Investigating the effect of probiotic supplementation on the gastrointestinal microbiota of preterm infants

Author: Erica Plummer

Submitted in total fulfilment of the requirements of the degree of

Master of Science (Bioinformatics)

November 2014

Faculty of Science

The University of Melbourne

Declaration

This is to certify that:

i. *the thesis comprises only my original work towards the masters except where indicated in the Preface,*
ii. *due acknowledgement has been made in the text to all other material used,*
iii. *the thesis is less than 15,000 words in length, exclusive of figures, tables, maps, bibliographies and appendices.*

Signed: ____Erica Plummer____

Acknowledgements

I would like to express my sincere thanks to my supervisors Dr. Jimmy Twin, Dr. Dieter Bulach, A/Prof. Sepehr Tabrizi and Prof. Suzanne Garland for their support of my Masters study. The guidance and teaching they provided has helped me immensely in the completion of this project and thesis.

Preface

Sample collection was conducted as part of *The ProPrems Trial*.

The Centre for Women's Infectious Diseases, The Royal Women's Hospital, Melbourne completed all sample preparation and molecular biology activities. This project would not have been possible without the many hours spent in the laboratory by Dr. Jimmy Twin and colleagues.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 3

PREFACE 4

LIST OF FIGURES 7

LIST OF TABLES 8

1. INTRODUCTION AND JUSTIFICATION 9

2. OBJECTIVES 10

2.1 RESEARCH QUESTION 10
2.2 PRIMARY OBJECTIVE 10
2.3 SECONDARY OBJECTIVES 10

3. LITERATURE REVIEW 10

3.1 GI MICROBIOTA - CHARACTERISATION OF COLONISATION 10
3.2 GI MICROBIOTA - HEALTH, DISEASE AND THERAPEUTIC TARGET 13
3.3 SEQUENCING APPROACHES 15
3.4 BIOINFORMATICS PIPELINES 21

4. METHODOLOGY 26

4.1 SAMPLE COLLECTION 26
4.2 SAMPLE PREPARATION AND DNA SEQUENCING 27
4.3 DNA QUANTIFICATION 27
4.4 TAXONOMIC ANALYSIS 27
4.4.1 QIIME 28
4.4.2 MOTHUR 29
4.4.3 MG-RAST 30
4.5 DIVERSITY ANALYSIS 31
4.6 MULTIPLE SEQUENCE ALIGNMENTS AND PHYLOGENETIC ANALYSIS OF PROBIOTIC GENERA 31
4.7 STATISTICAL ANALYSIS 32
4.7.1 EXPLORATORY DATA ANALYSIS 32
4.7.2 COMPARISON OF PIPELINES 32
4.7.3 TAXONOMIC AND DIVERSITY DIFFERENCES 32
4.7.4 BACTERIAL LOAD ANALYSIS 33

5. RESULTS 33

5.1 QUALITY CONTROL AND PRE-PROCESSING 33
5.2 COMPARISON OF PIPELINES 34
5.2.1 NUMBER OF ANNOTATIONS 34
5.2.2 TAXONOMIC COMPOSITION COMPARISONS 35
5.2.3 DIVERSITY ANALYSIS COMPARISONS 38
5.2.4 USABILITY COMPARISONS 39
5.3 TAXONOMIC ANALYSIS 40
5.4 DIVERSITY ANALYSIS 49

5.5 BACTERIAL LOAD 52
5.6 INVESTIGATION OF THE PROBIOTIC 53

6. DISCUSSION 54

7. CONCLUSION 61

7.1 WHAT DIDN'T WE ACHIEVE? 62

8. REFERENCES 63

9. APPENDICES 67

List of Figures

Figure 1 – Factors that Influence the GI Microbiota Colonisation 13
Figure 2 – Schematic Representation of the 16S rRNA Gene 16
Figure 3 – Two Different Sequencing Approaches for Analysing Bacterial Communities 18
Figure 4 – Copy Number Variation and 16S rRNA Gene Analysis. 19
Figure 5 – Overview of Workflows used by MG-RAST, QIIME and mothur.....25
Figure 6 – Overview of Study Methodology.....26
Figure 7 – Number of Reads Entering Each Pipeline versus Number of Reads Annotated..... 35
Figure 8 – Comparison of Taxonomic Classification by MG-RAST, QIIME and mothur at a Phylum Level 36
Figure 9 – Venn Diagram Showing the Number of Genera Identified by Each Pipeline 36
Figure 10 – Comparison of Unclassified Reads Generated by MG-RAST, QIIME and mothur..... 37
Figure 11 - Comparison of Diversity Measures Between MG-RAST, QIIME and mothur 38
Figure 12 – Abundance of Phyla in *ProPrems* Samples 40
Figure 13 - Bacterial Diversity in *ProPrems* Samples.....42
Figure 14 – Graphical Representation of Ordination Analysis Between Allocation Groups 43
Figure 15 – Graphical Representation of Ordination Analysis Between Age Groups 44
Figure 16 – Principal Component Analysis of Genus Abundance Data 45
Figure 17 – Principal Component Analysis of Genus Abundance Data post removal of *Bifidobacterium* ..47
Figure 18 – Effect of Allocation Group on Diversity Measures 49
Figure 19 – Effect of Age on Diversity 50
Figure 20 – The Effect of Age and Allocation Group on Beta Diversity.....51
Figure 21 – The Relationship Between Age and Bacterial Load.....52
Figure 22 – The Relationship Effective Number of Genera and Bacterial Load..... 53

List of Tables

Table 1 – Comparison of Three Bioinformatics Pipelines ……22
Table 2 – Relevant Sample Characteristics …… 33
Table 3 – Age Groups Selected for Analysis…… 34
Table 4 – Statistical significance (p-value) of Number of Reads Annotated by Each Pipeline…… 35
Table 5 – The Five Most Abundant Genera Detected by MG-RAST, QIIME and mothur …… 37
Table 6 – Statistical Significance (pair wise p-values) of Diversity Measures …… 39
Table 7 – Most Abundant Genera in Probiotic and Placebo Groups and Their Contribution to Dissimilarity ……46
Table 8 – Dissimilarity Between Age Groups ……46
Table 9 – Dissimilarity Between Age Groups Following Removal of *Bifidobacterium* ……48
Table 10 – Results of post Hoc Testing of Differences in Effective Number of Genera Between Age Groups ……50
Table 11 – Average Bacterial Load Across Different Age Groups …… 53

1. Introduction and justification

The ProPrems Trial (1, 2) was a large multi-centre, randomised, double blind, placebo controlled trial that aimed to investigate the effects of probiotic supplementation on late onset sepsis in very low birth weight (VLBW) premature infants. As part of *ProPrems*, VLBW premature infants (weighing below 1500g and born at less than 32 weeks gestation) were randomised to receive either a probiotic mixture (*Bifidobacterium longum* subsp. *Infantis* BB-02, *Streptococcus thermophilus* TH-4 and *B. animalis* subsp. *Lactis* BB-12) or placebo shortly after birth. The study recruited infants from Australia and New Zealand and collected up to eight longitudinal faecal samples from Victorian participants (1).

This current project aims to characterise and compare the overall early life bacterial colonisation pattern (i.e. how quickly, which organisms and in what abundance) of the gastrointestinal (GI) tract in premature infants between those who received the probiotic and those who received the placebo. We processed between two and seven faecal samples from 67 infants for a total of 300 samples. A total of 215 samples were included in the final analysis.

Our understanding of the human GI microbiota is limited; our understanding of the premature infant's GI microbiota is even more limited. We know that the GI microbiota has an important role in health and immunity, and that disrupting it can lead to disease. However, there is still much to learn about the GI microbiota composition, including how it varies between and within individuals, over a lifetime, and in different disease states. Furthermore, the use of probiotics for the treatment of conditions with associated dysbiosis, such as necrotising enterocolitis (NEC), also warrants investigation. NEC is an inflammatory disease of the GI tract. It has a high mortality rate and is one of the most common diseases seen in Neonatal Intensive Care Units (NICU). It is multifactorial and abnormal gut colonisation is hypothesised to be one of the contributing factors. Because of the linkage with bacterial colonisation there is interest in investigating the effect of probiotic prophylaxis in infants at risk of NEC (3, 4).

The exact mechanism by which probiotics work is not known, but the effects of using probiotics can be investigated by measuring changes in the bacterial communities that make up the GI microbiota. Methods for measuring these changes are well established and are based around high-throughput sequencing. By determining the sequence of a specific gene

(typically the 16S rRNA gene) in a sample of the GI microbiome, we can generate a taxonomic overview of the bacteria present in the sample.

This is important research that will address key questions about how GI microbiota changes over time and in response to probiotics in a cohort of preterm infants.

2. Objectives

2.1 Research Question

What is the effect of probiotic supplementation with *Bifidobacterium longum* subsp. *infantis*, *Streptococcus thermophilus* and *B. animalis* subsp. *lactis* on the GI microbiota of premature infants?

2.2 Primary Objective

GI microbiota samples previously collected over a period of one to two years commencing after birth from preterm infants from *The ProPrems Trial* (1, 2) will be analysed to determine a taxonomic profile for each sample. Comparisons of the profiles will be used to evaluate:

1. Bacterial colonisation patterns over the first one to two years of life,
2. Differences between infants randomised to probiotic and those randomised to placebo.

2.3 Secondary Objectives

1. To compare and refine methods for generating and analysing GI microbiota profiles.

3. Literature review

3.1 GI microbiota - characterisation of colonisation

Microbiota is defined as the microbial taxa present in a specific location; microbiome is defined as the collection of microorganisms that inhabit an environment, their genetic material, and their interactions with that environment (3). Distinct microbiotas exist in and on the human body (4). The GI microbiota is complex and is comprised of up to 10^{14} CFU/ml in the colon and more than 800 different microbial species (5, 6). In comparison, 10^{8} CFU/ml are found in vaginal secretions and between 10^{8} - 10^{9} CFU/ml are found in saliva (7, 8).

Due to the high variability in microbiota between individuals, there is no clear consensus of what defines a healthy or normal GI microbiota. A study by Turnbaugh et al. (9) attempted to quantitate the diversity of the GI microbiota and microbiome in twins and reported that no single phylotype was detected at an abundant frequency in all 154 individuals analysed. The participants included 31 monozygotic twin pairs and 23 dizygotic twin pairs.

It is however apparent that in the context of number of organisms present in the GI microbiota, Firmicutes and Bacteroidetes are the two main phyla present. Proteobacteria, Actinobacteria and Verrucomicrobia are also present, but in smaller abundance. Spirochaetes and Fusobacteria have also been observed in healthy adults (10-13).

The infant GI microbiota

The infant microbiota while developing, is limited in diversity and susceptible to both internal and external influences (14). It is generally accepted that the foetal gut is initially sterile and that colonisation begins rapidly after birth (15, 16). It is the general consensus that gut colonisation in the newborn begins with facultative anaerobes (FA) such as Enterobacteria, coliforms and lactobacilli. The FAs deplete the gut of oxygen, leading to colonisation of strict anaerobes (e.g. *Bifidobacterium*, *Bacteroides* and *Clostridium* spp.) in the days and weeks after the establishment of low oxygen environment (12, 17, 18). The bacteria that initially colonise the neonate are thought to strongly influence the composition of the resulting adult GI microbiota and potentially the future health (immediate and long term) of an individual, and thus they are critical to the establishment of a strong and healthy GI microbiota (18).

In vaginally delivered newborns, colonisation begins following passage through the birth canal during which the infant is exposed to maternal vaginal and faecal bacteria (17, 19). Infants born by caesarean section are thought to have delayed colonisation and have a higher abundance of bacteria found in the maternal skin microbiota in their gut (12). There is evidence to suggest that breast-fed neonates have a more diverse GI microbiota and harbour fewer potentially pathogenic bacteria than formula-fed neonates. This is thought to be due not only to the bacteria and substrates present in breast milk, but also because of presumably high exposure to the maternal skin microbiota due to increased contact with maternal skin during feeding (15, 20). Other factors are thought to impact the developing GI microbiota such as prolonged stays in NICU, where intravenous/intra-arterial catheters and assisted ventilation may be in use (thus bypassing the defences of these normally sterile sites), maternal and/or

neonate exposure with antibiotics, and contact with siblings and other family members (12, 15, 19).

The preterm infant GI microbiota

Preterm infants (less than 37 weeks gestation) are an extremely vulnerable population. They are at an increased risk of death, disease and lifelong health complications (21). Preterm babies have a different GI microbiota compared to the full term baby, which may be due to the delayed colonisation through reduced exposure to maternal flora. This delay may render the preterm infant more susceptible to the colonisation of potentially pathogenic bacteria (22). Differences in flora that have been observed include a reduced diversity of species, high variability between individuals, lower numbers of bacteria normally observed in healthy individuals, and often higher numbers of Enterobacteriaceae, *Clostridium* and *Klebsiella* than the full term infant (12, 23).

The other factors influencing the establishment of the preterm GI microbiota include delayed feeding and prolonged stays in NICU (which may result in increased exposure to either or both antibiotics and nosocomial bacteria) (24). The immaturity of the preterm gut may also influence microbiota development. For example, the immature gut may not respond appropriately to colonising bacteria, which could provide an opportunity for overgrowth of potentially pathogenic bacteria such as *Clostridium* (15). The internal and external factors that have been shown in previous studies to influence the colonisation and development of the GI microbiota are summarised in Figure 1. Cilieborg et al. (15) state that while all factors may influence the developing microbiota, the dominant influences are the internal host properties, delivery mode and exposure to antibiotics.

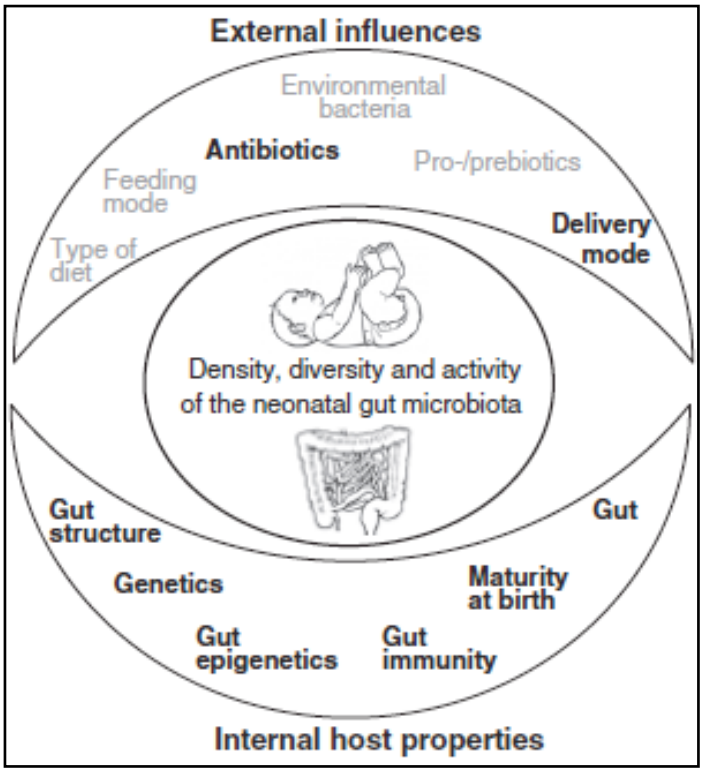


Figure 1 – Factors that Influence the GI Microbiota Colonisation

A wide range of both internal host properties and external factors influence the colonisation of the infant GI microbiota. The dominant factors influencing developing of the GI microbiota are bolded (15).

3.2 GI microbiota - health, disease and therapeutic target

The development of the GI microbiota continues until it begins to resemble that of an adult, which can be as early as one year of age and on average three years (3, 12, 25). The contemporary theory is that the GI microbiota remains highly dynamic and plastic over the entire lifetime of the host (10, 11).

Health

There are many beneficial functions of a healthy GI microbiota. It has digestive functions, provides energy to the body through metabolism of indigestible substrates, has a role in vitamin production, and has been shown to have a function in intestinal development and development of the immune system (11, 26). Arguably the most important role of the GI microbiota is protection of the host (14).

The GI microbiota protects the host in a variety of ways against colonisation and/or proliferation of potential pathogens. Commensal bacteria limit the availability of some nutrients through competition of substrates, and interaction between the immune system and

the GI microbiota can result in host tolerance towards constituents of the GI microbiota (11). The GI microbiota also teaches the immune system to discriminate between commensal and harmful bacteria (26). This is vital for the host in maintaining the balance between promoting healthy growth of symbiotic bacteria and preventing invasion and overgrowth of pathogenic species (26). Some of the bacteria that naturally inhabit the gut are potentially pathogenic (e.g. both *Enterococcus faecalis* and *Bacteroides fragilis* can cause disease in the gut of immunocompromised people) but a healthy immune system should work to protect the host against such species (11).

Disease

Mounting evidence indicates that different disease states are associated with dysbiosis. For example, obesity, cystic fibrosis, inflammatory bowel disease, periodontal disease, celiac disease, and NEC all have a unique or disrupted microbial signature (14, 26). Additionally, conditions such as asthma, eczema, and food allergies have been linked to deviations from the 'normal' GI microbiota (27).

NEC is a devastating inflammatory disease of the intestine that affects approximately 7% of preterm infants. Up to 30% of affected babies do not survive and those that do often suffer from neurodevelopmental delays and short bowel syndrome as they lose large sections of necrosed bowel (28, 29). The syndrome is associated with an enhanced immune response followed by necrosis of intestinal tissue that requires medical and sometimes surgical intervention (28, 30).

The exact aetiology of NEC is unknown, however it often occurs in clusters which suggests an infectious agent (31). Previous research has associated many bacterial signatures to NEC, as well as viruses, yet no causative pathogen(s) has been identified to date, nor the mechanisms for establishment of disease determined (1, 24, 32). There are also numerous non-microbial factors, such as intestinal immaturity, genetic predisposition, and hypoxia-ischemia, which are thought to contribute to the development of NEC. The current consensus is that the cause of NEC is likely to be multifactorial (28, 33).

Therapeutic target

Diseases and disorders that are associated with unusual bacterial signatures in the GI microbiota may respond well to bacterial supplements in the form of probiotics (14). This

approach is supported by current research, which suggests that probiotics given prophylactically may be useful in decreasing the incidence of NEC in preterm babies (2, 29). Probiotics are defined by the World Health Organisation as "live microorganisms which, when administered in adequate amounts, confer a health benefit on the host" (34). While the exact mechanism of action is unknown, there are several ways by which probiotics are hypothesised to work in NEC (29). These include modifying colonisation of the gut, influencing the immune response, reducing intestinal permeability (potentially via strengthening tight junctions in the gut), and improving the mucosal barrier of the gut (30, 34-36). Not only are probiotics thought to inhibit colonisation of pathogenic bacteria (through both antimicrobial effects and competition of substrates), but recent research suggests that colonisation of the probiotic strain may be enhanced (29).

Bacterial strains that have been administered as part of probiotic preparations that are used to reduce the incidence of NEC include: *Bifidobacterium breve, B. infantis, B. bifidus, B. lactis, B. longum, Lactobacillus rhamnosus, L. acidophilus, L. casei, Streptococcus thermophilus* and *Saccharomyces boulardii* (37). Bernardo et al. (35) conducted a meta-analysis of randomised controlled trials (RCTs) investigating the effect of probiotics on severe NEC. The analysis reviewed 11 RCTs in preterm and/or VLBW infants and concluded that probiotics were beneficial in reducing the incidence of severe NEC (incidence of severe NEC in the control group was 7.2% and in the probiotics group it was 3.2%). Furthermore, *The ProPrems Trial* (2) described a 54% absolute reduction in NEC following probiotic supplementation.

Despite the findings outlines above, doubt still exists around the efficacy of probiotics in the prophylactic management of NEC and in preterm infants. These doubts are based around the absence of any single probiotic formulation that provides better outcomes and limited long-term risk data associated with the probiotic treatment or prophylaxis of preterm and VLBW infants. Additionally, there is doubt as to whether colonisation of the probiotic strain always occurs in the patient (29, 30).

3.3 Sequencing approaches

Metagenome refers to the totality of gene sequences found in a specific environment. Metagenomics is the analysis of the collective genomes in such an environment (38). High

throughput sequencing techniques have dramatically increased our ability to investigate and define metagenomes. This has fuelled projects aimed at discovering more about the human microbiome and the interactions with its host, such as the Human Microbiome Project (HMP) coordinated by the National Institute of Health (4). There are two non-culture based approaches that have been used to investigate the GI microbiota and microbiome: targeted amplicon sequencing, most commonly 16S ribosomal RNA (rRNA) gene analysis, and whole metagenome shotgun (WMS) sequencing (39). This review will discuss both methods.

Targeted amplicon sequencing and the16S rRNA gene

The 16S rRNA gene is the most commonly used target for identification of bacteria and investigations of bacterial diversity. The 16S rRNA gene is 1.5 k.bp in length and encodes 16S rRNA, which is one of the three rRNAs that makes up the bacterial ribosome. The ribosome is required for protein synthesis and thus is essential for life (40). Because of the fundamental role of the ribosome, the 16S rRNA gene is present in all bacteria. The gene consists of highly conserved regions and hypervariable regions (Figure 2). The hypervariable regions (V1-V9) vary between bacteria and can be used to determine bacterial identity. The hypervariable regions are flanked by conserved regions which are used to design primers for polymerase chain reaction (PCR) amplification of the 16S rRNA gene from a broad range of genera (41). This structure of alternating conserved and hypervariable regions across the gene and the fact that the gene is omnipresent in bacteria are the key reasons underpinning the use of the 16S rRNA gene for classification studies.

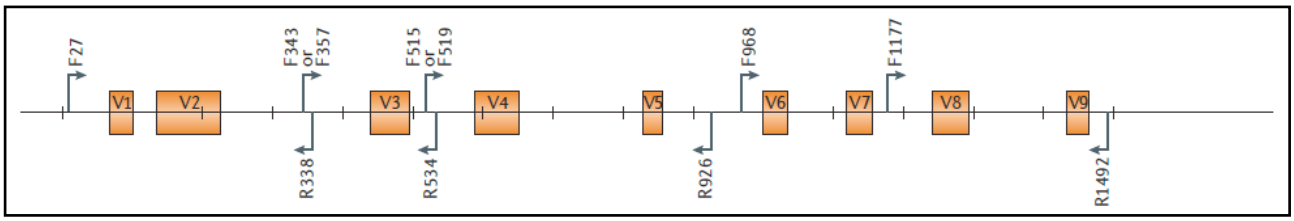


Figure 2 – Schematic Representation of the 16S rRNA Gene

The variable regions are coloured orange and the conserved regions are found along the straight line. Locations of common primers are shown on the forward (F) and reverse (R) strands (42).

A consequence of using a gene that is ubiquitous and has a conserved function across all bacteria is that there are a limited number of sequence differences between related bacteria. This means that no single hypervariable region can be used to differentiate between all bacterial species and often classification studies focus on two or more hypervariable regions. A study by Chakravorty et al. (43) inferred sequence similarity dendrograms for variable

regions V1-V8 for 110 different bacterial species. Results showed that each region had varying ability to discriminate between species. For example, V1 was shown to be effective in discriminating between *Streptococcus* spp., while V6 was most effective in discriminating between *Bacillus anthracis* and *B. cereus*. Furthermore, different regions were more effective at discriminating bacteria to a species level while other regions only discriminated to a genus level. The study determined that the combination of regions V2, V3 and V6 was best for discriminating between the 110 species. Another problem that potentially limits the utility of the 16S rRNA approach is the design of amplifying primers. For this approach to be effective, the primers should equally amplify all bacterial species present. Despite the broad conservation of sequences in the 'conserved' regions of the 16S rRNA gene, there are only a limited number of suitable primer sites (44). Commonly used primers are shown in Figure 2.

Understanding the limitations of the 16S rRNA approach to the analysis of mixed bacterial populations is particularly pertinent to this current study, as an important bacteria may go undetected (41). *Bifidobacterium* are common inhabitants of the GI microbiota, and two *Bifidobacterium* strains were present in the probiotic formula used in *The ProPrems Trial*. The detection of *Bifidobacterium* spp. with current primers is limited and their presence often underestimated in amplicon-based studies (43). Sim et al. (44) highlighted this fact by developing *Bifidobacterium* optimised primers for use in 16S rRNA gene analysis and compared results with results when using an existing primer set. Sim et al. found that when *Bifidobacterium* DNA made up over 90% of a DNA mixture, the maximum percentage of *Bifidobacterium* spp. detected using the existing primer set was 1.6%, while the optimised primer detected the proportion to be upwards of 60%. This shows how important it is to be aware of the properties of both the hypervariable regions and primers selected for amplicon studies.

Bacteria are clustered into Operational Taxonomic Units (OTUs) (through a process known as binning) based on sequence similarity of the 16S rRNA gene using reference databases such as the Ribosomal Database Project (RDP) (43) and SILVA (45). A threshold of 97% identity has historically been used for species classification (i.e. a new species is identified when the similarity score is less than 97%) (9, 46). A generalised summary of targeted amplicon analysis with the 16S rRNA gene is shown in Figure 3.

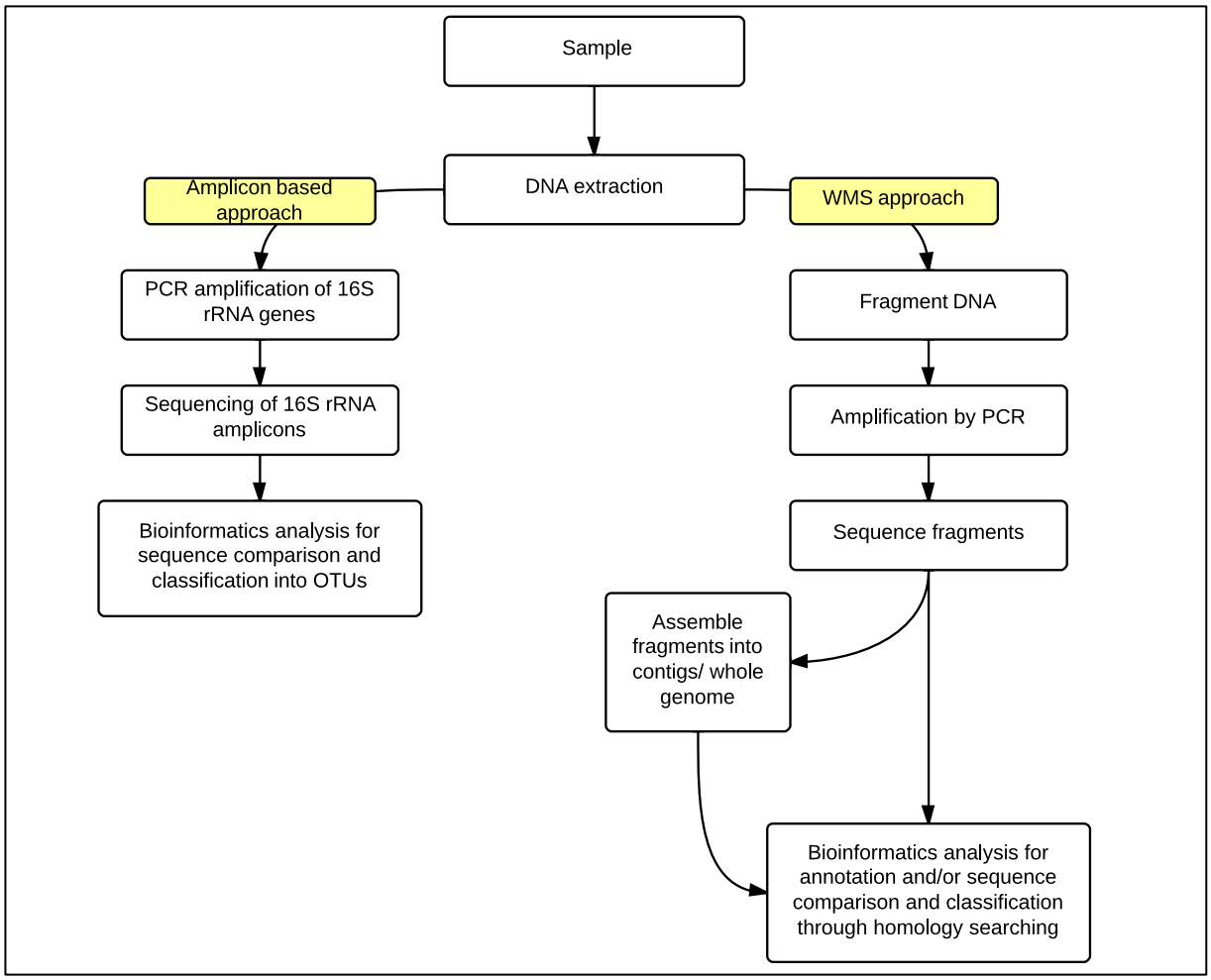


Figure 3 – Two Different Sequencing Approaches for Analysing Bacterial Communities
A summary and comparison of the steps involved in targeted amplicon based analysis (using the 16S rRNA gene) and WMS analysis. Sequencing data generated from WMS sequencing can be analysed as fragments or assembled into contigs for analysis. PCR = polymerase chain reaction, OTU = operational taxonomic unit, WMS = whole metagenome shotgun. Adapted from (42, 44, 46).

Some limitations of 16S rRNA gene analysis include:

1. Detection of the whole bacterial community: underrepresentation of particular species due to poor priming, not only as a result of lack of truly universal primers (as described above), but also due to PCR biases and sequencing errors (47).
2. Resolving power: Small number of differences in 16S rRNA gene means that there are situations where related species and strains cannot be resolved. For example, *Bacillus globisporus* and *B. psychrophilus* are greater than 99.5% similar in 16S rRNA gene sequence (48). In addition, PCR artefacts further confound analysis. For example, a chimera is a DNA sequence that has been formed by the incorrect joining of two or more unrelated sequences. During PCR, chimera formation can occur when incomplete

PCR products act as primers. The incomplete product incorrectly anneals to a sequence, extends during the PCR cycle and forms a chimeric sequence (42, 49).

3. Copy number: 16S rRNA gene copy number varies (one to 15 copies) depending on the particular species. It is important to be aware of copy number variations between species. As shown in Figure 4, copy number variation makes it difficult to estimate relative abundance of bacterial species in a sample (39, 50). This is particularly important to be aware of when using the 16S rRNA gene to quantify bacterial load, the number of copies of the 16S rRNA gene in a sample is not always equivalent to the number of bacteria present in a sample.
4. There are complications due to exceptional biology. For example, *Aeromonas veronii* can have up to six copies of the 16S rRNA gene and these copies can differ in sequence by upwards of 1.5% (48, 50).

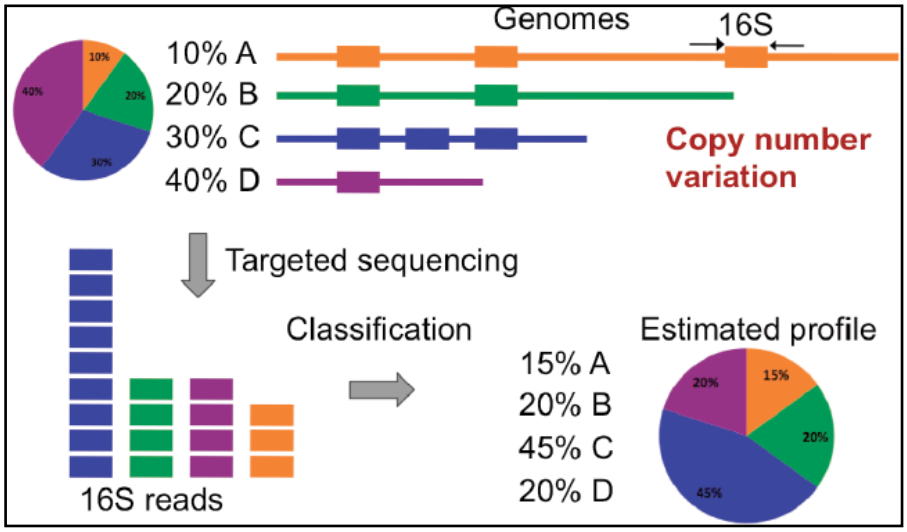


Figure 4 – Copy Number Variation and 16S rRNA Gene Analysis.
Copy number variation can influence classification and estimated profile of a microbiota. A microbiota sample may be predominately one species (e.g. species D), but because a less abundant species has multiple copies of the 16S rRNA gene in the genome (e.g. species C) it may appear to be in higher abundance (39).

Additionally, there is limited literature available to guide researchers on the number of reads per sample required for 16S rRNA analysis. A study conducted by Kuczynski et al. (51) investigated this issue and determined that a high read count is not always necessary for accurate classification. Kuczynski et al. found that while using less than approximately 100 reads per sample can lead to inaccurate taxonomic profiles, as the number of reads per sample increases above approximately 100, the additional reads add little to high-level overviews of

these patterns. The information gained by the additional reads is most important when elucidating finer details of ecological patterns and relationships.

Whole metagenome shotgun (WMS) sequencing

WMS involves sequencing the complete genomes of all organisms present in a sample. An overview of the process is shown in Figure 3. Shotgun sequences can either be analysed as is, or assembled into contigs/whole genomes and then analysed. Shotgun sequencing is free from detection biases as observed with 16S rRNA gene analysis, but this is potentially offset by lower sensitivity and a far more complicated approach to analysis of sequencing data (42, 46).

Some limitations to the WGS approach are outlined below:

1. Estimation of bacterial composition of a sample can be affected by differences in both genome sizes and abundance of species (39, 50, 52).
2. Producing high quality de novo assemblies is expensive in time and computer memory, and can be difficult due to issues such as complexity of the metagenomic sample and the requirement for high read depth (52, 53).
3. Large amounts of starting DNA is required and whole genome amplification may introduce biases such as chimerism with host and bacterial DNA (46).

Errors and biases can also be introduced during processing of samples and DNA sequencing. These errors need to be kept in mind when performing both targeted and shotgun analyses. For example:

1. There is always potential for contamination during the collection and processing of samples either from the host or scientists handling the samples (46).
2. The effects of storing complex bacterial samples for long periods of time are unknown, so it is recommended to extract the DNA soon after collection (42).
3. There are several different sequencing platforms, each of which has a unique error profile. For example, the Roche-454 platform is the most common platform used for 16S rRNA gene analysis and has an indel rate of 0.38 in 100 bases when calling homopolymers (54).

4. The method of DNA extraction and sequencing platform can affect the relative abundance of species present in a sample (55).

3.4 Bioinformatics pipelines

Studies such as the proposed require the analysis of a large number of samples. This adds a level of complexity to the analysis, and additional statistics are often needed to adequately compare samples. The large amount of data generated by amplicon and WMS is most often analysed using sophisticated bioinformatics tools and pipelines. There is limited literature comparing the functions and usability of bioinformatics pipelines for 16S rRNA gene and WMS analysis. The following presents a brief summary of three open source pipelines that can be used for comparing and investigating microbial communities: QIIME (56), mothur (57) and MG-RAST (58). Table 1 presents key features of these three pipelines.

Table 1 – Comparison of Three Bioinformatics Pipelines

	QIIME (56, 59)	mothur (57)	MG-RAST (58, 60)
License	Open-source	Open-source	Open-source
Implemented in	Python	C++	Perl
Current version	1.8.0	1.31.2	3.5
Website	http://qiime.org/	http://www.mothur.org/	http://metagenomics.anl.gov
Web-based interface	YES (http://www.n3phele.com/) Not supported by the QIIME team	NO	YES (at website above)
Usage	At the command line	At the command line	GUI at the website above
	QIIME (56, 59)	**mothur (57)**	**MG-RAST (58, 60)**
Amplicon analysis	YES	YES	YES
Shotgun analysis	YES – experimental at this stage	NO	YES
Quality control	YES	YES	YES
Databases searched	RDP, SILVA, Greengenes and custom databases accepted	RDP, SILVA and Greengenes and custom databases accepted	M5NR, The SEED, NCBI and eggNOG, for protein sequences and M5RNA, RDP, SILVA and Greengenes for ribosomal sequences
Taxonomic analysis	YES (UCLUST, RDP, BLAST, mothur)	YES (Wang/RDP approach)	YES (BLAT)
Clustering algorithm	YES (UCLUST, CD-HIT, mothur, BLAST)	YES (mothur, adapts DOTUR and CD-HIT)	YES (BLAT)
Beta and alpha diversity analysis	YES	YES	YES
Phylogenetic Tree	YES, FastTree is the default	YES, Clearcut algorithm is the default	YES
Chimera detection	YES (UCHIME, chimera slayer, BLAST)	YES (UCHIME, chimera slayer, and more)	No
Visualisation	YES, including PCA plots, OTU networks, heat maps	YES, including dendrograms, heat maps, Venn diagrams	YES, including PCA plots, heat maps, pie charts. Uses Krona and Circos for visualisation
User Support	Forum, tutorials, FAQs, help videos	Forum, SOPs, FAQs, user manual	Video tutorials, FAQs, user manual, 'How to' section on website

Where known the software used by each pipeline is named.
GUI= Graphical User Interface, NAST=Nearest Alignment Space Termination, SINA = SILVA Incremental Aligner, RDP = Ribosomal Database Project, CD-HIT=Cluster Database at High Identity with Tolerance, PyNAST= PythonNAST, BLAST= Basic Local Alignment Search Tool, PCA=Principal Coordinate Analysis, OTU=Operational Taxonomic Unit, BLAT=BLAST-Like Alignment Tool, M5NR= M5 Non redundant database, NCBI=National Center for Biotechnology Information, M5RNA=Non-redundant multisource ribosomal RNA annotation., FAQ = Frequently Asked Questions, SOPs= Standard Operating Procedures

QIIME (Quantitative Insights Into Microbial Ecology)

QIIME is primarily used to analyse data generated from amplicon-based studies, but it can also be used to analyse 16S rRNA gene fragments from WMS data (56, 59). QIIME wraps pre-existing software packages and provides users with a 'one-stop-shop' for analysing sequencing data. QIIME can be used for QC and pre-processing, assignment to OTUs (based on a search of a reference database, or based on sequence similarity using mothur), chimera detection, taxonomic assignment, phylogenetic analysis and analyses of diversity measures. The supplementary information provided by Caporaso et al. (55) and the QIIME website (58) list the numerous software packages that are incorporated in QIIME.

mothur

mothur was designed to analyse data generated by amplicon sequencing (57). Like QIIME, mothur was designed as a comprehensive resource that wraps pre-existing software including DOTUR, SONS, and TreeClimber. As mothur was designed as a single platform it can be used for almost all steps of data analysis, including reducing sequencing errors, removing chimeras and contaminants, quality control (QC), aligning user supplied sequences with reference databases, assignment to OTUs (based on sequence similarity), calculating phylogenetic distances, performing phylotype based analyses, diversity analyses and more. Schloss et al. (57) provide detailed reviews of the mothur features, and Schloss, Gevers and Westcott (47) provide protocols for analysis of 16S rRNA gene sequences generated by 454 pyrosequencing.

