16	July 22-24	July 13-16	July 22-24	July 13-16	July 22-24
REF1	64.33715	-95.96131	12.6	15.2	7.07	7.59	81.6	84.1	59.0	59.7	0.04	0.04	9.5	8.2
REF2	64.336836	-96.016599	14.8	16.1	7.38	7.72	112.3	60.4	80.0	43.0	0.06	0.03	9.8	9.1
REF3	64.33707	-96.02652	16.1	16.7	7.62	8.06	101.0	73.8	71.4	52.3	0.05	0.04	9.3	9.7
DWI	64.31557	-96.01727	10.1	11.1	6.94	7.31	108.2	31.0	77.1	22.0	0.05	0.02	10.9	11.0
HCO	64.31519	-96.01374	6.1	11.1	7.14	7.59	231	48.6	165	34.7	0.12	0.02	11.8	11.4
BLM	64.311844	-95.993918	6.1	9.5	6.96	7.47	158.6	31.0	113	22.0	0.08	0.02	12.4	10.9
BLP	64.30668	-95.96240	7.0	9.9	7.02	7.51	142.1	33.0	101	23.7	0.07	0.02	12.8	10.3
ACO	64.30960	-95.97852	8.3	9.2	7.04	7.48	149.7	34.3	107	24.4	0.08	0.02	12.3	11.6
ALO	64.31702	-95.97436	14.6	12.9	9.06	7.62	86.4	79.7	61.2	56.5	0.04	0.04	10.5	11.3
ALI	64.32044	-95.97424	16.2	14.5	9.06	7.73	94.6	82.7	67.4	58.8	0.05	0.04	10.7	10.5
FLI	64.32967	-95.99605	17.2	15.0	9.04	7.16	166.5	202	119	143	0.08	0.10	9.6	7.7
FLO	64.32796	-95.98515	17.4	15.1	9.64	9.13	164	215	117	152	0.08	0.11	10.1	10.8
LL	64.33113	-96.00263	16.1	17.6	9.40	8.95	188.0	289	134	205	0.09	0.14	14.3	10.3
WWL	64.32971	-96.00595	18.7	21.3	7.84	7.79	928	1125	670	794	0.47	0.56	3.3	6.7
BLC	64.31625	-96.04948	10.3	11.5	7.40	7.52	231	44.5	164	31.6	0.12	0.02	10.9	10.9
BL	64.31300	-96.02019	6.2	N/A	7.42	N/A	64.2	N/A	45.5	N/A	0.03	N/A	10.7	N/A
ALE	64.32290	-95.95353	14.7	N/A	8.15	N/A	78.6	N/A	55.8	N/A	0.04	N/A	9.7	N/A

Table A2: Manually assigned resistance classes for gene families identified by ShortBRED for antibiotic resistance gene (ARG) families based on CARD gene family classifications including NCBI and CARD identifiers.

Resistance Class	NCBI Identifier	CARD Identifier	ARG Family
Aminoglycoside Acetyltransferase	AAA03550_1	3002523	AAC(2_)-Ia
Aminoglycoside Acetyltransferase	AAC44793_1	3002524	AAC(2_)-Ib
Aminoglycoside Acetyltransferase	CCP42991_1	3002525	AAC(2_)-Ic
Aminoglycoside Acetyltransferase	AAB41701_1	3002526	AAC(2_)-Id
Aminoglycoside Acetyltransferase	APB03221_1	3003988	AAC(2_)-IIb
Aminoglycoside Acetyltransferase	AAG15269_1	3002528	AAC(3)-Ia
Aminoglycoside Acetyltransferase	AAL82588_1	3002600	AAC(3)-Ib_AAC(6_)-Ib__
Aminoglycoside Acetyltransferase	CAD53575_1	3002531	AAC(3)-Ic
Aminoglycoside Acetyltransferase	AAR21614_1	3002529	AAC(3)-Id
Aminoglycoside Acetyltransferase	AAA26548_1	3002534	AAC(3)-IIb
Aminoglycoside Acetyltransferase	CAA38525_1	3002535	AAC(3)-IIc
Aminoglycoside Acetyltransferase	CAA39184_1	3002536	AAC(3)-IIIa
Aminoglycoside Acetyltransferase	AAA25683_1	3002538	AAC(3)-IIIc
Aminoglycoside Acetyltransferase	ABB43029_1	3002539	AAC(3)-IV
Aminoglycoside Acetyltransferase	AAA25334_1	3002543	AAC(3)-IXa
Aminoglycoside Acetyltransferase	AAA16194_1	3002540	AAC(3)-VIa
Aminoglycoside Acetyltransferase	AAA88552_1	3002541	AAC(3)-VIIa
Aminoglycoside Acetyltransferase	AAA26685_1	3002542	AAC(3)-VIIIa
Aminoglycoside Acetyltransferase	BAA78619_1	3002544	AAC(3)-Xa
Aminoglycoside Acetyltransferase	AAK26252_2	3002583	AAC(6_)-29a
Aminoglycoside Acetyltransferase	CAE48335_2	3002599	AAC(6_)-30_AAC(6_)-Ib_
Aminoglycoside Acetyltransferase	CAK55563_1	3002585	AAC(6_)-31
Aminoglycoside Acetyltransferase	ABR10839_1	3002586	AAC(6_)-32
Aminoglycoside Acetyltransferase	APB03223_1	3003989	AAC(6_)-34
Aminoglycoside Acetyltransferase	AAP43642_1	3002588	AAC(6_)-I30
Aminoglycoside Acetyltransferase	AAA98298_1	3002545	AAC(6_)-Ia
Aminoglycoside Acetyltransferase	NP_460578_1	3002571	AAC(6_)-Iaa
Aminoglycoside Acetyltransferase	BAD12078_1	3002572	AAC(6_)-Iad
Aminoglycoside Acetyltransferase	BAD14386_1	3002573	AAC(6_)-Iae
Aminoglycoside Acetyltransferase	ACI28880_1	3002575	AAC(6_)-Iai
Aminoglycoside Acetyltransferase	BAM46120_1	3003677	AAC(6_)-Iaj
Aminoglycoside Acetyltransferase	BAO21229_1	3003199	AAC(6_)-Iak
Aminoglycoside Acetyltransferase	BAD11815_1	3002593	AAC(6_)-Ib-SK
Aminoglycoside Acetyltransferase	AKN19287_1	3002578	AAC(6_)-Ib7
Aminoglycoside Acetyltransferase	AAA26549_1	3002549	AAC(6_)-Ic
Aminoglycoside Acetyltransferase	NP_863643_1	3002597	AAC(6_)-Ie-APH(2__)-Ia
Aminoglycoside Acetyltransferase	CAA39038_1	3002553	AAC(6_)-If
Aminoglycoside Acetyltransferase	AAA21889_1	3002554	AAC(6_)-Ig

Aminoglycoside Acetyltransferase	AAB63533_1	3002556	AAC(6_)-Ii
Aminoglycoside Acetyltransferase	AAM92464_1	3002594	AAC(6_)-IIa
Aminoglycoside Acetyltransferase	AAA25680_1	3002595	AAC(6_)-IIb
Aminoglycoside Acetyltransferase	AAD46626_1	3002596	AAC(6_)-IIc
Aminoglycoside Acetyltransferase	CAE50925_1	3002589	AAC(6_)-Iid
Aminoglycoside Acetyltransferase	CAE50926_1	3002590	AAC(6_)-Iih
Aminoglycoside Acetyltransferase	AAC41392_1	3002557	AAC(6_)-Ij
Aminoglycoside Acetyltransferase	AAA87229_1	3002558	AAC(6_)-Ik
Aminoglycoside Acetyltransferase	AAK63041_1	3002559	AAC(6_)-Ip
Aminoglycoside Acetyltransferase	CAA91010_1	3002559	AAC(6_)-Ip
Aminoglycoside Acetyltransferase	AAC25500_1	3002560	AAC(6_)-Iq
Aminoglycoside Acetyltransferase	AAD03490_1	3002561	AAC(6_)-Ir
Aminoglycoside Acetyltransferase	AAD03491_1	3002562	AAC(6_)-Is
Aminoglycoside Acetyltransferase	BAD10948_2	3002563	AAC(6_)-Isa
Aminoglycoside Acetyltransferase	AAD03494_1	3002566	AAC(6_)-Iv
Aminoglycoside Acetyltransferase	AAD03495_1	3002567	AAC(6_)-Iw
Aminoglycoside Nucleotidyltransferase	AAU10334_1	3002628	aad(6)
Aminoglycoside Nucleotidyltransferase	YP_009081591_1	3002613	aadA13
Aminoglycoside Nucleotidyltransferase	CAI57696_1	3002614	aadA14
Aminoglycoside Nucleotidyltransferase	ACF17980_1	3002616	aadA16
Aminoglycoside Nucleotidyltransferase	AET15272_1	3003197	aadA25
Aminoglycoside Nucleotidyltransferase	AAF17880_1	3002605	aadA5
Aminoglycoside Nucleotidyltransferase	BAD00739_1	3002607	aadA7
Aminoglycoside Nucleotidyltransferase	ABG49324_1	3002609	aadA9
Aminoglycoside Nucleotidyltransferase	CAB14620_1	3002627	aadK
Gene Modulating Resistance	XP_753111_1	3003942	abcA
MATE Transporter	BAD89844_2	3000753	abeM
SMR Antibiotic Efflux	YP_002325052_1	3000768	abeS
Class C Beta-Lactamase	AAF86691_1	3001816	ACC-2
RND Antibiotic Efflux	NP_414995_1	3000216	acrB
RND Antibiotic Efflux	YP_490697_1	3000491	acrD
RND Antibiotic Efflux	AAC76297_1	3000499	acrE
RND Antibiotic Efflux	AAC76298_1	3000502	acrF
RND Antibiotic Efflux	AAC76296_1	3000656	acrS
Class C Beta-Lactamase	WP_004746565_1	3003848	ADC-2
RND Antibiotic Efflux	YP_002325610_1	3000774	adeA
RND Antibiotic Efflux	YP_002325611_1	3000775	adeB
RND Antibiotic Efflux	ALX99516_1	3003811	adeC
RND Antibiotic Efflux	YP_001706894_1	3000777	adeF
RND Antibiotic Efflux	YP_001706893_1	3000778	adeG
RND Antibiotic Efflux	YP_001713101_1	3000779	adeH