MG-RAST (Metagenomics-rapid annotation using subsystem technology)

MG-RAST is a high-throughput pipeline that allows the user to investigate function and composition of microbial communities. It was initially built for processing WMS data; however it can also be used to analyse amplicon sequence datasets (58). The user is taken through several steps including feature prediction (for example, looking for gene coding regions), similarity searching of protein and nucleotide databases for phylogenetic, functional and metabolic classifications, annotation, visualisation and statistical analysis of results (58, 61). MG-RAST was designed such that new analysis steps could easily be incorporated to the pathway if required. MG-RAST can be used to analyse unassembled reads, removing the requirement for assembly (58, 62). Future versions of MG-RAST will likely include an assembly pathway (60).

Unlike QIIME and mothur where the user completes all analysis steps, the MG-RAST team quality control of sequences after the user has uploaded them to the system. The order in which samples undergo quality control depends on the priority assigned to the job. Sequence data that are made publically accessible by the user is assigned the highest priority. Sequence data that are to remain indefinitely private are assigned the lowest priority and can take anywhere between one day to several weeks to complete.

Workflow

The workflows used by the three pipelines share similar steps, see Figure 5. An oversimplified summary of analysis used by the three pipelines is that sequence data undergoes quality control, clustering and taxonomic assignment. The most notable difference in the workflows is that mothur generates an alignment of the data by creating pairwise alignments between query sequences and reference sequences in a 16S rRNA gene reference database. The alignment is then cleaned to ensure all sequences overlap in the same region of the 16S rRNA gene and any columns within the alignment that don't contain sequence data are removed. This process is unique to the mothur pipeline and in a short commentary (63) Dr. Patrick Schloss (the person who initiated mothur) justifies the inclusion of these steps, stating that they increase the robustness of OTU assignment. A detailed overview of the workflows used by QIIME, mothur and MG-RAST as part of this project are provided in Section 4.4 of this document.

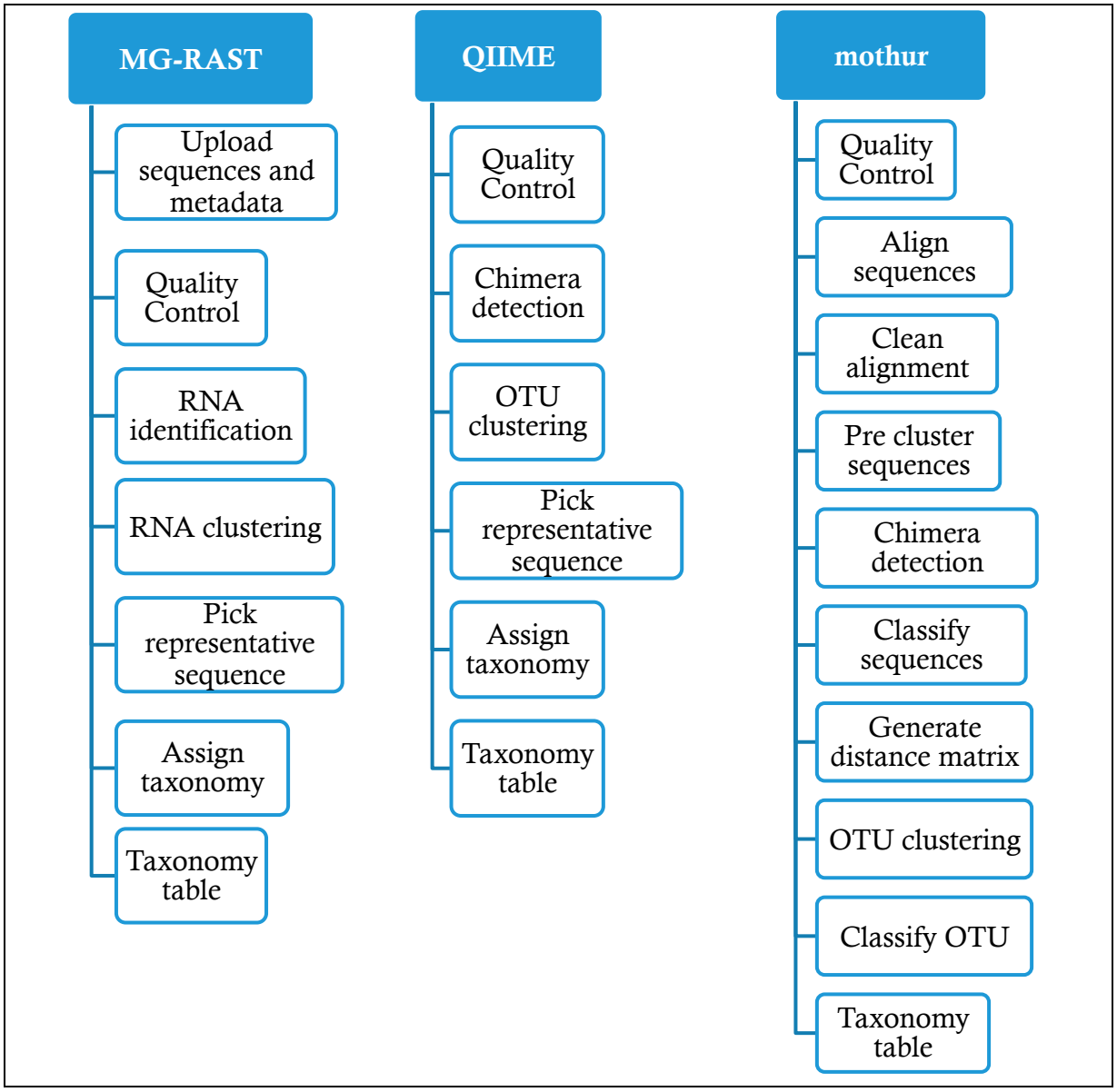


Figure 5 – Overview of Workflows used by MG-RAST, QIIME and mothur

Several steps are shared across the three pipelines (e.g. quality control, clustering, classification/assigning taxonomy). mothur has a unique step in which all sequences must be aligned to a template database and any sequences which do not overlap in the same space are removed from the analysis.

4. Methodology

An overview of the study methodology is show in Figure 6.

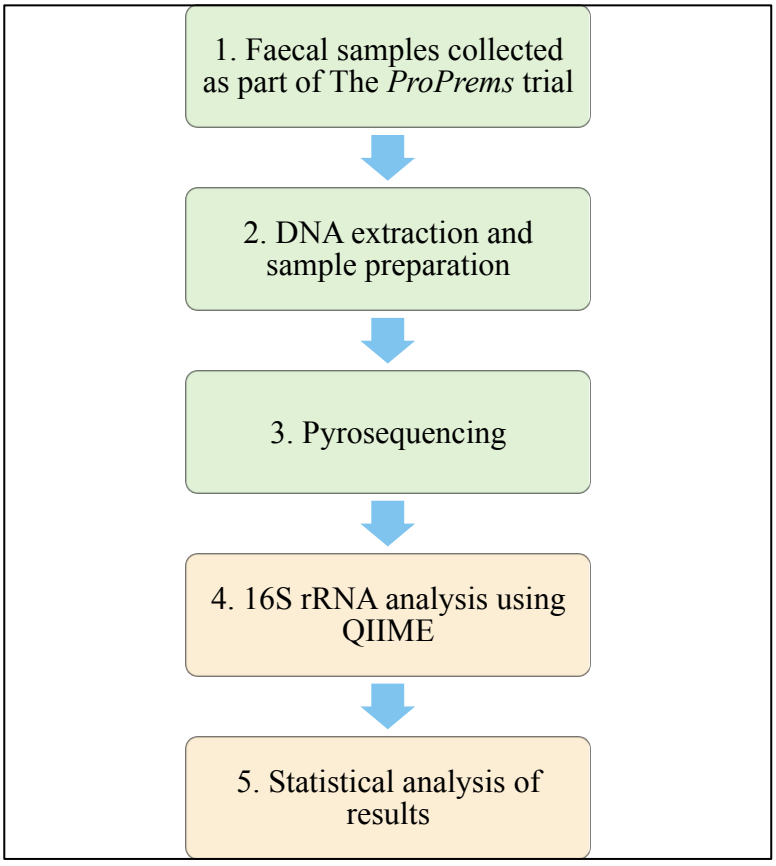


Figure 6 – Overview of Study Methodology
Steps 1-3 were completed/facilitated by The Centre for Women's Infectious Diseases. Steps 4 and 5 were completed by the Masters student.

4.1 Sample collection

The samples were collected as stool swabs (or perianal swabs) as per study protocol for *The ProPrems Trial* (1, 2). Depending on the location of the baby (hospital or the home), samples were collected by either hospital staff or a parent. A total of 300 samples were processed.

The sample size was smaller than initially anticipated. This is because samples were only included in this project if associated allergy data and blood samples had been collected from the study participant. This is to allow for future research, which will aim to investigate the impact of probiotics and composition of the GI microbiota on allergy outcomes.

4.2 Sample preparation and DNA sequencing

The Centre for Women's Infectious Diseases conducted all DNA extraction and sample preparation steps.

DNA was extracted from faecal samples using the MagNA Pure 96 System (Roche Diagnostics, Branchburg, NJ). The extracted DNA was used to generate an amplicon based library using *Bifidobacterium* optimised PCR primers 357F/926Rb (357F - CCTACGGGAGGCAGCAG, 926Rb - CCGTCAATTYMTTTRAGT) that target the V3–V5 hypervariable regions of the 16S rRNA gene as described by Sim et al (44). Each sample was barcoded with a Molecular Identifier (MID) tag (10 bp in length).

The V3-V5 regions of the 16S rRNA gene were amplified by PCR and a 2% agarose gel was used to confirm the presence of amplicons. Following PCR purification, samples were sent to Macrogen Inc. (South Korea) for pyrosequencing using the Roche 454 Genome Sequencer (GS FLX).

4.3 DNA quantification

The Centre for Women's Infectious Diseases used a modification of a PCR assay by Nikkari et al (64) with the incorporation of a *TaqMan* probe, 516F (5'-6FAM]TGCCAGCAGCCGCGGTAA[BHQ1]-3') to determine the total bacterial load of each sample (total bacterial load was defined as number of copies of 16S rRNA gene per reaction).

4.4 Taxonomic analysis

Initially, a small subset of samples (N=35) was analysed with QIIME (Version 1.8.0), mothur (Version 1.31.2) and MG-RAST (Version 3.3.7.3) to enable comparison of the three pipelines and to determine which pipeline should be used for the main study. The three pipelines were compared in terms of the abundance and type of detected phyla and genera. Sample diversity was also compared across the two pipelines. Diversity was compared at a genus level. Each pipeline can be used to calculate diversity measures, however to maintain consistency in data analysis, the Vegan package (65) in R was used to calculate the diversity measures for each dataset: genus richness, effective number of genera (as an expression of alpha diversity), and beta diversity. Details on diversity measures are provided in Section 4.5 of this document.

Based on the results of the comparative study, taxonomic analysis was then completed using QIIME.

4.4.1 QIIME

The protocol for sequence analysis was based on the 454 tutorial available from the QIIME website (http://qiime.org/tutorials/tutorial.html). Default settings were used unless otherwise specified. Sequences were de-multiplexed and underwent quality control using the *split_libraries.py* command. The following quality control parameters were used: sequences were removed if they were less than 250bp in length, contained greater than eight ambiguous bases, contained homopolymers greater than eight base pairs in length, and had an average quality score of less than 25. Chimera filtering on the resulting sequence file was performed using the *uchime_ref* command in UCHIME (66). The rRNA 16S Gold fasta file was used as the chimera checked reference database.

OTU (Operational Taxonomic Unit) picking was performed using the *pick_otus.py* command and the default UCLUST algorithm (67).

The UCLUST (67) algorithm uses the USEARCH algorithm to assign sequences to a cluster. The USEARCH algorithm works by searching a query sequence against target sequences and recording the k-mers in common between the two sequences. Rather than inferring sequence similarity as the number of matching k-mers between a query and target sequence, USEARCH arranges the target sequences in decreasing order of the number of unique k-mers shared between the two sequences. If a target sequence passes an identity threshold it is recorded as an 'accept', if not, it is recorded as a 'reject'. In the UCLUST algorithm, the query sequences are arranged into clusters. Each cluster is defined by a centroid. Each centroid shares a level of similarity below a set identity threshold level with each other centroid. The remaining query sequences are then assigned to a centroid (target sequence) based on identity threshold using the USEARCH algorithm described above. If the query sequence does not share similarity with a centroid above the threshold a new cluster is created.

Following this, a representative sequence was selected for each OTU using *pick_rep_set.py*. In QIIME, the representative sequence is identified as the most abundant sequence in each OTU. Using the representative sequences, *assign_taxonomy.py* was used to assign taxonomy. The default UCLUST consensus taxonomy assigner and SILVA reference database (Version 111) (45) were used. A 97% similarity cut-off was used for taxonomic classification.

Statistical analysis was then completed on the taxonomy tables generated.

A representative sequence from a selection of OTUs that could not be assigned an identity by QIIME were submitted to the RDP Naive Bayesian rRNA Classifier (68) to investigate whether these reads could be classified.

4.4.2 MOTHUR

The 454 analysis SOP outlined by Scholss et al. (47) was followed (mothur.org/wiki/454_SOP, accessed 30 May 2014).

Sequences were trimmed using the *trim.seqs* command. Sequences were trimmed on the same parameters as QIIME. The dataset was then simplified to reduce computational time using the *unique.seqs* command. We then generated an alignment with the *align.seqs* command using the SILVA template alignment. The alignment was cleaned using the *screen.seqs* and *filter.seqs* commands. We then ran the *pre.cluster* command to merge together any sequences that were within two base-pair-similarity of a more abundant sequence. Chimeric sequences were removed using the *chimera.uchime* command.

Sequences were assigned a taxonomic classification using the *classify.seqs* command using the SILVA reference database (Version 111) and the RDP Naive Bayesian rRNA Classifier (68). Sequences that could not be classified at a kingdom level were removed using the *remove.lineages* command. A distance matrix was built using the *dist.seqs* command (using the DNADIST algorithm (Felsenstein, 1993)). Distances greater than 0.15 were discarded to reduce the overall size of the distance matrix. We then ran the *cluster* command (which utilises the average neighbour algorithm) to cluster the sequences into OTUs based on the distance matrix. *Classify.otu* command was used to obtain consensus taxonomy for each OTU.

4.4.3 MG-RAST

Chimera screening was performed using the *uchime_ref* command in UCHIME (66) and the same reference database utilised in QIIME. Non-chimeric sequences were uploaded to MG-RAST for sequence analysis under the project ID 10404. The project remains private. The following MG-RAST pipeline options were used:

- Artificial replication sequences were removed,
- Sequences were screened for host contamination using H. Sapiens NCBI v36 database as a reference database,
- Sequences were filtered on length (reads greater than 2 standard deviations from median read length were discarded) and ambiguous bases were there was no quality score information available.

Data were analysed using the 'Best Hit Classification' option in MG-RAST. Best Hit Classification reports the highest scoring annotation/s for each read. In cases where there are two or more equally high scoring annotations, MG-RAST will report all annotations. This can result in a single read having multiple annotations. Despite displaying multiple annotations, the Best Hit Classification is preferred when searching for a specific species or genera (69) and thus was deemed most appropriate for this study, where a key focus is to identify presence/absence of the probiotic species.

MG-RAST uses the BLAT algorithm (70) to identify rRNA sequences by searching a reduced RNA database (69). The UCLUST algorithm (67) is then used to cluster identified rRNA sequences. In MG-RAST, the longest sequence of each cluster is identified as the representative sequence and it is then search against a reference RNA database using a second BLAT similarity search.

We searched the data against the SILVA reference database (Version unknown) (http://www.arb-silva.de). A maximum e-value cut-off of e-5 and a minimum alignment length of 250 base pairs were selected. A minimum percentage identity cut-off of 97% was used for taxonomic classification. Multiple annotations were identified using the pivot table function in Microsoft Excel 2010 (Microsoft Corporation, Redmond, USA), and were resolved using the RDP Naive Bayesian rRNA Classifier (68).

4.5 Diversity analysis

The following diversity measures were calculated using the Vegan package (65):

- Genera richness – measured as number of distinct genera present
- Effective Number of Genera – was used as an expression of alpha diversity. Alpha diversity was measured using the Shannon Diversity Index, which estimates the within sample diversity:

$H = -\sum_{i=1}^{S} p_i \ln p_i$, where S = Number of distinct genera, and p_i=proportion of the ith genera

Effective number of genera is calculated as exp(H)

- Beta diversity - measured by the Sørensen dissimilarity index. Dissimilarity ranges from 0 to 1, where 0 indicates complete overlap in genera between samples, and 1 indicates no similarity:

$\beta = \frac{b + c}{2a + b + c}$, where a= number of genera shared between two sites, b= number of genera unique to sample 1, and c= number of genera unique to sample two.

NOTE: raw read numbers (not proportion data) were used for diversity analysis as this is what is required by the Vegan package, and reads that were unable to be classified at the genus level were excluded from analysis to reduce any noise created by these reads.

4.6 Multiple Sequence Alignments and Phylogenetic Analysis of Probiotic Genera

Multiple sequence alignments (MSA) were generated using the Clustal Omega algorithm (71) in SeaView (Version 4.5.3) (72). We obtained the representative sequence from a selection of *Bifidobacterium* OTUs outputted from QIIME from both allocation groups and aligned them to *Bifidobacterium* type strains downloaded from the RDP Hierarchy Browser (73). We used MEGA6 (Release #6140220) (74) and the resulting MSA from SeaView to infer phylogenetic relationships. We used the Maximum Likelihood method based on the Jukes-Cantor model (75). All positions containing gaps and missing data were eliminated. The resulting tree was used to investigate:

1. Whether there were any differences in the sequence of *Bifidobacterium* reads between the allocation groups
2. Which *Bifidobacterium* type strains our *Bifidobacterium* reads were most similar to.

We also performed a MSA of *Streptococcus* spp. and aligned the PCR primers described in Section 4.2 of this document to ensure that detection of *Streptococcus* spp. was not inhibited by primer selection.

4.7 Statistical Analysis

Abundance tables were analysed in R (76).

4.7.1 Exploratory data analysis

A heat map and associated dendrograms were generated using proportion data and the heatmap.2 command in the gplots package (77) to obtain an overview of sample diversity and bacterial abundance at the order level. Clustering was performed with the Vegan package using average linkage clustering and the Bray-Curtis Dissimilarity distance measure. Reads that could not be classified to the genus level were excluded from further analysis.

Ordination analysis using Nonmetric Multidimensional Scaling (NMDS) was performed using the Picante package (78) (using the Bray-Curtis Dissimilarity distance measure) to investigate clustering of samples with resect to both allocation groups and age-groups. Further ordination analysis using Principal component analysis (PCA) was completed in PRIMER-E Version 6 (79) to identify which genera have the greatest effect on the clustering of samples.

4.7.2 Comparison of Pipelines

Where the data was normally distributed, statistical significance of differences between the pipelines were evaluated using Repeated Measures ANOVA and Pairwise comparisons using the paired t-test. The non-parametric Friedman rank sum test was used for analysis of non-normally distributed data. A significance level of alpha=0.05 was used for all tests.

4.7.3 Taxonomic and Diversity Differences

Welch's t-test and the non-parametric Wilcoxon Rank Sum test in R was used to identify differences in between the two allocation groups (Group-A and Group-B) with respect to the relative abundance of each genera present and diversity measures (76). ANOVA was used to

evaluate significance between age groups with respect to diversity and relative abundance (76). Where post hoc testing was required, Tukey's Honest Significant Difference (HSD) test was used. A significance level of alpha=0.05 was used for all tests, except where multiple testing was used, in which case, p-values were corrected for using Bonferroni's correction.

Similarity Percentage analysis (SIMPER) (80) on a bray-curtis dissimilarity measure was completed using PRIMER-6 to identify discriminating features of each age and allocation group and to measure the dissimilarity between each group.

4.7.4 Bacterial Load Analysis

Statistical analysis of bacterial load was completed on log-transformed data. We used R to investigate possible relationships between bacterial load and both allocation group and age using Welch's t-test and ANOVA respectively. Where post hoc testing was required, Tukey's Honest Significant Difference (HSD) test was used. A significance level of alpha=0.05 was used for all tests Pearson's correlation coefficient was used to investigate the relationship between bacterial load and the effective number of genera.

5. Results

5.1 Quality control and pre-processing

A total of 300 samples were processed from 67 individuals. Of these samples, 47 samples were not sent for sequencing due to insufficient PCR product. The remaining 253 samples, which included a total of 18 repeated samples, were processed through QIIME. A total of 1,064,333 raw reads were submitted to QIIME for analysis. 240 samples (a total of 787,416 reads) passed quality control.

Samples were excluded from analysis if they produced less than 100 sequences post quality control, and only one sample was included per faecal swab (i.e. no duplicate samples included in final analysis). The resulted was a total of 215 samples included in the final analysis from 66 individuals. A total of 730,861 reads were included in the final analysis. A summary of samples included in the analysis is included in Table 2.

Table 2 – Relevant Sample Characteristics

	Probiotic Group	Control Group
Number of individuals	38	28
Number of samples	124	91

Samples per baby, mean (SD)	3 (0.83)	3 (0.80)
Sequences per sample, mean (range)	3632 (110-17730)	3082 (123 – 7613)
Age of samples (days), mean (range)	96.73 (3-529)	90.99 (1 – 613)
Mean age in days probiotic/placebo started (SD)	3.59 (1.75)	4.82 (5.04)
Mean age in days probiotic/placebo stopped (SD)	68.42 (20.53)	75.96 (20.04)

Samples were divided into 6 main age groups to allow for comparisons between different ages. The number and age range of samples from each group has been summarised in Table 3.

Table 3 – Age Groups Selected for Analysis

Age Group	N (n in probiotic group)	Minimum Age	Maximum age	Key milestones/features expected based on literature review
1 to 9	28 (14)	1	9	FA should dominate and samples will be low in diversity (12, 17, 18)
10 to 19	36 (20)	10	18	All babies will have started probiotic/powder by this time
20 to 39	49 (29)	20	39	
40 to 99	53 (31)	40	96	Average Age at which probiotic supplementation stopped
100 to 299	35 (23)	228	297	All babies will have stopped probiotic treatment by this age and be at home
300 +	14 (7)	300	613	GI microbiota should start to resemble that of an adult (3, 12, 25)

Samples were assigned unique study identifiers in the format 10334_A_10 (SubjectID_Allocation_Age).

5.2 Comparison of Pipelines

A total number of 35 samples (and a total of 151,691 reads) were used for the comparative analysis of QIIME, MG-RAST and mothur.

5.2.1 Number of Annotations

The number of reads post de-multiplexing was consistent across the three pipelines (Figure 7). However, while the number of reads annotated was comparable in QIIME and MG-RAST (123,909 and 123,022, respectively), mothur annotated only 26,477 reads. The p-values from the pairwise comparisons using paired t-tests are summarised in Table 4. The difference between mothur and both QIIME and MG-RAST in terms of number of reads annotated was considered statistically significant.

Table 4 – Statistical significance (p-value) of Number of Reads Annotated by Each Pipeline

	MG-RAST	mothur
mothur	<0.001	-
QIIME	0.71	<0.001

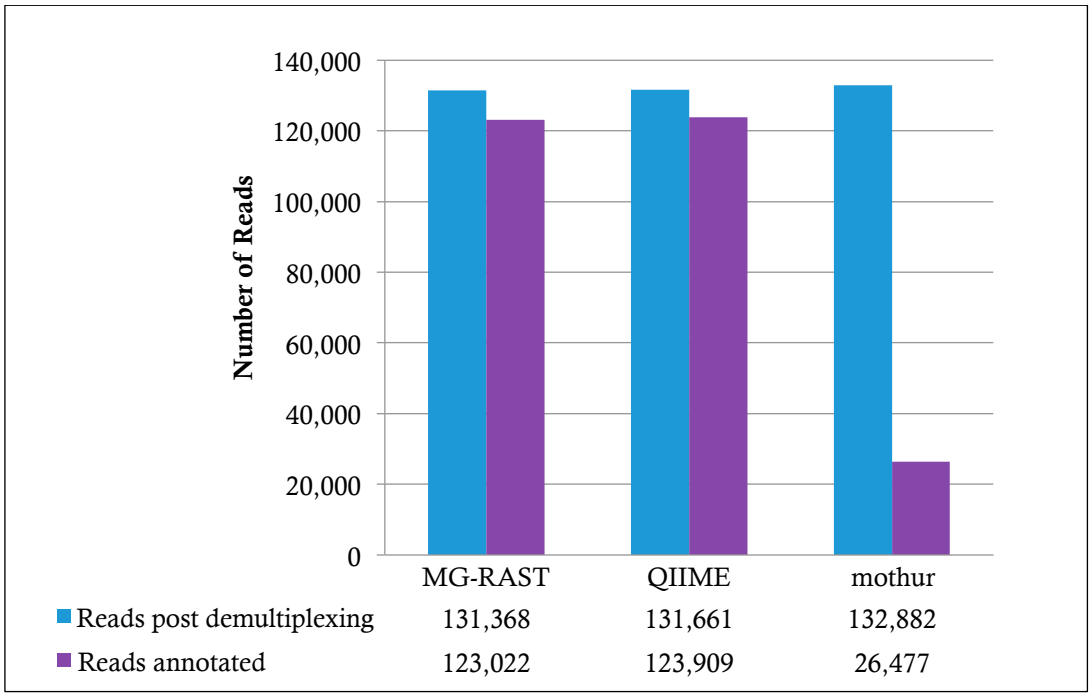


Figure 7 – Number of Reads Entering Each Pipeline versus Number of Reads Annotated
While the number of reads following de-multiplexing is comparable across pipelines, mothur annotated significantly fewer reads than both MG-RAST and QIIME.

5.2.2 Taxonomic Composition Comparisons

MG-RAST, QIIME and mothur detected the same phyla at similar abundances (Figure 8). The only difference in phyla was that Verrucomicrobia was detected at very low abundances by both QIIME and mothur (0.15% and 0.08%, respectively), but was not detected by MG-RAST. The most notable difference in abundance at the phylum level was that MG-RAST was unable to classify a higher proportion of reads (11.19%) than both QIIME and mothur (<1%). Additionally, Proteobacteria was detected in lower abundance by MG-RAST (30.44%) than by both QIIME and mothur (40.78% and 46.00%, respectively). Mothur detected Actinobacteria at slightly lower abundance (22.99%) than MG-RAST and QIIME (31.96% and 31.73%, respectively). The differences were not considered statistically significant (p=0.277).

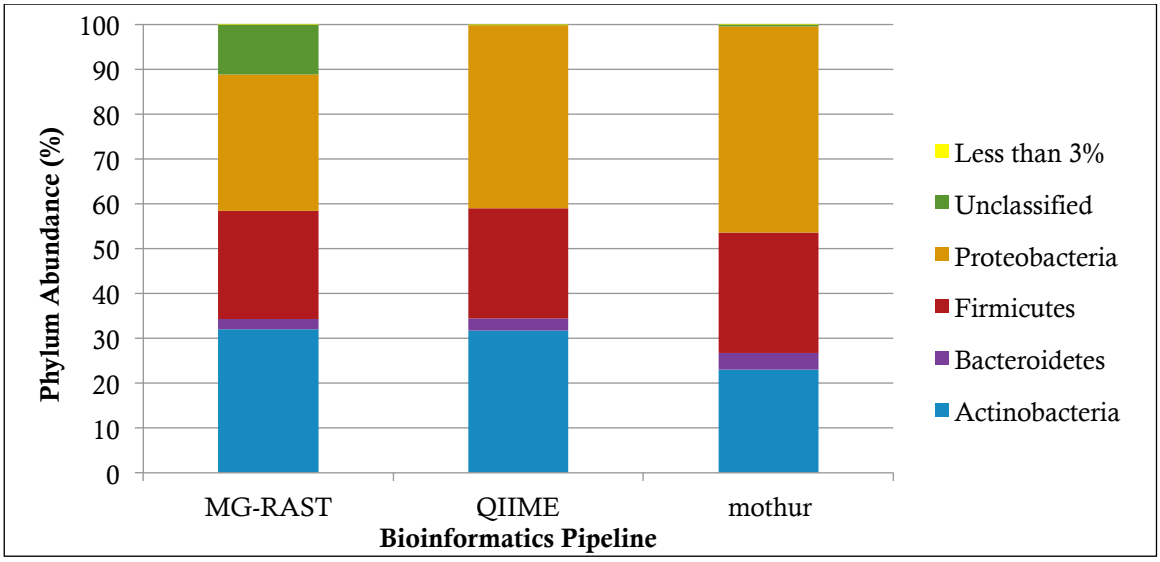


Figure 8 – Comparison of Taxonomic Classification by MG-RAST, QIIME and mothur at a Phylum Level
Compares the abundance of phyla detected by MG-RAST, QIIME and mothur. The four most abundance phyla (Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes) were detected in similar abundances across the three pipelines. The phyla present in the less than 3% category are: Verrucomicrobia and Fusobacteria.

A total of 84 distinct genera were identified across the three pipelines. QIIME identified 70 distinct genera, while mothur identified 47 and MG-RAST identified 57. Figure 9 shows the number of genera that were detected by each pipeline. MG-RAST and mothur were the least similar sharing only 35 genera, while QIIME shared 46 genera with mothur and 44 genera with MG-RAST. mothur identified one genus (*Acidovorax*) that was not detected by either MG-RAST or QIIME.

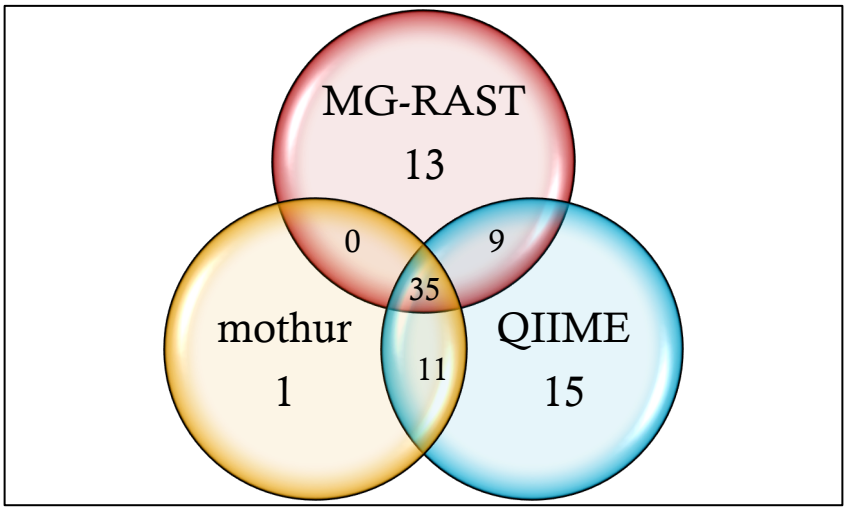


Figure 9 – Venn Diagram Showing the Number of Genera Identified by Each Pipeline
QIIME identified the highest number of unique genera (15), while mothur only identified one genus (*Acidovorax*) that was not detected by either MG-RAST or QIIME.

mothur performed poorly at genus level classification, leaving 44.61% of reads unclassified (11,812 reads). QIIME left 10.27% of reads unclassified (11,206 reads) and MG-RAST left 16.46% reads unclassified (20,253). The majority of reads unclassified by mothur at the genus level were from the Enterobacteriaceae family (80.36%). The percentage of reads unable to be classified at both the phylum and genus level is shown in Figure 10.

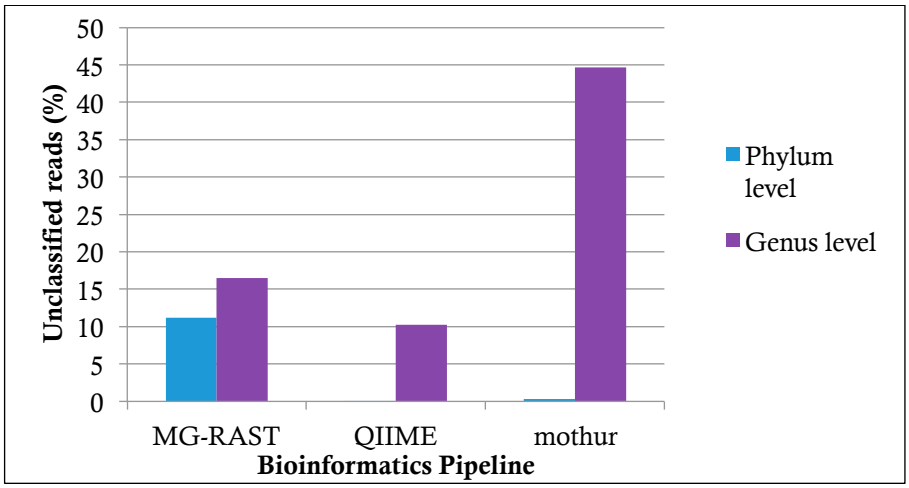


Figure 10 – Comparison of Unclassified Reads Generated by MG-RAST, QIIME and mothur
Compares the number of unclassified reads across the three pipelines. QIIME generated the smallest percentage of unclassified reads at both the phylum and genus levels. A total of 44.61% of reads could not be classified by mothur.

Bifidobacterium was the most abundant genus detected by the three pipelines (MG-RAST 36.96%, QIIME 35.79% and mothur 18.71%). There were some notable differences at genus level classifications. MG-RAST identified a high abundance of *Klebsiella* (12.20%), while QIIME and mothur detected *Klebsiella* at <1% abundance. QIIME identified a high abundance of *Enterobacter* (28.46%), which was identified in very low abundance by MG-RAST (2.59%) and not at all by mothur. A summary of the top five genera identified by each pipeline is presented in Table 5. A complete list of genera identified across the three pipelines is presented in Appendix A.

Table 5 – The Five Most Abundant Genera Detected by MG-RAST, QIIME and mothur

MG-RAST	QIIME	mothur
Bifidobacterium (36.96%)	*Bifidobacterium* (35.79%)	*Bifidobacterium* (18.71%)
Klebsiella (12.20%)	*Enterobacter* (28.46%)	*Escherichia-Shigella* (12.84%)
Escherichia-Shigella (7.20%)	*Enterococcus* (4.09%)	*Veillonella* (4.25%)
Enterococcus (4.17%)	*Clostridium* (3.44%)	*Staphylococcus* (4.09%)
Veillonella (3.57%)	*Staphylococcus* (3.41%)	*Lactobacillus* (3.31%)

It should be noted that the abundance of unclassified reads have been removed from this table. *Escherichia* and *Shigella* are difficult to resolve using 16S rRNA gene analysis and as such have been grouped together as one genus.

5.2.3 Diversity Analysis Comparisons

Figure 11 presents a comparison of three diversity measures (effective number of genera, genus richness and beta diversity) of the three pipelines.

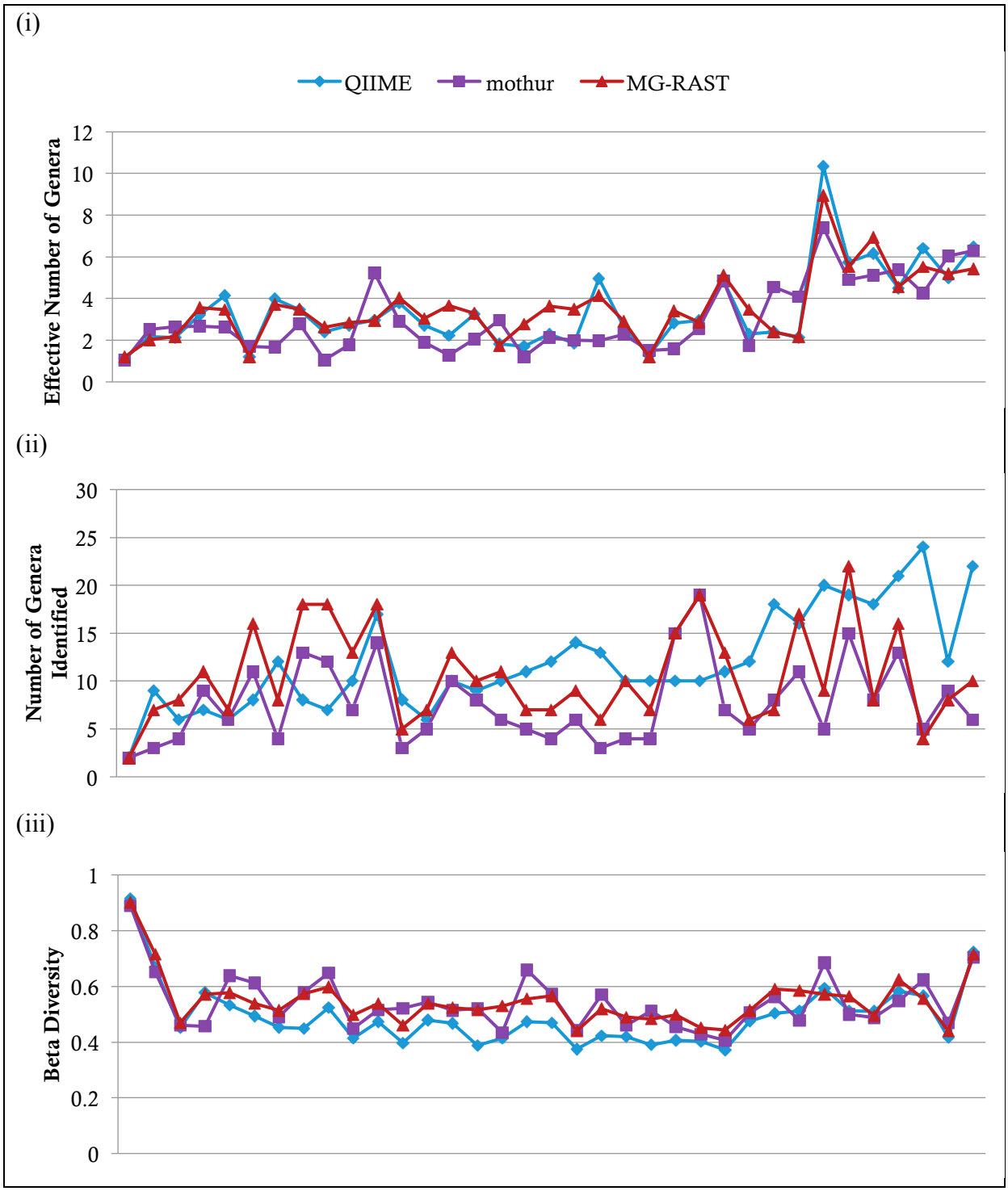


Figure 11 - Comparison of Diversity Measures Between MG-RAST, QIIME and mothur

Samples are on the x-axis of the three graphs and are arranged in order of increasing age. (i) Effective number of genera (as a measure of alpha diversity). (ii) Genus Richness (iii) The mean beta-diversity (Sørensen Dissimilarity Index) of samples. All diversity measures were calculated using genus level data.