RND Antibiotic Efflux	YP_002320475_1	3000780	adeI
RND Antibiotic Efflux	AAX14802_1	3000781	adeJ
RND Antibiotic Efflux	AAX14803_1	3000782	adeK
RND Antibiotic Efflux	ALH22601_1	3000620	adeL
RND Antibiotic Efflux	AGV28567_1	3000559	adeN
RND Antibiotic Efflux	ADM92605_1	3000553	adeR
RND Antibiotic Efflux	ADM92606_1	3000549	adeS
Class A Beta-Lactamase	AAC09015_1	3002481	AER-1
Class B Beta-Lactamase	CAQ53840_1	3000853	AIM-1
RND Antibiotic Efflux	NP_250709_1	3002982	amrA
Aminoglycoside Nucleotidyltransferase	AAC64365_1	3000230	ANT(2__)-Ia
Aminoglycoside Nucleotidyltransferase	AAL51021_2	3002598	ANT(3__)-Ii-AAC(6_)-Iid
Aminoglycoside Nucleotidyltransferase	EEX02086	3004089	ANT(3__)-IIa
Aminoglycoside Nucleotidyltransferase	ENU91137	3004090	ANT(3__)-IIb
Aminoglycoside Nucleotidyltransferase	ENU37733	3004091	ANT(3__)-IIc
Aminoglycoside Nucleotidyltransferase	AAO83986_1	3002623	ANT(4_)-Ia
Aminoglycoside Nucleotidyltransferase	YP_006938491_1	3003905	ANT(4_)-Ib
Aminoglycoside Nucleotidyltransferase	AAA25717_1	3002624	ANT(4_)-IIa
Aminoglycoside Nucleotidyltransferase	AAM76670_1	3002625	ANT(4_)-IIb
Aminoglycoside Nucleotidyltransferase	YP_008997281_1	3002626	ANT(6)-Ia
Aminoglycoside Nucleotidyltransferase	AIJ27543_1	3002629	ANT(6)-Ib
Aminoglycoside Nucleotidyltransferase	CBH51824_1	3002629	ANT(6)-Ib
Aminoglycoside Nucleotidyltransferase	AFU35063_1	3002630	ANT(9)-Ia
Gene Modulating Resistance	NP_415611_1	3004049	fabG
Gene Modulating Resistance	NP_415804_1	3004045	fabI
Aminoglycoside Phosphotransferase	AAX38178_1	3002634	APH(2__)-Ie
Aminoglycoside Phosphotransferase	AGV10818_1	3002669	APH(2__)-Ig
Aminoglycoside Phosphotransferase	AAK63040_1	3002635	APH(2__)-IIa
Aminoglycoside Phosphotransferase	AAB49832_1	3002636	APH(2__)-IIIa
Aminoglycoside Phosphotransferase	CAA37605_1	3002638	APH(3__)-Ia
Aminoglycoside Phosphotransferase	ABK33456_1	3002639	APH(3__)-Ib
Aminoglycoside Phosphotransferase	ABC68330_1	3002640	APH(3__)-Ic
Aminoglycoside Phosphotransferase	CAE51638_1	3002641	APH(3_)-Ia
Aminoglycoside Phosphotransferase	AAA26412_1	3002642	APH(3_)-Ib
Aminoglycoside Phosphotransferase	WP_000572405_1	3002644	APH(3_)-IIa
Aminoglycoside Phosphotransferase	CAA62365_1	3002645	APH(3_)-IIb
Aminoglycoside Phosphotransferase	ADQ43421_1	3002646	APH(3_)-IIc
Aminoglycoside Phosphotransferase	AGV10830_1	3002647	APH(3_)-IIIa
Aminoglycoside Phosphotransferase	CAA27061_1	3002648	APH(3_)-IVa
Aminoglycoside Phosphotransferase	ENV34035	3004087	APH(3_)-IX
Aminoglycoside Phosphotransferase	AAA26699_1	3002649	APH(3_)-Va

Aminoglycoside Phosphotransferase	AAC32025_1	3002650	APH(3_)-Vb
Aminoglycoside Phosphotransferase	AAB21326_1	3002651	APH(3_)-Vc
Aminoglycoside Phosphotransferase	AGI04227_1	3003687	APH(3_)-VI
Aminoglycoside Phosphotransferase	AAA76822_1	3002654	APH(3_)-VIIa
Aminoglycoside Phosphotransferase	EPF73263	3004086	APH(3_)-VIII
Aminoglycoside Phosphotransferase	CAA24743_1	3002655	APH(4)-Ia
Aminoglycoside Phosphotransferase	CAA52372_1	3002656	APH(4)-Ib
Aminoglycoside Phosphotransferase	CAA68516_1	3002657	APH(6)-Ia
Aminoglycoside Phosphotransferase	CAA29136_1	3002658	APH(6)-Ib
Aminoglycoside Phosphotransferase	CAA25854_1	3002659	APH(6)-Ic
Aminoglycoside Phosphotransferase	AAC23556_1	3002660	APH(6)-Id
Aminoglycoside Phosphotransferase	CAA27276_1	3002661	APH(7__)-Ia
Aminoglycoside Phosphotransferase	AAB58447_1	3002662	APH(9)-Ia
Aminoglycoside Phosphotransferase	AAB66655_1	3002663	APH(9)-Ib
Other ARG	CBL58181_1	3003918	apmA
Class C Beta-Lactamase	BAM76830_1	3002993	AQU-1
MFS Antibiotic Efflux	YP_001332362_1	3000838	arlR
MFS Antibiotic Efflux	YP_499945_1	3000839	arlS
rRNA Methyltransferase	ADC55560_1	3000858	armA
RND Antibiotic Efflux	NP_252408_1	3004056	ArmR
Other ARG	AAC75315_1	3002985	arnA
Other ARG	NP_252244	3002985	arnA
Rifampin Resistance	AAC05822_1	3002846	arr-1
Rifampin Resistance	AAC64366_1	3002847	arr-2
Rifampin Resistance	ABV26707_1	3002850	arr-5
Rifampin Resistance	CAZ48628_1	3002852	arr-7
Rifampin Resistance	AGC29882_1	3002853	arr-8
Gene Modulating Resistance	AAC76093_1	3002986	bacA
Gene Modulating Resistance	CAX52582_1	3003324	mprF
Macrolide Resistance	EEL41021_1	3003072	mphL
Macrolide Resistance	EOO80837_1	3003767	mphM
RND Antibiotic Efflux	YP_490321_1	3000828	baeR
RND Antibiotic Efflux	BAA15934_1	3000829	baeS
Other ARG	APB03218_1	3003984	BahA
Gene Modulating Resistance	AAL82403_1	3003297	gyrA
Gene Modulating Resistance	WP_005768149_1	3003302	gyrB
Other ARG	NP_253464_1	3003582	basR
Other ARG	AEX49906_1	3003583	basS
Class B Beta-Lactamase	AAA22562_1	3002878	BcII
Gene Modulating Resistance	AAA99504_1	3002987	bcrA
Gene Modulating Resistance	AAA99505_1	3002988	bcrB

Gene Modulating Resistance	AAA99503_1	3003250	bcrC
Class A Beta-Lactamase	ACT09140_1	3002387	BEL-3
Gene Modulating Resistance	YP_003971446	3003730	ileS
Class B Beta-Lactamase	BAL75272_1	3000856	BJP-1
Class A Beta-Lactamase	AAR20595_1	3000090	BLA1
Class B Beta-Lactamase	WP_024565805_1	3000579	BlaB
Class A Beta-Lactamase	ABU39978_1	3000160	blaI
Class A Beta-Lactamase	ABU39979_1	3000217	blaR1
MFS Antibiotic Efflux	AAC36944_1	3003006	blt
MFS Antibiotic Efflux	AAA22277_1	3003007	bmr
Gene Modulating Resistance	Q8FW76	3003772	mprF
Gene Modulating Resistance	YP_002344422_1	3003789	gyrA
Gene Modulating Resistance	EEK14408_1	3003931	gyrA
ABC-F Ribosomal Protection	AAC32027_1	3002817	carA
Class A Beta-Lactamase	CCW43444_1	3002255	CARB-16
Class A Beta-Lactamase	AHJ02283_1	3003186	CARB-23
Class A Beta-Lactamase	AAB19430_2	3002242	CARB-3
Class A Beta-Lactamase	AAP22374_1	3002248	CARB-9
Chloramphenicol Resistance	AAA22081_1	3002670	cat
Chloramphenicol Resistance	AAA23018_1	3002670	cat
Chloramphenicol Resistance	AAA25655_1	3002670	cat
Chloramphenicol Resistance	AAA26613_1	3002670	cat
Chloramphenicol Resistance	AAB23649_1	3002670	cat
Chloramphenicol Resistance	AAL08441_1	3002670	cat
Chloramphenicol Resistance	AAQ63644_1	3002670	cat
Chloramphenicol Resistance	CAA63498_1	3002670	cat
Chloramphenicol Resistance	AAB53259_1	3002671	cat-TC
Chloramphenicol Resistance	AAA22289_1	3002672	cat86
Chloramphenicol Resistance	AAA73865_1	3002674	catB
Chloramphenicol Resistance	CAI47810_1	3003110	catB10
Chloramphenicol Resistance	NP_848167_1	3002675	catB2
Chloramphenicol Resistance	NP_249397_1	3002679	catB7
Chloramphenicol Resistance	YP_009077553_1	3002680	catB8
Chloramphenicol Resistance	AAL68645_1	3002681	catB9
Chloramphenicol Resistance	CAA37806_1	3002684	catII
Chloramphenicol Resistance	CAA30695_1	3002685	catIII
Chloramphenicol Resistance	AAB51421_1; AAF66228_1	3002686	catP|catD
Chloramphenicol Resistance	AAA23215_1	3002687	catQ
Chloramphenicol Resistance	CAA52904_1	3002688	catS
Chloramphenicol Resistance	APB03217_1	3003983	CatU
Class B Beta-Lactamase	CAC87665_1	3000855	CAU-1

Class A Beta-Lactamase	ACT97415_1	3002999	CblA-1
Class B Beta-Lactamase	AAA22904_1	3000578	CcrA
MATE Transporter	CAE00499_1	3003835	cdeA
RND Antibiotic Efflux	AAB58160_1	3003009	ceoA
RND Antibiotic Efflux	AAB58161_1	3003010	ceoB
Class A Beta-Lactamase	AAA21532_1	3003559	cepA
Class C Beta-Lactamase	CAA56561_1	3003553	CepS
rRNA Methyltransferase	CAL64019_1	3003441	cfrA
Class A Beta-Lactamase	ACT97371_1	3003097	CfxA6
Class B Beta-Lactamase	ABS29619_1	3000841	CGB-1
Other ARG	AAN28945	3003785	murA
rRNA Methyltransferase	AAS79458_1	3001302	chrB
rRNA Methyltransferase	WP_015735625_1	3003907	cipA
rRNA Methyltransferase	YP_001420189_1	3002814	clbA
rRNA Methyltransferase	YP_002773985_1	3002815	clbB
rRNA Methyltransferase	YP_174574_1	3002816	clbC
Gene Modulating Resistance	Q0SSM7	3003773	mprF
Gene Modulating Resistance	WP_009887863_1	3003357	EF-Tu
Gene Modulating Resistance	YP_001086469_1	3003995	gyrA
RND Antibiotic Efflux	ABS43901_1	3000783	cmeA
RND Antibiotic Efflux	ABS43151_1	3000784	cmeB
RND Antibiotic Efflux	BAO79432_1	3000785	cmeC
RND Antibiotic Efflux	YP_002343805_1	3000526	cmeR
Chloramphenicol Resistance	AAB36568_1	3002700	cmlv
Class C Beta-Lactamase	BAJ17544_1	3002069	CMY-59
Class C Beta-Lactamase	AAD50818_2	3002019	CMY-8
Other ARG	APB03224_1	3003994	cpaA
Class B Beta-Lactamase	AJP77054	3003716	CPS-1
RND Antibiotic Efflux	NP_312864_1	3000830	cpxA
RND Antibiotic Efflux	BAE77933_1	3000518	CRP
Class A Beta-Lactamase	CBW46935_1	3001959	CTX-M-100
Class A Beta-Lactamase	AEM44648_1	3001970	CTX-M-110
Class A Beta-Lactamase	BAP34782_1	3002008	CTX-M-151
Class A Beta-Lactamase	ABD73290_1	3001920	CTX-M-59
Class A Beta-Lactamase	AAV97957_1	3001933	CTX-M-72
Glycopeptide Resistance	AAN62561_1	3003970	D-Ala-D-Ala
Glycopeptide Resistance	YP_007652797_1	3001205	BRP(MBL)
Gene Modulating Resistance	CAC19929_1	3002854	dfrA1
Gene Modulating Resistance	AHG97174_1	3003011	dfrA10
Gene Modulating Resistance	AHW42429_1	3002858	dfrA12
Gene Modulating Resistance	ACI32877_1	3002859	dfrA14

Gene Modulating Resistance	AHB39758_1	3003013	dfrA15
Gene Modulating Resistance	AAK60186_1	3003014	dfrA16
Gene Modulating Resistance	ABG91835_1	3002860	dfrA17
Gene Modulating Resistance	CAC81324_1	3003015	dfrA19
Gene Modulating Resistance	CAE53424_1	3003016	dfrA20
Gene Modulating Resistance	CAP69659_1	3003017	dfrA21
Gene Modulating Resistance	CAG34233_2	3003019	dfrA23
Gene Modulating Resistance	CAI99385_1	3002856	dfrA24
Gene Modulating Resistance	ABB71176_1	3003020	dfrA25
Gene Modulating Resistance	CAL48457_1	3002857	dfrA26
Gene Modulating Resistance	AAA25550_1	3003105	dfrA3
Gene Modulating Resistance	AHV80711_1	3002863	dfrA8
Gene Modulating Resistance	AAN41433_1	3002864	dfrB1
Gene Modulating Resistance	ACR57831_1	3003022	dfrB3
Gene Modulating Resistance	AAO04716_1	3002865	dfrC
Gene Modulating Resistance	AAA85213_1	3002866	dfrD
Gene Modulating Resistance	AAD01867_1	3002875	dfrE
Gene Modulating Resistance	AAD01868_1	3002867	dfrF
Gene Modulating Resistance	BAE15963_1	3002868	dfrG
Class C Beta-Lactamase	AIT76100_1	3002152	DHA-21
Class B Beta-Lactamase	AGC92784_1	3000848	DIM-1
Class B Beta-Lactamase	AAN32638_1	3000842	EBR-1
Other ARG	AHH86051_1	3004063	EdeQ
MFS Antibiotic Efflux	BAG75524_1	3003954	efmA
MFS Antibiotic Efflux	NP_217362_1	3003955	efpA
Gene Modulating Resistance	CDO61513_1	3003948	efrA
Gene Modulating Resistance	CDO61516_1	3003949	efrB
MATE Transporter	BAC11911_1	3003551	emeA
MFS Antibiotic Efflux	BAA16547_1	3000027	emrA
MFS Antibiotic Efflux	AAC75733_1	3000074	emrB
MFS Antibiotic Efflux	EFF04178_1	3000309	emrD
MFS Antibiotic Efflux	BAA11236_1	3000206	emrK
MFS Antibiotic Efflux	NP_417169_1	3000516	emrR
MFS Antibiotic Efflux	BAA11237_1	3000254	emrY
General Bacterial Porin	AAK11270_1	3003385	Omp36
RND Antibiotic Efflux	ABG77965_1	3004042	acrA
Gene Modulating Resistance	AFK58561_1	3003079	liaS
Gene Modulating Resistance	AFK58562_1	3003078	liaR
Gene Modulating Resistance	ALL09868	3003092	cls
Gene Modulating Resistance	YP_006374661_1	3003438	EF-Tu
Macrolide Resistance	AAC78336_1	3002826	EreA2