A significant difference was observed between QIIME both mothur and MG-RAST in terms of beta diversity. No difference was observed in effective number of genera detected by each pipeline (p=0.108); however, a significant difference was observed between mothur and both QIIME and MG-RAST with respect to genus richness. Pairwise comparison p-values are presented in Table 6.

Table 6 – Statistical Significance (pair wise p-values) of Diversity Measures

Beta Diversity			Genus Richness		
	MG-RAST	**mothur**		**MG-RAST**	**mothur**
mothur	0.603	-	**mothur**	<0.001	-
QIIME	<0.001	<0.001	**QIIME**	0.250	<0.001

5.2.4 Usability Comparisons

Analysis with MG-RAST is straightforward. MG-RAST does not require the user to input any commands, and it has fewer analysis options than QIIME and mothur, making it more suitable for researchers without bioinformatics or command line experience. The website is easy to navigate and the analysis options are clear and well explained. However, the data produced by MG-RAST requires a lot of cleaning due to the multiple annotation of reads. Though it is not difficult to do, cleaning the data is time consuming and would be challenging to complete in a timely manner for large data sets.

The most common type of multiple annotation observed was where reads were annotated as both a specific bacterial species and 'unclassified'. Other multiple annotations commonly observed were where a read is annotated as:

- Two or more related bacterial strains (i.e. *Clostridium difficile* CD196 and *C. difficile* R20291),
- A specific species and one or more related subspecies (i.e. *B. longum* and *B. longum* subsp. *Infantis*),
- The same organism multiple times, and
- Two completely different organisms (i.e. *Escherichia coli* and *Klebsiella pneumoniea*)

Analysis with both QIIME and mothur require command line experience. The support information provided on both the QIIME and mothur websites are very comprehensive and provide clear instructions for data analysis and script writing. mothur has the most extensive quality control process, and it takes more to optimise the scripts in mothur than in QIIME.

Analysis of the 35 samples took approximately one hour of computational time in QIIME, 12 hours of computational time in mothur and approximately two days of manual data cleaning in MG-RAST. Furthermore, while analysis can begin immediately in QIIME and mothur, samples must undergo quality control by the MG-RAST team prior to being available for analysis. This can add up to several weeks to the data analysis time.

The taxonomic summary tables generated by QIIME are the most user friendly and easiest to adapt to downstream analyses in R. However, determining what each read has been annotated as is much easier in MG-RAST than either mothur or QIIME.

5.3 Taxonomic Analysis

Proteobacteria was identified as the most abundant phylum across our study samples. It was detected at an abundance of at least 50% in 87 of the 215 samples, and at any abundance in 200 of the 215 samples. Actinobacteria and Firmicutes were also highly abundant (detected at an abundance of at least 50% in 61 and 49 samples respectively). Three unique samples were identified. Two samples, 10063_B_613 and 10141_A_3, contained a high abundance of Bacteroidetes (75.48% and 66.52% respectively) and one sample, 10345_B_283, contained Verrucomicrobia at an abundance of 74.35%. Figure 12 shows the total abundance of phyla present.

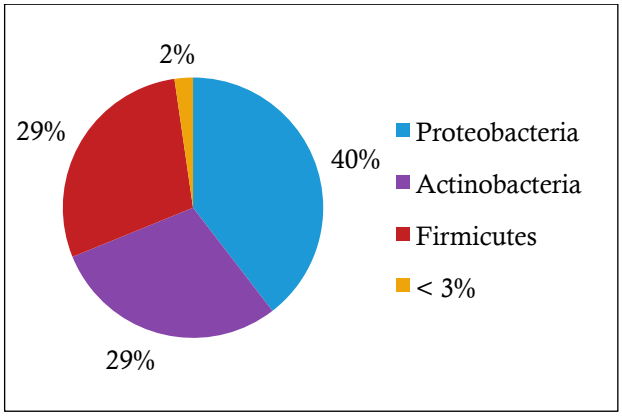


Figure 12 – Abundance of Phyla in *ProPrems* Samples
The most abundant phylum detected is Proteobacteria, followed by both Actinobacteria and Firmicutes. Phyla present in the "<3%" category are Verrucomicrobia, Fusobacteria, Cyanobacteria, and reads that could not be classified.

An overview of sample composition and diversity at the order level is provided in the heat map in Figure 13. Similarity between samples is identified by the dendrogram at the top of the figure. The dendrogram shows that the samples group into multiple clusters. The most distinct samples are those that form the cluster at the far left of the heat map. These samples have a high abundance of Bacillales. The heat map quite clearly shows that bacteria from one or a couple of orders dominate each sample. Furthermore, most orders are present in low abundance, with Enterobacteriales, Bifidobacteriales, Bacillales and Lactobacillales being the most abundant bacterial orders.

The heat map also indicates that there is no clear clustering between the allocation groups. However, there does appear to be more samples from the probiotic group with a high abundance of Bifidobacteriales.

A total of 96,873 reads (13.26%) could not be assigned to a genus by QIIME; this included 400 reads that could not be assigned any identity. Most reads that could not be classified to a genus level came from the Enterobacteriaceae family (58.17%). We submitted 95 representative sequences from OTUs that could not be assigned any identity to the RDP classifier. 94 OTUs were classified as bacterial sequences and 60 were assigned to a phylum level. Of these 20 representative sequences were classified as Enterobacteriaceae.

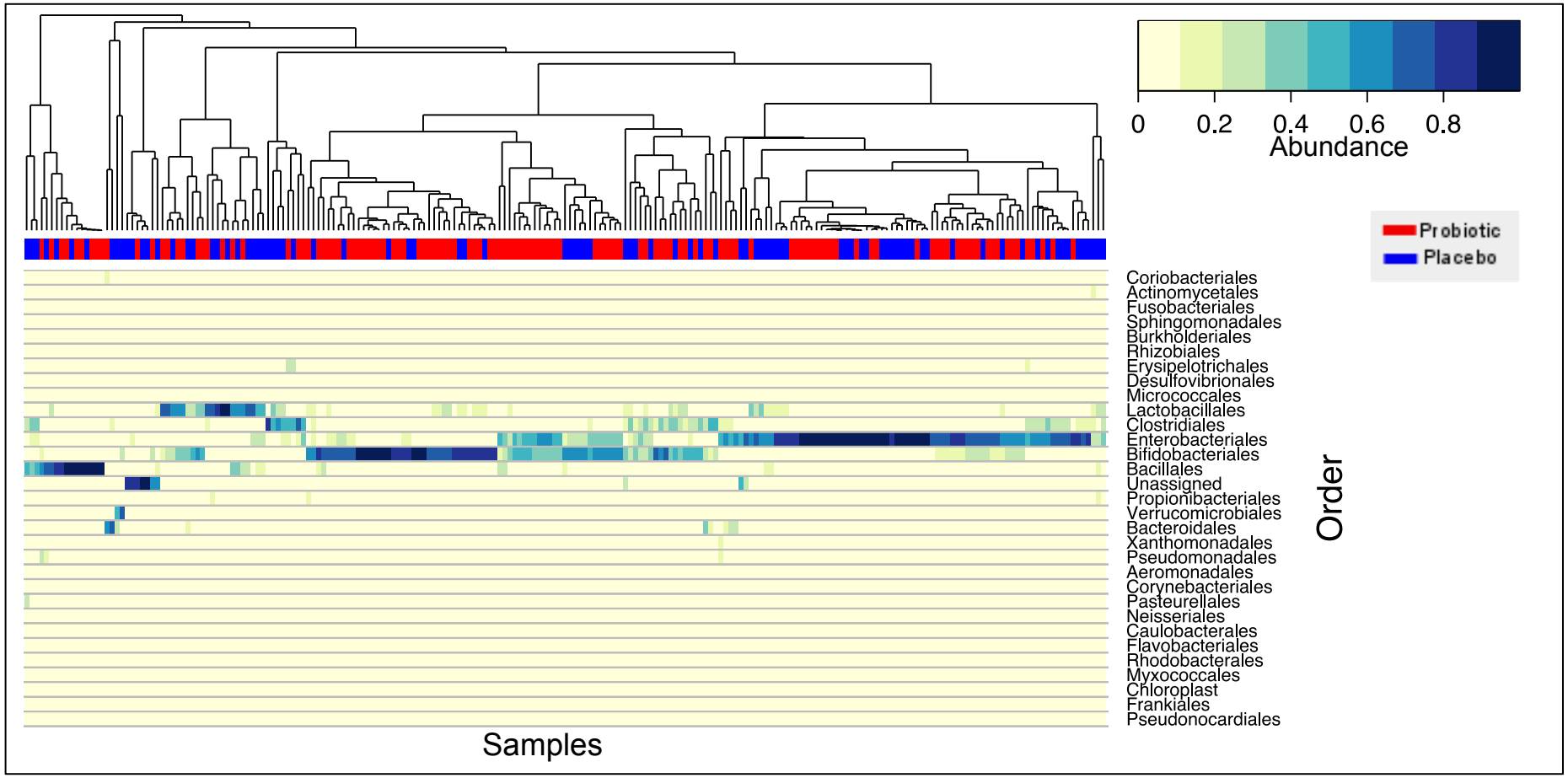


Figure 13 - Bacterial Diversity in *ProPrems* Samples

Columns in the heat map show samples, and rows show taxonomy assignment (order level). Orders in high abundance are coloured dark blue. The dendrogram shows clustering of study samples. There is no clear clustering of samples in the different allocation groups. Clearly, there are very few bacterial orders that are present in high abundance.

At a genus level, *Enterobacter* was present in the most number of samples (172), at an average abundance of 16.42% (SD=24.85). *Bifidobacterium* was present in fewer samples but at higher abundance (161 samples at an average abundance of 28.40% [SD=31.77]). A total of 121 different genera were identified and of these, 86 were detected in 10 or fewer samples. Only six genera were identified in more than half of the study samples. See Appendix B, which provides abundance data for each sample at the genus level.

The ordination analysis suggests that there is little difference between samples. As shown in Figure 14 there is one large cluster and there is a lot of overlap between samples from the probiotic and placebo allocation groups.

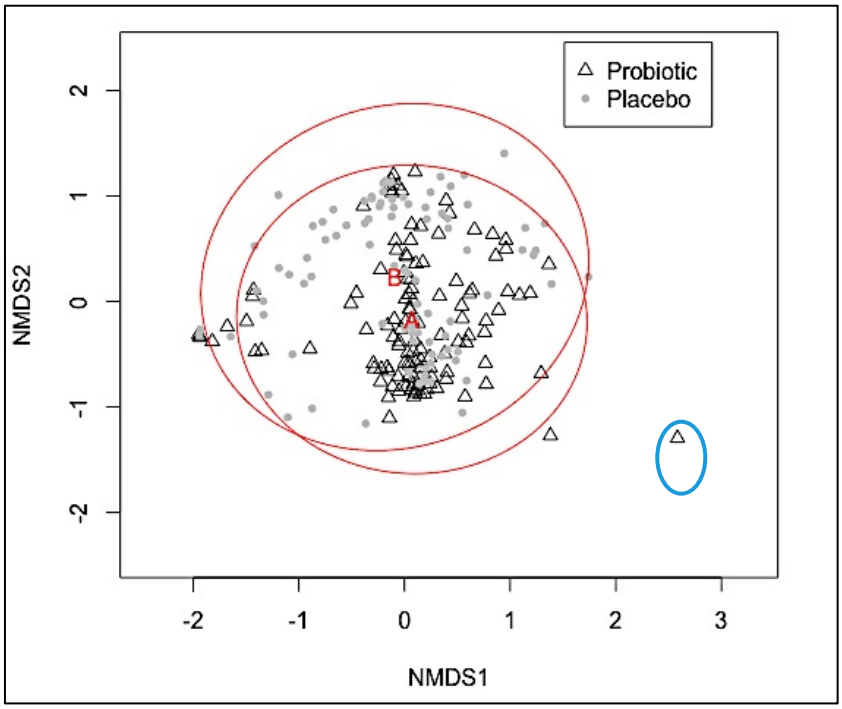


Figure 14 – Graphical Representation of Ordination Analysis Between Allocation Groups
The majority of samples form one large cluster. As can be seen by the red circles, there is a lot of overlap between the samples from group A (probiotic) and B (placebo). A noticeable outlier has been circled in blue. Distance measure used to generate the plot was the Bray-Curtis Distance Measure. (NMDS - Nonmetric multidimensional scaling)

The NMDS was repeated using age as a factor. As shown in Figure 15, the majority of samples form one large cluster and there is a high degree of overlap between samples from the different age groups. However, several samples from the youngest age group cluster outside of the main group.

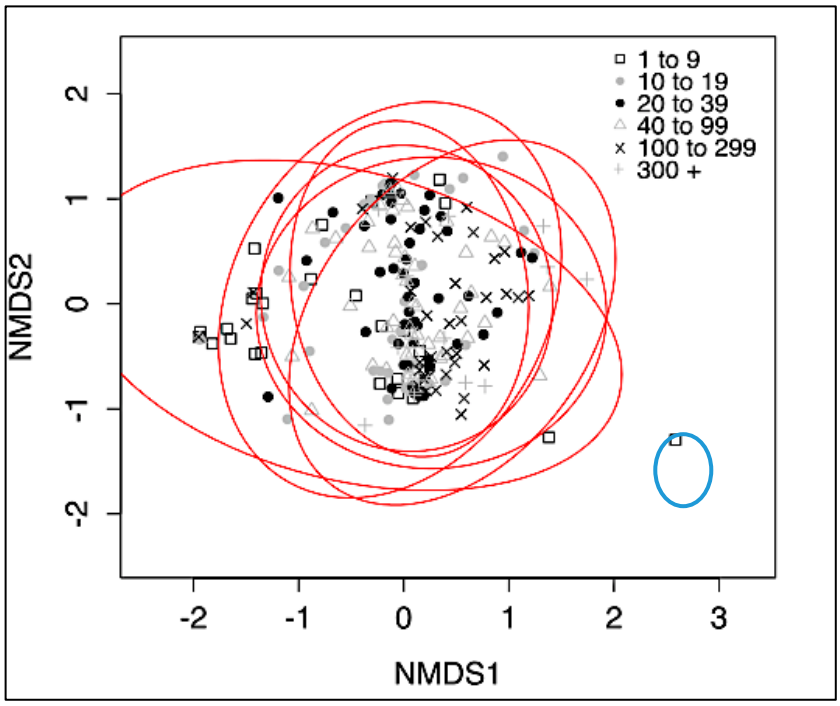


Figure 15 – Graphical Representation of Ordination Analysis Between Age Groups
The red circles represent the individual age groups. As can be seen, the majority of samples form one large cluster in the centre of the figure. Several samples from the 1 to 9 Day age group cluster outside of the main cluster indicating that these are different from the other age groups. A noticeable outlier has been circled in blue. Note, the red circles cluster around samples from the different age groups. Distance measure used to generate the plot was the Bray-Curtis Distance Measure. (NMDS - Nonmetric multidimensional scaling)

One sample failed to cluster with the main group in both PCA plots. This sample (10141_A_3) has been circled in both Figure 14 and 15 and its taxonomic composition is quite distinct from the other samples. 10141_A_3 was identified as having a high abundance of Bacteroidetes. 10141_A_3 contained a high abundance of *Prevotella* (66.52%) and *Atopobium* at an abundance of 17.65%. The next highest abundance of *Prevotella* and *Atopobium* were 31.36% and 3.79% respectively. *Prevotella* and *Atopobium* were more commonly identified in older samples, so 10141_A_3 is very unique given it was collected at day three and is dominated by these two genera. Interestingly, neither *Prevotella* nor *Atopobium* were found in any of the later samples collected from infant 10141.

The first two principal components explain over 50% of the variation in the data and show that the samples separate into three main groups, each characterised by a high abundance of *Bifidobacterium*, *Enterobacter* or *Staphylococcus*. From the PCA plot shown in Figure 16, it appears that there are more samples from the probiotic group clustering in the direction of *Bifidobacterium*. An additional PCA plot that has the samples factored by age range instead

of allocation group has been provided in Appendix C (Figure C1). Very few samples cluster in the direction of *Staphylococcus*, and they appear to come from the youngest age group.

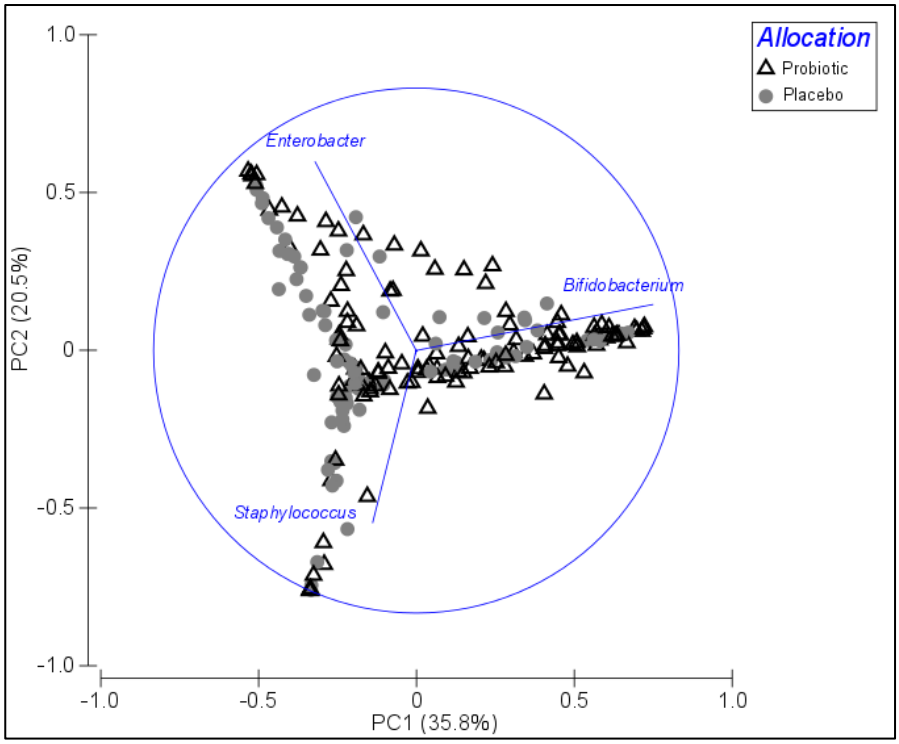


Figure 16 – Principal Component Analysis of Genus Abundance Data
The first and second PCA components show separation of samples into three main groups, one group with a high abundance of *Enterobacter*, another group with a high abundance of *Bifidobacterium* and a final group with a high abundance of *Staphylococcus.*

The results of the SIMPER analysis of dissimilarity revealed that the average dissimilarity between samples from the two allocation groups was 79.13% and that *Bifidobacterium* contributed the most to the dissimilarity between the allocation groups (26.97%), followed by *Enterobacter*, the combined *Escherichia-Shigella* genus and *Staphylococcus*. The contribution of various genera to the total dissimilarity between allocation groups is presented in Table 7. Table 7 also shows that the most abundant genus in the probiotic group was *Bifidobacterium*, while *Enterobacter* was the most abundant genus in the placebo group.

Despite there being several genera contributing to the overall dissimilarity between the probiotic and placebo groups, *Bifidobacterium* was the only genus confirmed as being significantly different between the two groups (adjusted $p<0.001$). The abundance of *Bifidobacterium* was significantly higher in two age groups, *10 to 19 days* and *20 to 39 days* ($p<0.001$).

Table 7 – Most Abundant Genera in Probiotic and Placebo Groups and Their Contribution to Dissimilarity

Genus	Average Abundance (Probiotic)	Average Abundance (Placebo)	Cumulative contribution to dissimilarity (%)
Bifidobacterium	0.42	0.21	26.97
Enterobacter	0.17	0.23	45.05
Escherichia-Shigella	0.10	0.09	55.33
Staphylococcus	0.09	0.09	65.31
Enterococcus	0.04	0.11	73.02
Streptococcus	0.05	0.03	78.04
Veillonella	0.03	0.04	81.78
Clostridium	0.01	0.04	84.24
Lactobacillus	0.01	0.02	86.22
Citrobacter	0.01	0.02	88.03
Bacteroides	0.01	0.01	89.17
Akkermansia	0.00	0.01	90.19

Abundance is indicated as a proportion.

We repeated the above analysis for the age groups. The most abundant genus across all age groups was *Bifidobacterium* (average abundance ranged from 0.27 to 0.31), except in the age-group *1 to 9 days* where *Staphylococcus* (abundance of 0.29) was identified as most abundant genus. Full details of the most abundant genera across age groups are presented in Appendix D, Table D1.

The SIMPER analysis revealed that the most similar groups with respect to bacterial abundance were groups *20 to 39 days* and *40 to 99 days*. The highest levels of dissimilarity were observed between samples from *1 to 9 days* and the other groups, with the greatest dissimilarity between samples from *1 to 9 days* and *300 plus days*. The percentage dissimilarity for all groups is presented in Table 8.

Table 8 – Dissimilarity Between Age Groups

	1 to 9	10 to 19	20 to 39	40 to 99	100 to 299
1 to 9	-	-	-	-	-
10 to 19	80.91	-	-	-	-
20 to 39	83.53	74.17	-	-	-
40 to 99	82.67	75.71	71.2	-	-
100 to 299	84.16	79.26	75.29	75.9	-
300 +	86.35	80.82	77.61	78.27	77.96

Dissimilarity is measured as percentage dissimilarity between groups.

We then removed *Bifidobacterium* from the abundance table and repeated the PCA and SIMPER analyses. Again, the first two components of the PCA explain over 50% of the variation between the samples. Figure 17 shows that samples cluster into three groups, each with a high abundance of *Enterobacter, Escherichia-Shigella* or *Staphylococcus*. It is these three genera that describe most of the variation between the samples. While there was no statistically significant difference in the abundance of these genera between the allocation groups, there were some differences identified between the age groups. In Figure 17 it can be seen that the majority of samples branching in the direction of *Staphylococcus* were from the youngest age group (this is also clear from Figure A of Appendix 3). The removal of *Bifidobacterium* from the abundance table increased the dissimilarity between age groups, as shown in Table 9. An additional PCA plot showing the effect of removing *Bifidobacterium* on clustering of the allocation groups has been provided in Appendix 3.

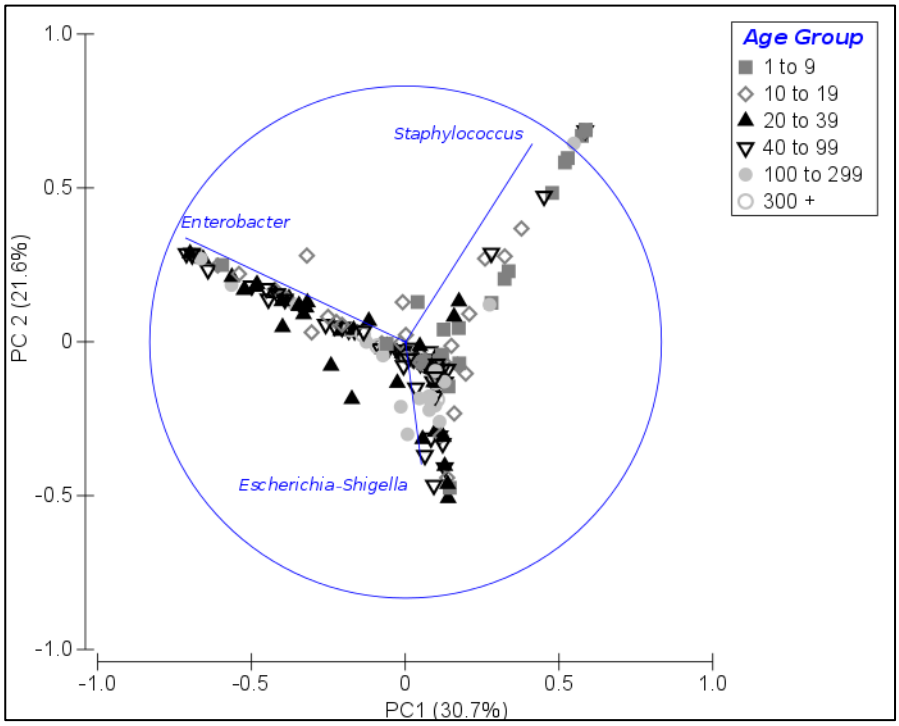


Figure 17 – Principal Component Analysis of Genus Abundance Data post removal of *Bifidobacterium*
The first and second PCA components show separation of samples into three main groups, one group with a high abundance of *Enterobacter*, another group with a high abundance of *Staphylococcus* and a final group with a high abundance of the combined genus *Escherichia-Shigella*. As can be seen, the younger samples cluster primarily in the *Staphylococcus* group.

Table 9 – Dissimilarity Between Age Groups Following Removal of *Bifidobacterium*

	1 to 9	10 to 19	20 to 39	40 to 99	100 to 299
1 to 9	-	-	-	-	-
10 to 19	86.42	-	-	-	-
20 to 39	90.28	82.17	-	-	-
40 to 99	88.76	83.25	78.77	-	-
100 to 299	91.06	88.3	84.35	84.16	-
300 +	94.38	91.15	88.41	88.34	87.48

Dissimilarity is measured as percentage dissimilarity between groups.

ANOVA confirmed that there were six genera that were present in significantly different abundances in one or more of the age groups. Post hoc testing confirmed the following differences ($p<0.001$):

- *Staphylococcus* abundance was significantly higher in the youngest age group as compared to the other age groups
- Four genera (*Anaerostipes*, *Barnesiella*, *Faecalibacterium* and *Parasutterella*) were found in significantly higher abundance in the oldest age group as compared to the other age groups
- *Bacteroides* was found in significantly higher abundance in the oldest group as compared to all other groups, except the 100 to 299 day age group

The p-values and confidence intervals for these findings are presented in Appendix E.

The SIMPER analysis was repeated separately on both allocation groups to investigate any differences between the allocation groups at particular ages. *Bifidobacterium* was the most abundant genus in all age groups that received the probiotic, except for the youngest babies where *Staphylococcus* was dominant. In the placebo group, *Staphylococcus* was the most abundant genus in the youngest age group (*1 to 9 days*), *Enterobacter* was most abundant from days 10 to 39, and *Bifidobacterium* was most abundant in babies aged 40 days and older. Interestingly, the abundance of *Bifidobacterium* is higher in the placebo group at age 100 to 299 Days (0.36) compared to the same age group who received the probiotic (0.22). Full details of the most abundant genera in each allocation group across the different age groups are in Appendix D, Tables D2 and D3.

Known pathogens from the *Clostridium* and *Klebsiella* genera were not consistently present in high abundance. *Clostridium* was detected at greater than 10% abundance in nine samples.

Five of these were late samples (> 230 days) and in the four early samples (< 59 days), presence of *Clostridium* did not persist. *Klebsiella* was detected at greater than 10% abundance in only two samples.

To demonstrate the variability between and within individuals, three babies were selected from each allocation group and their colonisation pattern is presented in Appendix F.

5.4 Diversity Analysis

There were no statistically significant differences between the allocation groups in terms of effective number of genera (p= 0.231, 95%*CI:*[-0.6507,0.1580]) or genus richness (p= 0.531, 95%*CI:*[-0.8805,1.7037]). Box plots showing the various diversity measures for each allocation group are provided in Figure 18.

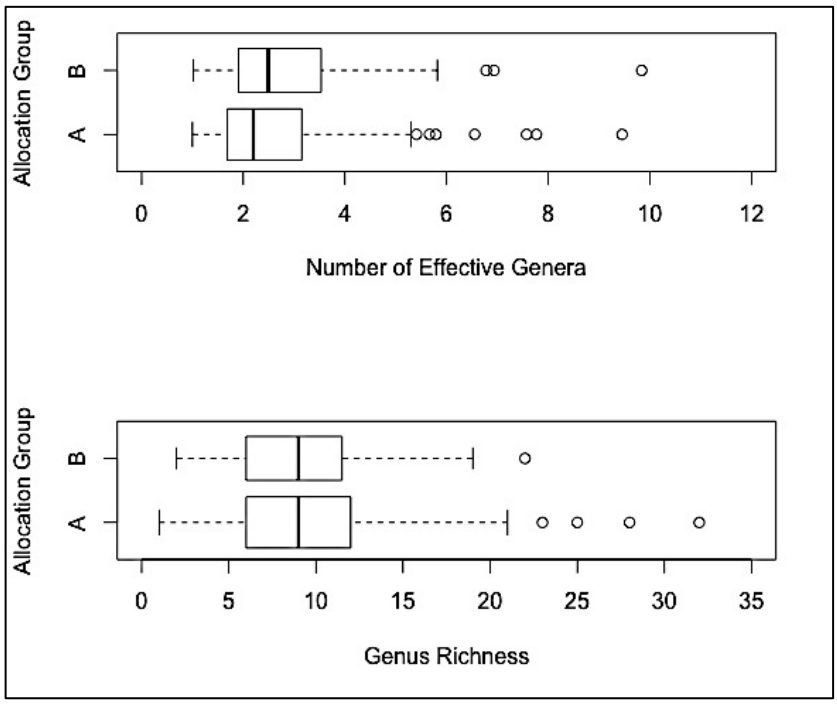


Figure 18 – Effect of Allocation Group on Diversity Measures

The allocation group (A=Probiotic, B=Placebo) has little effect on the effective number of genera (i.e. alpha diversity) or genus richness. The solid black line indicates the median value, the borders of the box give the interquartile range and outliers are displayed as circles

ANOVA revealed significant differences between the various age groups with respect to the effective number of genera (p<0.001). Tukey's HSD post hoc test was used to find which age range/s differ at a 95% family-wise confidence level. The results are summarised in Table 10. Significant differences were observed between the older age groups (*100 to 299 days*, and *300*

plus days) and the younger age groups, except that there was no significant difference in effective number of genera between aged *40 to 99 days* and *300 plus days*.

Table 10 – Results of post Hoc Testing of Differences in Effective Number of Genera Between Age Groups

Pair-wise comparisons	Adjusted p-value	95% CI
100 to 299 - 1 to 9	< 0.001	[0.9561,2.9089]
300+ - 1 to 9	< 0.001	[0.6267,3.1477]
100 to 299 - 10 to 19	< 0.001	[0.7817,2.6100]
300+ - 10 to 19	0.002	[0.4376,0.8634]
20 to 39 - 100 to 299	< 0.001	[-2.2268,0.5223]
40 to 99 - 100 to 299	0.005	[-1.8966,0.2191]
300+ - 20 to 39	0.015	[0.1622,2.4962]

NB: '100 to 299 - 1 to 9' means the pairwise comparison was made between Age groups s '*100 to 299 days'* and *'1 to 9 days'*

A similar pattern was observed with respect to genus richness, which is not unexpected given the relationship between effective number of genera and genus richness. Both alpha diversity and genus richness increased with age as shown in Figure 19 (although the mean effective number of genera was marginally lower in the *300 plus days* age group [3.80] compared to the *100 to 299 day* group [3.84]).

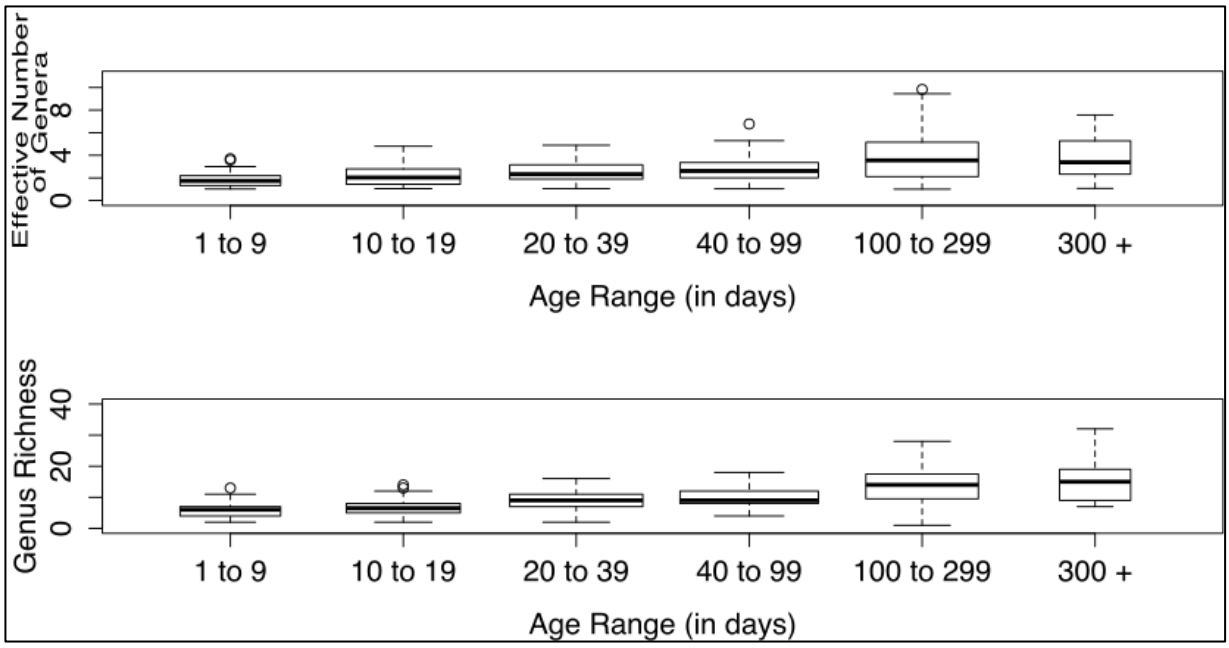


Figure 19 – Effect of Age on Diversity

Diversity appears to increase linearly with age. The solid black line indicates the median value, the borders of the box give the interquartile range and outliers are displayed as circles. Note: the width of the boxplot increases as the number of samples increase.

The average beta diversity for the various age groups and the allocation groups is shown in Figure 20. As can be seen from the plots, samples in the *40 to 99 day* group are most similar (i.e. have a lower beta diversity) and the samples from group *1 to 9 days* are the least similar. ANOVA confirmed significant differences between exist between many of the age groups ($p<0.001$). Post hoc testing confirmed that beta diversity was lower in the *40 to 99dDay* group compared to every other group bar the *20 to 39 day* group ($p<0.001$). The observed trend is that between sample diversity is at its highest when the baby is young, decreases at the one to three month age and then begins to increase again. Other statistically significant differences were found between age groups during post hoc testing and these results are provided in Appendix E.

There was no significant difference in beta diversity between the allocation groups. Additionally, after correction for multiple testing, no statistically significant differences were found in effective number of genera or species richness between the allocation groups at the different age ranges, see Appendix G for figures. This indicates that probiotic supplementation had little effect on bacterial diversity.

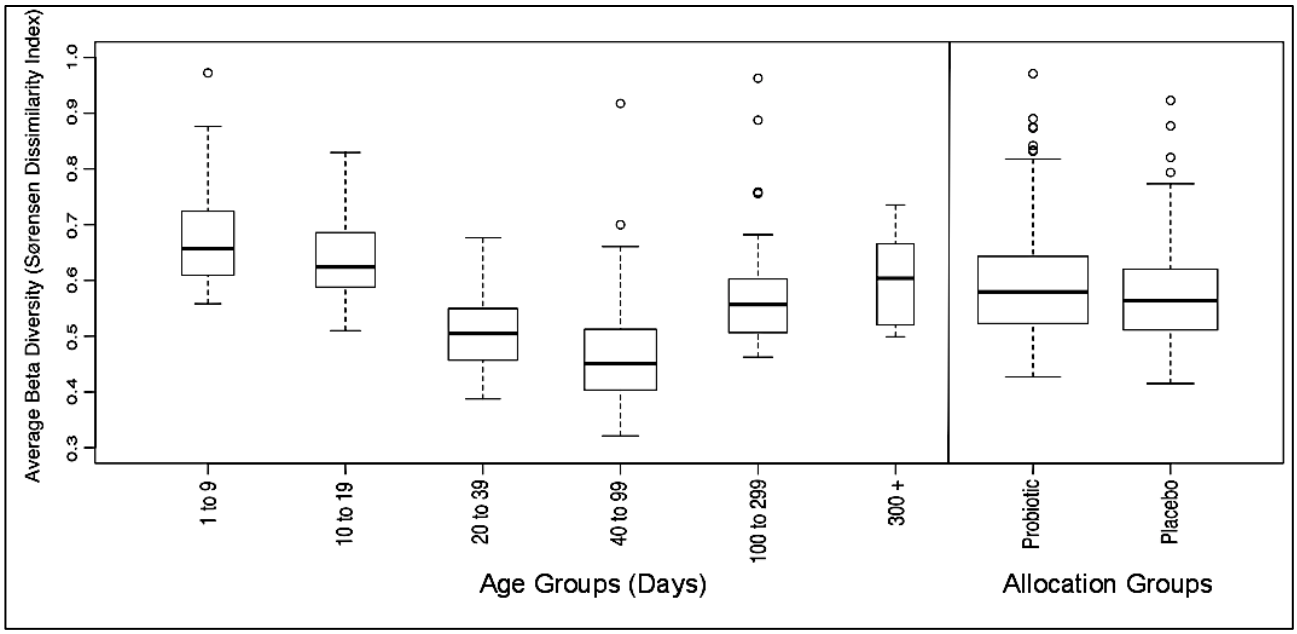


Figure 20 – The Effect of Age and Allocation Group on Beta Diversity

Beta diversity (as measured by the Sorensen Dissimilarity index 0 = samples are identical, 1 = samples share no genera in common) decreases from birth for a few months, after which it starts to increase again. Beta diversity is the highest in the youngest samples and is at its lowest in the *40 to 99 day* age group. Allocation group does not appear to effect beta diversity. The solid black line indicates the median value, the borders of the box give the interquartile range and outliers are displayed as circles. Note: the width of the boxplot increases as the number of samples increase.

5.5 Bacterial Load

We quantified the bacterial load for 178 of the 215 samples that were included in the final taxonomic analysis. Eleven samples from a wide range of different ages (mean 91.45 days, SD 125.38) failed, as such data from 167 samples were included in the bacterial load analysis. The failed samples were likely technical failures, rather than failing due to low total bacterial content, as each sample produced a usable number of reads from 454 sequencing.

Log transformed bacterial load over time is represented in Figure 21 and mean values (and standard deviation) for each age group are summarised in Table 11. Figure 21 and Table 11 show that there is little difference between the age groups with respect to bacterial load. A statistically significant difference in bacterial load was observed between age groups *1 to 9 days* and *20 to 39 days* (p=0.027, (95%CI:[-0.071,1.954]), but not between any other age groups. Bacterial load was most variable in the *100 to 299 days* group.

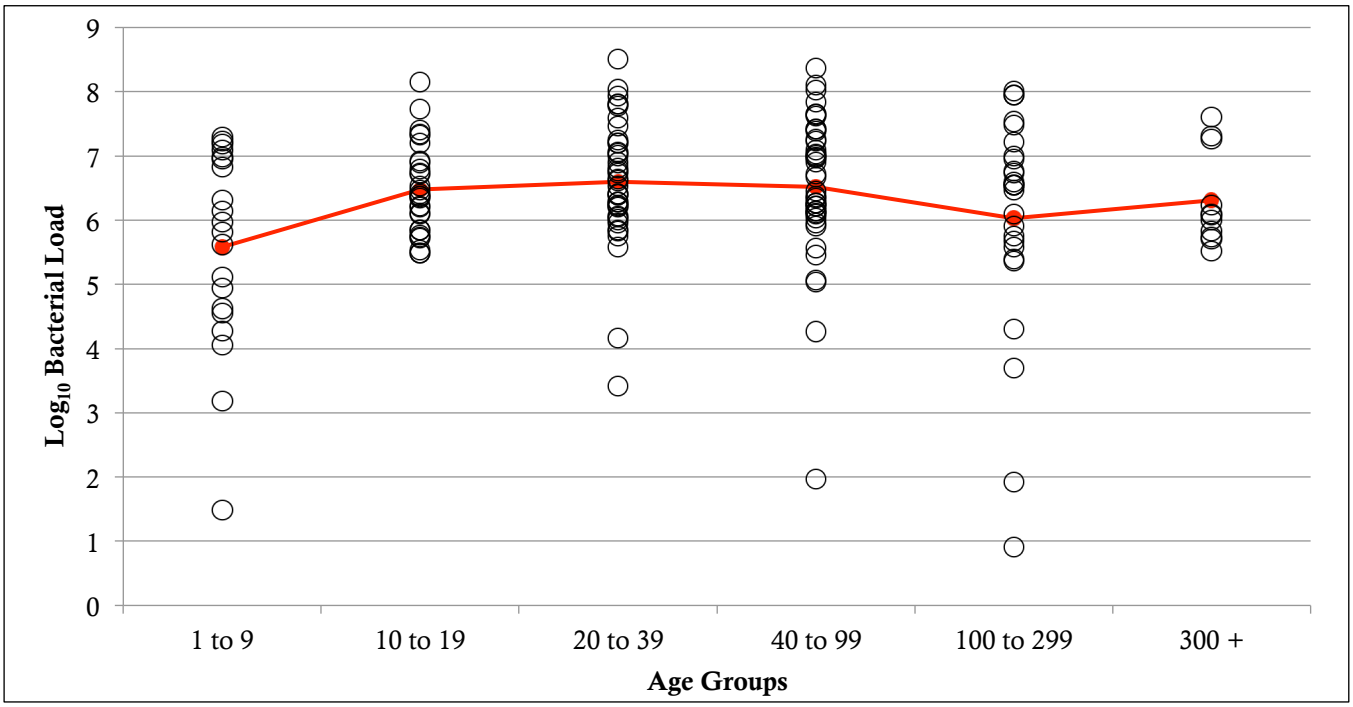


Figure 21 – The Relationship Between Age and Bacterial Load

For each sample, bacterial content was estimated by TaqMan real-time PCR with universal bacterial primers. Bacterial load is calculated as the number 16S rRNA gene copies per PCR reaction. Bacterial load remains relatively constant over the six age groups. The red line shows the mean bacterial load over for each age group. Age groups *1 to 9 days* and *100 to 299 days* show the highest variability in bacterial load.

Table 11 – Average Bacterial Load Across Different Age Groups

Age Range (days)	1 to 9	10 to 19	20 to 39	40 to 99	100 to 299	300 +
N	20	29	40	41	26	11
Log_{10} bacterial load, mean (SD)	5.58 (1.56)	6.47 (0.70)	6.59 (0.97)	6.52 (1.14)	6.03 (1.72)	6.30 (0.73)

There was no statistically significant difference in bacterial load between the allocation groups overall (p=0.374, 95%CI:[-0.545,0.206]), or between the age at different age groups (p>0.05) suggesting that the probiotic does not effect bacterial load. Box plots presenting bacterial load at the various age ranges by allocation group is provided in Appendix H.

We found a weak positive correlation between bacterial load and effective number of genera (correlation coefficient=0.170, p=0.028, 95%CI:[0.019,0.314]), see Figure 22.

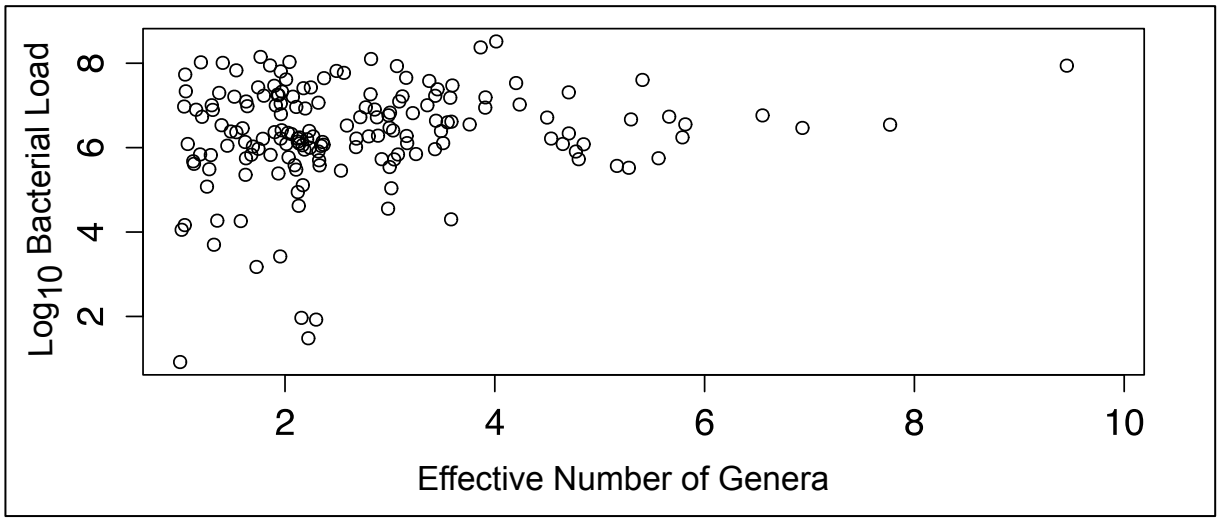


Figure 22 – The Relationship Effective Number of Genera and Bacterial Load
There was a significant but weak positive correlation between effective number of genera and total bacterial load. Note: total bacterial load was measured as the number of 16S rRNA gene copies per PCR reaction.

5.6 Investigation of the Probiotic

We collated all samples collected after administration of the probiotic or placebo had ceased and investigated differences in terms of genera abundance between the allocation groups. There was no statistically significant difference in the abundance of any genera after prophylaxis had stopped. The abundance of *Bifidobacterium* in both groups post ceasing prophylaxis was very variable. The average abundance of *Bifidobacterium* in the placebo

group was 26.90% (ranged from 0 to 93.61%), while in the probiotic group it was 28.69% (ranged from 0 to 98.90%).

The MSA and phylogenetic tree of representative sequences from *Bifidobacterium* OTUs did not reveal any obvious differences between sequences from the two allocation groups (see Appendix I). Additionally, while some of the sequences outputted from QIIME clustered with *Bifidobacterium* type strains, the majority of sequences from both allocation groups did not group with any type strains.

The other organism present in the probiotic was *Streptococcus thermophilus*. *Streptococcus* was detected in 141 of the 214 samples, however it was detected at a relatively low abundance (average abundance of 3.96% (SD), range 0.02-92.29%). The MSA (Appendix J) shows that the *Bifidobacterium* optimised primers should be successful in amplifying *Streptococcus*. It is possible that some *Streptococcus* sequences were miss-classified as unassigned because:

- A BLAST (81) search of 15 representative sequences from OTUs assigned as *Streptococcus* returned significant alignments with both uncultured and *Streptococcus* bacteria
- Of the 94 unassigned OTUs classified as bacteria, three were classified as belonging to the Bacilli class. However, we do not know what genus/genera these OTUs belong to and three out of 94 OTUs represents only a small proportion of total unclassified reads.

6. Discussion

This study explored the GI microbiota of VLBW preterm babies who were recruited to *The ProPrems Trial*. We investigated whether there is any generalisable order in which bacterial colonisation of the gut takes place in preterm babies and what effect probiotics have on the colonisation process. We found that while *Bifidobacterium* was detected in higher abundance in babies receiving the probiotic, age has a much greater impact on the bacterial composition and diversity of the gut than prophylaxis with a probiotic.

The literature states that FAs are the first to colonise the gut of a newborn followed by strict anaerobes (12) and this is consistent with what we found. We found that the GI microbiota of a newborn preterm baby is initially low in diversity with the majority of babies having a

predominance of *Staphylococcus* spp. during the first week of life. *Staphylococcus* has previously been found in preterm babies and is mentioned in current literature as an early coloniser of the gut, particularly in the absence of 'normal' early gut colonisers such as members from the Enterobacteriaceae family and *Enterococcus* (14, 82, 83). While the presence of *Staphylococcus* in this age group is not unexpected, it is possible that the high abundance of *Staphylococcus* spp. we observed in our cohort was actually a result of sampling technique (where the faecal swab collected more skin than faeces) as *Staphylococcus epidermidis* is one of the most abundant bacteria found on the skin (84).

Bifidobacterium, *Enterococcus* and genera from the Enterobacteriaceae family were also found in high abundance in some individuals during the first ten days of life. We observed a high degree of variability between the GI microbiota of individuals during this time. In fact, beta diversity was at its highest in the youngest samples. This variability is highlighted by the fact that while *Staphylococcus* was the dominant genera in many samples, it was not detected or detected at only very low abundance in other samples. It is further highlighted by unique samples and the presence of otherwise absent or typically low abundance genera appearing as the dominant genus in a sample. For example, one individual (10276_A_7) had a very high abundance of *Streptococcus* (92.30%) during the first week of life, and another baby (10141_A_3) had a higher than usual abundance of both *Prevotella* and *Atopobium* (66.52% and 17.65% respectively).

We found that over the next 10 days, the presence of *Staphylococcus* decreases and members from the Enterobacteriaceae family (particularly *Enterobacter*) begin to increase in abundance, particularly in the placebo group. *Bifidobacterium* asserts itself as the most dominant genus in the probiotic group between days ten and 19, and limited changes occur in the probiotic group over the next two to three months. The abundance of *Bifidobacterium* slowly increases in the placebo group and becomes the dominant genus between days 40 to 99. While alpha diversity increases over the second and third months of life, the beta diversity is at its lowest during this time.

Bifidobacterium remains the most abundant genus in most samples, but new genera (including *Bacteroides* spp. and *Lactobacillus* spp.) begin to appear at medium to high abundances in some samples after about three months. Interestingly, the literature states that *Bifidobacterium* is often reduced in preterm babies (14, 82) but this was not observed in our

cohort. *Bifidobacterium* was highly abundant from day three in selected babies from both allocation groups. The overall abundance of *Bifidobacterium* was higher in the probiotic group, which is not unexpected given that two of the probiotic strains come from the *Bifidobacterium* genus. However, *Bifidobacterium* was also found in high abundance in the placebo group, and over the *100 to 299 day* period, the abundance of *Bifidobacterium* was actually higher in the placebo group. The discrepancy between this study and previous studies with regards to the abundance of *Bifidobacterium* could be a result of the *Bifidobacterium* optimized primers that we used for 16S rRNA gene amplification.

Some papers have reported higher abundances of potentially pathogenic bacteria, such as Enterobacteriaceae (particularly *Klebsiella pneumoniae*) and *Clostridium* in the preterm infant compared to the full term infant (12). However, while we did see high numbers of bacteria from the Enterobacteriaceae family, we did not find a consistently high abundance of either *Clostridium* or *Klebsiella* in our cohort.

Beta diversity begins to increase after three months of life, and is high in the older samples. Alpha diversity and species richness increase from day one and reach their maximum in the later samples. It is not unexpected that diversity is very high in the older babies, because it is at this time that the microbiota is starting to resemble that of an adult and high inter-individual variability in the adult GI microbiota (even between related individuals) has been reported in the literature (9). Even at its lowest, beta diversity was not that low in our cohort. Our findings, combined with that of Turnbaugh et al. (9) suggest that high variability between the GI microbiota of different individuals exists over a lifetime.

After the first nine days of life, bacterial load does not increase. This is consistent with data from Palmer et al. (85) which reported that in a cohort of 13 babies, total bacterial content increases after birth and then reaches a plateau anywhere between two to ten days of life. The pattern we observed in bacterial load over the first week of life supports the theory that colonisation of the gut begins rapidly after birth (15). Dramatic drops in bacterial load were observed at various time points over the first year of life in some babies by Palmer et al. (85), and this is also consistent with our data which shows significant outliers particularly in the *100 to 299 Day* group. Though the correlation between bacterial load and effective number of genera was weakly positive, the pattern observed in Figures 19 and 21 show that in a large number of our samples even though the alpha diversity increases with age, bacterial load

doesn't increase much past the second week of life. This suggests that the density of existing bacteria decreases to compensate for new incoming species, thus maintaining homeostasis of the GI microbiota and a constant bacterial load over time. However, care needs to be taken when interpreting bacterial load data, as the number of 16S rRNA gene copies per PCR reaction is not always equivalent to the total number of bacteria present (as different bacteria may have more than one copy of the 16S rRNA gene).

The above presents a generalised overview of how bacterial colonisation takes place in the preterm gut. There are no two babies that exhibit the exact same colonisation pattern in either allocation group, which highlights the high beta diversity between our samples. Additionally, very few samples have the exact same composition. This highlights the dynamic nature of the GI microbiota; it constantly affected by and has to adapt to internal and external factors such as method of delivery, illness and use of antibiotics. Only six genera (*Enterobacter*, *Bifidobacterium*, *Escherichia-Shigella, Enterococcus, Streptococcus and Staphylococcus*) were identified in the majority of samples and often they were found in very high abundance. The other 115 genera identified were present in a minority of samples and at varying abundances. This diversity and dynamicity makes it difficult to define a 'normal' GI microbiota at any particular time point. It is not immediately obvious why an individual may have an aberrantly high abundance of one of these 115 genera or the absence of one of the six dominant bacteria, however a full examination of the metadata associated with this cohort (which includes factors such as those mentioned above) would give insight into what is driving the diversity that we see. For example, as reported above, infant 10141 had unusually high abundance of both *Prevotella* and *Atopobium*. These genera have been reported as part of the vaginal flora of some women of reproductive age (86). Thus it is possible that baby 10141 was vaginally delivered and analysis of the metadata could clarify this.

As mentioned above, we found little difference in the GI microbiota between babies from the two allocation groups. Diversity measures were similar and even the taxonomic composition was similar between the two groups. The only statistically significant difference was that *Bifidobacterium* was detected in higher abundance in babies aged ten to 39 days receiving the probiotic prophylaxis. Prophylaxis on average started at four days of age and stopped at age 68 days so it is not unexpected that *Bifidobacterium* is in higher abundance in the probiotic babies of this age range as they are actually ingesting the bacteria. This is further supported

by the fact that we did not find any statistically significant difference in the abundance of *Bifidobacterium* (or in fact, any genera) post prophylaxis with the probiotic.

The other organism present in the probiotic was *S. thermophilus.* Curiously we didn't find any difference in *Streptococcus* abundance between the two allocation groups, and while it was detected in a large number of samples it was detected at only low abundance. This could be because *Bifidobacterium* spp. are more commonly found as normal inhabitants of the infant gut than *Streptococcus* spp. As a result, while the *Bifidobacterium* spp. in the probiotic are able to grow and thrive (particularly *B. longum* subsp. *Infantis* which is commonly found in the infant gut (87, 88)), *S. thermophilus* cannot.

There is existing doubt as to whether colonisation of probiotic species always occurs in individuals receiving prophylaxis (29, 30). As such, it is important to understand why *Bifidobacterium* was found in higher abundance in the probiotic group, and whether the probiotic strains *B. longum* subsp. *Infantis* BB-02 and *B. animalis* subsp. *Lactis* BB-12 colonise in the gut or if they are simply being detected as they transiently pass through following ingestion. The fact that there was no difference in abundance of *Bifidobacterium* post prophylaxis with the probiotic indicates supports the latter point. Phylogenetic analysis suggested no discernable differences between the *Bifidobacterium* sequences obtained from the two allocation groups. This may be because there is no difference between the *Bifidobacterium* organisms present in each group. But it is just as likely that we can't tell the difference between the organisms using the 16S rRNA gene. In fact, phylogenetic analysis provided little insight into what *Bifidobacterium* spp. were present in our samples, as many of our representative sequences did not cluster with any of the type strains. This may suggest that our representative sequences were *Bifidobacterium* spp. for which we did not include any type strain to the MSA. It may also be because our representative sequences were much shorter than the region spanning the two primer sites (and thus much shorter than the type strains) and this may have confused the MSA.

The reality is that distinguishing bacterial species and strains is difficult using 16S rRNA gene analysis due to the low resolution of the 16S rRNA gene. Potentially there is no difference in the V3-V5 region of the 16S rRNA gene between different *B. longum* or *B. animalis* strains making it impossible for us to determine whether or not the probiotic strains have colonised using 16S rRNA gene analysis. In fact, even genus identification can be unreliable. For

example, QIIME groups together bacteria from the *Escherichia* and *Shigella* genera because they cannot be distinguished by their 16S rRNA gene sequence. There is no doubt that the analysis of high-throughput community sequencing data is made much easier by the use of pipelines such as QIIME. However, strong quality control and checking of taxonomic assignments generated by such pipelines is necessary. Using additional alignment or search tools to aid in classification of unassigned reads or to correct dubious classifications is important for obtaining additional information and improving the accuracy of taxonomic assignments using 16S rRNA gene analysis.

As with many areas of bioinformatics, there are multiple methods to choose from for each step of 16S rRNA gene analysis, and careful consideration is needed when deciding what method or tool to use. Our comparative analysis revealed that while QIIME, mothur and MG-RAST all work in different ways and utilise different algorithms, phylum level taxonomic classifications do not differ greatly between the three pipelines.

At face value MG-RAST is the easiest pipeline to use, and it is an excellent tool available for researchers with limited informatics skills. However, the issue of multiple annotations is not small. Cleaning the multiple annotations is straightforward but time consuming, and potentially inaccurate conclusions about the bacterial community could be made by not resolving these annotations. Annotating a read multiple times as the one bacterium can lead to inflated abundance counts, while annotating a read as two or more distinct bacteria can lead researchers to conclude the presence of unique or potentially pathogenic bacteria in a 'normal' sample. For example, several of our reads were annotated in MG-RAST as both *Escherichia coli*, a common inhabitant of the human gut, and *Shigella spp.* (*S. dysenteriae* and *S. boydii),* which are common pathogens. Interestingly, very few articles that cite the use MG-RAST comment about the issue of multiple annotations and how it was resolved.

MG-RAST is superior in that it is very easy to find what taxonomic classification each sequence has been assigned. This is important for downstream analysis. For example, you may want to perform phylogenetic analysis on all reads in a sample that were annotated as *Bifidobacterium*. While it is a straightforward task to identify such reads using MG-RAST, it is much more complicated and time consuming with QIIME and mothur.

Analysis time takes the longest in MG-RAST. This is due to both the time required to clean the annotations and the time required for the MG-RAST team to quality control the data. It is not always in the best interest of the research team to make their sequencing data public through MG-RAST and unfortunately, it is not feasible to wait days (or weeks) for data to pass through quality control when this can be performed by mothur or QIIME in a matter of minutes.

The quality control employed by mothur is a lot stricter than that of either QIIME or MG-RAST, and from the 454 analysis SOP (mothur.org/wiki/454_SOP) and articles written by Patrick D. Schloss (47, 63, 89, 90), it is clear that excellent sequence quality is the top priority of this pipeline. However, it's not clear whether having such high quality data adds more to the classification than what is lost through filtering out over 75% of the reads, especially given over 40% of the reads that passed quality control could not be classified to a genus level. Each step in the mothur pipeline has several parameters that can be changed by the user and it should be noted that the classification we achieved with mothur may improve following further tweaking of these parameters.

Differences between the three pipelines are most notable at the genus level, particularly in the classification of members from the Enterobacteriaceae family. We also observed that the majority of reads that could not be classified by both QIIME and mothur belonged to the Enterobacteriaceae family. This may be a result of the variable regions (V3-V5) of the 16S rRNA gene we chose to target. Though there is no single variable region that can distinguish closely related Enterobacteriaceae bacteria, research by Chakravorty et al. (43) suggests that the combined V3 and V6 regions may be best for distinguishing these bacteria. Additionally, the difficulties we observed in being able to distinguish *Bifidobacterium* spp. and certain genera leads us to question whether the traditional cut-off of 97% similarity is appropriate for species (or even genus) level classification.

Unfortunately our comparison does not tell us which pipeline produces the most accurate classifications. This would only be possible using a sample of known composition. In the end, QIIME was recommended as the analysis pipeline of choice due to its quick analysis time, ease of use and the low numbers of unclassified assignments.

7. Conclusion

The results of the phylogenetic analysis and the differences observed in genus level classification between the three bioinformatics pipelines highlight the key limitation of 16S rRNA gene analysis, that it does not provide the necessary resolution for detailed metagenomic analysis. While 16S rRNA gene analysis is an excellent method for obtaining an overview of sample composition, we should look to other methods, such as whole metagenome shotgun sequencing, for detailed species and strain specific information. A metagenomic approach that investigates strain specific markers will be essential for determining 1) if the *Bifidobacterium* and *Streptococcus* strains present in the gut differ between babies from the two allocation groups, and 2) if colonisation of the probiotic strains took place in babies from the probiotic group.

We have detailed metadata collected from participants in *The ProPrems Trial* and there are many questions we can investigate with our dataset. We can look at the impact of external factors such as antibiotic use, type of feeding, and mode of delivery on bacterial colonisation of the gut. We also have detailed allergy and immunological data for our cohort, so we can investigate not only the impact the GI microbiota composition and probiotic supplementation has on development of allergies, but also if probiotics have any effect on the immune response.

This is a unique study with rich data and it presents a rare opportunity to investigate how probiotics work. While this study indicates probiotic prophylaxis does not have a significant impact on the GI microbiota in terms of diversity or taxonomic composition, the story is not yet complete. Future research on this cohort may provide insight into whether probiotics work through immunomodulatory effects. The knowledge gained about probiotics may have far reaching implications, because probiotics are currently being used in a wide range of different health conditions. Importantly, by understanding how probiotics may work, we can advise the community on how best to use probiotics. Additionally, increasing our knowledge of GI colonisation in the preterm infant, and the effects of external factors on GI microbiota colonisation, is vital for improving the health outcomes of this very vulnerable population.

7.1 What didn't we achieve?

Initially we had wanted to investigate differences in the taxonomic profile of infants that developed NEC and those that remained healthy. However, this was not possible as there were no samples from infants who developed NEC included in the final analysis. Infant selection for inclusion in this study was driven by the presence of allergy and immunological metadata.

8. References

1. Garland SM, Tobin JM, Pirotta M, Tabrizi SN, Opie G, Donath S, et al. The ProPrems trial: investigating the effects of probiotics on late onset sepsis in very preterm infants. BMC infectious diseases. 2011;11:210.
2. Jacobs SE, Tobin JM, Opie GF, Donath S, Tabrizi SN, Pirotta M, et al. Probiotic effects on late-onset sepsis in very preterm infants: a randomized controlled trial. Pediatrics. 2013;132(6):1055-62.
3. Ursell LK, Metcalf JL, Parfrey LW, Knight R. Defining the human microbiome. Nutrition reviews. 2012;70 Suppl 1:S38-44.
4. Proctor LM. The Human Microbiome Project in 2011 and beyond. Cell host & microbe. 2011;10(4):287-91.
5. Rauch M, Lynch SV. The potential for probiotic manipulation of the gastrointestinal microbiome. Current opinion in biotechnology. 2012;23(2):192-201.
6. Ojetti V, Gigante G, Ainora ME, Fiore F, Barbaro F, Gasbarrini A. Microflora imbalance and gastrointestinal diseases. Digestive and Liver Disease Supplements. 2009;3:35-9.
7. Coolen MJ, Post E, Davis CC, Forney LJ. Characterization of microbial communities found in the human vagina by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. Applied and environmental microbiology. 2005;71(12):8729-37.
8. Gu F, Li Y, Zhou C, Wong DTW, Ho CM, Qi F, et al. Bacterial 16S rRNA/rDNA Profiling in the Liquid Phase of Human Saliva. The Open Dentistry Journal. 2009;3:80-4.
9. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. Nature. 2009;457(7228):480-4.
10. Candela M, Biagi E, Maccaferri S, Turroni S, Brigidi P. Intestinal microbiota is a plastic factor responding to environmental changes. Trends in microbiology. 2012;20(8):385-91.
11. Hooper LV, Macpherson AJ. Immune adaptations that maintain homeostasis with the intestinal microbiota. Nature reviews Immunology. 2010;10(3):159-69.
12. Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, de La Cochetiere MF. Development of intestinal microbiota in infants and its impact on health. Trends in microbiology. 2013;21(4):167-73.
13. Consortium THMP. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14.
14. Buccigrossi V, Nicastro E, Guarino A. Functions of intestinal microflora in children. Current opinion in gastroenterology. 2013;29(1):31-8.
15. Cilieborg MS, Boye M, Sangild PT. Bacterial colonization and gut development in preterm neonates. Early human development. 2012;88 Suppl 1:S41-9.
16. Indrio F, Neu J. The intestinal microbiome of infants and the use of probiotics. Current opinion in pediatrics. 2011;23(2):145-50.
17. Gerritsen J, Smidt H, Rijkers GT, de Vos WM. Intestinal microbiota in human health and disease: the impact of probiotics. Genes & nutrition. 2011;6(3):209-40.
18. Martin R, Nauta AJ, Ben Amor K, Knippels LM, Knol J, Garssen J. Early life: gut microbiota and immune development in infancy. Beneficial microbes. 2010;1(4):367-82.
19. Moore TA, Hanson CK, Anderson-Berry A. Colonization of the Gastrointestinal Tract in Neonates: A Review. ICAN: Infant, Child, & Adolescent Nutrition. 2011;3(5):291-5.
20. Mshvildadze M, Neu J. The infant intestinal microbiome: friend or foe? Early human development. 2010;86 Suppl 1:67-71.
21. Chang HH, Larson J, Blencowe H, Spong CY, Howson CP, Cairns-Smith S, et al. Preventing preterm births: analysis of trends and potential reductions with interventions in 39 countries with very high human development index. The Lancet. 2013;381(9862):223-34.
22. Mugambi MN, Musekiwa A, Lombard M, Young T, Blaauw R. Probiotics, prebiotics infant formula use in preterm or low birth weight infants: a systematic review. Nutrition Journal. 2012;11(58):1-18.
23. Morrow AL, Lagomarcino AJ, Schibler KR, Taft DH, Yu Z, Wang B, et al. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. Microbiome. 2012;1(13).
24. Torrazza RM, Neu J. The altered gut microbiome and necrotizing enterocolitis. Clinics in perinatology. 2013;40(1):93-108.
25. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.
26. Fujimura KE, Slusher NA, Cabana MD, Lynch SV. Role of the gut microbiota in defining human health. Expert review of anti-infective therapy. 2010;8(4):435-54.
27. Penders J, Thijs C, van den Brandt PA, Kummeling I, Snijders B, Stelma F, et al. Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study. Gut. 2007;56(5):661-7.
28. Neu J, Walker WA. Necrotizing enterocolitis. The New England journal of medicine. 2011;364(3):255-64.
29. Sherry A. Luedtke JTY, and Heather E. Wild. Probiotics and Necrotizing Enterocolitis: Finding the Missing Pieces
of the Probiotic Puzzle. The Journal of Pediatric Pharmacology and Therapeutics. 2012;17(4):308–28.

30. Ganguli K, Walker WA. Treatment of necrotizing enterocolitis with probiotics. Gastroenterology clinics of North America. 2012;41(4):733-46.
31. Wendelboe AM, Smelser C, Lucero CA, McDonald LC. Cluster of necrotizing enterocolitis in a neonatal intensive care unit: New Mexico, 2007. American journal of infection control. 2010;38(2):144-8.
32. Carlisle EM, Morowitz MJ. The intestinal microbiome and necrotizing enterocolitis. Current opinion in pediatrics. 2013;25(3):382-7.
33. Gephart SM, McGrath JM, Effken JA, Halpern MD. Necrotizing enterocolitis risk: state of the science. Advances in neonatal care : official journal of the National Association of Neonatal Nurses. 2012;12(2):77-87; quiz 8-9.
34. Torrazza RM, Neu J. The developing intestinal microbiome and its relationship to health and disease in the neonate. Journal of perinatology : official journal of the California Perinatal Association. 2011;31 Suppl 1:S29-34.
35. Bernardo WM, Aires FT, Carneiro RM, Sa FP, Rullo VE, Burns DA. Effectiveness of probiotics in the prophylaxis of necrotizing enterocolitis in preterm neonates: a systematic review and meta-analysis. Jornal de pediatria. 2013;89(1):18-24.
36. Williams NT. Probiotics. American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists. 2010;67(6):449-58.
37. Chen C-C, Walker WA. Clinical applications of probiotics in gastrointestinal disorders in children. The National Medical Journal of India. 2011;24(3):153-60.
38. Wilke A, Wilkening J, Glass EM, Desai NL, Meyer F. An experience report: porting the MG-RAST rapid metagenomics analysis pipeline to the cloud. Concurrency and Computation: Practice and Experience. 2011;23(17):2250-7.
39. Liu B, Gibbons T, Ghodsi M, Treangen T, Pop M. Accurate and fast estimation of taxonomic profiles from metagenomic shotgun sequences. BMC genomics. 2011;12(Suppl 2):S4.
40. Han XY. Bacterial Identification Based on 16S Ribosomal RNA Gene Sequence Analysis. 2006. In: Advanced Techniques in Diagnostic Microbiology [Internet]. New York, USA: Springer; [323].
41. Wang YaQ, PY. Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. PloS one. 2009;4(10):e7401.
42. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, et al. Experimental and analytical tools for studying the human microbiome. Nature reviews Genetics. 2012;13(1):47-58.
43. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. Journal of microbiological methods. 2007;69(2):330-9.
44. Sim K, Cox MJ, Wopereis H, Martin R, Knol J, Li MS, et al. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. PloS one. 2012;7(3):e32543.
45. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic acids research. 2013;41(Database issue):D590-6.
46. Petrosino JF, Highlander S, Luna RA, Gibbs RA, Versalovic J. Metagenomic pyrosequencing and microbial identification. Clinical chemistry. 2009;55(5):856-66.
47. Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PloS one. 2011;6(12):e27310.
48. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. Journal of clinical microbiology. 2007;45(9):2761-4.
49. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome research. 2011;21(3):494-504.
50. Vetrovsky T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PloS one. 2013;8(2):e57923.
51. Kuczynski J, Liu Z, Lozupone C, McDonald D, Fierer N, Knight R. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nat Methods. 2010;7(10):813-9.
52. Scholz MB, Lo CC, Chain PS. Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Current opinion in biotechnology. 2012;23(1):9-15.
53. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, et al. Metagenomic analysis of the human distal gut microbiome. Science. 2006;312(5778):1355-9.
54. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. Nature biotechnology. 2012;30(5):434-9.
55. Shah N, Tang H, Doak TG, Ye Y. Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. Pacific Symposium on Biocomputing. 2011:165-76.
56. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6.
57. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology. 2009;75(23):7537-41.

58. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC bioinformatics. 2008;9:386.
59. Team Q. QIIME Quantitative Insights Into Microbial Ecology 2013 [21 August 2013]. Available from: http://qiime.org/.
60. Wilke A, Glass EM, Bischof J, Braithwaite D, D'Souza M, Gerlach W, et al. MG-RAST Manual for version 3.3.6, revision 9. 2014 5 May 2014. Report No.
61. Glass EM, Meyer F. Analysis of Metagenomics Data. 2012. In: Bioinformatics for High Throughput Sequencing [Internet]. Springer New York; [219-29].
62. Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microbial Informatics and Experimentation. 2012;2(3).
63. Schloss PD. Secondary structure improves OTU assignments of 16S rRNA gene sequences. The ISME journal. 2013;7(3):457-60.
64. Nikkari S, Lopez FA, Lepp PW, Cieslak PR, Ladd-Wilson S, Passaro D, et al. Broad-Range Bacterial Detection and the Analysis of Unexplained Death and Critical Illness. Emerging Infectious Diseases. 2002;8(2):188-94.
65. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. R package version 2.0-10. 2013.
66. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011;27(16):2194-200.
67. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1.
68. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology. 2007;73(16):5261-7.
69. Wilke A, Glass EM, Bischof J, Braithwaite D, D'Souza M, Gerlach W, et al. MG-RAST Technical Report and Manual for version 3.3.6, revision 5. 2014.
70. Kent WJ. BLAT---The BLAST-Like Alignment Tool. Genome research. 2002;12(4):656-64.
71. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. 2011;7:539.
72. Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution. 2010;27(2):221-4.
73. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic acids research. 2014;42(Database issue):D633-42.
74. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution. 2013;30:2725-9.
75. Jukes TH, Cantor CR. Evolution of protein molecules. HN IM, editor. Academic Press, New York.1969. 21-132 p.
76. R RCT. A language and environment for statistical computing. . Vienna, Austria.: R Foundation for Statistical Computing; 2014.
77. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, et al. gplots: Various R programming tools for plotting data. R package version 2.13.0 ed2014.
78. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, et al. Picante: R tools for integrating phylogenies and ecology. Bioinformatics. 2010;26:1463-4.
79. Clarke K, Gorley R. PRIMER v6: User Manual/Tutorial. PRIMER-E: Plymouth; 2006.
80. Clarke K. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology. 1993;18:117-43.
81. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10.
82. Hallab JC, Leach ST, Zhang L, Mitchell HM, Oei J, Lui K, et al. Molecular characterization of bacterial colonization in the preterm and term infant's intestine. Indian journal of pediatrics. 2013;80(1):1-5.
83. Adlerberth I, Lindberg E, Aberg N, Hesselmar B, Saalman R, Strannegard IL, et al. Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? Pediatric research. 2006;59(1):96-101.
84. Sanford JA, Gallo RL. Functions of the skin microbiota in health and disease. Seminars in immunology. 2013;25(5):370-7.
85. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the Human Infant Intestinal Microbiota. PLOS Biology. 2007;5(7):1556-73.
86. Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, Davis RW. Microbes on the human vaginal epithelium. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(22):7952-7.
87. Favier CF, de Vos WM, Akkermans AD. Development of bacterial and bifidobacterial communities in feces of newborn babies. Anaerobe. 2003;9(5):219-29.

88. Rautava S. Potential uses of probiotics in the neonate. Seminars in fetal & neonatal medicine. 2007;12(1):45-53.
89. Schloss PD. A high-throughput DNA sequence aligner for microbial ecology studies. PloS one. 2009;4(12):e8230.
90. Schloss PD, Westcott SL. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Applied and environmental microbiology. 2011;77(10):3219-26.

9. Appendices

The appendices have been provided as a separate document. A list of appendices has been provided below:

Appendix A

Table A1 – Abundance of Genera Detected by mothur, QIIME and MG-RAST

Appendix B

Table B1 – Abundance Table Generated by QIIME

Appendix C

Figure C1 - Principal Component Analysis of Genus Abundance Data
Figure C2 - Principal Component Analysis of Genus Abundance Data post removal of *Bifidobacterium*

Appendix D

Table D1 – Top Five Most Abundant Genera Across the Different Age Groups Across All Samples
Table D2 - Top Five Most Abundant Genera Across the Different Age Groups in the Placebo Group
Table D3 - Top Five Most Abundant Genera Across the Different Age Groups in the Probiotic Group

Appendix E

Table E1 - Results of post Hoc Testing of Genera Differences Between Age Groups
Table E2 – Results of post Hoc Testing of Differences in Beta Diversity Between Age Groups

Appendix F

Figures F1 to F6 – Abundance Profiles of Selected Individuals

Appendix G

Figure G1 – The Effect of Age and Allocation Group on Effective Number of Genera and Genus Richness

Appendix H

Figure H1 – The Effect of Age and Allocation Group on Bacterial Load

Appendix I

Figure I1 - Phylogenetic Tree of *Bifidobacterium* Type Strains and Representative Sequences from *Bifidobacterium* OTUs.