Macrolide Resistance	CAA01212_1	3000363	EreB
rRNA Methyltransferase	AAC69328_1	3001265	Erm(30)
rRNA Methyltransferase	AAC69327_1	3000598	Erm(31)
rRNA Methyltransferase	NP_899170_1	3000599	Erm(33)
rRNA Methyltransferase	AAP74657_1	3000600	Erm(34)
rRNA Methyltransferase	AAK07612_1	3000604	Erm(35)
rRNA Methyltransferase	AAL68827_1	3000605	Erm(36)
rRNA Methyltransferase	CCP44758_1	3000392	Erm(37)
rRNA Methyltransferase	AAN86837_1	3000601	Erm(38)
rRNA Methyltransferase	AAR92235_1	3000602	Erm(39)
rRNA Methyltransferase	ABW06859_1	3000603	Erm(41)
rRNA Methyltransferase	CBY77552_1	3003106	Erm(42)
rRNA Methyltransferase	CCF55073_1	3003205	Erm(43)
rRNA Methyltransferase	CUU67654_1	3003971	erm(44)
rRNA Methyltransferase	ANZ79476_1	3003908	Erm(47)
rRNA Methyltransferase	YP_001315192_1	3000347	ErmA
rRNA Methyltransferase	AAF86219_1	3000375	ErmB
rRNA Methyltransferase	AAA98296_1	3000250	ErmC
rRNA Methyltransferase	AAA22597_1	3000495	ErmD
rRNA Methyltransferase	CAB60001_1	3000326	ErmE
rRNA Methyltransferase	AAA88675_1	3000498	ErmF
rRNA Methyltransferase	AAC37034_1	3000522	ErmG
rRNA Methyltransferase	AAC32026_1	3002823	ErmH
rRNA Methyltransferase	CAA66307_1	3000592	ErmN
rRNA Methyltransferase	AAA26779_1	3001303	ErmO
rRNA Methyltransferase	AAC36915_1	3000593	ErmQ
rRNA Methyltransferase	ALX06067_1	3000594	ErmR
rRNA Methyltransferase	AAA26742_1	3001304	ErmS
rRNA Methyltransferase	AAA98096_1	3000595	ErmT
rRNA Methyltransferase	CAA44667_1	3001305	ErmU
rRNA Methyltransferase	AAB51440_1	3002824	ErmV
rRNA Methyltransferase	BAA03402_1	3001306	ErmW
rRNA Methyltransferase	AAG03357_1	3000596	ErmX
rRNA Methyltransferase	BAB20748_1	3002825	ErmY
RND Antibiotic Efflux	AAC74603_2	3003378	marA
Gene Modulating Antibiotic Efflux	AAC77033_1	3003381	soxR
Gene Modulating Resistance	AAN82549_1	3003369	EF-Tu
Other ARG	AIL15701	3003775	murA
Gene Modulating Resistance	BAE77595_1	3003303	gyrB
SMR Antibiotic Efflux	CAA77936_1	3004039	emrE
Gene Modulating Resistance	CDJ72593	3003889	GlpT

Gene Modulating Resistance	CDJ73208	3003890	UhpT
Gene Modulating Resistance	NP_312937_1	3003288	rpoB
RND Antibiotic Efflux	NP_414997_1	3003807	acrR
Gene Modulating Resistance	NP_415372_1	3003751	nfsA
Gene Modulating Antibiotic Efflux	NP_418486_1	3003511	soxS
General Bacterial Porin	YP_489201_1	3003390	ompF
Gene Modulating Resistance	YP_491362_1	3003386	folP
Class B Beta-Lactamase	AJP77085	3003717	ESP-1
Gene Modulating Antibiotic Efflux	NP_311275_1	3000832	evgA
Gene Modulating Antibiotic Efflux	AAC75429_1	3000833	evgS
Class A Beta-Lactamase	AAA26775_1	3003564	EXO
MFS Antibiotic Efflux	AFK80333_1	3001313	facT
MFS Antibiotic Efflux	NP_273367_1	3003961	farA
MFS Antibiotic Efflux	NP_273368_1	3003962	farB
Class B Beta-Lactamase	CAB96921_1	3000606	FEZ-1
MFS Antibiotic Efflux	AAG16656_1	3002705	floR
Other ARG	BAA32493_1	3000423	FomA
Other ARG	BAA32494_1	3000449	FomB
Fosfomycin Resistance	NP_249820_1	3000149	FosA
Fosfomycin Resistance	ACC85616_1	3002804	FosA2
Fosfomycin Resistance	BAP18892_1	3003210	FosA4
Fosfomycin Resistance	AJE60855_1	3003209	FosA5
Fosfomycin Resistance	NP_831795_1	3000172	FosB
Fosfomycin Resistance	YP_001715981_1	3000172	FosB
Fosfomycin Resistance	ADX95999_1	3002873	FosB3
Fosfomycin Resistance	CAA83855_1	3000380	FosC
Fosfomycin Resistance	BAJ10053_1	3002874	FosC2
Fosfomycin Resistance	BAO79518_1	3003207	FosK
Fosfomycin Resistance	CWV56762_1	3000198	FosX
Class C Beta-Lactamase	CAA71325_1	3002156	FOX-2
Other ARG	AAL12234_1	3003552	fusB
Other ARG	WP_011303797	3003731	fusD
Other ARG	CAA90432_1	3003026	fusH
RND Antibiotic Efflux	NP_417969_1	3003834	gadE
RND Antibiotic Efflux	ANK04027_1	3003838	gadW
RND Antibiotic Efflux	YP_491919_1	3000508	gadX
Class A Beta-Lactamase	ADZ48685_1	3002346	GES-17
Class B Beta-Lactamase	AIY26289_1	3003194	GIM-2
Macrolide Resistance	CAA11707_1	3000463	gimA
Class B Beta-Lactamase	AAF04458_1	3000850	GOB-1
RND Antibiotic Efflux	NP_459349_1	3000504	golS

Gene Modulating Antibiotic Efflux	NP_309766_1	3000676	H-NS
Gene Modulating Resistance	WP_014065640_1	3003924	gyrA
Gene Modulating Resistance	WP_041918279_1	3003925	parC
Class B Beta-Lactamase	AMY61250_1	3004092	HMB-1
MATE Transporter	WP_014550864_1	3003953	hmrM
MFS Antibiotic Efflux	NP_207972_1	3003964	hp1181
MATE Transporter	NP_207975_1	3003965	hp1184
Class A Beta-Lactamase	ACX71212_1	3001860	IMI-3
Class B Beta-Lactamase	CAA71441_1	3003095	imiS
Class B Beta-Lactamase	ABC88434_1	3002213	IMP-22
Class B Beta-Lactamase	AEH41427_1	3002218	IMP-27
Class B Beta-Lactamase	AGS82587_1	3002222	IMP-31
Class B Beta-Lactamase	BAM38093_1	3002225	IMP-34
Class B Beta-Lactamase	AFP97028_1	3002228	IMP-37
Class B Beta-Lactamase	AIA58910_1	3002236	IMP-45
Class B Beta-Lactamase	AIT76110_1	3002239	IMP-48
Class B Beta-Lactamase	ADK25051_1	3002268	IND-12
Class B Beta-Lactamase	AAG29761_2	3002259	IND-3
Class B Beta-Lactamase	AAG29765_2	3002260	IND-4
Class B Beta-Lactamase	ACZ65153_1	3002265	IND-9
Rifampin Resistance	AAB41059_1	3002884	iri
Class B Beta-Lactamase	AAK38324_1	3000840	JOHN-1
Other ARG	NP_415222_1	3003841	kdpE
Class B Beta-Lactamase	BAF91108_1	3000847	KHM-1
RND Antibiotic Efflux	ABR75897_1	3003373	acrR
RND Antibiotic Efflux	ABR76005_1	3003380	ramR
RND Antibiotic Efflux	CAC41008_1	3004041	acrA
Gene Modulating Resistance	CDO13981_1	3003585	PhoP
General Bacterial Porin	YP_005226137_1	3003966	OmpK35
General Bacterial Porin	YP_005228001_1	3003968	OmpK36
Class A Beta-Lactamase	ACE62798_1	3002317	KPC-7
Class B Beta-Lactamase	CAB75346_1	3000582	L1
Class D Beta-Lactamase	CAA40146_1	3002997	LCR-1
MFS Antibiotic Efflux	NP_414618_4	3003843	leuO
Gene Modulating Resistance	NP_465220	3003770	mprF
rRNA Methyltransferase	APB03216_1	3003982	LlmA
Gene Modulating Resistance	CAB12062_1	3003028	lmrA
Gene Modulating Resistance	KIX81495_1	3002813	lmrB
ABC-F Ribosomal Protection	ABF66011_1	3002881	lmrC
Gene Modulating Resistance	ABF66027_1	3002882	lmrD
MFS Antibiotic Efflux	ABF33001_1	3003969	lmrP

Lincosamide Nucleotidyltransferase	CAL44992_1	3002835	lnuA
Lincosamide Nucleotidyltransferase	AGI42804_1	3002836	lnuB
Lincosamide Nucleotidyltransferase	AAY32951_1	3002837	lnuC
Lincosamide Nucleotidyltransferase	ABR14060_1	3002838	lnuD
Lincosamide Nucleotidyltransferase	AGT57825	3003762	lnuE
Lincosamide Nucleotidyltransferase	CAD91132_1	3002839	lnuF
Lincosamide Nucleotidyltransferase	WP_002333496_1	3004085	lnuG
Other ARG	AJF82049_1	3003573	LpxA
Other ARG	AJF83452_1	3003574	LpxC
Class A Beta-Lactamase	ACH58980_1	3002482	LRA-1
Class C Beta-Lactamase	ACH58999_1	3002489	LRA-10
Class B Beta-Lactamase	ACH58990_1	3002511	LRA-12
Class C Beta-Lactamase	ACH58991_1	3002484	LRA-13
Class B Beta-Lactamase	ACH58994_1	3002512	LRA-17
Class C Beta-Lactamase	ACH58997_1	3002492	LRA-18
Class B Beta-Lactamase	ACH59005_1	3002513	LRA-19
Class B Beta-Lactamase	ACH58985_1	3002485	LRA-2
Class B Beta-Lactamase	ACH58987_1	3002510	LRA-3
Class A Beta-Lactamase	ACH59002_1	3002483	LRA-5
Class B Beta-Lactamase	ACH58998_1	3002486	LRA-7
Class B Beta-Lactamase	ACH58988_1	3002487	LRA-8
MFS Antibiotic Efflux	AAC43550_1	3003967	lrfA
ABC-F Ribosomal Protection	AAT46077_1	3000300	lsaA
ABC-F Ribosomal Protection	NP_899166_1	3003111	lsaB
ABC-F Ribosomal Protection	AEA37904_1	3003112	lsaC
ABC-F Ribosomal Protection	AFU35065_1	3003206	lsaE
Gene Modulating Resistance	AAV85981_1	3000533	macA
Gene Modulating Resistance	AAV85982_1	3000535	macB
Gene Modulating Antibiotic Efflux	YP_489794_1	3000263	marA
Other ARG	AKF16168	3003689	MCR-1
MFS Antibiotic Efflux	AFH35853_1	3001328	mdfA
RND Antibiotic Efflux	NP_459347_2	3000789	mdsA
RND Antibiotic Efflux	NP_459346_1	3000790	mdsB
RND Antibiotic Efflux	NP_459345_2	3000791	mdsC
RND Antibiotic Efflux	AAC75135_2	3000792	mdtA
RND Antibiotic Efflux	AAC75136_1	3000793	mdtB
RND Antibiotic Efflux	AAC75137_1	3000794	mdtC
RND Antibiotic Efflux	AAC75138_1	3001330	mdtD
RND Antibiotic Efflux	AAC76539_1	3000796	mdtF
MFS Antibiotic Efflux	YP_489321_1	3001329	mdtG
MFS Antibiotic Efflux	AAC74149_2	3001216	mdtH

MATE Transporter	AML99881_1	3001327	mdtK
MFS Antibiotic Efflux	AAC76733_1	3001215	mdtL
MFS Antibiotic Efflux	AAC77293_1	3001214	mdtM
MFS Antibiotic Efflux	BAE78084_1	3003548	mdtN
MFS Antibiotic Efflux	BAE78083_1	3003549	mdtO
MFS Antibiotic Efflux	BAE78082_1	3003550	mdtP
Penicillin-binding Protein	AGC51118_1	3000617	mecA
Penicillin-binding Protein	BAI83385_1	3003440	mecB
Penicillin-binding Protein	WP_000725529_1	3001209	mecC
Penicillin-binding Protein	NP_373280_1	3000124	mecI
Penicillin-binding Protein	YP_001245420_1	3000215	mecR1
ABC-F Ribosomal Protection	YP_008997285_1	3000615	mefA
MFS Antibiotic Efflux	ACJ63262_1	3003107	mefB
MFS Antibiotic Efflux	BAL43360	3003745	mefC
MFS Antibiotic Efflux	NP_358565_1	3000614	mefE
ABC-F Ribosomal Protection	YP_002743273_1	3000616	mel
MATE Transporter	AAU95768_1	3000026	mepA
MATE Transporter	YP_001440920_1	3000746	mepR
RND Antibiotic Efflux	NP_249116_1	3000377	MexA
RND Antibiotic Efflux	AAA74437_1	3000378	MexB
RND Antibiotic Efflux	AAB41956_1	3000800	MexC
RND Antibiotic Efflux	AAB41957_1	3000801	MexD
RND Antibiotic Efflux	NP_251183_1	3000803	MexE
RND Antibiotic Efflux	NP_251184_1	3000804	MexF
RND Antibiotic Efflux	NP_252894_1	3000806	mexG
RND Antibiotic Efflux	NP_252895_1	3000807	mexH
RND Antibiotic Efflux	NP_252896_1	3000808	mexI
RND Antibiotic Efflux	NP_252367_1	3003692	mexJ
RND Antibiotic Efflux	AAG07064_1	3003693	mexK
RND Antibiotic Efflux	NP_252368_1	3003710	mexL
RND Antibiotic Efflux	BAE06005_1	3003704	mexM
RND Antibiotic Efflux	BAE06006_1	3003705	mexN
RND Antibiotic Efflux	BAE06007_1	3003698	mexP
RND Antibiotic Efflux	BAE06008_1	3003699	mexQ
RND Antibiotic Efflux	NP_249115_1	3000506	mexR
RND Antibiotic Efflux	ADT64081_1	3000813	MexS
RND Antibiotic Efflux	NP_251182_1	3000814	MexT
RND Antibiotic Efflux	AAG07762_1	3003030	mexV
RND Antibiotic Efflux	AAG07763_1	3003031	mexW
RND Antibiotic Efflux	BAA34300_1	3003033	mexY
RND Antibiotic Efflux	NP_250710_1	3003709	mexZ

Quinolone Resistance	NP_415632_1	3003844	mfd
Quinolone Resistance	CCP46182_1	3003035	mfpA
Gene Modulating Antibiotic Efflux	YP_003281576_1	3000815	mgrA
Macrolide Resistance	ABA28305_2	3000462	mgtA
Class A Beta-Lactamase	AIT76113_1	3002174	MIR-9
Gene Modulating Resistance	YP_007503908_1	3003306	gyrB
Class A Beta-Lactamase	ACS44784_1	3002185	MOX-6
Class D Beta-Lactamase	WP_042649345	3002191	MOX-9
Macrolide Resistance	BAA03776_1	3000316	mphA
Macrolide Resistance	BAA12910_1	3000318	mphB
Macrolide Resistance	CAJ51085_1	3000319	mphC
Macrolide Resistance	ANP63073_1	3003741	mphE
Macrolide Resistance	WP_010550189_1	3003071	mphE
Macrolide Resistance	BAL43359	3003742	mphG
Macrolide Resistance	APB03226_1	3003991	mphI
Macrolide Resistance	AAS13767_1	3003839	Mrx
Gene Modulating Resistance	NP_415434_1	3003950	msbA
Class B Beta-Lactamase	AJP77057	3003718	MSI-1
Class B Beta-Lactamase	AJP77058	3003719	MSI-OXA
ABC-F Ribosomal Protection	CCQ20328_1	3000251	msrA
ABC-F Ribosomal Protection	NP_416292_1	3002818	msrB
ABC-F Ribosomal Protection	AAK01167_1	3002819	msrC
ABC-F Ribosomal Protection	YP_724476_1	3003109	msrE
RND Antibiotic Efflux	CCP46065_1	3000816	mtrA
RND Antibiotic Efflux	NP_274719_1	3000810	mtrC
RND Antibiotic Efflux	NP_274718_1	3000811	mtrD
RND Antibiotic Efflux	CAA64891_1	3000812	mtrE
RND Antibiotic Efflux	YP_002002225_1	3000817	mtrR
Class B Beta-Lactamase	WP_063860852	3003842	MUS-2
RND Antibiotic Efflux	NP_251218_1	3004073	MuxA
RND Antibiotic Efflux	NP_251217_1	3004074	MuxB
RND Antibiotic Efflux	NP_251216_1	3004075	MuxC
RND Antibiotic Efflux	NP_253005_1	3004069	MvaT
Gene Modulating Resistance	AAK44936_1	3003395	rpsL
Gene Modulating Resistance	AAK46002_1	3003445	tlyA
Gene Modulating Resistance	AAK48336_1	3003458	ethA
Gene Modulating Resistance	AFN51819_1	3003327	embC
Gene Modulating Resistance	AFN51820_1	3003453	embA
Gene Modulating Resistance	CAA55486_1	3003459	gyrB
Gene Modulating Resistance	CAC29732_1	3003389	folP
Gene Modulating Resistance	CAC30845_1	3003284	rpoB