Appendix J

Figure J1 – Multiple Sequence Alignment of *Streptococcus* Type Strains and *Bifidobacterium* Optimised Primers

Appendix A

Table A1 - Abundance of Genera Detected by mothur, QIIME and MG-RAST

Genus	mothur (%)	QIIME (%)	MG-RAST (%)
*Bifidobacterium**	18.71436	35.7940	36.96168
*Escherichia-Shigella**	12.83756	1.28804	7.19790
Veillonella	4.25275	3.23947	3.56765
Staphylococcus	4.09034	3.41299	2.71659
Lactobacillus	3.30853	2.17418	2.22237
Bacteroides	2.92707	1.52370	1.09086
Clostridium	2.30011	3.44043	3.51401
Enterococcus	2.18303	4.08687	4.16999
Streptococcus	1.20104	1.04996	0.64379
Prevotella	1.11040	0.60609	0.63647
Acinetobacter	0.44189	0.15415	0.04958
Faecalibacterium	0.40035	0.51893	0.57632
*Klebsiella**	0.16618	0.44872	12.20188
Providencia	0.16241	0.06537	0.06665
Blautia	0.13597	0.08070	0.00244
Haemophilus	0.13219	0.06134	0.02764
Parabacteroides	0.12086	0.03067	0.02845
Eggerthella	0.11708	0.05488	0.05934
Actinomyces	0.10953	0.05326	0.05284
Megamonas	0.09820	0.19773	0.00000
Dysgonomonas	0.09442	0.12671	0.00000
Propionibacterium	0.08687	0.05407	0.01463
Aeromonas	0.05288	0.01937	0.02032
Parasutterella	0.03777	0.00968	0.00000
Raoultella	0.03777	0.00081	0.00000
Akkermansia	0.03399	0.03228	0.00000
Megasphaera	0.03399	0.02260	0.00244
Granulicatella	0.03021	0.01049	0.00163
Neisseria	0.02266	0.00646	0.00000
Morganella	0.01888	0.01130	0.01870
Anaerococcus	0.01511	0.00968	0.00406
Corynebacterium	0.01511	0.07506	0.05365
Finegoldia	0.01511	0.01211	0.01219
Collinsella	0.01133	0.01049	0.01219
Anaerostipes	0.00755	0.00404	0.00163
Catenibacterium	0.00755	0.00646	0.00650
Desulfovibrio	0.00755	0.00404	0.00000
Peptoniphilus	0.00755	0.00404	0.00488
Rothia	0.00755	0.00161	0.00081
Sutterella	0.00755	0.00404	0.00000
Acidovorax	0.00378	0.00000	0.00000
Actinobacillus	0.00378	0.02663	0.00000

Genus	mothur (%)	QIIME (%)	MG-RAST (%)
Eubacterium	0.00378	0.00323	0.01138
Fusobacterium	0.00378	0.00242	0.00244
Gardnerella	0.00378	0.00081	0.00081
Nesterenkonia	0.00378	0.00081	0.00000
Varibaculum	0.00378	0.00081	0.00000
Abiotrophia	0.00000	0.00000	0.00813
Atopobium	0.00000	0.00081	0.00081
Barnesiella	0.00000	0.00081	0.00000
Bradyrhizobium	0.00000	0.00161	0.00000
Burkholderia	0.00000	0.00000	0.00325
Cedecea	0.00000	0.00081	0.00000
Cellulosilyticum	0.00000	0.00242	0.00000
Citrobacter	0.00000	2.29765	0.04877
Coprococcus	0.00000	0.00565	0.00000
Cronobacter	0.00000	0.01211	0.00081
Dialister	0.00000	0.00081	0.00081
*Enterobacter**	0.00000	28.4596	2.58815
Epulopiscium	0.00000	0.00161	0.00000
Erwinia	0.00000	0.00000	0.07560
Erysipelatoclostridium	0.00000	0.00000	0.32515
Flavonifractor	0.00000	0.03147	0.00000
Gemella	0.00000	0.00646	0.00650
Geobacillus	0.00000	0.00000	0.08698
Howardella	0.00000	0.00161	0.00000
Intestinibacter	0.00000	0.00000	0.02195
Kluyvera	0.00000	0.03470	1.21198
Lachnoclostridium	0.00000	0.00000	0.00163
Methylocystis	0.00000	0.00000	0.00081
Pantoea	0.00000	0.00726	0.54950
Pectobacterium	0.00000	0.00000	0.05202
Peptoclostridium	0.00000	0.00000	0.05202
Planomicrobium	0.00000	0.00242	0.00000
Pseudobutyrivibrio	0.00000	0.00081	0.00000
Pseudomonas	0.00000	0.00000	0.01382
Ruminococcus	0.00000	0.00081	2.34511
Salmonella	0.00000	0.00000	0.11461
Sarcina	0.00000	0.00081	0.00000
Subdoligranulum	0.00000	0.03147	0.00000
Tatumella	0.00000	0.00807	0.00000
Trabulsiella	0.00000	0.07748	0.00000
Variovorax	0.00000	0.00242	0.00000
Vibrio	0.00000	0.00000	0.07397
Unclassified*	44.61231	10.2688	16.46291

Data presented as percentage abundance data. * Indicates genera that were found at an abundance of at least 5% by one or more pipeline. *Escherichia* and *Shigella* are combined into one genus as they cannot be distinguished by their 16S rRNA gene sequence.

Appendix B

Table B1 – Abundance Table Generated by QIIME

Sample ID	Acidaminococcus	Acidovorax	Acinetobacter	Actinobacillus	Actinobaculum	Actinomyces	Adlercreutzia	Aeromicrobium	Aeromonas	Akkermansia	Alistipes	Alkanindiges	Anaerococcus
10002_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10002_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_256	0	0	0	0	0	0.000508647	0	0	0	0	0	0	0
10006_A_3	0	0	0.000160128	0	0	0	0	0	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0	0	0	0	0.003700479	0	0	0	0	0	0	0
10013_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0.245354858	0	0	0	0	0	0	0	0	0	0
10020_A_285	0	0	0	0	0	0.000573394	0	0	0	0	0	0	0
10020_A_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_285	0	0	0	0	0	0.00295421	0	0	0	0	0	0	0
10024_B_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10025_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_266	0	0	0	0	0	0.000711997	0	0	0	0	0	0	0
10029_A_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_529	0	0	0.000170619	0	0	0	0.000170619	0	0	0	0	0	0.000341239
10029_A_59	0	0	0	0	0	0.00306925	0	0	0	0	0	0	0.000575484
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0	0	0	0.00122449	0	0	0	0	0	0	0.001632653
10042_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0	0	0	0	0	0.0746875	0	0	0	0	0	0	0
10059_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0.000341413
10059_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0	0	0.111374408	0	0	0	0	0	0	0
10063_B_36	0	0	0.000386847	0	0	0.012959381	0	0	0	0	0	0	0.011411992
10063_B_57	0	0	0	0	0	0.038321559	0	0	0	0	0	0	0.542496253
10063_B_613	0	0	0.003831418	0	0	0	0	0	0	0	0.148877942	0	0
10078_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Anaerostipes	Anaerotruncus	Aquabacterium	Atopobium	Bacillus	Bacteroides	Barnesiella	Bifidobacterium	Bilophila	Blautia	Bradyrhizobium	Brevundimonas	Burkholderia
10002_A_32	0	0	0	0	0	0	0	0.000552079	0	0	0	0	0
10002_A_9	0	0	0	0	0.001102941	0	0	0	0	0	0	0	0
10006_A_256	0	0	0	0	0	0	0	0.034079349	0	0.007121058	0	0	0
10006_A_3	0	0	0	0	0	0	0	0.983186549	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0	0.537885096	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0	0.324584896	0	0	0	0	0
10009_B_35	0	0	0	0	0	0	0	0.192574902	0	0	0	0	0
10009_B_59	0	0	0	0	0	0	0	0.626168224	0	0	0	0	0
10013_A_294	0	0	0	0.000653026	0	0	0	0.5711798	0	0.001741402	0	0	0
10013_A_32	0	0	0	0	0	0	0	0.749130022	0	0	0	0	0
10013_A_58	0	0	0	0	0	0	0	0.674408509	0	0.001193835	0	0	0
10013_A_9	0	0	0	0	0	0	0	0.000799361	0	0	0	0	0
10020_A_18	0	0	0	0	0	0	0	0.12762717	0	0	0	0	0
10020_A_285	0	0	0	0	0	0.152522936	0.002293578	0.198394495	0	0.112385321	0	0	0
10020_A_29	0	0	0	0	0	0	0	0.108712849	0	0.000452028	0	0	0
10020_A_60	0	0	0	0	0	0	0	0.7203125	0	0.0015625	0	0	0
10024_B_285	0	0	0	0	0	0	0	0.003692762	0.040620384	0.010339734	0	0	0
10024_B_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0.025839793	0	0	0	0	0
10025_B_58	0	0	0	0	0	0	0	0.638157895	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0.614551607	0	0	0	0	0
10029_A_266	0	0	0	0	0	0	0	0.061587754	0	0.032751869	0	0	0
10029_A_38	0	0	0	0	0	0	0	0.290563972	0	0	0	0	0
10029_A_529	0.003924245	0.000170619	0	0	0.002047432	0.004436103	0	0.000170619	0	0.062958539	0	0	0
10029_A_59	0	0	0	0	0	0	0	0.37694226	0	0	0	0	0
10041_B_2	0	0	0	0	0	0	0	0.002824859	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0.031976744	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0	0.000816327	0	0	0	0.341632653	0	0	0	0	0
10042_A_32	0	0	0	0	0	0	0	0.639388489	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0.004605642	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0.191011236	0	0.002808989	0	0	0
10042_A_9	0	0	0	0	0	0	0	0.02796875	0	0	0	0	0
10059_A_32	0	0	0	0	0	0	0	0.164902697	0	0	0	0	0
10059_A_58	0	0	0	0	0	0	0	0.114384749	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0.037914692	0	0.068720379	0	0.109004739	0	0.011848341	0	0	0
10063_B_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_57	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_613	0	0.000547345	0	0	0	0.483305966	0.007662835	0.001642036	0	0.006568144	0	0	0
10078_A_13	0	0	0	0	0	0	0	0.160503241	0	0.001906214	0	0	0

Sample ID	Catenibacterium	Cedecea	Cellulomonas	Cellulosilyticum	Chloroplast	Christensenella	Citrobacter	Clostridium	Collinsella	Coprobacillus	Coprococcus	Corynebacterium	Cronobacter
10002_A_32	0	0	0	0	0	0	0.000184026	0	0	0	0	0	0
10002_A_9	0	0	0	0	0	0	0	0	0	0	0	0.020588235	0
10006_A_256	0	0	0	0	0	0	0	0.01017294	0	0	0.211088505	0	0
10006_A_3	0	0	0	0	0	0	0	0.000960769	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0	0.001606856	0	0	0	0	0
10009_B_35	0	0	0	0	0	0	0	0.021059488	0	0	0	0.000434216	0
10009_B_59	0	0	0	0	0	0	0.001038422	0.001038422	0	0	0	0	0
10013_A_294	0	0	0	0.000217675	0.000217675	0	0	0.000217675	0	0	0	0	0
10013_A_32	0	0	0	0	0	0	0	0.001581778	0	0	0	0	0
10013_A_58	0	0	0	0	0	0	0	0.000651183	0	0	0	0	0
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0.007993605	0
10020_A_18	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_285	0	0	0	0	0	0	0	0.007454128	0.004013761	0	0.001146789	0	0
10020_A_29	0	0	0	0	0	0	0	0	0	0	0	0.000452028	0
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_285	0	0	0	0	0	0	0.000738552	0.005908419	0	0	0.014032496	0	0
10024_B_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0.023255814	0	0	0	0	0
10025_B_58	0	0	0	0	0	0	0	0.006578947	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_266	0	0	0	0.003203987	0	0	0	0.00249199	0	0	0.043431826	0	0
10029_A_38	0	0	0	0	0	0	0	0.000912575	0	0	0	0	0
10029_A_529	0	0	0	0	0	0	0	0.000511858	0.020303702	0	0.001706193	0	0
10029_A_59	0	0	0	0	0	0	0.000191828	0.001918281	0	0	0	0	0
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0.035110533	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0.002422481	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0	0	0	0	0	0.002040816	0	0	0	0	0
10042_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0	0	0.00015625	0	0	0	0	0	0	0	0	0.00109375	0
10059_A_32	0	0	0	0	0	0	0.000341413	0.002048481	0	0	0	0	0
10059_A_58	0	0	0	0	0	0	0	0.013864818	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0	0	0	0.007109005	0.052132701	0	0	0	0	0
10063_B_36	0	0	0	0	0	0	0.027079304	0	0	0	0	0	0
10063_B_57	0	0	0	0	0	0	0	0.000214087	0	0	0	0	0
10063_B_613	0	0	0	0	0	0	0.001642036	0	0	0	0.004378763	0	0
10078_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Desulfovibrio	Dialister	Dorea	Dysgonomonas	Eggerthella	Elizabethkingia	Enterobacter	Enterococcus	Enterorhabdus	Epulopiscium	Erwinia	Escherichia-Shigella	Eubacterium
10002_A_32	0	0	0	0	0	0	0.599006257	0.000184026	0	0	0	0.001472212	0
10002_A_9	0	0	0	0	0	0	0	0.591911765	0	0	0	0	0
10006_A_256	0	0	0	0	0.002034588	0	0.004577823	0.006612411	0	0	0	0.086978637	0
10006_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0.001665279	0.005828476	0	0	0	0.084929226	0
10009_B_12	0	0	0	0	0	0	0.587663709	0.177017322	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0.001071237	0.001071237	0	0	0	0.01178361	0
10009_B_35	0	0	0	0	0	0	0.557099436	0.130264872	0	0	0	0.000434216	0
10009_B_59	0	0	0	0	0	0	0.126687435	0.016614746	0	0	0.001038422	0.001038422	0
10013_A_294	0	0	0	0	0.006312582	0	0.002612103	0.007183283	0	0	0	0	0
10013_A_32	0	0	0	0	0	0	0.119582411	0	0	0	0	0.000158178	0
10013_A_58	0	0	0	0	0.00010853	0	0.061536792	0.009008031	0	0	0	0	0
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0	0	0	0	0.006701188	0	0	0	0	0	0
10020_A_285	0	0	0	0	0	0	0.002293578	0.012614679	0	0	0.000573394	0.001720183	0
10020_A_29	0	0	0	0	0	0	0.278675557	0.008023505	0	0	0	0	0
10020_A_60	0	0	0	0	0	0	0.05625	0	0	0	0	0	0
10024_B_285	0	0	0	0	0.015509601	0	0.097488922	0.058345643	0	0	0	0.000738552	0
10024_B_33	0	0	0	0	0	0	0.017002519	0.001259446	0	0	0	0.003778338	0
10024_B_58	0	0	0	0	0	0	0.180878553	0.07751938	0	0	0	0.010335917	0
10025_B_58	0	0	0	0	0	0	0.151315789	0.013157895	0	0	0	0.026315789	0
10029_A_14	0	0	0	0	0	0	0.368527919	0	0	0	0	0.002763677	0
10029_A_266	0	0	0	0	0	0	0.002135991	0.000711997	0	0	0	0.466002136	0
10029_A_38	0	0	0	0	0	0	0.145099471	0.001277605	0	0	0	0.221755795	0
10029_A_529	0	0	0	0	0.000853097	0	0.004436103	0.000682477	0.000341239	0	0	0.213444805	0.000511858
10029_A_59	0	0	0	0	0	0	0.004603875	0.001342797	0	0	0	0.512564742	0
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0.012678804	0.017880364	0	0	0	0.894993498	0
10041_B_58	0	0	0	0	0	0	0.084786822	0.038275194	0	0	0	0.209302326	0
10041_B_9	0	0	0	0	0	0	0.07576769	0	0	0	0	0.015020027	0
10042_A_272	0	0.000408163	0	0	0	0	0.531428571	0.02122449	0	0	0	0.016326531	0
10042_A_32	0	0	0	0	0	0	0	0	0	0	0	0.00044964	0
10042_A_58	0	0	0	0	0	0	0.91378814	0	0	0	0	0.012377663	0
10042_A_61	0	0	0	0	0	0	0.634831461	0	0	0	0	0.058988764	0
10042_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_32	0	0	0	0.003072721	0	0	0.677364288	0.012632298	0	0	0	0.030385797	0
10059_A_58	0	0	0	0	0	0	0.419410745	0.079722704	0	0	0	0.060658579	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0	0.004739336	0	0.021327014	0.059241706	0	0	0	0.267772512	0
10063_B_36	0	0	0	0	0	0	0.028239845	0.008704062	0	0	0	0.451450677	0
10063_B_57	0	0	0	0	0	0	0.002569043	0.108113894	0	0	0	0.07921216	0
10063_B_613	0	0	0	0	0	0	0.012588944	0	0	0	0	0.001642036	0
10078_A_13	0	0	0	0	0	0	0.009721693	0.001715593	0	0	0	0.000381243	0

Sample ID	Faecalibacterium	Finegoldia	Flavonifractor	Fusobacterium	Gardnerella	Gemella	Gordonibacter	Granulicatella	Haemophilus	Hafnia-Obesumbacterium	Holdemania	Howardella	Klebsiella
10002_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0.000368053
10002_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_256	0	0	0.004577823	0	0	0	0	0	0	0	0	0	0
10006_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0.008027038
10009_B_322	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_35	0	0.001953973	0	0	0	0	0	0	0	0	0	0	0.003690838
10009_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0.001306051	0	0	0	0.001959077	0.000217675	0.011972138	0	0	0	0	0
10013_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0.015501424
10013_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0.007163013
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_285	0	0	0	0	0	0	0	0.000573394	0	0	0	0	0
10020_A_29	0	0.001243078	0	0	0	0.000904057	0	0	0	0	0	0	0.000113007
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_285	0	0	0.29689808	0	0	0.000738552	0	0.000738552	0	0	0	0	0
10024_B_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0.100775194
10025_B_58	0	0	0	0	0.006578947	0	0	0	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0	0	0	0	0	5.64016E-05
10029_A_266	0	0	0.027055892	0	0	0	0	0	0	0	0	0	0
10029_A_38	0	0.000547545	0	0	0	0	0	0	0	0	0	0	0.00109509
10029_A_529	0.046067224	0.000682477	0.003412387	0	0	0	0	0	0	0	0.000682477	0	0
10029_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0.000191828
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0.000325098	0	0	0	0	0	0	0	0	0	0	0.000325098
10041_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_32	0	0.00022482	0	0	0	0	0	0	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_32	0	0.003755548	0	0	0	0	0	0	0	0.008193923	0	0	0
10059_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0.016587678	0	0	0	0.016587678	0	0	0	0	0	0.004739336
10063_B_36	0	0.000193424	0	0	0	0	0	0	0	0	0	0	0
10063_B_57	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_613	0.002736727	0	0.010399562	0	0	0	0	0	0	0	0	0	0
10078_A_13	0	0.008387343	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Kluyvera	Lachnospira	Lactobacillus	Lactococcus	Massilia	Megamonas	Megasphaera	Methylobacterium	Microbacterium	Morganella	Negativicoccus	Neisseria	Nesterenkonia
10002_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10002_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_256	0	0	0	0.002543235	0	0	0	0	0	0	0	0	0
10006_A_3	0	0	0.001120897	0	0	0	0	0	0	0	0	0	0
10006_A_74	0	0	0.029975021	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_322	0	0	0.614890198	0	0	0	0	0	0	0	0	0	0
10009_B_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0	0.000217675	0.00043535	0	0	0	0	0	0	0	0	0
10013_A_32	0	0	0.002372667	0	0	0	0	0	0	0	0	0	0
10013_A_58	0	0	0.000542652	0	0	0	0	0	0	0	0	0	0
10013_A_9	0	0	0.002398082	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0.0001523	0	0.0001523	0	0	0	0	0	0	0	0
10020_A_285	0	0.014334862	0.001720183	0	0	0	0	0	0	0	0	0	0
10020_A_29	0.001017064	0	0.000452028	0	0	0	0	0	0	0	0	0	0
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_285	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_33	0	0	0.000629723	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10025_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_266	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_38	0	0	0.002372696	0	0	0	0	0	0	0	0	0	0
10029_A_529	0	0.003071148	0	0.000170619	0	0	0	0	0	0	0	0	0
10029_A_59	0	0	0.000383656	0	0	0	0	0	0	0	0.023978515	0	0
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0.000650195	0	0	0	0	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0.000408163	0	0	0	0	0	0	0	0	0	0
10042_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0	0	0.00140625	0	0	0	0	0	0	0	0	0	0
10059_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_9	0	0	0.001930502	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0.061611374	0	0	0	0	0	0	0	0	0	0
10063_B_36	0	0	0.068471954	0	0	0	0	0	0	0	0.223210832	0	0
10063_B_57	0	0	0.172768144	0	0	0	0	0	0	0	0.001284522	0	0
10063_B_613	0	0	0	0	0	0	0	0	0	0	0	0	0
10078_A_13	0	0	0	0	0	0	0	0	0	0	0.000571864	0	0

Sample ID	Novosphingobium	Odoribacter	Pantoea	Parabacteroides	Parasutterella	Pediococcus	Pelomonas	Peptococcus	Peptoniphilus	Peptostreptococcus	Planomicrobium	Porphyromonas	Prevotella
10002_A_32	0	0	0.000920132	0	0	0	0	0	0	0	0	0	0
10002_A_9	0	0	0	0	0	0	0	0	0	0	0.000735294	0	0
10006_A_256	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_74	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0	0	0	0	0	0	0	0	0	0.00043535	0	0
10013_A_32	0	0	0.004903512	0	0	0.000158178	0	0	0	0	0	0	0
10013_A_58	0	0	0.003038854	0	0	0	0	0	0	0	0.002713262	0	0
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0	0	0	0.007767286	0	0	0	0	0	0	0
10020_A_285	0	0	0	0.001146789	0	0	0	0	0	0	0	0	0
10020_A_29	0	0	0	0	0	0	0	0	0	0	0.000678043	0	0
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_285	0	0	0	0	0	0	0	0	0	0	0.036189069	0	0.000738552
10024_B_33	0	0	0.967884131	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0.025839793	0	0	0	0	0	0	0	0.165374677	0	0
10025_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_14	0	0	0.000112803	0	0	0	0	0	0	0	0	0	0
10029_A_266	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_38	0	0	0.000182515	0	0	0	0	0	0	0	0	0	0
10029_A_529	0	0	0	0.006824774	0	0	0	0	0	0	0	0.000170619	0
10029_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0	0.000816327	0	0	0	0	0	0	0	0.017142857	0	0
10042_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_32	0	0	0	0	0	0	0	0	0.002048481	0	0	0	0
10059_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_57	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_613	0	0.003831418	0	0.028461959	0	0	0	0	0	0	0	0	0.082101806
10078_A_13	0	0	0	0	0	0	0	0	0.000381243	0.007052993	0	0	0

Sample ID	Propionibacterium	Proteus	Providencia	Pseudobutyrivibrio	Pseudomonas	Pseudonocardia	Ralstonia	Raoultella	Rhizobacter	Rhizobium	Roseburia	Rothia	Ruminococcus
10002_A_32	0	0	0	0	0.000184026	0	0	0	0	0	0	0	0
10002_A_9	0	0	0	0	0.000735294	0	0	0	0	0	0	0	0
10006_A_256	0	0	0	0	0	0	0	0	0	0	0	0.000508647	0
10006_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10006_A_74	0.029975021	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_12	0	0	0	0	0.000422476	0	0	0	0	0	0	0	0
10009_B_322	0	0	0	0	0	0	0	0	0	0	0	0	0
10009_B_35	0.000217108	0	0	0	0.000217108	0	0	0	0	0	0	0	0
10009_B_59	0.038421599	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0	0	0	0	0	0	0	0	0	0	0.000217675	0
10013_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10020_A_18	0	0	0	0	0.000609199	0	0	0	0	0	0	0	0
10020_A_285	0	0	0	0.001146789	0	0	0	0	0	0	0.030389908	0	0
10020_A_29	0.010622669	0	0	0	0	0	0	0.000565036	0	0	0	0	0
10020_A_60	0	0	0	0	0	0	0	0	0	0	0	0.0015625	0
10024_B_285	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0	0	0	0	0.002583979	0
10025_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_14	0	0	0	0	0	0	0	0	0	0	0	5.64016E-05	0
10029_A_266	0	0	0	0	0	0	0	0	0	0	0	0	0
10029_A_38	0.003285271	0	0	0	0	0	0	0	0	0	0	0.00036503	0
10029_A_529	0	0	0	0.015014503	0	0	0	0	0	0	0.003412387	0	0.140249104
10029_A_59	0.000767312	0	0	0	0	0	0	0	0	0	0	0.000575484	0
10041_B_2	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10041_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_272	0	0.000408163	0	0	0	0	0	0	0	0	0	0	0
10042_A_32	0.00067446	0.006070144	0	0	0.00022482	0	0	0	0	0	0	0	0
10042_A_58	0	0.000719632	0	0	0	0	0	0	0	0	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10042_A_9	0.01484375	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_32	0.001365654	0	0	0	0	0	0	0.000341413	0	0	0	0	0
10059_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10059_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0.007109005	0	0	0	0	0	0	0	0	0
10063_B_36	0.000386847	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_57	0.000214087	0	0	0.00278313	0	0	0	0	0	0	0.000214087	0	0
10063_B_613	0	0	0	0	0	0	0	0	0	0	0	0	0
10078_A_13	0.000190621	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Salmonella	Sarcina	Selenomonas	Serratia	Sneathia	Sphingomonas	Sporosarcina	Staphylococcus	Stenotrophomonas	Streptococcus	Subdoligranulum	Sutterella	Tatumella
10002_A_32	0	0	0	0	0	0	0	0.000184026	0	0.017666544	0	0	0.000368053
10002_A_9	0	0	0	0	0	0	0	0.377573529	0	0	0	0	0
10006_A_256	0	0	0	0	0	0	0	0	0	0.002034588	0	0	0
10006_A_3	0	0	0	0	0	0	0	0.012009608	0	0.002401922	0	0	0
10006_A_74	0	0	0	0	0	0	0	0	0	0.013322231	0	0	0
10009_B_12	0	0	0	0.007182087	0	0	0	0.163075623	0	0	0	0	0.000844951
10009_B_322	0	0	0	0	0	0	0	0	0	0.01981789	0.002678093	0	0
10009_B_35	0	0	0	0	0	0	0	0.000651324	0	0	0	0	0
10009_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10013_A_294	0	0	0	0	0	0	0	0	0	0.367871136	0	0	0
10013_A_32	0	0	0	0	0	0	0	0.000158178	0	0.020879469	0	0	0
10013_A_58	0	0	0	0	0	0	0	0.00010853	0	0.183742132	0	0	0
10013_A_9	0	0	0	0	0	0	0	0.752997602	0	0.235811351	0	0	0
10020_A_18	0	0	0	0	0	0	0	0.579195857	0	0.031830643	0	0	0
10020_A_285	0	0	0	0	0	0	0	0	0	0.003440367	0	0.001146789	0
10020_A_29	0	0	0	0	0	0	0	0.053678382	0	0.042942705	0	0	0
10020_A_60	0	0	0	0	0	0	0	0.0140625	0	0.1015625	0	0	0
10024_B_285	0	0	0	0	0	0	0	0	0	0.015509601	0	0	0
10024_B_33	0	0	0	0	0	0	0	0.003148615	0	0	0	0	0
10024_B_58	0	0	0	0	0	0	0	0	0	0.090439276	0	0	0
10025_B_58	0	0	0	0	0	0	0	0.013157895	0	0.019736842	0	0	0
10029_A_14	0	0	0	0	0	0	0	0.000846024	0	0.00783982	0	0	0
10029_A_266	0	0	0	0	0	0	0	0	0	0.000711997	0	0	0
10029_A_38	0	0	0	0	0	0	0	0.000912575	0	0.00036503	0	0	0
10029_A_529	0	0	0	0	0	0	0	0	0	0.021156799	0.072683842	0	0
10029_A_59	0	0	0	0	0	0	0	0	0	0.038941109	0	0	0
10041_B_2	0	0	0	0	0	0	0	0.997175141	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0	0	0.018530559	0	0	0
10041_B_58	0	0	0	0	0	0	0	0	0	0.000484496	0	0	0
10041_B_9	0	0	0	0	0	0	0	0.000667557	0	0	0	0	0
10042_A_272	0	0	0	0	0	0	0	0	0	0.040408163	0	0	0
10042_A_32	0	0	0	0	0	0	0	0.002697842	0	0	0	0	0
10042_A_58	0	0	0	0	0	0	0	0.001583189	0	0.001583189	0	0	0
10042_A_61	0	0	0	0	0	0	0	0	0	0.109550562	0	0	0
10042_A_9	0	0	0	0	0	0	0	0.87984375	0	0	0	0	0
10059_A_32	0	0	0	0	0	0	0	0.002048481	0	0.000341413	0	0	0
10059_A_58	0	0	0	0	0	0	0	0	0	0.001733102	0	0	0
10059_A_9	0	0	0	0	0	0	0	0.998069498	0	0	0	0	0
10063_B_286	0	0	0	0	0	0	0	0	0	0.007109005	0	0	0
10063_B_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10063_B_57	0	0	0	0	0	0.000214087	0	0	0	0	0	0	0
10063_B_613	0	0	0	0	0	0	0	0	0.003831418	0.000547345	0	0	0
10078_A_13	0	0	0	0	0	0	0	0.002096836	0	0.775066717	0	0	0

Sample ID	Tessaracoccus	Trabulsiella	Turicibacter	Varibaculum	Variovorax	Veillonella	Weissella	Unclassified
10002_A_32	0	0	0	0	0	0.000368053	0	0.37854251
10002_A_9	0	0	0	0	0	0	0	0.007352941
10006_A_256	0	0	0	0	0	0.000508647	0	0.626653103
10006_A_3	0	0	0	0	0	0.000160128	0	0
10006_A_74	0	0	0	0	0	0.293089092	0	0.003330558
10009_B_12	0	0	0	0	0	0	0	0.055766793
10009_B_322	0	0	0.000535619	0	0	0.017675415	0	0.004284949
10009_B_35	0	0	0	0	0	0.028224056	0	0.063178463
10009_B_59	0	0	0	0	0	0.013499481	0	0.174454829
10013_A_294	0	0	0	0	0	0.000653026	0.000217675	0.020461471
10013_A_32	0	0	0	0	0	0	0	0.085574185
10013_A_58	0	0	0	0	0	0.000217061	0	0.055567614
10013_A_9	0	0	0	0	0	0	0	0
10020_A_18	0	0	0	0	0	0	0	0.000609199
10020_A_285	0	0	0	0	0	0.010894495	0	0.439220183
10020_A_29	0	0	0	0	0	0.008588541	0	0.482879421
10020_A_60	0	0	0	0	0	0.0140625	0	0.090625
10024_B_285	0	0	0	0	0	0.005169867	0	0.393648449
10024_B_33	0	0	0	0	0	0	0	0.006297229
10024_B_58	0	0	0	0	0	0	0	0.297157623
10025_B_58	0	0.006578947	0	0	0	0	0	0.118421053
10029_A_14	0	0	0	0	0	0	0	0.005245347
10029_A_266	0	0	0	0	0	0.005695977	0	0.353506586
10029_A_38	0	0	0	0	0	0.280890673	0	0.050374156
10029_A_529	0	0	0	0	0	0.001023716	0	0.368196553
10029_A_59	0	0	0	0	0	0.01323614	0	0.020717437
10041_B_2	0	0	0	0	0	0	0	0
10041_B_30	0	0	0	0	0	0	0	0.019505852
10041_B_58	0	0	0	0	0	0	0	0.632751938
10041_B_9	0	0	0	0	0	0	0	0.908544726
10042_A_272	0	0	0	0	0	0.00122449	0	0.022857143
10042_A_32	0	0	0	0	0	0	0	0.350269784
10042_A_58	0	0	0	0	0	0	0	0.065342545
10042_A_61	0	0	0	0	0	0	0	0.002808989
10042_A_9	0	0	0	0	0	0	0	0
10059_A_32	0	0	0	0	0	0.003072721	0	0.087743257
10059_A_58	0	0	0	0	0	0	0	0.310225303
10059_A_9	0	0	0	0	0	0	0	0
10063_B_286	0	0	0	0.002369668	0	0.007109005	0	0.125592417
10063_B_36	0	0	0	0.001547389	0	0.011798839	0	0.154158607
10063_B_57	0	0	0	0.000428174	0	0.000214087	0	0.050952687
10063_B_613	0	0	0	0	0	0	0	0.195402299
10078_A_13	0	0	0	0	0	0.006290507	0	0.025733892

Sample ID	Acidaminococcus	Acidovorax	Acinetobacter	Actinobacillus	Actinobaculum	Actinomyces	Adlercreutzia	Aeromicrobium	Aeromonas	Akkermansia	Alistipes	Alkanindiges	Anaerococcus
10078_A_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10078_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_28	0	0	0	0	0	0	0	0	0	0	0	0	0.001699235
10088_A_346	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_304	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_228	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_41	0	0	0	0	0	0	0	0	0	0	0	0	0.001221001
10135_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_79	0	0	0	0	0	0	0	0	0	0.477542109	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_27	0	0	0	0	0	0	0	0	0	0	0	0	0.002767209
10141_A_292	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_3	0	0	0	0	0	0.040723982	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0.007821229	0	0	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_336	0	0	0	0.007825601	0	0.000558971	0	0	0	0	0	0	0
10148_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_3	0	0	0	0	0	0	0	0	0	0	0	0.023094688	0
10153_B_40	0	0	0	0	0	0.001956947	0	0	0	0	0	0	0
10153_B_522	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0	0	0	0	0	0	0	0	0	0.04962406	0	0	0
10171_A_70	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0	0	0	0	0	0
10173_B_237	0	0	0	0	0	0.009186646	0	0	0	0	0	0	0.000448129
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0.00022007
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Anaerostipes	Anaerotruncus	Aquabacterium	Atopobium	Bacillus	Bacteroides	Barnesiella	Bifidobacterium	Bilophila	Blautia	Bradyrhizobium	Brevundimonas	Burkholderia
10078_A_27	0	0	0	0	0	0	0	0.044585987	0	0	0	0	0
10078_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_28	0	0	0	0	0	0	0	0.291418862	0	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0.738135593	0	0.002542373	0	0	0
10088_A_62	0	0	0	0	0	0	0	0.22602195	0	0.001431525	0	0	0
10088_A_9	0	0	0	0	0	0	0	0.563125298	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_304	0.004334691	0	0	0	0	0	0	0.000262709	0.000131354	0.038618153	0	0	0
10090_B_59	0	0	0	0	0	0	0	0	0	0.000814996	0	0	0
10132_A_16	0	0	0	0	0	0	0	0.583592939	0	0	0	0	0
10132_A_228	0	0	0	0	0	0.002147239	0	0.053680982	0	0.000306748	0	0	0
10132_A_41	0	0	0	0	0	0	0	0.401302401	0	0	0	0	0
10135_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0.011976048	0	0	0	0	0
10135_B_79	0	0	0	0	0	0	0	0.016375546	0	0	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0.416767922	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0.079754601	0	0	0	0	0
10141_A_27	0	0	0	0	0	0	0	0.736769284	0	0	0	0	0
10141_A_292	0.000358423	0	0	0	0	0	0	0.414336918	0	0	0	0	0
10141_A_3	0	0	0	0.176470588	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0.006145251	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_336	0	0	0	0	0	0.071268865	0	0.264393516	0	0	0	0	0
10148_A_36	0	0	0	0	0	0	0	0.303191489	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0.254697286	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_3	0	0	0.018475751	0	0	0	0	0	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0	0.337573386	0	0	0	0	0
10153_B_522	0.01843318	0	0	0	0	0	0	0.073732719	0	0.016129032	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0.000302572	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0.162371134	0	0	0	0	0
10171_A_260	0	0	0	0	0	0.180451128	0.001503759	0.063157895	0	0	0	0	0
10171_A_70	0	0	0	0	0	0	0	0.522569859	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0	0.501939058	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0.360041623	0	0	0	0	0
10173_B_237	0	0	0	0	0	0	0	0.790499664	0	0	0	0	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Catenibacterium	Cedecea	Cellulomonas	Cellulosilyticum	Chloroplast	Christensenella	Citrobacter	Clostridium	Collinsella	Coprobacillus	Coprococcus	Corynebacterium	Cronobacter
10078_A_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10078_A_58	0	0	0	0	0	0	0	0.000383583	0	0	0	0	0
10088_A_28	0	0	0	0	0	0	0	0.049277825	0	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0.003389831	0	0	0	0	0
10088_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0	0	0	0	0	0.067727609	0	0	0	0	0
10090_B_304	0	0	0	0	0	0	0.002889794	0.006042296	0	0	0.175095232	0	0
10090_B_59	0	0	0	0	0	0	0	0.512632437	0	0	0	0	0
10132_A_16	0	0	0	0	0	0	0.000519211	0	0	0	0	0	0
10132_A_228	0	0	0	0	0	0	0.000613497	0	0.001840491	0	0	0	0
10132_A_41	0	0	0	0	0	0	0.001017501	0.005901506	0	0	0	0.0002035	0
10135_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0.001122754	0	0	0	0	0
10135_B_79	0	0	0	0	0	0	0.000935745	0.041016843	0	0	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0.041919806	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0.05398773	0.000408998	0	0	0	0	0
10141_A_27	0	0.000345901	0	0	0	0	0.012106538	0.01279834	0	0	0	0.000345901	0.000691802
10141_A_292	0	0	0	0	0	0	0	0.017204301	0	0	0	0	0
10141_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0.008130081	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0.001837753	0.008926227	0	0	0	0	0
10148_A_336	0.003353829	0	0	0	0	0	0	0	0.000558971	0	0	0	0
10148_A_36	0	0	0	0	0	0	0	0.000967118	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0.003131524	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0.009966777	0	0	0	0	0
10153_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0	0.009784736	0	0	0	0	0
10153_B_522	0	0	0	0	0	0	0	0.110599078	0	0	0.002304147	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0.000302572	0	0	0.05204236	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0.012886598	0	0	0	0	0
10171_A_260	0	0	0	0	0	0	0.05112782	0	0.006015038	0	0	0	0
10171_A_70	0	0	0	0	0	0	0	0.000238834	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0.008426452	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0.00166205	0.135457064	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0.033818939	0	0	0	0	0
10173_B_237	0	0	0	0	0	0	0.000224065	0.002912839	0	0	0	0	0
10173_B_37	0	0	0	0	0	0	0.005941901	0.002860915	0	0	0	0	0
10175_B_14	0	0	0	0	0	0	0.045708648	0	0	0	0	0	0

Sample ID	Desulfovibrio	Dialister	Dorea	Dysgonomonas	Eggerthella	Elizabethkingia	Enterobacter	Enterococcus	Enterorhabdus	Epulopiscium	Erwinia	Escherichia-Shigella	Eubacterium
10078_A_27	0	0	0	0	0	0	0.350318471	0	0	0	0	0.044585987	0
10078_A_58	0	0	0	0	0	0	0.935749904	0	0	0	0	0.018411968	0
10088_A_28	0	0	0	0	0	0	0.066270178	0.026338148	0	0	0.003398471	0.002548853	0
10088_A_346	0	0	0	0	0.000847458	0	0	0.049152542	0	0	0	0	0
10088_A_62	0	0	0	0	0.005089868	0	0.00683951	0.001431525	0	0	0	0.684110068	0
10088_A_9	0	0	0	0	0	0	0.054787994	0	0	0	0.000952835	0.000476417	0
10090_B_10	0	0	0	0	0	0	0.816742688	0.010195868	0	0	0	0.001073249	0
10090_B_24	0	0	0	0	0	0	0.503330866	0.001665433	0	0	0	0.000185048	0
10090_B_304	0	0	0	0	0.000788126	0	0.008275319	0.006436359	0	0	0	0.007881256	0
10090_B_59	0	0	0	0	0	0	0.135289324	0.06193969	0	0	0	0.000814996	0
10132_A_16	0	0	0	0	0	0	0.141744548	0	0	0	0.001038422	0.001038422	0
10132_A_228	0	0	0	0	0	0	0.770552147	0.002760736	0	0	0	0.093251534	0
10132_A_41	0	0	0	0	0	0	0.112128612	0.162800163	0	0	0.001017501	0.001831502	0
10135_B_10	0	0	0	0	0	0	0.628739694	0.326737338	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0.511227545	0.340194611	0	0	0	0.000187126	0
10135_B_79	0	0	0	0	0.000311915	0	0.006394261	0.027448534	0	0	0	0.000779788	0
10136_A_13	0	0	0	0	0	0	0.195512821	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0.106318348	0.014580802	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0.302249489	0.021676892	0	0	0	0	0
10141_A_27	0	0	0	0	0	0	0.084053961	0.001383604	0	0	0	0.01279834	0
10141_A_292	0	0	0	0	0.003584229	0	0.003225806	0.004121864	0	0.000179211	0	0.000537634	0
10141_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0.070391061	0	0	0	0	0.017877095	0
10146_B_15	0	0	0	0	0	0	0.083410138	0.011059908	0	0	0.000460829	0.017511521	0
10146_B_28	0	0	0	0	0	0	0.195121951	0	0	0	0	0.040650407	0
10146_B_62	0	0	0	0	0	0	0.322656865	0.006300866	0	0	0.000262536	0.018902599	0
10148_A_336	0	0.000279486	0	0	0	0	0	0	0	0	0	0.001676914	0
10148_A_36	0	0	0	0	0	0	0.054158607	0.067698259	0	0	0.001450677	0	0
10148_A_59	0	0	0	0	0	0	0.044885177	0.014613779	0	0	0	0.007306889	0
10153_B_15	0	0	0	0	0	0	0.003322259	0.279069767	0	0	0	0.262458472	0
10153_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0.015655577	0.004892368	0	0	0	0.003913894	0
10153_B_522	0	0	0.011520737	0	0	0	0	0	0	0	0	0	0.002304147
10156_B_14	0	0	0	0	0	0	0.059229002	0.00102119	0	0	0	0.015062548	0
10156_B_30	0	0	0	0	0	0	0.036913767	0.011497731	0	0	0	0.012405446	0
10156_B_4	0	0	0	0	0	0	0.062135922	0	0	0	0	0.011650485	0
10171_A_10	0	0	0	0	0	0	0.476804124	0.046391753	0	0	0	0.164948454	0
10171_A_260	0	0	0	0	0.001503759	0	0.257142857	0.003007519	0	0	0	0.034586466	0
10171_A_70	0	0	0	0	0	0	0.379030332	0.016240745	0	0	0	0.03940769	0
10172_B_17	0	0	0	0	0	0	0.557794468	0	0	0	0	0.000915919	0
10172_B_28	0	0	0	0	0	0	0.168975069	0.000277008	0	0	0	0.003601108	0
10172_B_73	0	0	0	0	0	0	0.200832466	0.013527575	0	0	0	0.028616025	0
10173_B_237	0	0	0	0	0.004033162	0	0.003136903	0.013443872	0	0	0	0.000224065	0
10173_B_37	0	0	0	0	0	0	0.870378521	0.004841549	0	0	0	0.000440141	0
10175_B_14	0	0	0	0	0	0	0.247287077	0.523840842	0	0	0	0.000657678	0

Sample ID	Faecalibacterium	Finegoldia	Flavonifractor	Fusobacterium	Gardnerella	Gemella	Gordonibacter	Granulicatella	Haemophilus	Hafnia-Obesumbacterium	Holdemania	Howardella	Klebsiella
10078_A_27	0	0.356687898	0	0	0	0	0	0	0	0	0	0	0
10078_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_28	0	0.010195412	0	0	0	0	0	0	0.000849618	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0.002146499
10090_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0.001480385
10090_B_304	0.000131354	0	0.026008144	0	0	0	0	0	0.000788126	0	0	0	0
10090_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_228	0	0	0	0	0	0	0	0	0	0	0	0	0.000306748
10132_A_41	0	0.001221001	0	0	0	0	0	0	0	0	0	0	0
10135_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0.002120141
10135_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0.001684132
10135_B_79	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0.001602564
10136_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0.000607533
10136_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_27	0	0.003804912	0	0	0	0	0	0	0	0	0	0	0
10141_A_292	0	0	0.00125448	0.000358423	0	0.000179211	0	0.001075269	0.005555556	0	0	0.000358423	0
10141_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_336	0.182504192	0	0.000279486	0	0	0	0	0	0.006987144	0	0	0	0
10148_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_522	0.032258065	0	0.002304147	0	0	0	0	0.002304147	0	0	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0.003007519	0	0.022556391	0	0	0	0	0	0	0	0	0	0
10171_A_70	0	0	0	0	0	0	0	0	0.001194172	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0.038834951
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0.022437673
10172_B_73	0	0	0	0	0	0	0	0	0	0	0	0	0.000520291
10173_B_237	0	0.000896258	0	0	0	0	0	0	0	0	0	0	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0.012544014
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0.000328839

Sample ID	Kluyvera	Lachnospira	Lactobacillus	Lactococcus	Massilia	Megamonas	Megasphaera	Methylobacterium	Microbacterium	Morganella	Negativicoccus	Neisseria	Nesterenkonia
10078_A_27	0	0	0	0	0	0	0	0	0	0	0.038216561	0	0
10078_A_58	0	0	0.009014193	0	0	0	0	0	0	0	0	0	0
10088_A_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_346	0	0	0.165254237	0	0	0	0	0	0	0	0	0	0
10088_A_62	0	0	0.001113409	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0.002775722	0	0	0	0	0	0	0	0	0	0
10090_B_304	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_59	0	0	0.186634067	0	0	0	0	0	0	0	0	0	0
10132_A_16	0	0	0.000259605	0	0	0	0	0	0	0	0	0	0
10132_A_228	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_41	0	0	0.028286528	0	0	0	0	0	0	0	0.021367521	0	0
10135_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_79	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_292	0	0	0.238351254	0	0	0	0	0	0	0	0	0.001433692	0
10141_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_336	0	0	0	0	0	0.06176635	0	0	0	0	0	0.001117943	0
10148_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0.003322259	0	0	0	0	0
10153_B_3	0	0	0	0	0.006928406	0	0	0.002309469	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_522	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0	0	0	0	0	0.048120301	0.036090226	0	0	0.021052632	0	0	0
10171_A_70	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0	0	0	0	0	0
10173_B_237	0	0	0.044812906	0	0	0	0.000672194	0	0	0	0	0	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0.00022007
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Novosphingobium	Odoribacter	Pantoea	Parabacteroides	Parasutterella	Pediococcus	Pelomonas	Peptococcus	Peptoniphilus	Peptostreptococcus	Planomicrobium	Porphyromonas	Prevotella
10078_A_27	0	0	0	0	0	0	0	0	0.050955414	0.01910828	0	0	0
10078_A_58	0	0	0.009589567	0	0	0	0	0	0	0	0	0	0
10088_A_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0	0	0	0.005084746	0	0
10088_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0.004561309	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0.001295337	0	0	0	0	0	0	0	0	0	0
10090_B_304	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_228	0	0	0.000613497	0	0	0	0	0	0	0	0	0	0
10132_A_41	0.0002035	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_10	0	0	0	0.000706714	0	0	0	0	0	0	0	0	0
10135_B_24	0	0	0.000748503	0.00130988	0	0	0	0	0	0	0	0	0
10135_B_79	0	0	0	0.272145976	0	0	0	0	0.000155958	0	0.000155958	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_27	0	0	0	0	0	0	0	0	0.000691802	0	0	0	0
10141_A_292	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_3	0	0	0	0	0	0	0	0	0.009049774	0.063348416	0	0	0.665158371
10141_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_15	0	0	0.000460829	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0.008130081	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0.003150433	0	0	0	0	0	0.005250722	0	0	0	0
10148_A_336	0	0	0	0.003633315	0	0	0	0	0	0	0	0	0.313583007
10148_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_40	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_522	0	0	0	0	0.002304147	0	0	0	0	0	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0	0	0	0.042105263	0	0	0	0	0	0	0	0	0
10171_A_70	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0	0	0	0	0	0
10173_B_237	0	0	0	0	0	0	0	0	0.000448129	0	0	0	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Propionibacterium	Proteus	Providencia	Pseudobutyrivibrio	Pseudomonas	Pseudonocardia	Ralstonia	Raoultella	Rhizobacter	Rhizobium	Roseburia	Rothia	Ruminococcus
10078_A_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10078_A_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_28	0.005947324	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_62	0.005726101	0	0	0	0	0	0	0	0	0	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_304	0	0	0	0.002495731	0	0	0	0	0	0	0.000656771	0	0
10090_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_228	0	0	0	0	0	0	0	0	0	0	0	0	0
10132_A_41	0.003866504	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_10	0	0	0	0	0.000471143	0.000235571	0	0	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0
10135_B_79	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0.000607533	0	0	0	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0.010224949	0	0	0	0	0
10141_A_27	0.002075406	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_292	0	0	0	0	0	0	0	0	0	0	0	0.000179211	0
10141_A_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_336	0	0	0	0	0	0	0	0	0	0	0	0	0.003633315
10148_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10148_A_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_15	0.156146179	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_3	0	0	0	0	0.050808314	0	0	0	0.004618938	0	0	0	0
10153_B_40	0.04109589	0	0	0	0	0	0	0	0	0	0	0	0
10153_B_522	0	0	0	0.002304147	0	0	0	0	0	0	0.00921659	0	0.02764977
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0	0	0.121804511	0	0	0	0	0	0	0	0	0	0
10171_A_70	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_17	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_73	0.002081165	0	0	0	0	0	0	0	0	0	0	0	0
10173_B_237	0	0	0	0	0	0	0	0	0	0	0	0.000224065	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Salmonella	Sarcina	Selenomonas	Serratia	Sneathia	Sphingomonas	Sporosarcina	Staphylococcus	Stenotrophomonas	Streptococcus	Subdoligranulum	Sutterella	Tatumella
10078_A_27	0	0	0	0	0	0	0	0	0	0.006369427	0	0	0
10078_A_58	0	0	0	0	0	0	0	0	0	0.009205984	0	0	0
10088_A_28	0	0	0	0	0	0	0	0.002548853	0	0	0	0	0
10088_A_346	0	0	0	0	0	0	0	0	0	0.023728814	0	0	0
10088_A_62	0	0	0	0	0	0	0	0.000159058	0	0.003658343	0	0	0
10088_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10090_B_10	0	0	0	0	0	0	0	0.000268312	0	0	0	0	0.001341562
10090_B_24	0	0	0	0	0	0	0	0	0	0	0	0	0.000370096
10090_B_304	0	0	0	0	0	0	0	0	0	0.002233022	0	0	0
10090_B_59	0	0	0	0	0	0	0	0.003259984	0	0.016299919	0	0	0
10132_A_16	0	0	0	0	0	0	0	0.029854621	0	0	0	0	0.000519211
10132_A_228	0	0	0	0	0	0	0	0	0	0.001533742	0	0	0
10132_A_41	0	0	0	0	0	0	0	0.027269027	0	0.001831502	0	0	0
10135_B_10	0	0	0	0	0	0	0	0.000471143	0	0	0	0	0
10135_B_24	0	0	0	0	0	0	0	0.00879491	0	0.005988024	0	0	0.000187126
10135_B_79	0	0	0	0	0	0	0	0.000467873	0	0.0051466	0	0	0
10136_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10136_A_66	0	0	0	0	0	0	0	0.000817996	0	0.030265849	0	0	0
10141_A_27	0	0	0	0	0	0	0	0.05499827	0	0.002421308	0	0	0
10141_A_292	0	0	0	0	0	0	0	0	0	0.004301075	0	0.000537634	0
10141_A_3	0	0	0	0	0.031674208	0	0	0	0	0	0	0.013574661	0
10141_A_6	0	0	0	0	0	0	0	0.00726257	0	0.00726257	0	0	0
10146_B_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10146_B_62	0	0	0	0	0	0	0	0.000787608	0	0	0	0	0.000525072
10148_A_336	0	0	0	0	0	0	0	0	0	0.0039128	0	0.000279486	0
10148_A_36	0	0	0	0	0	0	0	0.250483559	0	0.002417795	0	0	0
10148_A_59	0	0	0	0	0	0	0	0.049060543	0	0.003131524	0	0	0
10153_B_15	0	0	0	0.033222591	0	0	0	0.139534884	0	0.006644518	0	0	0
10153_B_3	0	0	0	0	0	0.009237875	0	0.861431871	0	0	0	0	0
10153_B_40	0	0	0	0.074363992	0	0	0	0	0	0.167318982	0	0	0
10153_B_522	0	0	0	0	0	0	0	0	0	0.368663594	0	0	0
10156_B_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10156_B_30	0	0	0	0	0	0	0	0.001815431	0	0.002420575	0	0	0
10156_B_4	0	0	0	0	0	0	0	0.151456311	0	0.203883495	0	0	0
10171_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10171_A_260	0	0	0	0	0	0	0	0	0	0.001503759	0	0	0
10171_A_70	0	0	0	0	0	0	0	0.001194172	0	0.012897062	0	0	0
10172_B_17	0	0	0	0	0	0	0	0.000732735	0	0	0	0	0.000183184
10172_B_28	0	0	0	0	0	0	0	0	0	0	0	0	0
10172_B_73	0	0	0	0	0	0	0	0.000520291	0	0.054630593	0	0	0
10173_B_237	0	0	0	0	0	0	0	0	0	0.087609231	0	0	0
10173_B_37	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_14	0	0	0	0	0	0	0	0.149950674	0	0	0	0	0

Sample ID	Tessaracoccus	Trabulsiella	Turicibacter	Varibaculum	Variovorax	Veillonella	Weissella	Unclassified
10078_A_27	0	0	0	0	0	0.063694268	0	0.025477707
10078_A_58	0	0	0	0	0	0.000191791	0	0.017453011
10088_A_28	0	0	0	0.000849618	0	0	0	0.538657604
10088_A_346	0	0	0	0	0	0.003389831	0	0.008474576
10088_A_62	0	0	0	0.000477175	0	0.020041355	0	0.043900111
10088_A_9	0	0	0	0	0	0	0	0.380657456
10090_B_10	0	0	0	0	0	0	0	0.163670512
10090_B_24	0	0	0	0	0	0.214100666	0	0.207068838
10090_B_304	0	0	0	0	0	0.017076054	0	0.69985551
10090_B_59	0.004889976	0	0	0	0	0.001629992	0	0.075794621
10132_A_16	0	0	0	0	0	0	0	0.241433022
10132_A_228	0	0	0	0	0	0	0	0.072392638
10132_A_41	0	0	0	0.001628002	0	0.015873016	0	0.211029711
10135_B_10	0	0	0	0	0	0	0	0.040518257
10135_B_24	0	0	0	0	0	0	0	0.116579341
10135_B_79	0	0	0	0	0	0.002027449	0	0.149095446
10136_A_13	0	0	0	0	0	0	0	0.802884615
10136_A_35	0	0	0	0	0	0	0	0.419198056
10136_A_66	0	0	0	0	0	0	0	0.500613497
10141_A_27	0	0	0	0	0	0.030093393	0	0.04185403
10141_A_292	0	0	0	0	0	0.135125448	0	0.167741935
10141_A_3	0	0	0	0	0	0	0	0
10141_A_6	0	0	0	0	0	0	0	0.883240223
10146_B_15	0	0	0	0	0	0	0	0.887096774
10146_B_28	0	0	0	0	0	0	0	0.74796748
10146_B_62	0	0	0	0	0	0.00603833	0	0.625360987
10148_A_336	0	0	0	0	0	0.06372275	0	0.008664058
10148_A_36	0	0	0	0	0	0.004352031	0	0.315280464
10148_A_59	0	0	0	0	0	0.019832985	0	0.603340292
10153_B_15	0	0	0	0	0	0	0	0.106312292
10153_B_3	0	0	0	0	0	0	0	0.023094688
10153_B_40	0	0	0	0	0	0.10665362	0	0.236790607
10153_B_522	0	0	0	0	0	0.023041475	0	0.294930876
10156_B_14	0	0	0	0	0	0	0	0.924687261
10156_B_30	0	0	0	0	0	0	0	0.882299546
10156_B_4	0	0	0	0	0	0	0	0.570873786
10171_A_10	0	0	0	0	0	0	0	0.136597938
10171_A_260	0	0.001503759	0	0	0	0	0	0.054135338
10171_A_70	0	0	0	0	0	0.001671841	0	0.02555529
10172_B_17	0	0	0	0	0	0	0	0.393112292
10172_B_28	0	0	0	0	0	0	0	0.16565097
10172_B_73	0	0	0	0	0	0.28616025	0	0.01925078
10173_B_237	0	0	0	0.000224065	0	0.035850325	0	0.005153484
10173_B_37	0	0.000660211	0	0	0	0.014964789	0	0.086927817
10175_B_14	0	0.005919106	0	0	0	0	0	0.026307136

Sample ID	Acidaminococcus	Acidovorax	Acinetobacter	Actinobacillus	Actinobaculum	Actinomyces	Adlercreutzia	Aeromicrobium	Aeromonas	Akkermansia	Alistipes	Alkanindiges	Anaerococcus
10175_B_261	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_302	0	0	0.042998649	0	0	0.000225124	0	0	0.005402972	0.001575867	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_250	0	0	0	0	0	0.003466205	0	0	0	0	0	0	0
10183_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_262	0	0	0	0	0	0.000234907	0	0	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_302	0	0	0	0	0	0.000552029	0	0	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_45	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_473	0	0.000200521	0	0	0	0	0	0	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_96	0	0	0	0	0	0	0	0	0	0	0	0	0.032571912
10206_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0.00362757
10209_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_281	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_286	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Anaerostipes	Anaerotruncus	Aquabacterium	Atopobium	Bacillus	Bacteroides	Barnesiella	Bifidobacterium	Bilophila	Blautia	Bradyrhizobium	Brevundimonas	Burkholderia
10175_B_261	0	0	0	0	0	0.06522717	0	0.301244564	0	0.000449843	0	0	0
10175_B_59	0	0	0	0	0	0	0	0.000213129	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0.747764416	0	0	0	0	0
10182_A_302	0	0	0	0	0	0.207113913	0	0.020261144	0	0.000900495	0	0	0
10182_A_31	0	0	0	0	0	0	0	0.411580595	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0.256410256	0	0	0	0	0
10183_A_20	0	0	0	0	0	0	0	0.089628297	0	0	0	0	0
10183_A_250	0	0	0	0	0	0	0	0.832659349	0	0	0	0	0
10183_A_36	0	0	0	0	0	0	0	0.833560523	0	0	0	0	0
10183_A_66	0	0	0	0	0	0	0	0.959780622	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_262	0.000704722	0	0	0.000234907	0	0	0	0.303500117	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0.968296139	0	0	0	0	0
10196_A_302	0	0	0	0	0	0	0	0.389456252	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0.005298013	0	0	0	0	0
10202_B_45	0	0	0	0	0	0	0	0.00021725	0	0	0	0	0
10202_B_473	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0.000628141	0	0	0.001256281	0	0
10202_B_96	0	0	0	0	0	0	0	0.557106599	0	0.000423012	0	0	0
10206_A_279	0	0	0	0	0	0	0	0.186321419	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0	0.580904802	0	0	0	0	0
10206_A_61	0	0	0	0	0	0	0	0.251511487	0	0	0	0	0
10209_A_12	0	0	0	0	0	0	0	0.257142857	0	0	0	0	0
10209_A_41	0	0	0	0	0	0	0	0.068391867	0	0	0	0	0
10209_A_68	0	0	0	0	0	0	0	0.451381414	0	0	0	0	0
10215_A_16	0	0	0	0	0	0	0	0.426448171	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0.000860338	0	0.001147118
10215_A_33	0	0	0	0	0	0	0	0.562457221	0	0.000171116	0	0	0
10217_B_281	0	0	0	0	0	0	0	0.936153545	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0	0.000278552	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_15	0	0	0	0	0.000335683	0	0	0.710976838	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0.00456621	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0.766708385	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0.527272727	0	0	0	0.009090909	0
10224_A_18	0	0	0	0	0	0	0	0.381283422	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0	0.259541985	0	0	0	0	0
10224_A_60	0	0	0	0	0.000612933	0	0	0.904076004	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0	0	0	0	0	0	0	0.000374532	0	0	0	0	0
10235_B_286	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Catenibacterium	Cedecea	Cellulomonas	Cellulosilyticum	Chloroplast	Christensenella	Citrobacter	Clostridium	Collinsella	Coprobacillus	Coprococcus	Corynebacterium	Cronobacter
10175_B_261	0	0	0	0	0	0	0	0.246663668	0	0	0	0	0
10175_B_59	0	0	0	0	0	0	0.000852515	0.003623188	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_302	0	0	0	0	0	0	0.282305268	0.000450248	0	0	0.001125619	0	0.002476362
10182_A_31	0	0	0	0	0	0	0.000782473	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0.001079622	0.000269906	0	0	0	0	0
10183_A_20	0	0	0	0	0	0	0	0.00029976	0	0	0	0	0
10183_A_250	0	0	0	0	0	0	0	0.002695937	0	0	0.000385134	0	0
10183_A_36	0	0	0	0	0	0	0	0.001772083	0	0	0	0	0
10183_A_66	0	0	0	0	0	0	0	0.004265692	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0.066920152	0.42851711	0	0	0	0	0
10185_B_262	0	0	0	0.000704722	0	0	0.000939629	0.030772845	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0.027027027	0
10196_A_12	0	0	0	0	0	0	0.000708466	0	0	0	0	0	0
10196_A_302	0	0	0	0	0	0	0.002760144	0.256693348	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0.160706402	0	0	0	0	0	0.000441501
10202_B_45	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_473	0	0	0	0	0	0	0.000200521	0	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0	0	0	0	0.029522613	0
10202_B_96	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_279	0	0	0	0	0	0	0	0.015198424	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0.025811823	0	0	0	0	0	0
10206_A_61	0	0	0	0	0	0	0.111245466	0	0	0	0	0	0
10209_A_12	0	0	0	0	0	0	0.092009685	0	0	0	0	0	0
10209_A_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_68	0	0	0	0	0	0	0	0.000358809	0	0	0	0	0
10215_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0.00152439
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0.069113851	0
10215_A_33	0	0	0	0	0	0	0.001711157	0.000855578	0	0	0	0.000342231	0.000171116
10217_B_281	0	0	0	0	0	0	0	0.007833921	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0.000965251	0	0	0	0	0
10219_A_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0.000500626	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0	0.000534759	0	0	0	0	0	0	0	0	0	0.001069519	0
10224_A_279	0	0	0	0	0	0	0	0.030534351	0	0	0	0	0
10224_A_60	0	0.000919399	0	0	0	0	0	0	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0.065	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0	0	0	0	0	0	0	0.007490637	0	0	0	0	0
10235_B_286	0	0	0	0	0	0	0.000547645	0	0	0	0	0	0

Sample ID	Desulfovibrio	Dialister	Dorea	Dysgonomonas	Eggerthella	Elizabethkingia	Enterobacter	Enterococcus	Enterorhabdus	Epulopiscium	Erwinia	Escherichia-Shigella	Eubacterium
10175_B_261	0.000149948	0	0	0	0.002099265	0	0.001649423	0.020992653	0	0	0	0.039586145	0
10175_B_59	0	0	0	0	0	0	0.857630009	0.002344416	0	0	0	0.062659847	0
10182_A_10	0	0	0	0	0	0	0	0.155103299	0	0	0	0	0
10182_A_302	0	0	0	0	0.000450248	0	0.213192256	0.001800991	0	0	0	0.023863125	0
10182_A_31	0	0	0	0	0	0	0.48200313	0	0	0	0	0.000978091	0
10182_A_66	0	0	0	0	0.003778677	0	0.591093117	0.003238866	0	0	0	0.000269906	0
10183_A_20	0	0	0	0	0	0	0.724220624	0.025179856	0	0	0	0.087230216	0
10183_A_250	0	0	0	0	0.000770268	0	0.046793761	0.003273638	0	0.000192567	0	0.003466205	0
10183_A_36	0	0	0	0	0	0	0.025081788	0.12431843	0	0	0	0.002453653	0
10183_A_66	0	0	0	0	0	0	0.009140768	0.009140768	0	0	0	0.000609385	0
10185_B_10	0	0	0	0	0	0	0.058174905	0	0	0	0	0	0
10185_B_262	0	0	0	0.036880432	0.001409443	0	0.105238431	0.022081278	0	0	0	0.013859525	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0.007793128	0	0	0	0	0	0
10196_A_302	0	0	0	0	0	0	0.030085564	0.073695832	0	0	0	0.002208115	0
10196_A_32	0	0	0	0	0	0	0.762472406	0	0	0	0	0.030022075	0
10202_B_45	0	0	0	0	0	0	0.627851401	0.010862481	0	0	0	0.000434499	0
10202_B_473	0	0	0	0	0	0	0.308802888	0.005614598	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0.684673367	0	0	0	0	0
10202_B_96	0	0	0	0	0	0	0.014805415	0.035956007	0	0	0	0.116751269	0
10206_A_279	0	0	0	0	0	0	0.14354067	0.039684773	0	0	0	0.574162679	0
10206_A_33	0	0	0	0	0	0	0.175964474	0.013877324	0	0	0	0.001942825	0
10206_A_61	0	0	0	0	0	0	0.345828295	0.015719468	0	0	0	0.016928658	0
10209_A_12	0	0	0	0	0	0	0.345439871	0.000322841	0	0	0	0.000161421	0
10209_A_41	0	0	0	0	0	0	0.029574861	0.003080715	0	0	0	0.000616143	0
10209_A_68	0	0	0	0	0	0	0.01722282	0.020810908	0	0	0	0.20739146	0
10215_A_16	0	0	0	0	0	0	0.400914634	0	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0.653283625	0	0	0	0	0
10215_A_33	0	0	0	0	0	0	0.302361396	0	0	0	0	0.001368925	0
10217_B_281	0	0	0	0	0	0	0	0.020955738	0	0	0	0.013317665	0
10217_B_34	0	0	0	0	0	0	0.349303621	0.003342618	0	0	0	0.503064067	0
10217_B_6	0	0	0	0	0	0	0.014614344	0.119350474	0	0	0	0.840866035	0
10217_B_62	0	0	0	0	0	0	0.697876448	0.000965251	0	0	0	0.022200772	0
10219_A_15	0	0	0	0	0	0	0	0.254112118	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0.00913242	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0.156946183	0	0	0	0.032290363	0
10219_A_65	0	0	0	0	0.018181818	0	0	0.390909091	0	0	0	0	0
10224_A_18	0	0	0	0	0	0	0.000534759	0	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0.022900763	0.007633588	0	0	0	0.198473282	0
10224_A_60	0	0	0	0	0	0	0.000919399	0.032178976	0	0	0	0.017775054	0
10227_B_10	0	0	0	0	0	0	0.746949842	0.140985088	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0.58	0.035	0	0	0	0.05	0
10227_B_8	0	0	0	0	0	0	0.808039747	0.073170732	0	0	0	0.001355014	0
10235_B_12	0	0	0	0	0	0	0.006367041	0	0	0	0	0.760674157	0
10235_B_286	0	0	0	0	0	0	0.200438116	0.001916758	0	0	0	0.083515882	0

Sample ID	Faecalibacterium	Finegoldia	Flavonifractor	Fusobacterium	Gardnerella	Gemella	Gordonibacter	Granulicatella	Haemophilus	Hafnia-Obesumbacterium	Holdemania	Howardella	Klebsiella
10175_B_261	0.000449843	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0.002344416
10182_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_302	0.003601981	0	0.003601981	0	0	0.000450248	0	0.000450248	0	0	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0.008802817
10182_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0.012955466
10183_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_250	0	0	0.000192567	0	0	0.000962835	0	0.000577701	0	0	0	0	0
10183_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0.003422053
10185_B_262	0	0	0	0	0	0	0	0	0.001879258	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_302	0	0	0	0.000276014	0	0	0	0	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0.003090508
10202_B_45	0	0	0	0	0	0	0	0	0	0	0	0	0.00021725
10202_B_473	0	0	0	0	0	0	0	0	0	0	0	0	0.003208342
10202_B_8	0	0	0	0	0.000628141	0	0	0	0	0	0	0	0
10202_B_96	0	0.008883249	0	0	0	0	0	0	0	0	0	0	0
10206_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0.002814523
10206_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0.01582015
10206_A_61	0	0.00120919	0	0	0	0	0	0	0	0	0	0	0.00120919
10209_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0.000322841
10209_A_41	0	0	0	0	0	0	0	0	0	0	0	0	0.000308071
10209_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0.000358809
10215_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0.004449008
10217_B_281	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0.027576602
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0.015444015
10219_A_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_31	0	0.000250313	0	0	0	0	0	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0.069588793
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0.105
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0.071364047
10235_B_12	0	0.000749064	0	0	0	0	0	0	0	0	0	0	0
10235_B_286	0	0.000273823	0	0	0	0	0	0	0	0	0	0	0.0449069

Sample ID	Kluyvera	Lachnospira	Lactobacillus	Lactococcus	Massilia	Megamonas	Megasphaera	Methylobacterium	Microbacterium	Morganella	Negativicoccus	Neisseria	Nesterenkonia
10175_B_261	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_302	0	0	0.000225124	0	0	0	0	0	0	0	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_20	0	0	0.000899281	0	0	0	0	0	0	0	0	0	0
10183_A_250	0	0	0.044675525	0	0	0	0	0	0	0	0	0	0
10183_A_36	0	0	0.00368048	0	0	0	0	0	0	0	0	0	0
10183_A_66	0	0	0.004875076	0	0	0	0	0	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_262	0	0	0.083626967	0	0	0	0.000234907	0	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0.000177117	0	0	0	0	0	0	0	0	0	0
10196_A_302	0	0	0.147943693	0	0	0	0	0	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_45	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_473	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_96	0	0	0	0	0	0	0	0	0	0	0.014805415	0	0
10206_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_12	0.006941082	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_41	0.000308071	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_68	0	0	0	0	0.000358809	0	0	0	0	0	0	0	0
10215_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_281	0	0	0.008617313	0	0	0	0	0	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_15	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0	0	0.010160428	0	0	0	0	0	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0	0	0	0	0.000374532	0	0	0	0	0	0.002247191	0	0
10235_B_286	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Novosphingobium	Odoribacter	Pantoea	Parabacteroides	Parasutterella	Pediococcus	Pelomonas	Peptococcus	Peptoniphilus	Peptostreptococcus	Planomicrobium	Porphyromonas	Prevotella
10175_B_261	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_59	0	0	0.001491901	0	0	0	0	0	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0	0	0	0.000308356	0	0
10182_A_302	0	0	0	0	0.002701486	0	0	0	0	0	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_250	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_36	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_262	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_302	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_45	0	0	0	0	0	0	0	0	0.00021725	0	0	0	0
10202_B_473	0	0	0	0	0	0	0	0	0	0	0.000200521	0	0
10202_B_8	0	0	0	0	0	0	0	0	0	0	0.001256281	0	0
10202_B_96	0	0	0	0	0	0	0	0	0.009729272	0	0	0	0
10206_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_61	0	0	0.083434099	0	0	0	0	0	0	0	0	0	0
10209_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_16	0	0	0.001143293	0	0	0	0	0	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_33	0	0	0.000513347	0	0	0	0	0	0	0	0	0	0
10217_B_281	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_34	0	0	0	0	0	0	0	0	0.007520891	0	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_15	0	0	0	0	0	0	0	0	0	0	0.000671366	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_60	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_286	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Propionibacterium	Proteus	Providencia	Pseudobutyrivibrio	Pseudomonas	Pseudonocardia	Ralstonia	Raoultella	Rhizobacter	Rhizobium	Roseburia	Rothia	Ruminococcus
10175_B_261	0	0	0	0	0	0	0	0	0	0	0	0	0
10175_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_302	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_20	0.000899281	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_250	0	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_36	0.004634678	0	0	0	0	0	0	0	0	0	0	0	0
10183_A_66	0.007921999	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_262	0	0	0	0	0	0	0	0	0	0	0	0	0
10185_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_302	0	0	0	0	0	0	0	0	0	0	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_45	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_473	0	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_8	0.001884422	0	0	0	0	0	0	0	0	0	0	0	0
10202_B_96	0.005076142	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_279	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10206_A_61	0.00120919	0	0	0	0	0	0	0	0	0	0	0	0
10209_A_12	0	0	0	0	0	0	0	0.000161421	0	0	0	0	0
10209_A_41	0.001540357	0	0	0	0	0	0	0.007393715	0	0	0	0	0
10209_A_68	0.001435235	0	0	0	0	0	0	0.002511661	0	0	0	0.000358809	0
10215_A_16	0.002286585	0	0	0	0	0	0	0.000381098	0	0	0	0	0
10215_A_288	0	0	0	0	0	0	0	0	0	0	0	0	0
10215_A_33	0.016769336	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_281	0.011163337	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_34	0.001392758	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10217_B_62	0.000965251	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_15	0.001007049	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_274	0	0	0	0	0	0	0	0	0	0	0	0.050228311	0
10219_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_18	0.043850267	0	0	0	0	0	0	0.4	0	0	0	0	0
10224_A_279	0.007633588	0	0	0	0	0	0	0	0	0	0	0	0
10224_A_60	0.003677597	0	0	0	0	0	0	0.024210849	0	0	0	0	0
10227_B_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10227_B_8	0.000451671	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_12	0.000749064	0	0	0	0	0	0	0	0	0	0	0	0
10235_B_286	0.000547645	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Salmonella	Sarcina	Selenomonas	Serratia	Sneathia	Sphingomonas	Sporosarcina	Staphylococcus	Stenotrophomonas	Streptococcus	Subdoligranulum	Sutterella	Tatumella
10175_B_261	0	0	0	0	0	0	0	0	0	0.004198531	0.000899685	0	0
10175_B_59	0	0	0	0	0	0	0	0.000426257	0	0.010869565	0	0	0.000852515
10182_A_10	0	0	0	0	0	0	0	0.093432007	0	0.002466852	0	0	0
10182_A_302	0	0	0	0	0	0	0	0	0	0.025889239	0	0	0
10182_A_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10182_A_66	0	0	0	0	0	0	0	0.000269906	0	0	0	0	0
10183_A_20	0	0	0	0	0	0	0	0	0	0.00029976	0	0	0
10183_A_250	0	0	0	0	0	0	0	0	0	0.034854612	0	0	0
10183_A_36	0	0	0	0	0	0	0	0	0	0.000817884	0	0	0
10183_A_66	0	0	0	0	0	0	0	0.001218769	0	0.000609385	0	0	0
10185_B_10	0	0	0	0	0	0	0	0.424714829	0	0	0	0	0
10185_B_262	0	0	0	0	0	0	0	0	0	0.017852948	0	0	0
10185_B_4	0	0	0	0	0	0	0	0.972972973	0	0	0	0	0
10196_A_12	0	0	0	0	0	0	0	0.013637974	0	0.00053135	0	0	0
10196_A_302	0	0	0	0	0	0	0	0	0	0.006900359	0	0	0
10196_A_32	0	0	0	0	0	0	0	0	0	0.002649007	0	0	0.000441501
10202_B_45	0	0	0	0	0	0	0	0.008038236	0	0	0	0	0.00021725
10202_B_473	0	0	0	0	0	0	0	0.003208342	0	0	0	0	0
10202_B_8	0	0	0	0	0	0	0	0.277638191	0	0	0	0	0
10202_B_96	0	0	0	0	0	0	0	0.013113367	0	0.115482234	0	0	0
10206_A_279	0	0	0	0	0	0	0	0	0	0.001407261	0	0	0
10206_A_33	0	0	0	0	0	0	0	0	0	0.000277546	0	0	0
10206_A_61	0.015719468	0	0	0	0	0	0	0.00120919	0	0	0	0	0
10209_A_12	0	0	0	0	0	0	0	0	0	0.006133979	0	0	0.000322841
10209_A_41	0	0	0	0	0	0	0	0.000924214	0	0.020332717	0	0	0
10209_A_68	0	0	0	0	0	0	0	0.003946896	0	0.051309652	0	0	0
10215_A_16	0	0	0	0	0	0	0	0.003429878	0	0.002667683	0	0	0
10215_A_288	0	0	0	0	0	0	0	0.269572699	0	0	0	0	0
10215_A_33	0	0	0	0	0	0.000171116	0	0.001882272	0	0.006160164	0	0	0
10217_B_281	0	0	0	0	0	0	0	0.000195848	0	0.001566784	0	0	0
10217_B_34	0	0	0	0	0	0	0	0.001114206	0	0.069916435	0	0	0
10217_B_6	0	0	0	0	0	0	0	0.008930988	0	0	0	0	0
10217_B_62	0	0	0	0	0	0	0	0.001930502	0	0.141891892	0	0	0
10219_A_15	0	0	0	0	0	0	0	0.026854649	0	0.001342732	0	0	0
10219_A_274	0	0	0	0	0	0	0	0.936073059	0	0	0	0	0
10219_A_31	0	0	0	0	0	0	0	0	0	0.033291615	0	0	0
10219_A_65	0	0	0	0	0	0	0	0	0	0.009090909	0	0	0
10224_A_18	0	0	0	0	0	0	0	0.028877005	0	0.000534759	0	0	0
10224_A_279	0	0	0	0	0	0	0	0	0	0.106870229	0	0	0
10224_A_60	0	0	0	0	0	0	0	0.002758198	0	0.002451732	0	0	0
10227_B_10	0	0	0	0	0	0	0	0.001807501	0	0	0	0	0
10227_B_31	0	0	0	0	0	0	0	0.005	0	0	0	0	0
10227_B_8	0	0	0	0	0	0	0	0.00767841	0	0	0	0	0
10235_B_12	0	0	0	0	0	0	0	0.004868914	0	0.001498127	0	0	0
10235_B_286	0	0	0.000273823	0	0	0	0	0.001642935	0	0.050383352	0	0	0