Other ARG	CCE36834	3003784	murA
Gene Modulating Resistance	CCP42728_1	3003295	gyrA
Gene Modulating Resistance	CCP43072_1	3003448	iniA
Gene Modulating Resistance	CCP43073_1	3003451	iniC
Gene Modulating Resistance	CCP44023_1	3003455	embR
Gene Modulating Resistance	CCP44244_1	3003393	inhA
Gene Modulating Resistance	CCP44620_1	3003461	ndh
Gene Modulating Resistance	CCP44816_1	3003394	pncA
Gene Modulating Resistance	CCP45025_1	3003463	kasA
Gene Modulating Resistance	CCP46624_1; AAK48268_1	3003326;3003465	embB
Gene Modulating Resistance	CCP46748_1	3003470	gidB
Gene Modulating Resistance	NP_216424_1	3003392	katG
Gene Modulating Resistance	CAC29514_1	3003298	gyrA
Gene Modulating Resistance	AKJ52802_1	3003310	parC
rRNA Methyltransferase	BAA03674_1	3001300	myrA
RND Antibiotic Efflux	NP_252410_1	3000818	nalC
RND Antibiotic Efflux	NP_252264_1	3000819	nalD
Class B Beta-Lactamase	BAO79439_1	3002362	NDM-12
General Bacterial Porin	AAB57788_1	3000464	por
Penicillin-binding Protein	NP_273462_1	3003937	PBP2
Gene Modulating Resistance	YP_207769_1	3003928	gyrA
Gene Modulating Resistance	YP_208330_1	3003929	parC
RND Antibiotic Efflux	NP_253290_1	3000820;3004059;3004060	nfxB|TypeA NfxB|TypeB NfxB
Class A Beta-Lactamase	CAA79966_1	3003665	NmcR
MFS Antibiotic Efflux	AAS68233_1	3000391	norA
MFS Antibiotic Efflux	CCQ22388_1	3000421	norB
Gene Modulating Resistance	AAF67494_2	3002522	novA
rRNA Methyltransferase	BAF80809_1	3002665	npmA
Class D Beta-Lactamase	CAA33795_1	3003563	NPS
Class C Beta-Lactamase	ABF50909_1	3002521	OCH-8
Class A Beta-Lactamase	CAP12359_2	3002452	OKP-B-19
ABC-F Ribosomal Protection	AAA50325_1	3003036	oleB
Gene Modulating Resistance	AAA26793	3003748	oleC
Other ARG	ABA42119_1	3000865	oleD
Other ARG	ABA42118_2	3000866	oleI
RND Antibiotic Efflux	AAC43969_1	3003037	opcM
RND Antibiotic Efflux	NP_251215_1	3004072	OpmB
RND Antibiotic Efflux	NP_252897_1	3000809	opmD
RND Antibiotic Efflux	BAE06009_1	3003700	opmE
RND Antibiotic Efflux	NP_253661_1	3003682	OpmH

RND Antibiotic Efflux	BAM10414_1	3003039	oprA
RND Antibiotic Efflux	AAB41958_1	3000802	OprJ
RND Antibiotic Efflux	NP_249118_1	3000379	OprM
RND Antibiotic Efflux	NP_251185_1	3000805	OprN
RND Antibiotic Efflux	YP_001693237_1	3003922	oqxA
RND Antibiotic Efflux	YP_001693238	3003923	oqxB
Tetracycline Resistance	CAA37477_1	3002891	otr(A)
MFS Antibiotic Efflux	AAD04032_1	3002892	otr(B)
Gene Modulating Resistance	AAR96051_1	3002894	otrC
Class D Beta-Lactamase	AIA58911_1	3001396	OXA-1
Class D Beta-Lactamase	ABW70410_1	3001644	OXA-113
Class D Beta-Lactamase	AAN41427_1	3001775	OXA-119
Class D Beta-Lactamase	AAA83417_1	3001407	OXA-12
Class D Beta-Lactamase	AAC46344_1	3001408	OXA-13
Class D Beta-Lactamase	ACI28281_1	3001779	OXA-146
Class D Beta-Lactamase	AAB58555_1	3001413	OXA-18
Class D Beta-Lactamase	AFO09968_1	3001476	OXA-184
Class D Beta-Lactamase	ADZ54048_1	3001766	OXA-192
Class D Beta-Lactamase	ADT70779_1	3001805	OXA-198
Class D Beta-Lactamase	AEM66528_1	3001809	OXA-209
Class D Beta-Lactamase	AEV91550_1	3001710	OXA-211
Class D Beta-Lactamase	AEV91554_1	3001714	OXA-215
Class D Beta-Lactamase	AAD12233_1	3001417	OXA-22
Class D Beta-Lactamase	AFQ90085_1	3001610	OXA-243
Class D Beta-Lactamase	AGC60012_1	3001786	OXA-244
Class D Beta-Lactamase	CCE73593_2	3001503	OXA-258
Class D Beta-Lactamase	AAG35608_1	3001421	OXA-26
Class D Beta-Lactamase	CAC35728_1	3001424	OXA-29
Class D Beta-Lactamase	AAC41449_1	3001398	OXA-3
Class D Beta-Lactamase	AET35493_1	3001777	OXA-347
Class D Beta-Lactamase	AGW83449_1	3001538	OXA-351
Class D Beta-Lactamase	AHA11126_1	3001550	OXA-363
Class D Beta-Lactamase	AAG33665_1	3001431	OXA-37
Class D Beta-Lactamase	AIN56719_1	3002496	OXA-418
Class D Beta-Lactamase	BAP28835_1	3003116	OXA-420
Class D Beta-Lactamase	CAD32565_1	3001770	OXA-43
Class D Beta-Lactamase	CAD58780_1	3001794	OXA-45
Class D Beta-Lactamase	CAA41211_1	3001400	OXA-5
Class D Beta-Lactamase	AAQ76277_1	3001796	OXA-50
Class D Beta-Lactamase	AAR03105_1	3001813	OXA-55
Class D Beta-Lactamase	AAQ08905_1	3001808	OXA-60

Class D Beta-Lactamase	AAT01092_1	3001773	OXA-61
Class D Beta-Lactamase	AAR32134_1	3001792	OXA-62
Class D Beta-Lactamase	AAP69916_1	3001780	OXA-85
Class D Beta-Lactamase	AAA98406_1	3001404	OXA-9
Class A Beta-Lactamase	AAL78275_1	3002391	OXY-1-3
Gene Modulating Resistance	NP_417544_5	3000024	patA
Gene Modulating Resistance	NP_358969_1	3000025	patB
Class A Beta-Lactamase	YP_001569085_1	3000621	PC1
Class C Beta-Lactamase	ACQ82810_1	3002502	PDC-5
Class B Beta-Lactamase	AJP77059	3003670	PEDO-1
Class B Beta-Lactamase	AJP77071	3003714	PEDO-2
Class B Beta-Lactamase	AJP77076	3003715	PEDO-3
Class A Beta-Lactamase	CAA79968_1	3002363	PER-1
Other ARG	BAG33043_1	3003920	pgpB
MATE Transporter	NP_250052_1	3004077	PmpM
MFS Antibiotic Efflux	NP_358469_1	3000822	pmrA
Other ARG	BAE78116_1	3003576	PmrC
Other ARG	AAC75089_1	3003577	PmrE
Other ARG	AAC75314_1	3003578	PmrF
Gene Modulating Resistance	WP_002530866_1	3003974	gyrA
Gene Modulating Resistance	BAA37152_1	3003702	gyrA|parC
Other ARG	NP_249649_1	3003686	oprD
Gene Modulating Resistance	NP_249870_1	3003895	PhoP
Gene Modulating Resistance	NP_249871_1	3003896	PhoQ
Gene Modulating Resistance	NP_251858_1	3003684	gyrA
Gene Modulating Resistance	NP_253654_1	3003685	parE
SMR Antibiotic Efflux	NP_253677_1	3004038	emrE
RND Antibiotic Efflux	SIP52035_1	3004054	CpxR
Other ARG	AAM15533_1	3003688	PvrR
MFS Antibiotic Efflux	BAJ09383_1	3003046	qacA
SMR Antibiotic Efflux	AAZ42322_1	3003836	qacH
MFS Antibiotic Efflux	AEZ36150_1	3000448	qepA
Quinolone Resistance	ABI50486_1	3002707	QnrA1
Quinolone Resistance	AEL31272_1	3002756	QnrB41
Quinolone Resistance	ACK75961_1	3002787	QnrC
Quinolone Resistance	ACG70184_1	3002788	QnrD1
Quinolone Resistance	AEG74319_1	3002794	QnrS5
Quinolone Resistance	ADI81040_1	3002800	QnrVC3
Quinolone Resistance	AJA36815_1	3003193	QnrVC7
Class A Beta-Lactamase	CAA37699_1	3003565	r39
Gene Modulating Antibiotic Efflux	AFK13828_1	3000823	ramA