Sample ID	Tessaracoccus	Trabulsiella	Turicibacter	Varibaculum	Variovorax	Veillonella	Weissella	Unclassified
10175_B_261	0	0	0	0.000149948	0	0.001649423	0	0.314589894
10175_B_59	0	0.002131287	0	0	0	0.002344416	0	0.052216539
10182_A_10	0	0	0	0	0	0	0	0.000925069
10182_A_302	0	0	0	0	0	0.000225124	0	0.158712292
10182_A_31	0	0	0	0	0	0	0	0.095852895
10182_A_66	0	0	0	0	0	0	0	0.130634278
10183_A_20	0	0	0	0	0	0.018884892	0	0.052458034
10183_A_250	0	0	0	0	0	0.003466205	0	0.021567495
10183_A_36	0	0	0	0	0	0.000272628	0	0.003407852
10183_A_66	0	0	0	0	0	0	0	0.002437538
10185_B_10	0	0.000380228	0	0	0	0	0	0.017870722
10185_B_262	0	0	0	0	0	0.055203195	0	0.324641767
10185_B_4	0	0	0	0	0	0	0	0
10196_A_12	0	0.007616011	0	0	0	0	0	0.001239816
10196_A_302	0	0	0	0	0	0.083908363	0	0.005520287
10196_A_32	0	0.001324503	0	0	0	0	0	0.033554084
10202_B_45	0	0.00021725	0	0	0.000651749	0.333260917	0	0.017814469
10202_B_473	0	0	0	0	0	0.278524163	0	0.400040104
10202_B_8	0	0	0	0	0	0	0	0.002512563
10202_B_96	0	0	0	0	0	0.073604061	0	0.001692047
10206_A_279	0	0	0	0	0	0.000562905	0	0.036307346
10206_A_33	0	0.000832639	0	0	0	0	0	0.184568415
10206_A_61	0	0	0	0	0	0.07617896	0	0.07496977
10209_A_12	0	0.001614205	0	0	0	0	0	0.289426957
10209_A_41	0	0	0	0	0	0	0	0.867529267
10209_A_68	0	0	0	0	0	0	0	0.242554718
10215_A_16	0	0	0	0	0	0	0	0.161204268
10215_A_288	0	0	0	0	0.000286779	0	0	0.005735589
10215_A_33	0	0	0	0	0	0.038501027	0	0.06211499
10217_B_281	0	0	0	0	0	0	0	0.000195848
10217_B_34	0	0	0	0	0	0.009470752	0	0.027019499
10217_B_6	0	0	0	0	0	0	0	0.01623816
10217_B_62	0	0	0	0	0	0.002895753	0	0.114864865
10219_A_15	0	0	0	0	0	0	0	0.004699564
10219_A_274	0	0	0	0	0	0	0	0
10219_A_31	0	0	0	0	0	0.007509387	0	0.002503129
10219_A_65	0	0	0	0	0	0	0	0.045454545
10224_A_18	0	0	0	0	0	0	0	0.13315508
10224_A_279	0	0	0	0	0	0.34351145	0	0.022900763
10224_A_60	0	0	0	0	0	0.000306466	0	0.010113393
10227_B_10	0	0	0	0	0	0	0	0.040668775
10227_B_31	0	0	0	0	0	0.07	0	0.09
10227_B_8	0	0	0	0	0	0	0	0.037940379
10235_B_12	0	0	0	0	0	0	0	0.214606742
10235_B_286	0	0	0	0	0	0.000821468	0	0.614731654

Sample ID	Acidaminococcus	Acidovorax	Acinetobacter	Actinobacillus	Actinobaculum	Actinomyces	Adlercreutzia	Aeromicrobium	Aeromonas	Akkermansia	Alistipes	Alkanindiges	Anaerococcus
10236_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0.000583771
10236_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0.000278668	0	0	0	0	0	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0.000984252
10251_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_486	0.000880127	0	0	0	0	0.000704101	0	0	0	0.018130611	0	0	0.000176025
10256_A_67	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_297	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0	0	0	0	0	0	0	0	0	0.029828851	0	0	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0.028920308	0	0	0	0	0	0.005784062	0.000642674	0	0	0
10276_A_7	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0	0	0	0.000509165	0	0	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Anaerostipes	Anaerotruncus	Aquabacterium	Atopobium	Bacillus	Bacteroides	Barnesiella	Bifidobacterium	Bilophila	Blautia	Bradyrhizobium	Brevundimonas	Burkholderia
10236_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0.158690801	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0.006421483	0	0	0	0	0
10236_A_9	0	0	0	0	0	0	0	0.809057528	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0.214119214	0	0	0	0	0
10237_A_8	0	0	0	0	0	0	0	0.937994723	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0.255555556	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0.187681159	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0.262295082	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0.513505664	0	0	0	0	0
10239_B_9	0	0	0	0	0	0	0	0.583748563	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0.003298969	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0.988992615	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0.090375056	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0.000311236	0	0	0	0	0	0	0.455960162	0	0	0	0	0
10249_B_38	0	0	0	0	0	0	0	0.278156997	0	0	0	0	0
10251_A_13	0	0	0	0	0	0	0	0.682689382	0	0	0	0	0
10251_A_33	0	0	0	0	0	0	0	0.355807087	0	0.00246063	0	0	0
10251_A_81	0.00871731	0	0	0.00124533	0.00747198	0	0	0.02615193	0	0.00124533	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0	0.608570682	0	0	0	0	0
10252_A_30	0	0	0	0	0	0	0	0.865966682	0	0	0	0	0
10252_A_8	0	0	0	0	0.002906977	0	0	0.831395349	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0.459330144	0	0	0	0	0
10253_B_29	0	0	0	0	0.000181951	0	0	0.858078603	0	0	0	0	0
10253_B_6	0	0	0	0	0.003982477	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0.729751403	0	0	0	0	0
10256_A_12	0	0	0	0	0	0	0	0.643099788	0	0	0	0	0.000212314
10256_A_38	0	0	0	0	0	0	0	0.687192118	0	0	0	0	0
10256_A_486	0.00228833	0	0	0	0	0.006864989	0.006336913	0.686146805	0.000176025	0.000176025	0	0	0
10256_A_67	0	0	0	0	0	0	0	0.839636914	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0.000698568	0	0.00174642	0	0	0.000698568
10261_B_29	0	0	0	0	0.000518672	0	0	0	0	0	0	0	0
10261_B_297	0	0	0	0	0	0.081395349	0	0.397674419	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0.314806055	0	0	0	0	0
10264_A_295	0	0	0	0	0	0.192665037	0	0.047432763	0	0	0	0	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0.000746269
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0	0	0	0.054627249	0	0.01285347	0	0	0	0	0
10276_A_7	0	0	0	0	0	0	0	0	0	0	0.001117006	0	0
10301_B_301	0	0	0	0	0	0	0	0.608452138	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0	0.004485109	0	0	0	0	0

Sample ID	Catenibacterium	Cedecea	Cellulomonas	Cellulosilyticum	Chloroplast	Christensenella	Citrobacter	Clostridium	Collinsella	Coprobacillus	Coprococcus	Corynebacterium	Cronobacter
10236_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0.001167542	0	0	0	0	0.001167542
10236_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0.001665002
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0.006666667	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0.000724638	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0	0	0	0.000311236	0	0	0.002178649	0.01774043	0	0	0.002178649	0	0
10249_B_38	0	0	0	0	0	0	0.000853242	0	0	0	0	0	0
10251_A_13	0	0	0	0	0	0	0	0.003346517	0	0	0	0	0
10251_A_33	0	0	0	0	0	0	0	0.020669291	0	0	0	0	0
10251_A_81	0	0	0	0	0	0	0.03611457	0.00373599	0	0	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0	0.002351712	0	0	0	0	0
10252_A_30	0	0	0	0	0	0	0	0.00029329	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0.014534884	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0.001510954	0.262402418	0	0	0	0	0
10253_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0	0.001231527	0	0	0	0	0
10256_A_486	0	0	0	0	0	0.019890864	0.000352051	0	0	0.000176025	0	0	0
10256_A_67	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_297	0	0	0	0	0	0	0	0	0.051162791	0	0	0	0
10261_B_58	0	0	0	0	0	0	0.000525348	0.405568689	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0	0	0	0	0	0	0	0.000488998	0.002444988	0	0	0	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0.00261194	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0.00931677	0
10276_A_255	0	0	0	0	0	0	0.10218509	0.001285347	0	0	0.001285347	0	0
10276_A_7	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0	0	0	0	0.000254582	0.000509165	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0.000179404	0	0	0	0	0	0

Sample ID	Desulfovibrio	Dialister	Dorea	Dysgonomonas	Eggerthella	Elizabethkingia	Enterobacter	Enterococcus	Enterorhabdus	Epulopiscium	Erwinia	Escherichia-Shigella	Eubacterium
10236_A_248	0	0	0	0	0	0	0.176837537	0	0	0	0	0.000542446	0
10236_A_39	0	0	0	0	0	0	0.006446814	0.029258616	0	0	0	0.795189685	0
10236_A_81	0	0	0	0	0	0	0.061879743	0.008756567	0	0	0	0.848219498	0
10236_A_9	0	0	0	0	0	0	0.000734394	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0.081918082	0.010989011	0	0	0	0.553446553	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0.051111111	0.028888889	0	0	0	0.515555556	0
10239_B_300	0	0	0	0	0.000724638	0	0.266666667	0.00326087	0	0	0	0.001086957	0
10239_B_32	0	0	0	0	0	0	0.666666667	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0.085100203	0.015974441	0	0	0	0.000580889	0
10239_B_9	0	0	0	0	0	0	0.173246455	0	0	0	0	0.015714833	0
10240_A_22	0	0	0	0	0	0	0.92	0	0	0	0	0.006597938	0
10240_A_300	0	0	0	0	0	0	0.004040686	0	0	0	0	0.000278668	0
10240_A_65	0	0	0	0	0	0	0.553547221	0.001807501	0	0	0	0.000451875	0
10249_B_12	0	0	0	0	0	0	0	0.051995163	0	0	0	0	0
10249_B_260	0	0	0	0.013694367	0.000622471	0	0.000311236	0.003734827	0	0	0	0.100529101	0
10249_B_38	0	0	0	0	0	0	0.499573379	0.003412969	0	0	0	0.053327645	0
10251_A_13	0	0	0	0	0	0	0	0.0422878	0	0	0	0	0
10251_A_33	0	0	0	0	0	0	0	0.018208661	0	0	0	0	0
10251_A_81	0	0	0	0	0	0	0	0.00498132	0	0	0	0.00124533	0
10252_A_233	0	0	0	0	0	0	0.879298121	0	0	0	0	0.006912442	0
10252_A_29	0	0	0	0	0	0	0.001567808	0.035536974	0	0	0	0.087535929	0
10252_A_30	0	0	0	0	0	0	0.000527921	0.024519005	0	0	0	0.097548099	0
10252_A_8	0	0	0	0	0	0	0	0.063953488	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0.004029212	0.047343238	0	0	0	0.018383279	0
10253_B_29	0	0	0	0	0	0	0.010189229	0.010007278	0	0	0	0.111717613	0
10253_B_6	0	0	0	0	0	0	0	0.957387495	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0.183640738	0.034081796	0	0	0	0.002405774	0
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0.000821018	0	0	0	0	0	0
10256_A_486	0	0.002464355	0	0	0.000528076	0	0.000176025	0.003168456	0	0	0	0.008801267	0
10256_A_67	0	0	0	0	0	0	0.002118003	0.102269289	0	0	0	0.000302572	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0.080134855	0.554979253	0	0	0	0.019450207	0
10261_B_297	0	0	0	0	0	0	0.006976744	0.004651163	0	0	0	0.034883721	0
10261_B_58	0	0	0	0	0	0	0.063829787	0	0	0	0	0.041239821	0
10264_A_20	0	0	0	0	0	0	0.00614948	0	0	0	0	0.672658467	0
10264_A_295	0	0	0	0	0	0	0.073838631	0.00195599	0	0	0	0.26405868	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0	0	0.000642674	0	0.039203085	0	0	0	0	0.316838046	0
10276_A_7	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0	0	0.006364562	0	0.034368635	0.023167006	0	0	0	0.226323829	0
10301_B_31	0	0	0	0	0	0	0.586831719	0	0	0	0	0.22335845	0

Sample ID	Faecalibacterium	Finegoldia	Flavonifractor	Fusobacterium	Gardnerella	Gemella	Gordonibacter	Granulicatella	Haemophilus	Hafnia-Obesumbacterium	Holdemania	Howardella	Klebsiella
10236_A_248	0	0	0	0	0	0	0	0	0.00081367	0	0	0	0.053430974
10236_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_9	0	0	0	0	0	0	0	0	0.001713586	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0.10955711	0	0	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0.065942029
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0.016393443
10239_B_65	0	0	0	0	0	0	0	0	0	0	0	0	0.002033111
10239_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0.000383289
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0.02185567
10240_A_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0.022141889
10249_B_12	0	0	0	0	0	0	0	0	0.246070133	0	0	0	0
10249_B_260	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_38	0	0	0	0	0	0	0	0	0.027730375	0	0	0	0.007679181
10251_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_33	0	0.003937008	0	0.000492126	0	0	0	0	0	0	0	0	0
10251_A_81	0.01867995	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0.012406948
10252_A_29	0	0.001567808	0	0	0	0	0	0	0	0	0	0	0
10252_A_30	0	0.000117316	0	0	0	0	0	0	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0	0.000251826	0	0	0	0.000503651
10253_B_29	0	0.000363901	0	0	0	0	0	0	0	0	0	0	0.000545852
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0.002405774
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_486	0	0.005456786	0.001408203	0	0	0	0.000704101	0	0	0	0	0	0
10256_A_67	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0.000259336	0	0	0	0	0	0	0	0	0	0	0.00129668
10261_B_297	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0.000525348
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0.001466993	0	0.009779951	0	0	0	0	0	0	0	0	0	0.002444988
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0.002570694	0	0.004498715	0	0	0	0	0	0	0	0	0	0.000642674
10276_A_7	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0	0	0	0	0	0	0	0	0	0	0.000254582
10301_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0.003946896

Sample ID	Kluyvera	Lachnospira	Lactobacillus	Lactococcus	Massilia	Megamonas	Megasphaera	Methylobacterium	Microbacterium	Morganella	Negativicoccus	Neisseria	Nesterenkonia
10236_A_248	0	0	0.000271223	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0.003996004	0	0	0	0	0	0	0	0	0	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0.02	0	0	0	0	0	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0	0	0.068783069	0	0	0	0.000311236	0	0	0	0	0	0
10249_B_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_33	0	0	0.022145669	0	0	0	0	0	0	0	0	0	0
10251_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0.003396917	0	0	0	0	0	0
10252_A_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_233	0	0	0.117098968	0	0	0	0	0	0	0	0	0	0
10253_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_12	0	0	0.002972399	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0.003694581	0	0	0	0	0	0	0	0	0	0
10256_A_486	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_67	0	0	0.042965204	0	0	0	0	0	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_297	0	0	0.260465116	0	0	0	0.002325581	0	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0	0	0	0	0	0.022493888	0.032762836	0	0	0.030317848	0	0	0
10264_A_66	0	0	0	0	0	0	0	0.001865672	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_7	0	0	0.068695895	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0.058299389	0	0	0	0	0	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Novosphingobium	Odoribacter	Pantoea	Parabacteroides	Parasutterella	Pediococcus	Pelomonas	Peptococcus	Peptoniphilus	Peptostreptococcus	Planomicrobium	Porphyromonas	Prevotella
10236_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0	0	0	0.000247954	0	0
10236_A_81	0	0	0.000583771	0	0	0	0	0	0	0	0	0	0
10236_A_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_9	0	0	0.001149866	0	0	0	0	0	0	0	0	0	0
10240_A_22	0	0	0.001237113	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_65	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_38	0	0	0.000426621	0	0	0	0	0	0	0	0	0	0
10251_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_33	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_81	0	0	0	0	0	0	0	0.00249066	0	0	0	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0	0	0	0	0.000261301	0	0
10252_A_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_486	0	0	0	0	0	0	0	0	0.002992431	0	0	0	0
10256_A_67	0	0	0	0	0	0	0	0	0.003025719	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_297	0	0	0	0.039534884	0	0	0	0	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0	0	0	0.052811736	0	0	0	0	0	0	0	0	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0	0	0.000642674	0	0	0	0	0	0	0	0
10276_A_7	0	0	0	0	0	0	0	0	0.001117006	0.005864284	0	0	0
10301_B_301	0	0	0.000509165	0	0	0	0	0	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Propionibacterium	Proteus	Providencia	Pseudobutyrivibrio	Pseudomonas	Pseudonocardia	Ralstonia	Raoultella	Rhizobacter	Rhizobium	Roseburia	Rothia	Ruminococcus
10236_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_81	0	0	0	0	0	0	0	0	0	0	0	0	0
10236_A_9	0.001958384	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_248	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_82	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0.004356666	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_9	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_65	0.007230005	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_260	0	0	0	0	0	0	0	0	0	0	0	0	0
10249_B_38	0	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_13	0.001216915	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_33	0.000492126	0	0	0	0	0	0	0	0	0	0	0	0
10251_A_81	0	0	0	0	0	0	0	0	0	0	0.00249066	0	0.00373599
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0.112882153	0	0	0	0	0	0	0	0	0	0	0.005226026	0
10252_A_30	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_8	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_29	0.000181951	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10253_B_9	0	0	0	0	0	0	0	0.000801925	0	0	0	0	0
10256_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_38	0.000410509	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_486	0	0	0	0	0	0	0	0	0	0	0	0	0
10256_A_67	0.005143722	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_11	0.15228781	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_29	0.000259336	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_297	0	0	0	0	0	0	0	0	0	0	0	0	0
10261_B_58	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_295	0	0	0.136430318	0	0	0	0	0	0	0	0	0	0
10264_A_66	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_255	0	0	0	0	0	0	0	0	0	0	0	0	0
10276_A_7	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_301	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Salmonella	Sarcina	Selenomonas	Serratia	Sneathia	Sphingomonas	Sporosarcina	Staphylococcus	Stenotrophomonas	Streptococcus	Subdoligranulum	Sutterella	Tatumella
10236_A_248	0	0	0.000542446	0	0	0	0	0.003525902	0	0.003254679	0	0	0.001356116
10236_A_39	0	0	0	0	0	0	0	0	0	0.001239772	0	0	0
10236_A_81	0.016929364	0	0	0.001751313	0	0	0	0	0	0.020431991	0	0	0
10236_A_9	0	0	0	0.012239902	0	0	0	0.166462668	0	0.00122399	0	0	0
10237_A_248	0	0	0	0	0	0	0	0.99970033	0	0	0	0	0
10237_A_39	0	0	0	0	0	0	0	0	0	0	0	0	0
10237_A_8	0	0	0	0.003392386	0	0	0	0.055220505	0	0.003015454	0	0	0
10237_A_82	0	0	0	0	0	0	0	0	0	0.026666667	0	0	0
10239_B_300	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_32	0	0	0	0	0	0	0	0	0	0	0	0	0
10239_B_65	0	0	0	0	0	0	0	0	0	0.000290444	0	0	0
10239_B_9	0.000766577	0	0	0	0	0	0	0.194327328	0	0	0	0	0
10240_A_22	0	0	0	0	0	0	0	0	0	0	0	0	0
10240_A_300	0	0	0	0	0	0	0	0.005434025	0	0.000418002	0	0	0.000139334
10240_A_65	0	0	0	0	0	0	0	0.000451875	0	0.001355626	0	0	0
10249_B_12	0	0	0	0	0	0	0	0.469165659	0	0	0	0	0
10249_B_260	0	0	0	0	0	0.000311236	0	0	0	0.01276066	0	0	0
10249_B_38	0.002133106	0	0	0	0	0	0	0.000426621	0	0.070819113	0	0	0
10251_A_13	0	0	0	0	0	0	0	0.22695467	0	0.008822635	0	0	0
10251_A_33	0	0	0	0	0	0	0	0.006889764	0	0.542814961	0	0	0
10251_A_81	0	0	0	0	0	0	0.00996264	0	0	0.727272727	0.00498132	0	0
10252_A_233	0	0	0	0	0	0	0	0	0	0	0	0	0
10252_A_29	0	0	0	0	0	0	0	0.000783904	0	0.111836948	0	0	0
10252_A_30	0	0	0	0	0	0	0	0.000117316	0	0.003108869	0	0	0
10252_A_8	0	0	0	0	0	0	0	0.029069767	0	0	0	0	0
10253_B_233	0	0	0	0	0	0	0	0	0	0.004029212	0	0	0
10253_B_29	0	0	0	0	0	0	0	0.002729258	0	0.002365357	0	0	0
10253_B_6	0	0	0	0	0	0	0	0.026284349	0	0	0	0	0
10253_B_9	0.001202887	0	0	0	0	0	0	0.006014435	0	0.000400962	0	0	0
10256_A_12	0	0	0	0	0	0	0	0.005732484	0	0	0	0	0
10256_A_38	0	0	0	0	0	0	0	0.002873563	0	0.000821018	0	0	0
10256_A_486	0	0	0	0	0	0	0	0	0	0.004400634	0.000528076	0	0
10256_A_67	0	0	0	0	0	0	0	0.001210287	0	0	0	0	0
10261_B_11	0	0	0	0	0	0	0	0.076493189	0	0.768075445	0	0	0
10261_B_29	0	0	0	0	0	0	0	0.18879668	0	0	0	0	0
10261_B_297	0	0	0	0	0	0	0	0	0	0.062790698	0	0.002325581	0
10261_B_58	0	0	0	0	0	0	0	0.469661151	0	0	0	0	0
10264_A_20	0	0	0	0	0	0	0	0	0	0.000236518	0	0	0
10264_A_295	0	0	0	0	0	0	0	0	0	0	0	0	0
10264_A_66	0	0	0	0	0	0	0	0.994776119	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0.990444338	0	0	0	0	0
10276_A_255	0	0	0	0	0	0	0	0	0	0.014138817	0	0	0
10276_A_7	0	0	0	0	0	0	0	0	0	0.922926557	0	0	0
10301_B_301	0.000509165	0	0	0	0	0	0	0	0	0.000509165	0	0	0
10301_B_31	0	0	0	0	0	0	0	0	0	0.002511661	0	0	0.000179404

Sample ID	Tessaracoccus	Trabulsiella	Turicibacter	Varibaculum	Variovorax	Veillonella	Weissella	Unclassified
10236_A_248	0	0	0	0	0	0.305668565	0	0.453756442
10236_A_39	0	0	0	0	0	0.00396727	0	0.004959088
10236_A_81	0	0	0	0	0	0	0	0.032107414
10236_A_9	0	0	0	0	0	0	0	0.006609547
10237_A_248	0	0	0	0	0	0	0	0.00029967
10237_A_39	0	0	0	0	0	0.004662005	0	0.01964702
10237_A_8	0	0	0	0	0	0	0	0.000376932
10237_A_82	0	0	0	0	0	0.013333333	0	0.082222222
10239_B_300	0	0	0	0	0	0	0	0.473913043
10239_B_32	0	0	0	0	0	0	0	0.054644809
10239_B_65	0	0	0	0	0	0.375544583	0	0.002613999
10239_B_9	0	0	0	0	0	0	0	0.030663089
10240_A_22	0	0	0	0	0	0	0	0.047010309
10240_A_300	0	0	0	0	0	0	0	0.000418002
10240_A_65	0	0	0	0	0	0.267962042	0	0.054676909
10249_B_12	0	0	0	0	0	0.23216445	0	0.000604595
10249_B_260	0	0	0	0	0	0.041705571	0	0.278555867
10249_B_38	0	0	0	0	0	0.006825939	0	0.048634812
10251_A_13	0	0	0	0	0	0	0	0.034682081
10251_A_33	0	0	0	0	0	0.017716535	0	0.00738189
10251_A_81	0	0	0	0	0	0	0	0.139476961
10252_A_233	0	0	0	0	0	0	0	0.101382488
10252_A_29	0	0	0	0.019858897	0	0.006532532	0	0.00209041
10252_A_30	0	0	0	0	0	0.00363679	0	0.004164711
10252_A_8	0	0	0	0	0	0	0	0.058139535
10253_B_233	0	0	0	0	0	0.074540418	0	0.010576681
10253_B_29	0	0	0	0	0	0.001455604	0	0.002183406
10253_B_6	0	0	0	0	0	0	0	0.012345679
10253_B_9	0	0	0	0	0	0	0	0.039294306
10256_A_12	0	0	0	0	0	0.347558386	0	0.000424628
10256_A_38	0	0.001231527	0	0	0	0.281198686	0	0.020525452
10256_A_486	0	0	0	0.000176025	0	0	0	0.226896673
10256_A_67	0	0	0	0.000907716	0	0	0	0.002420575
10261_B_11	0	0	0	0	0	0	0	0
10261_B_29	0	0	0	0	0	0	0	0.154304979
10261_B_297	0	0	0	0	0	0.037209302	0	0.018604651
10261_B_58	0	0	0	0	0	0	0	0.018649856
10264_A_20	0	0	0	0	0	0	0	0.00614948
10264_A_295	0	0	0	0	0	0	0	0.098777506
10264_A_66	0	0	0	0	0	0	0	0
10276_A_12	0	0	0	0	0	0	0	0.000238892
10276_A_255	0	0	0	0	0	0.001285347	0	0.411953728
10276_A_7	0	0	0	0	0	0	0	0.000279252
10301_B_301	0	0	0	0	0	0.000763747	0	0.039205703
10301_B_31	0	0	0	0	0	0	0	0.178507356

Sample ID	Acidaminococcus	Acidovorax	Acinetobacter	Actinobacillus	Actinobaculum	Actinomyces	Adlercreutzia	Aeromicrobium	Aeromonas	Akkermansia	Alistipes	Alkanindiges	Anaerococcus
10301_B_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_230	0	0	0	0	0.000347584	0.005213764	0	0	0	0	0	0	0.013903372
10315_B_3	0	0	0	0	0	0	0	0	0	0.008663366	0	0	0
10315_B_61	0	0	0.165178571	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0	0.040398126	0	0	0	0	0	0	0
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0	0	0.119098504	0	0	0.001360016	0	0	0	0.000582864	0.001748591	0	0.000194288
10325_B_1	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_13	0	0	0	0	0	0.001503114	0	0	0	0	0	0	0
10325_B_4	0	0	0	0	0	0.004504505	0	0	0	0	0	0	0.000750751
10326_A_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10326_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0	0	0	0.001569859
10334_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_252	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_64	0	0	0	0	0	0	0	0	0	0	0	0	0.000480538
10338_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0.00338696	0	0	0	0	0	0	0
10338_A_80	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0	0	0	0	0.743461451	0	0	0
10345_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0.00019861
10345_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0.001797843
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0.002520357	0	0	0	0	0	0	0
10357_A_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Anaerostipes	Anaerotruncus	Aquabacterium	Atopobium	Bacillus	Bacteroides	Barnesiella	Bifidobacterium	Bilophila	Blautia	Bradyrhizobium	Brevundimonas	Burkholderia
10301_B_41	0	0	0	0	0	0	0	0.844602978	0	0	0	0	0
10301_B_68	0	0	0	0	0	0	0	0.649359561	0	0	0	0	0
10315_B_230	0	0	0	0	0	0	0	0.356621481	0	0	0	0	0
10315_B_3	0.002475248	0	0	0	0	0.032178218	0	0.508663366	0	0.017326733	0	0	0
10315_B_61	0	0	0	0	0	0	0	0.080357143	0	0	0	0	0
10318_A_263	0	0	0	0.00019516	0	0	0	0.757611241	0	0	0	0	0.00078064
10318_A_68	0	0	0	0	0	0	0	0.175141243	0	0	0	0	0
10319_A_256	0.005440062	0	0	0	0.000194288	0.00466291	0	0.080629493	0	0.000194288	0	0	0
10325_B_1	0	0	0	0	0	0	0	0.00807502	0	0	0	0	0
10325_B_13	0	0	0	0	0	0	0	0.910886837	0	0	0	0	0
10325_B_4	0	0	0	0	0	0	0	0.811561562	0	0	0	0	0
10326_A_11	0.000668003	0	0	0	0	0	0	0.498663995	0	0	0	0	0
10326_A_6	0	0	0	0	0	0	0	0.074484256	0	0	0	0	0
10333_A_16	0	0	0	0	0	0	0	0.861805445	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0.368916797	0	0	0	0	0
10334_A_10	0	0	0	0	0	0	0	0.885876295	0	0	0	0	0
10334_A_252	0	0.001808318	0	0	0	0	0	0.004520796	0	0.002712477	0	0	0
10334_A_4	0	0	0	0	0.001299039	0	0	0	0	0	0	0	0.000259808
10337_B_31	0	0	0	0	0	0	0	0.000450349	0	0	0	0	0
10337_B_64	0	0	0	0	0	0	0	0	0	0.000720807	0	0	0
10338_A_13	0	0	0	0	0	0	0	0.756183746	0	0	0	0	0
10338_A_294	0	0	0	0	0	0.00084674	0	0.289585097	0	0.00169348	0	0	0
10338_A_80	0	0	0	0	0	0	0	0.019433883	0	0	0	0	0
10343_B_16	0	0	0	0	0.000658472	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0.002501706	0	0.241528315	0	0	0	0	0
10345_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_21	0	0	0	0	0	0	0	0.887637758	0	0.000479157	0	0	0
10349_A_35	0	0	0	0	0	0	0	0.99270073	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0.002039776	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0.506234414	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0.876696394	0	0	0	0	0
10357_A_34	0	0	0	0	0.000265534	0	0	0.483536909	0	0	0	0	0
10357_A_62	0	0	0	0.014804469	0	0	0	0.897765363	0	0	0	0	0

Sample ID	Catenibacterium	Cedecea	Cellulomonas	Cellulosilyticum	Chloroplast	Christensenella	Citrobacter	Clostridium	Collinsella	Coprobacillus	Coprococcus	Corynebacterium	Cronobacter
10301_B_41	0	0	0	0	0	0	0	0.000775434	0	0	0	0	0
10301_B_68	0	0	0	0	0	0	0.000228728	0.000457457	0	0	0	0	0
10315_B_230	0	0	0	0	0	0	0.010079944	0.118178658	0	0	0	0.000347584	0
10315_B_3	0	0	0	0	0	0	0	0	0	0	0.004950495	0	0
10315_B_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0	0	0.00058548	0.011709602	0	0	0	0	0
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0	0	0	0	0	0	0.067417913	0	0	0	0	0	0.000777152
10325_B_1	0	0	0	0	0	0	0	0	0	0	0	0.015889555	0
10325_B_13	0	0	0	0	0	0	0.000214731	0.001073653	0	0	0	0	0
10325_B_4	0	0	0	0	0	0	0	0.000750751	0	0	0	0	0
10326_A_11	0	0	0	0	0	0	0	0.009686039	0	0	0	0	0
10326_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_252	0	0	0	0	0	0	0.012658228	0.005424955	0	0	0	0	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0	0	0	0	0	0	0.90114839	0	0	0	0	0	0.000225175
10337_B_64	0	0	0	0	0	0	0.33421432	0.001922153	0	0	0	0	0
10338_A_13	0	0	0	0	0	0	0.010600707	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0.00169348	0	0	0	0	0
10338_A_80	0	0	0	0	0	0	0	0.000422476	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0.000377644	0	0	0	0	0	0
10343_B_62	0	0	0	0	0	0	0.001118068	0.002236136	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0	0	0.000909711	0.003866272	0	0	0	0
10345_B_34	0	0	0	0	0	0	0.090764647	0	0	0	0	0	0
10345_B_59	0	0	0	0	0	0	0.013184179	0	0	0	0	0	0
10349_A_21	0	0	0	0	0	0	0	0.000958313	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Desulfovibrio	Dialister	Dorea	Dysgonomonas	Eggerthella	Elizabethkingia	Enterobacter	Enterococcus	Enterorhabdus	Epulopiscium	Erwinia	Escherichia-Shigella	Eubacterium
10301_B_41	0	0	0	0	0	0	0.000155087	0.125930521	0	0	0	0.017369727	0
10301_B_68	0	0	0	0	0	0	0.077538884	0.021500457	0	0	0	0.004345837	0
10315_B_230	0	0.005561349	0	0	0.001390337	0	0.038234272	0.03163017	0	0	0	0.242961418	0
10315_B_3	0	0	0	0	0.002475248	0	0	0.065594059	0	0	0	0.002475248	0
10315_B_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0.00117096	0	0.003317721	0.033567525	0	0.00156128	0	0.020686963	0
10318_A_68	0	0	0	0	0	0	0.084745763	0.005649718	0	0	0	0.65819209	0
10319_A_256	0	0	0.001165728	0	0.000388576	0.000388576	0.008354381	0.02914319	0	0	0	0.373032835	0.000194288
10325_B_1	0	0	0	0	0	0	0.000260485	0.558218286	0	0	0	0	0
10325_B_13	0	0	0	0	0.000214731	0	0	0.002362036	0	0	0	0.022117243	0
10325_B_4	0	0	0	0	0.00975976	0	0	0.009009009	0	0	0	0.000750751	0
10326_A_11	0	0	0	0	0.006012024	0	0	0.001002004	0	0.000668003	0	0.004008016	0
10326_A_6	0	0	0	0	0	0	0.039739414	0.000217155	0	0	0	0.004777416	0
10333_A_16	0	0	0	0	0	0	0.05015125	0	0	0	0	0.002069734	0
10333_A_64	0	0	0	0	0.003139717	0	0	0.59811617	0	0	0	0	0
10334_A_10	0	0	0	0	0	0	0.041647657	0	0	0	0	0.00299811	0
10334_A_252	0	0	0	0	0.003616637	0	0.152802893	0.009041591	0	0	0	0.393309222	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0.000259808	0
10337_B_31	0	0	0	0	0	0	0.001801396	0.047286647	0	0	0	0.000225175	0
10337_B_64	0	0	0	0	0	0	0.273906776	0.022585296	0	0	0	0.000480538	0
10338_A_13	0	0	0	0	0	0	0.141342756	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0	0	0	0	0.00508044	0
10338_A_80	0	0	0	0	0	0	0.91339248	0.002746092	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0.002414399	0.685250219	0	0	0	0.303555751	0
10343_B_27	0	0	0	0	0	0	0.405589124	0	0	0	0	0.327416918	0
10343_B_62	0	0	0	0	0	0	0.064177102	0.879472272	0	0	0	0.017218247	0
10345_B_13	0	0	0	0	0	0	0.049004882	0	0	0	0	0	0
10345_B_283	0	0	0	0	0.000227428	0	0	0.0040937	0	0	0	0.000682283	0
10345_B_34	0	0	0	0	0	0	0.500496524	0	0	0	0	0.001588878	0
10345_B_59	0	0	0	0	0.00019976	0	0.141630044	0.281262485	0	0	0	0.100079904	0
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0.00729927	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0.048954615	0.410504844	0	0	0	0.499745028	0
10357_A_14	0	0	0	0	0	0	0.007065669	0	0	0	0	0.467996675	0
10357_A_256	0	0	0	0	0	0	0.000193874	0.000193874	0	0	0	0.009112059	0
10357_A_34	0	0	0	0	0.000796601	0	0.003186405	0.003983006	0	0	0	0.486192246	0
10357_A_62	0	0	0	0	0.000558659	0	0.000139665	0.001536313	0	0	0	0.061592179	0

Sample ID	Faecalibacterium	Finegoldia	Flavonifractor	Fusobacterium	Gardnerella	Gemella	Gordonibacter	Granulicatella	Haemophilus	Hafnia-Obesumbacterium	Holdemania	Howardella	Klebsiella
10301_B_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_68	0	0.000228728	0	0	0	0	0	0	0	0	0	0	0.002058554
10315_B_230	0	0.00243309	0	0.000347584	0	0	0	0	0	0	0	0	0
10315_B_3	0	0	0.008663366	0	0	0	0	0.001237624	0	0	0	0	0
10315_B_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0	0.00058548	0	0	0	0	0	0	0.0009758
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0.002914319	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_1	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_13	0	0	0	0	0	0.000644192	0	0.001288383	0	0	0	0	0
10325_B_4	0	0.000750751	0	0	0	0	0	0	0	0	0	0	0
10326_A_11	0	0	0.000668003	0	0	0	0	0.001002004	0.002672011	0	0	0	0
10326_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0.000868621
10333_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_10	0	0	0	0	0	6.51763E-05	0	0	0	0	0	0	0
10334_A_252	0	0	0.006329114	0	0	0	0	0	0	0	0	0	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_64	0	0	0	0	0	0	0	0	0	0.001201346	0	0	0.000240269
10338_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_294	0	0.00084674	0	0	0	0.00084674	0	0.00254022	0	0	0	0	0
10338_A_80	0	0	0	0	0	0	0	0	0	0	0	0	0.005703422
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0.000377644
10343_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0.000397219
10345_B_59	0	0.22932481	0	0	0	0	0	0	0	0	0	0	0.00019976
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0.005609383	0	0	0	0.001019888
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0.004265219	0.013958899	0	0	0	0
10357_A_34	0	0.000265534	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0	0	0	0	0	0	0.003351955	0	0	0	0