Other ARG	ADV91011_1	3000245	RbpA
Rifampin Resistance	AFO53532_1	3002883	rgt1438
rRNA Methyltransferase	AJD73064_1	3001301	RlmA(II)
Class B Beta-Lactamase	AGU01679_2	3003894	Rm3
rRNA Methyltransferase	BAC20579_1	3000859	rmtA
rRNA Methyltransferase	YP_001816610_1	3000860	rmtB
rRNA Methyltransferase	AIA09786_1	3000861	rmtC
rRNA Methyltransferase	ABY64751_1	3002667	rmtD
rRNA Methyltransferase	AFJ11385_1	3002666	rmtF
rRNA Methyltransferase	AGE00988_1	3002668	rmtG
rRNA Methyltransferase	AGH19769_1	3003198	rmtH
Class A Beta-Lactamase	CAA37052_1	3002995	ROB-1
RND Antibiotic Efflux	AFK13827_1	3000825	robA
MFS Antibiotic Efflux	AAC60781_1	3003048	rosA
MFS Antibiotic Efflux	AAC60780_1	3003049	rosB
Rifampin Resistance	AIA08936_1	3000444	rphA
Rifampin Resistance	APB03222_1	3003992	rphB
Tetracycline Resistance	YP_208874_1	3003930	rpsJ
ABC-F Ribosomal Protection	AGN74946	3003749	salA
RND Antibiotic Efflux	ACH50230_1	3003379	ramR
Gene Modulating Resistance	NP_461214_1	3003926	gyrA
Gene Modulating Resistance	NP_462089_1	3003939	parC
Gene Modulating Resistance	NP_462096_1	3003317	parE
Other ARG	BAD95494_1	3002895	SAT-1
Other ARG	CAA88265_1	3002898	SAT-3
Other ARG	AAB53445_1	3002897	SAT-4
Gene Modulating Resistance	YP_186749_1	3000489	sav1866
RND Antibiotic Efflux	NP_460903_1	3000826	sdiA
Class A Beta-Lactamase	AAK63223_1	3003561	Sed1
Class B Beta-Lactamase	AAT90847_1	3003557	SFB-1
Class B Beta-Lactamase	AAF09244_1	3000849	SFH-1
rRNA Methyltransferase	WP_063978071_1	3000862	sgm
Class A Beta-Lactamase	CAQ03505_1	3001338	SHV-100
Class B Beta-Lactamase	ACT66697_1	3000846	SIM-1
Class B Beta-Lactamase	AAT90846_1	3003556	SLB-1
Class B Beta-Lactamase	BAL14456_1	3000854	SMB-1
Class A Beta-Lactamase	AAG29813_1	3002380	SME-2
RND Antibiotic Efflux	AAD51344_1	3003051	smeA
RND Antibiotic Efflux	AAD51345_1	3003052	smeB
RND Antibiotic Efflux	AAD51346_1	3003053	smeC
RND Antibiotic Efflux	CAC14594_1	3003055	smeD

RND Antibiotic Efflux	CAC14595_1	3003056	smeE
RND Antibiotic Efflux	CAC14596_1	3003057	smeF
RND Antibiotic Efflux	AAD51348_1	3003066	smeR
RND Antibiotic Efflux	AAD51347_1	3003067	smeS
Aminoglycoside Nucleotidyltransferase	AGW81558_1	3002631	spd
Class B Beta-Lactamase	AJP77080	3003720	SPG-1
Class B Beta-Lactamase	CAD37801_1	3003793	SPM-1
ABC-F Ribosomal Protection	CAA45050_1	3002828	srmB
Class A Beta-Lactamase	BAA23130_1	3002493	SRT-1
Gene Modulating Resistance	ADJ67256	3003319	mprF
Gene Modulating Resistance	WP_001025093	3003323	pgsA
Gene Modulating Resistance	YP_039482_1	3003301	gyrB
Gene Modulating Resistance	YP_039483_1	3003296	gyrA
Gene Modulating Resistance	YP_039996_1	3003285;3003287	rpoB
Gene Modulating Resistance	YP_039997_1	3003291	rpoC
Gene Modulating Resistance	YP_040771_1	3003315	parE
Gene Modulating Resistance	YP_500802_1	3003074	cls
Gene Modulating Resistance	AAK74984_1	3003311	parC
Gene Modulating Resistance	AAL97684_1	3003387	folP
Penicillin-binding Protein	AFC91828_1	3003041	PBP1a
Penicillin-binding Protein	AFC91898_1	3003043	PBP2x
Penicillin-binding Protein	NP_359110_1	3003042	PBP2b
Gene Modulating Resistance	Q8DWT2	3003774	mprF
Gene Modulating Resistance	AAO47226_2	3003318	parY
Gene Modulating Resistance	CAA67349_1	3003359	EF-Tu
Gene Modulating Resistance	AEJ33969_1	3000410	sul1
Gene Modulating Resistance	AAL59753_1	3000412	sul2
Gene Modulating Resistance	ACJ63260_1	3000413	sul3
Gene Modulating Resistance	APB03219_1	3003986	TaeA
MFS Antibiotic Efflux	CAA03986_1	3000343	tap
MFS Antibiotic Efflux	AAA67509_1	3003554	tcmA
MFS Antibiotic Efflux	BAA07390_1	3002893	tcr3
Class A Beta-Lactamase	AHA80960_1	3001388	TEM-211
MFS Antibiotic Efflux	AAD09860_1	3000561	tet(30)
MFS Antibiotic Efflux	CAC80727_1	3000476	tet(31)
MFS Antibiotic Efflux	CAD12227_1	3000478	tet(33)
Gene Modulating Resistance	AAK37619_1	3000481	tet(35)
MFS Antibiotic Efflux	AAV80464_1	3000565	tet(38)
MFS Antibiotic Efflux	AAW66497_1	3000566	tet(39)
MFS Antibiotic Efflux	AFK31666_1	3000567	tet(40)
MFS Antibiotic Efflux	AAP93922_1	3000569	tet(41)

MFS Antibiotic Efflux	ACD35503_1	3000572	tet(42)
MFS Antibiotic Efflux	ACS83748_1	3000573	tet(43)
MFS Antibiotic Efflux	ADE08374_2	3003196	tet(45)
MFS Antibiotic Efflux	YP_007503840_1	3000165	tet(A)
MFS Antibiotic Efflux	AAO16462_1	3000167	tet(C)
MFS Antibiotic Efflux	CAE51745_1	3000168	tet(D)
MFS Antibiotic Efflux	AAA71915_1	3000173	tet(E)
MFS Antibiotic Efflux	AAD25538_1	3000174	tet(G)
MFS Antibiotic Efflux	CAA75663_1	3000175	tet(H)
MFS Antibiotic Efflux	AAD12753_1	3000177	tet(J)
MFS Antibiotic Efflux	YP_003283625_1	3000178	tet(K)
MFS Antibiotic Efflux	AAA22851_1	3000179	tet(L)
MFS Antibiotic Efflux	AAB84282_1	3000181	tet(V)
MFS Antibiotic Efflux	AAC72341_1	3000182	tet(Y)
MFS Antibiotic Efflux	AAD25063_1	3000183	tet(Z)
Tetracycline Resistance	CAC41371_1	3000196	tet32
Tetracycline Resistance	BAB59035_1	3002870	tet34
Tetracycline Resistance	CAD55718_1	3000197	tet36
Tetracycline Resistance	AAN28721_1	3002871	tet37
Tetracycline Resistance	CBH51823_1	3000556	tet44
Gene Modulating Resistance	AET10444_1	3004032	tetA(46)
MFS Antibiotic Efflux	APB03214_1	3003980	tetA(48)
Gene Modulating Resistance	ANZ79240_1	3004035	tetA(60)
MFS Antibiotic Efflux	AAA20116_1	3000180	tetA(P)
Gene Modulating Resistance	AET10445_1	3004033	tetB(46)
MFS Antibiotic Efflux	APB03215_1	3003981	tetB(48)
Gene Modulating Resistance	ANZ79241_1	3004036	tetB(60)
Tetracycline Resistance	AAA20117_1	3000195	tetB(P)
Tetracycline Resistance	CAJ67339_1	3000186	tetM
Tetracycline Resistance	AAA23033_1	3000190	tetO
Tetracycline Resistance	CAA79727_1	3000191	tetQ
MFS Antibiotic Efflux	CAD09823_1	3003479	tetR
Tetracycline Resistance	AAA25293_1	3000192	tetS
Tetracycline Resistance	AAF01499_1	3000193	tetT
Tetracycline Resistance	ACA23185_1	3000194	tetW
Tetracycline Resistance	AAA27471_1	3000205	tetX
Class B Beta-Lactamase	CAC33832_1	3000851	THIN-B
Class A Beta-Lactamase	AAD37403_1	3003202	TLA-1
Class A Beta-Lactamase	CAG27800_1	3003203	TLA-2
Class A Beta-Lactamase	AAA19882_1	3003562	TLE
rRNA Methyltransferase	AAD12162_1	3001299	tlrB

ABC-F Ribosomal Protection	AAA26832_1	3002827	tlrC
Other ARG	CAB12108_2	3003059	tmrB
Gene Modulating Resistance	ACN32294_1	3000237	tolC
RND Antibiotic Efflux	NP_248846_1	3003679	TriA
RND Antibiotic Efflux	NP_248847_1	3003680	TriB
RND Antibiotic Efflux	NP_248848_1	3003681	TriC
rRNA Methyltransferase	CCP44409_1	3003060	tsnr
Class B Beta-Lactamase	AAN63648_1	3000844	TUS-1
Gene Modulating Resistance	WP_004025678_1	3003305	gyrB
Gene Modulating Resistance	WP_010891786_1	3003309	parC
Glycopeptide Resistance	AAA65956_1	3000010	vanA
Glycopeptide Resistance	AHH83938_1	3000013	vanB
Glycopeptide Resistance	AAA24786_1	3000368	vanC
Glycopeptide Resistance	AAM09849_1	3000005	vanD
Glycopeptide Resistance	AAL27442_1	3002907	vanE
Glycopeptide Resistance	AAF36803_1	3002908	vanF
Glycopeptide Resistance	ABA71731_1	3002909	vanG
Glycopeptide Resistance	AAA65955_1	3002942	vanHA
Glycopeptide Resistance	AAB05626_1	3002943	vanHB
Glycopeptide Resistance	AAM09850_1	3002944	vanHD
Glycopeptide Resistance	AAF36802_1	3002945	vanHF
Glycopeptide Resistance	ACL82960_1	3002947	vanHM
Glycopeptide Resistance	AHA41499_1	3002948	vanHO
Glycopeptide Resistance	AEP96393_1	3003723	vanI
Glycopeptide Resistance	NP_627787	3002914	vanJ
Glycopeptide Resistance	WP_011461306	3003727	vanKI
Glycopeptide Resistance	ABX54687_1	3002910	vanL
Glycopeptide Resistance	ACL82961_1	3002911	vanM
Glycopeptide Resistance	AEP40500_1	3002912	vanN
Glycopeptide Resistance	AHA41500_1	3002913	vanO
Glycopeptide Resistance	AAA65953_1	3002919	vanRA
Glycopeptide Resistance	AAB05622_1	3002921	vanRB
Glycopeptide Resistance	AAF86641_1	3002922	vanRC
Glycopeptide Resistance	AAM09851_1	3002923	vanRD
Glycopeptide Resistance	AAL27445_1	3002924	vanRE
Glycopeptide Resistance	AAR84672_1	3002925	vanRF
Glycopeptide Resistance	ABA71727_1	3002926	vanRG
Glycopeptide Resistance	WP_011461303	3003728	vanRI
Glycopeptide Resistance	ABX54691_1	3002927	vanRL
Glycopeptide Resistance	ACL82957_1	3002928	vanRM
Glycopeptide Resistance	AEP40503_1	3002929	vanRN

Glycopeptide Resistance	AHA41505_1	3002930	vanRO
Glycopeptide Resistance	AAA65954_1	3002931	vanSA
Glycopeptide Resistance	AAB05623_1	3002932	vanSB
Glycopeptide Resistance	AAF86642_1	3002933	vanSC
Glycopeptide Resistance	ACM47284	3002934	vanSD
Glycopeptide Resistance	AAL27446_1	3002935	vanSE
Glycopeptide Resistance	AAR84673_1	3002936	vanSF
Glycopeptide Resistance	ABA71728_1	3002937	vanSG
Glycopeptide Resistance	WP_011461302_1	3003726	vanSI
Glycopeptide Resistance	ABX54692_1	3002938	vanSL
Glycopeptide Resistance	ACL82958_1	3002939	vanSM
Glycopeptide Resistance	AEP40504_1	3002940	vanSN
Glycopeptide Resistance	AHA41504_1	3002941	vanSO
Glycopeptide Resistance	AAD22403_1	3002970	vanTC
Glycopeptide Resistance	AAL27444_1	3002971	vanTE
Glycopeptide Resistance	ABA71733_1	3002972	vanTG
Glycopeptide Resistance	ABX54689_1	3002973	vanTmL
Glycopeptide Resistance	AEP40502_2	3002975	vanTN
Glycopeptide Resistance	ABX54690_1	3002974	vanTrL
Glycopeptide Resistance	ABA71726_1	3000575	vanU
Glycopeptide Resistance	AAO82019_1	3002916	vanV
Glycopeptide Resistance	AAB05625_1	3002964	vanWB
Glycopeptide Resistance	ABA71730_1	3002965	vanWG
Glycopeptide Resistance	WP_005813024_1	3003724	vanWI
Glycopeptide Resistance	AAA65957_1	3002949	vanXA
Glycopeptide Resistance	AAB05628_1	3002950	vanXB
Glycopeptide Resistance	AAM09852_1	3003070	vanXD
Glycopeptide Resistance	AAF36804_1	3002952	vanXF
Glycopeptide Resistance	WP_015943580_1	3003725	vanXI
Glycopeptide Resistance	ACL82962_1	3002953	vanXM
Glycopeptide Resistance	AHA41501_1	3002954	vanXO
Glycopeptide Resistance	AAF61331_1	3002966	vanXYC
Glycopeptide Resistance	AAL27443_1	3002967	vanXYE
Glycopeptide Resistance	ABA71732_1	3003069	vanXYG
Glycopeptide Resistance	ABX54688_1	3002968	vanXYL
Glycopeptide Resistance	AEP40501_1	3002969	vanXYN
Glycopeptide Resistance	AAA65958_1	3002955	vanYA
Glycopeptide Resistance	AAB05624_1	3002956	vanYB
Glycopeptide Resistance	AAM09853_1	3002957	vanYD
Glycopeptide Resistance	AAF36805_1	3002958	vanYF
Glycopeptide Resistance	ABA71729_1	3002959	vanYG1

Glycopeptide Resistance	ACL82959_1	3002961	vanYM
Glycopeptide Resistance	AAA65959_1	3002962	vanZA
Glycopeptide Resistance	AAF36806_1	3002963	vanZF
Other ARG	AAA26683_1	3002840	vatA
Other ARG	AAA86871_1	3002841	vatB
Other ARG	AAC61671_1	3002842	vatC
Other ARG	AAK84316_1	3002843	vatD
Other ARG	AAF86220_1	3002844	vatE
Other ARG	AAF63432	3003744	vatF
Other ARG	ACX92987_1	3002845	vatH
Other ARG	APB03220_1	3003987	VatI
Class A Beta-Lactamase	ALU64000	3003713	VCC-1
Class A Beta-Lactamase	AAK14293_1	3002378;3003711	VEB-9|VEB-1a
ABC-F Ribosomal Protection	AGN33258_1	3002829	vgaA
ABC-F Ribosomal Protection	AAB95639_1	3000118	vgaB
ABC-F Ribosomal Protection	AMP35312_1	3002831	vgaC
ABC-F Ribosomal Protection	ACX92986_2	3002832	vgaD
ABC-F Ribosomal Protection	CBY88983_1	3002833	vgaE
Other ARG	AAA98349_1	3001307	VgbA
Other ARG	AAC61670_1	3001308	VgbB
Other ARG	APB03225_1	3003990	VgbC
Class B Beta-Lactamase	AAN52134_1	3002275	VIM-5
Class B Beta-Lactamase	CAD61201_1	3002277	VIM-7
Other ARG	CAA26235_1	3003061	vph
Class A Beta-Lactamase	AAX55643_1	3003558	y56
SMR Antibiotic Efflux	CAB13166_1	3003063	ykkC
SMR Antibiotic Efflux	CAB13167_1	3003064	ykkD
Gene Modulating Resistance	NP_416715_1	3003952	YojI