Sample ID	Kluyvera	Lachnospira	Lactobacillus	Lactococcus	Massilia	Megamonas	Megasphaera	Methylobacterium	Microbacterium	Morganella	Negativicoccus	Neisseria	Nesterenkonia
10301_B_41	0	0	0.00294665	0	0	0	0	0	0	0	0	0	0
10301_B_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_230	0	0	0.028849496	0	0	0	0.003128259	0	0	0	0.001390337	0	0
10315_B_3	0	0	0.003712871	0	0	0	0	0	0	0	0	0	0
10315_B_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0.00019516	0	0	0	0	0	0	0	0	0
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_1	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_13	0	0	0.024694009	0	0	0	0	0	0	0	0	0	0
10325_B_4	0	0	0.035285285	0	0	0	0	0	0	0	0	0	0
10326_A_11	0	0	0.245490982	0	0	0	0	0	0	0	0	0.000334001	0
10326_A_6	0.000217155	0	0.000217155	0	0	0	0	0	0	0	0	0	0
10333_A_16	0	0	0.000636841	0	0	0	0	0	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_10	0	0	0.000325882	0	0	0	0	0	0	0	0	0	0
10334_A_252	0	0	0	0	0	0	0	0	0.025316456	0	0	0	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0.001576222	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_64	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_80	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_59	0	0	0	0	0	0	0	0	0	0	0.001598082	0	0
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0.002793296	0	0	0	0	0	0	0	0	0	0

Sample ID	Novosphingobium	Odoribacter	Pantoea	Parabacteroides	Parasutterella	Pediococcus	Pelomonas	Peptococcus	Peptoniphilus	Peptostreptococcus	Planomicrobium	Porphyromonas	Prevotella
10301_B_41	0	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_230	0	0	0.001737921	0	0	0	0	0	0.006604101	0	0	0	0.059784498
10315_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_61	0	0	0	0	0	0.008928571	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0	0	0.00019516	0	0	0	0	0	0
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0	0	0.001748591	0	0	0	0	0	0.000388576	0	0	0	0
10325_B_1	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10326_A_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10326_A_6	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_252	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_64	0	0	0.000240269	0	0	0	0	0	0	0	0	0	0
10338_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_80	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0.000658472	0	0
10343_B_27	0	0	0.016238671	0	0	0	0	0	0	0	0	0	0
10343_B_62	0	0	0.002012522	0	0	0	0	0	0	0	0.000447227	0	0
10345_B_13	0	0	0.000751033	0	0	0	0	0	0	0	0	0	0.000187758
10345_B_283	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_34	0	0	0.000595829	0	0	0	0	0	0	0	0	0	0
10345_B_59	0	0	0	0	0	0	0	0	0	0	0.000399521	0	0
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample ID	Propionibacterium	Proteus	Providencia	Pseudobutyrivibrio	Pseudomonas	Pseudonocardia	Ralstonia	Raoultella	Rhizobacter	Rhizobium	Roseburia	Rothia	Ruminococcus
10301_B_41	0.004497519	0	0	0	0	0	0	0	0	0	0	0	0
10301_B_68	0.033394328	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_230	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_3	0	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_61	0	0	0	0	0	0	0	0	0	0	0	0	0
10318_A_263	0	0	0	0	0	0	0	0	0	0	0	0.00156128	0
10318_A_68	0	0	0	0	0	0	0	0	0	0	0	0	0
10319_A_256	0	0	0	0	0	0	0	0.000194288	0	0	0	0	0.064115018
10325_B_1	0.000260485	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10325_B_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10326_A_11	0	0	0	0	0	0	0	0	0	0	0	0	0
10326_A_6	0.002823018	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10333_A_64	0.001569859	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_10	0	0	0	0	0	0	0	0	0	0	0	0	0
10334_A_252	0	0	0	0	0	0	0	0	0	0	0	0	0.009041591
10334_A_4	0	0	0	0	0	0	0	0	0	0	0	0	0
10337_B_31	0.000450349	0	0	0	0	0	0	0.002026571	0	0	0	0	0
10337_B_64	0	0	0	0	0	0	0	0.000961076	0	0	0	0	0
10338_A_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0	0	0	0	0	0
10338_A_80	0.000211238	0	0	0	0	0	0	0	0	0	0	0.000211238	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_62	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_283	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_34	0	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_59	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_21	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0	0	0	0	0.000775494	0
10357_A_34	0.000265534	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_62	0	0	0	0	0	0	0	0	0	0.000139665	0	0.000698324	0

Sample ID	Salmonella	Sarcina	Selenomonas	Serratia	Sneathia	Sphingomonas	Sporosarcina	Staphylococcus	Stenotrophomonas	Streptococcus	Subdoligranulum	Sutterella	Tatumella
10301_B_41	0	0	0	0	0	0	0	0	0	0.000620347	0	0	0
10301_B_68	0	0	0	0	0	0	0	0	0	0.001829826	0	0	0
10315_B_230	0.00729927	0	0	0	0	0	0	0	0	0	0	0	0
10315_B_3	0	0	0	0	0	0	0	0	0	0.006188119	0	0	0
10315_B_61	0	0	0	0	0	0	0	0.71875	0	0.022321429	0	0	0
10318_A_263	0	0	0	0	0	0	0	0	0	0.031811085	0	0	0
10318_A_68	0	0	0	0	0	0	0	0	0	0.070621469	0	0	0
10319_A_256	0.007965805	0	0	0	0	0	0	0	0.152127453	0.005051486	0	0	0
10325_B_1	0	0	0	0	0	0	0	0.409481636	0	0	0	0	0
10325_B_13	0	0	0	0	0	0	0	0	0	0.020399399	0	0	0
10325_B_4	0	0	0	0	0	0	0	0	0	0.089339339	0	0	0
10326_A_11	0	0	0	0	0	0	0	0	0	0.001336005	0	0	0
10326_A_6	0	0	0	0	0	0	0	0.0082519	0	0.008469055	0	0	0
10333_A_16	0	0	0	0	0	0	0	0.001273683	0	0.009393409	0	0	0
10333_A_64	0	0	0	0	0	0	0	0	0	0.017268446	0	0	0
10334_A_10	0	0	0	0	0	0	0	0.013230789	0	0.00527928	0	0	0
10334_A_252	0.04159132	0	0	0	0	0	0	0	0	0.002712477	0	0	0
10334_A_4	0	0	0	0	0	0	0	0.998181346	0	0	0	0	0
10337_B_31	0.000900698	0	0	0	0	0	0	0	0	0.001801396	0	0	0
10337_B_64	0.066554541	0	0	0	0	0	0	0.01609803	0	0.043969246	0	0	0
10338_A_13	0.010600707	0	0	0	0	0	0	0.06360424	0	0	0	0	0
10338_A_294	0	0	0	0	0	0	0	0	0	0.66977138	0	0	0
10338_A_80	0.001267427	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_16	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_27	0	0	0	0	0	0	0	0	0	0	0	0	0
10343_B_62	0.000447227	0	0	0	0	0	0	0	0	0	0	0	0
10345_B_13	0.7827638	0	0	0	0	0	0	0	0	0	0	0	0.000187758
10345_B_283	0	0	0	0	0	0	0	0	0	0.00204685	0	0	0
10345_B_34	0.084806356	0	0	0	0	0	0	0.001986097	0	0	0	0	0
10345_B_59	0.024570515	0	0	0	0	0	0	0	0	0.000599281	0	0	0
10349_A_21	0	0	0	0	0	0	0	0.002395783	0	0.108528989	0	0	0
10349_A_35	0	0	0	0	0	0	0	0	0	0	0	0	0
10349_A_72	0.002039776	0	0	0	0	0	0	0	0	0.001529832	0	0	0
10357_A_14	0	0	0	0	0	0	0	0	0	0	0	0	0
10357_A_256	0	0	0	0	0	0	0	0	0	0.092089957	0	0	0
10357_A_34	0	0	0	0	0	0	0	0.007966012	0	0.001858736	0	0	0
10357_A_62	0	0	0	0	0	0	0	0.000837989	0	0.01424581	0	0	0

Sample ID	Tessaracoccus	Trabulsiella	Turicibacter	Varibaculum	Variovorax	Veillonella	Weissella	Unclassified
10301_B_41	0	0	0	0	0	0.000155087	0	0.00294665
10301_B_68	0	0	0	0	0	0.015096066	0	0.193961574
10315_B_230	0	0	0	0	0	0.053180396	0	0.010775113
10315_B_3	0	0	0	0	0	0	0	0.33539604
10315_B_61	0	0	0	0	0	0	0	0.004464286
10318_A_263	0	0	0	0.00039032	0	0.00058548	0	0.092115535
10318_A_68	0	0	0	0	0	0	0	0.005649718
10319_A_256	0	0	0.000194288	0	0	0	0	0.070332232
10325_B_1	0	0	0	0	0	0	0	0.007814535
10325_B_13	0	0	0	0	0	0.002147305	0	0.01245437
10325_B_4	0	0	0	0	0	0.034534535	0	0.003003003
10326_A_11	0	0	0	0	0	0.110554442	0	0.117234469
10326_A_6	0	0	0	0	0	0.000651466	0	0.859283388
10333_A_16	0	0	0	0	0	0	0	0.074669639
10333_A_64	0	0	0	0	0	0	0	0.009419152
10334_A_10	0	0	0	0	0	0	0	0.05057681
10334_A_252	0	0	0	0	0	0.002712477	0	0.326401447
10334_A_4	0	0	0	0	0	0	0	0
10337_B_31	0	0	0	0	0	0	0	0.042107633
10337_B_64	0	0	0	0	0	0.024747717	0	0.211677078
10338_A_13	0	0	0	0	0	0	0	0.017667845
10338_A_294	0	0	0	0	0	0.00254022	0	0.021168501
10338_A_80	0	0	0	0	0	0.000633714	0	0.055978031
10343_B_16	0	0	0	0	0	0	0	0.007462687
10343_B_27	0	0	0	0	0	0.179380665	0	0.070619335
10343_B_62	0	0	0	0	0	0.018783542	0	0.014087657
10345_B_13	0	0	0	0	0	0.001126549	0	0.16597822
10345_B_283	0	0	0	0	0	0	0	0.000682283
10345_B_34	0	0	0	0	0	0.275868918	0	0.043296922
10345_B_59	0	0	0	0	0	0.047742709	0	0.157411107
10349_A_21	0	0	0	0	0	0	0	0
10349_A_35	0	0	0	0	0	0	0	0
10349_A_72	0	0	0	0	0	0	0	0.028556859
10357_A_14	0	0	0	0	0	0	0	0.018703242
10357_A_256	0	0	0	0	0	0	0	0.000193874
10357_A_34	0	0	0	0	0	0	0	0.011683484
10357_A_62	0	0	0	0	0	0.00027933	0	0.001256983

Appendix C

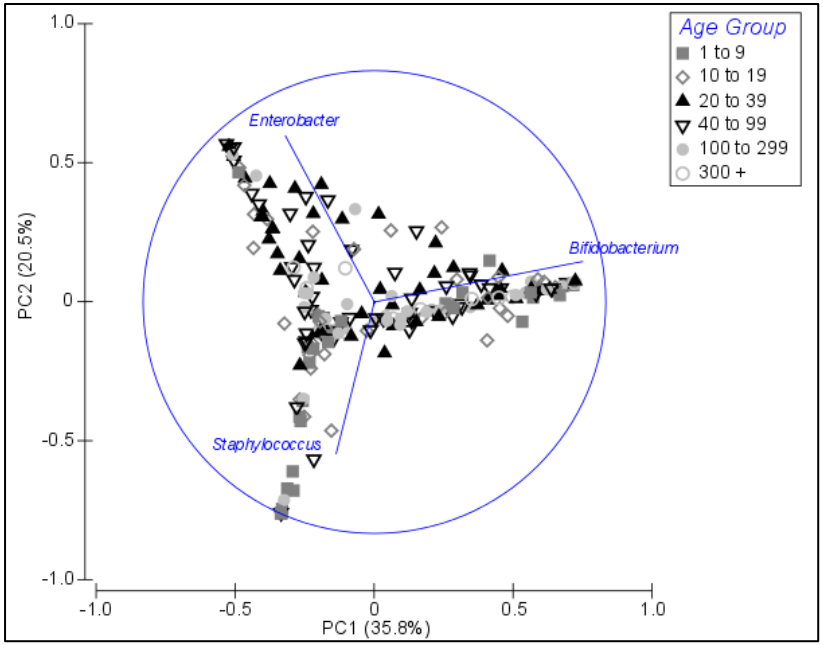


Figure C1 - Principal Component Analysis of Genus Abundance Data

The first and second PCA components show separation of samples into three main groups, one group with a high abundance of *Enterobacter*, another group with a high abundance of *Bifidobacterium* and a final group with a high abundance of the combined genus *Staphylococcus*. As can be seen, the younger samples cluster primarily in the *Staphylococcus* group. The first two components describe over 50% of the data.

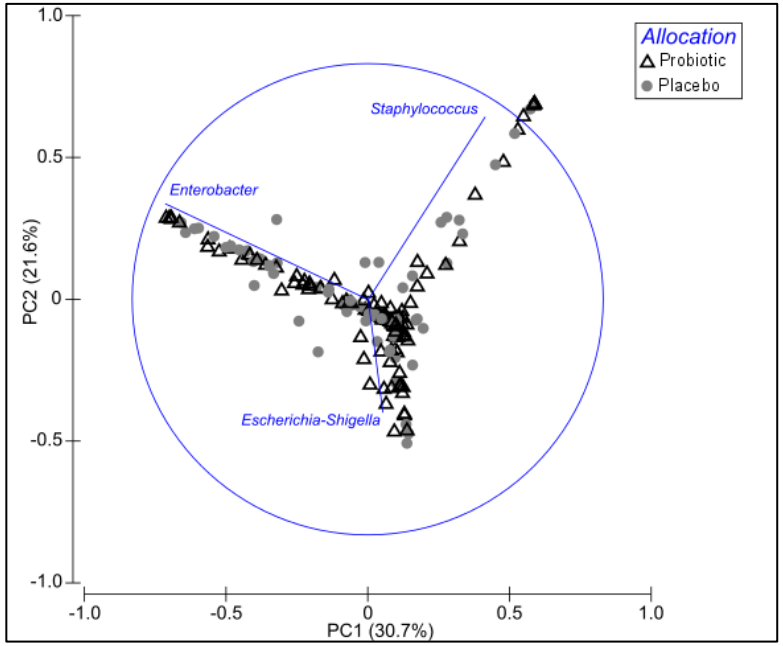


Figure C2 - Principal Component Analysis of Genus Abundance Data post removal of *Bifidobacterium*

The first and second PCA components show separation of samples into three main groups, one group with a high abundance of *Enterobacter*, another group with a high abundance of *Staphylococcus* and a final group with a high abundance of the combined genus *Escherichia-Shigella*. After removal of *Bifidobacterium* there is no clear separation of samples from the two allocation groups. The first two components describe over 50% of the data.

Appendix D

Table D1 – Top Five Most Abundant Genera Across the Different Age Groups Across All Samples

1 to 9	10 to 19	20 to 39	40 to 99	100 to 299	300 +
Staphylococcus (0.29)	*Bifidobacterium* (0.30)	*Bifidobacterium* (0.30)	*Bifidobacterium* (0.28)	*Bifidobacterium* (0.27)	*Bifidobacterium* (0.31)
Bifidobacterium (0.25)	*Enterobacter* (0.17)	*Enterobacter* (0.26)	*Enterobacter* (0.20)	*Escherichia-Shigella* (0.11)	*Lactobacillus* (0.07)
Enterococcus (0.11)	*Staphylococcus* (0.10)	*Escherichia-Shigella* (0.12)	*Escherichia-Shigella* (0.10)	*Enterobacter* (0.11)	*Enterobacter* (0.06)
Enterobacter (0.05)	*Enterococcus* (0.08)	*Enterococcus* (0.04)	*Enterococcus* (0.07)	*Staphylococcus* (0.07)	*Bacteroides* (0.06)
Streptococcus (0.03)	*Escherichia-Shigella* (0.06)	*Veillonella* (0.03)	*Staphylococcus* (0.07)	*Streptococcus* (0.05)	*Escherichia-Shigella* (0.04)

NB: Average abundance as a proportion in parentheses

Table D2 - Top Five Most Abundant Genera Across the Different Age Groups in the Placebo Group

1 to 9	10 to 19	20 to 39	40 to 99	100 to 299	300 +
Staphylococcus (0.28)	*Enterobacter* (0.24)	*Enterobacter* (0.33)	*Bifidobacterium* (0.21)	*Bifidobacterium* (0.36)	*Bifidobacterium* (0.17)
Bifidobacterium (0.19)	*Enterococcus* (0.14)	*Escherichia-Shigella* (0.13)	*Enterobacter* (0.19)	*Escherichia-Shigella* (0.07)	*Lactobacillus* (0.10)
Enterococcus (0.18)	*Staphylococcus* (0.09)	*Bifidobacterium* (0.11)	*Enterococcus* (0.08)	*Clostridium* (0.06)	*Enterobacter* (0.09)
Enterobacter (0.09)	*Escherichia-Shigella* (0.09)	*Enterococcus* (0.06)	*Veillonella* (0.06)	*Akkermansia* (0.06)	*Bacteroides* (0.07)
Escherichia-Shigella (0.06)	*Bifidobacterium* (0.06)	*Citrobacter* (0.05)	*Staphylococcus* (0.06)	*Lactobacillus* (0.06)	*Streptococcus* (0.06)

NB: Average abundance as a proportion in parentheses

Table D3 – Top Five Most Abundant Genera Across the Different Age Groups in the Probiotic Group

1 to 9	10 to 19	20 to 39	40 to 99	100 to 299	300 +
Staphylococcus (0.31)	*Bifidobacterium* (0.50)	*Bifidobacterium* (0.43)	*Bifidobacterium* (0.33)	*Bifidobacterium* (0.22)	*Bifidobacterium* (0.44)
Bifidobacterium (0.30)	*Enterobacter* (0.11)	*Enterobacter* (0.21)	*Enterobacter* (0.21)	*Enterobacter* (0.14)	*Prevotella* (0.04)
Streptococcus (0.08)	*Staphylococcus* (0.10)	*Escherichia-Shigella* (0.11)	*Escherichia-Shigella* (0.14)	*Escherichia-Shigella* (0.12)	*Lactobacillus* (0.04)
Prevotella (0.05)	*Streptococcus* (0.04)	*Streptococcus* (0.03)	Enterococcus (0.06)	*Staphylococcus* (0.10)	*Bacteroides* (0.04)
Enterococcus (0.05)	*Escherichia-Shigella* (0.04)	*Veillonella* (0.03)	*Streptococcus* (0.05)	*Streptococcus* (0.06)	*Citrobacter* (0.04)

NB: Average abundance as a proportion in parentheses

Appendix E

Table E1 - Results of Post Hoc Testing of Genera Differences Between Age Groups

Genera (ANOVA p-value corrected)	Pairwise comparison	Adjusted p-value	95% CI
Staphylococcus (<0.001)	10 to 19-1 to 9	0.004	[-0.3469 , -0.0437]
	100 to 299-1 to 9	<0.001	[-0.3822 , -0.0771]
	20 to 39-1 to 9	<0.001	[-0.423 , -0.1379]
	300 +-1 to 9	<0.001	[-0.4892 , -0.0952]
	40 to 99-1 to 9	<0.001	[-0.3891 , -0.1079]
Anaerostipes (0.023)	300 +-1 to 9	<0.001	[0.0006 , 0.0033]
	300 +-10 to 19	<0.001	[0.0008 , 0.0033]
	300 +-100 to 299	<0.001	[0.0006 , 0.0032]
	300 +-20 to 39	<0.001	[0.0008 , 0.0033]
	40 to 99-300 +	<0.001	[-0.0031 , -0.0007]
Bacteroides (0.003)	300 +-1 to 9	<0.001	[0.0165 , 0.0916]
	300 +-10 to 19	<0.001	[0.0191 , 0.0914]
	300 +-20 to 39	<0.001	[0.0204 , 0.09]
	40 to 99-300 +	<0.001	[-0.0897 , -0.0207]
Barnesiella (0.005)	300 +-1 to 9	<0.001	[0.0004 , 0.0016]
	300 +-10 to 19	<0.001	[0.0004 , 0.0016]
	300 +-100 to 299	<0.001	[0.0003 , 0.0015]
	300 +-20 to 39	<0.001	[0.0004 , 0.0016]
	40 to 99-300 +	<0.001	[-0.0016 , -0.0004]
Parasutterella (0.003)	300 +-1 to 9	<0.001	[0.0001 , 0.0006]
	300 +-10 to 19	<0.001	[0.0001 , 0.0006]
	300 +-100 to 299	<0.001	[0.0001 , 0.0005]
	300 +-20 to 39	<0.001	[0.0002 , 0.0006]
	40 to 99-300 +	<0.001	[-0.0006 , -0.0002]
Faecalibacterium (0.003)	300 +-1 to 9	<0.001	[0.0075 , 0.0307]
	300 +-10 to 19	<0.001	[0.0079 , 0.0303]
	300 +-100 to 299	<0.001	[0.0076 , 0.03]
	300 +-20 to 39	<0.001	[0.0083 , 0.0298]
	40 to 99-300 +	<0.001	[-0.0294 , -0.0081]

NB: 10 to 19-1 to 9 means the pairwise comparison was made between age groups ’10 to 19 Days’ and ‘1 to 9 Days’.

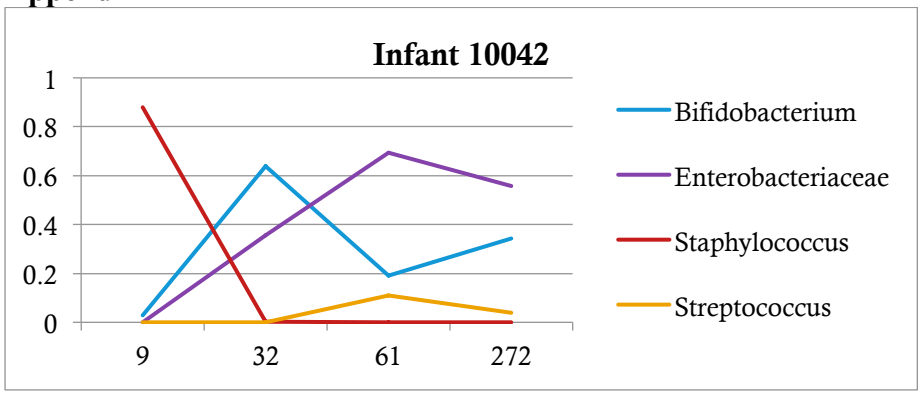
Table E2 – Results of Post Hoc Testing of Differences in Beta Diversity Between Age Groups

Pairwise comparison	Adjusted p-value	95% CI
10 to 19-100 to 299	0.049*	[0.0001 , 0.1262]
20 to 39-100 to 299	0.009*	[-0.1294 , -0.0119]
1 to 9-100 to 299	<0.001*	[0.0374 , 0.1721]
300-100 to 299	0.967	[-0.0606 , 0.1074]
40 to 99-100 to 299	<0.001*	[-0.168 , -0.0523]
20 to 39-10 to 19	<0.001*	[-0.1921 , -0.0755]
1 to 9-10 to 19	0.476	[-0.0253 , 0.1085]
300-10 to 19	0.746	[-0.1234 , 0.0439]
40 to 99-10 to 19	<0.001*	[-0.2307 , -0.116]

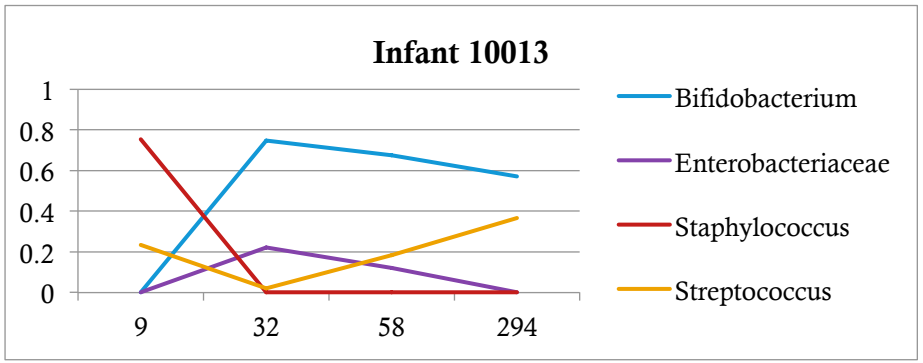
Pairwise comparison	Adjusted p-value	95% CI
1 to 9-20 to 39	<0.001*	[0.1125 , 0.2384]
300-20 to 39	0.012*	[0.0136 , 0.1745]
40 to 99-20 to 39	0.261	[-0.0922 , 0.0131]
300-1 to 9	0.081	[-0.1683 , 0.0056]
40 to 99-1 to 9	<0.001*	[-0.277 , -0.1529]
40 to 99-300	<0.001*	[-0.2134 , -0.0538]

* statistically significant difference

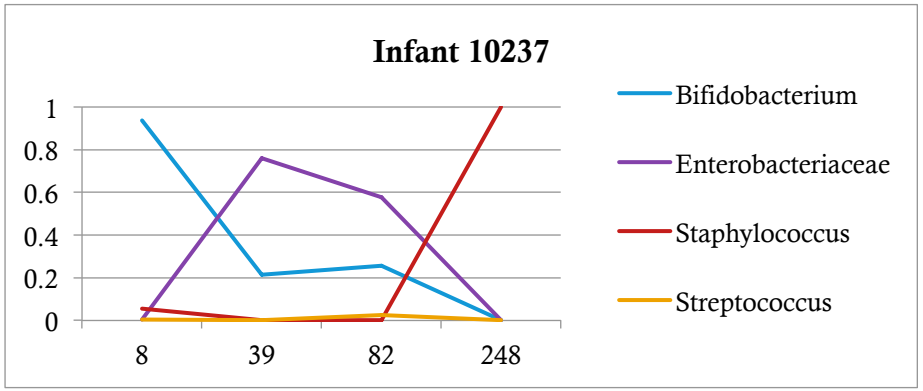
Appendix F



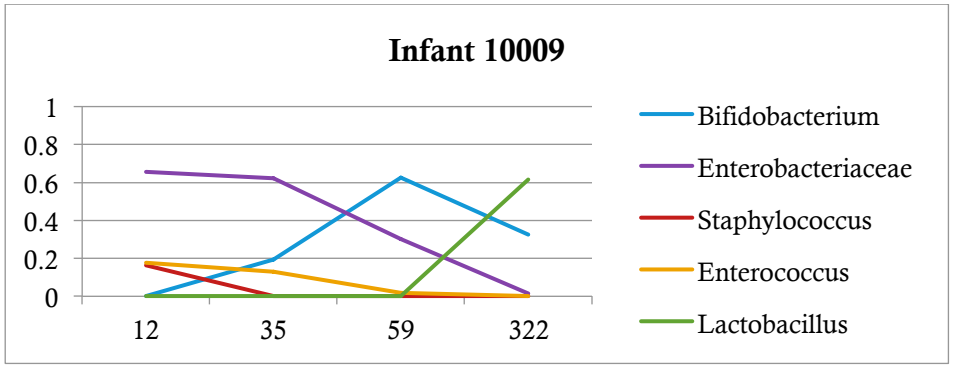
F1



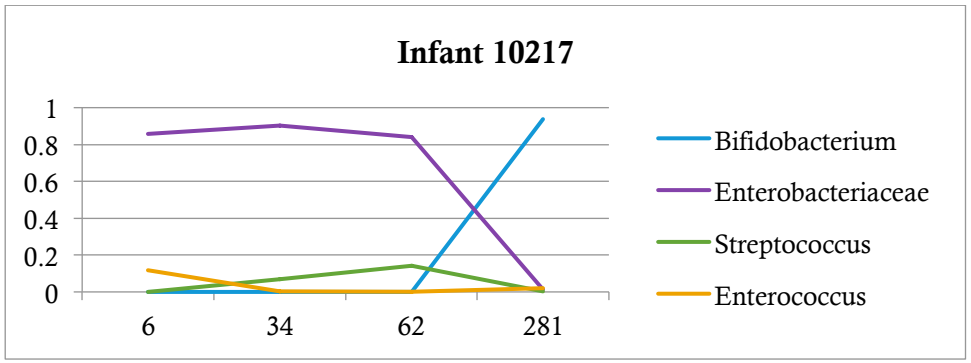
F2



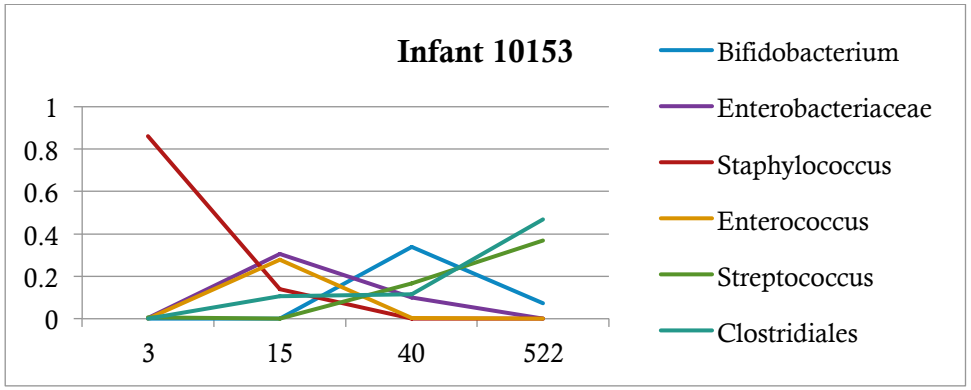
F3



F4



F5



F6

Figures F1 to F6 – Abundance Profiles of Selected Individuals

Abundance as a proportion is on the y-axis, the baby's age (in days) is on the x-axis. Infants in Figures F1 to F3 received the probiotic, while infants in Figures F4 to F6 received the placebo. The six figures shows a high degree of diversity between the babies and how dynamic the gut flora is over the first 18 months of life. *Staphylococcus* is often the most dominant genera in the youngest age group, however as observed in Figure F3 *Staphylococcus* spp. are also present later in life. Some babies had a consistently high abundance of one genus of bacterial family (infants in Figures F2 and F5), while other babies (Figure F6) show much more alpha diversity.

Appendix G

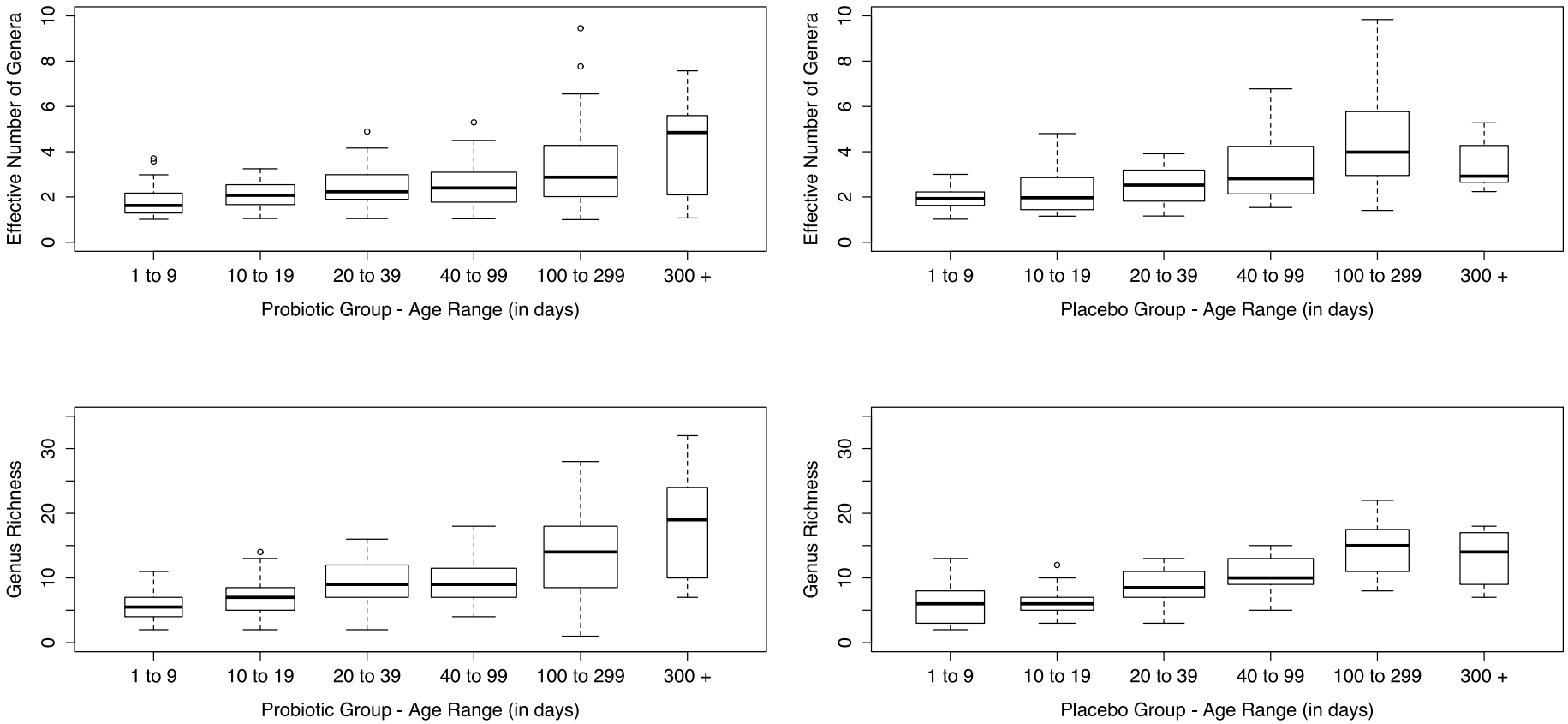


Figure G1 – The Effect of Age and Allocation Group on Effective Number of Genera and Genus Richness

Allocation group has very little effect on either effective number of genera or genus richness. The top two box plots show that effective number of genera increases with age in both allocation groups. The median number of effective genera was slightly lower in babies in the placebo group aged older than 300 days, but this difference was not statistically significant. The same pattern is observed for genus richness. It increases with age, and is slightly lower in babies from the placebo group aged older than 300 days.

Appendix H

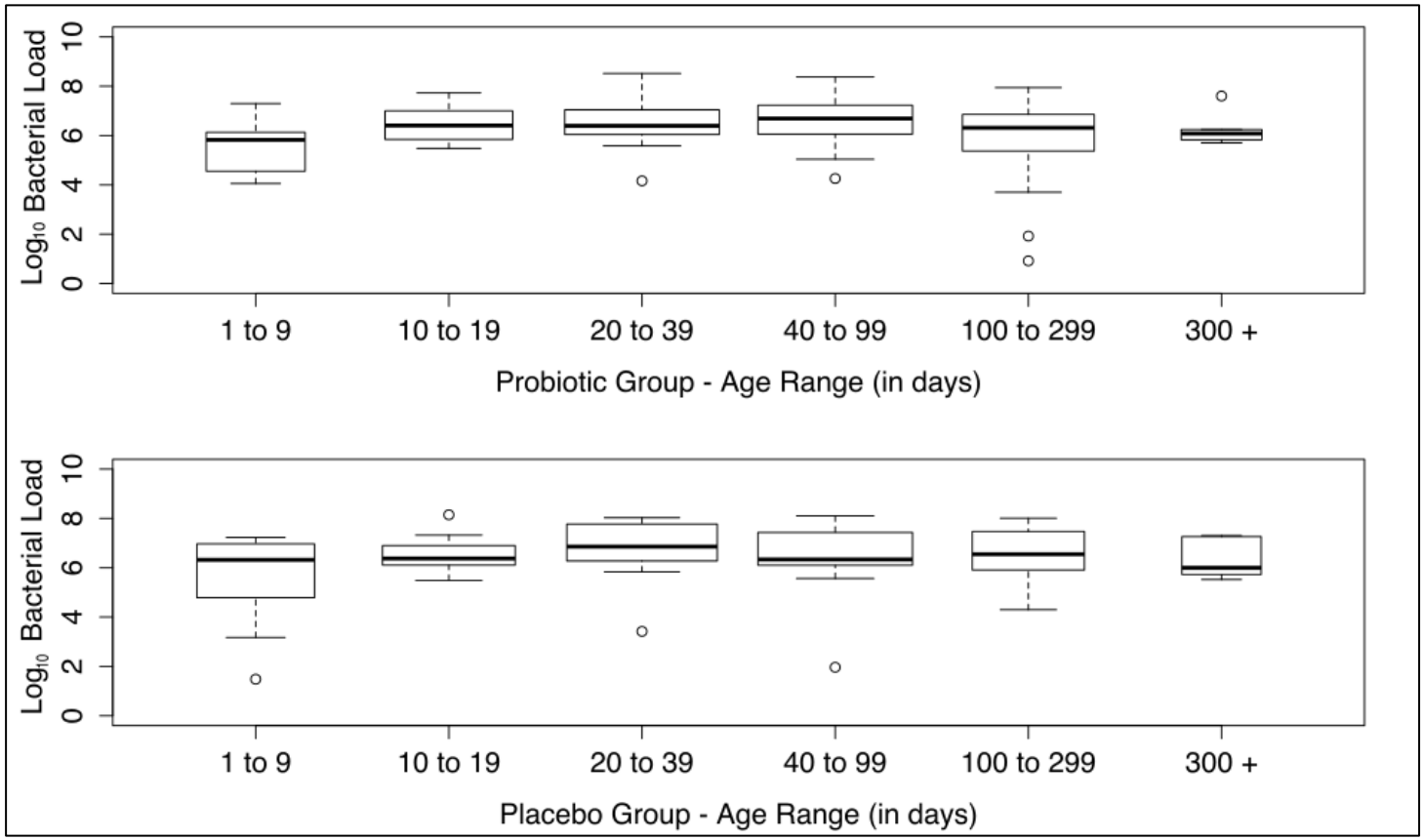


Figure H1 – The Effect of Age and Allocation Group on Bacterial Load

The probiotic did not have any effect on total bacterial load. The only statistically significant difference in bacterial load was between babies aged *1 to 9 days* and babies aged *20 to 39 days*. The y-axis is log transformed bacterial load (as measured by the number of copies of 16S rRNA gene per PCR reaction). NB: the width of the boxplots varies with sample number.

Appendix I

Figure I1 - Phylogenetic Tree of *Bifidobacterium* Type Strains and Representative Sequences from *Bifidobacterium* OTUs.

The tree was created using the Maximum Likelihood method based on the Jukes-Cantor model (76). Bootstrap values greater than 50 are shown above the tree branches. All positions containing gaps and missing data were omitted. The phylogenetic analysis was completed using SeaView (74) and MEGA6 (75). The tree shows little separation between respresentative sequences from the two allocation groups. The tree also shows that many of our representative sequences do not cluster with any of the *Bifidobacterium* type strains. This could indicate that the Bifidobacterium spp. in our cohort are distinct the type strains, or that our sequences have a high error rate, or that there was an error in the MSA on which this tree was based.

Appendix J

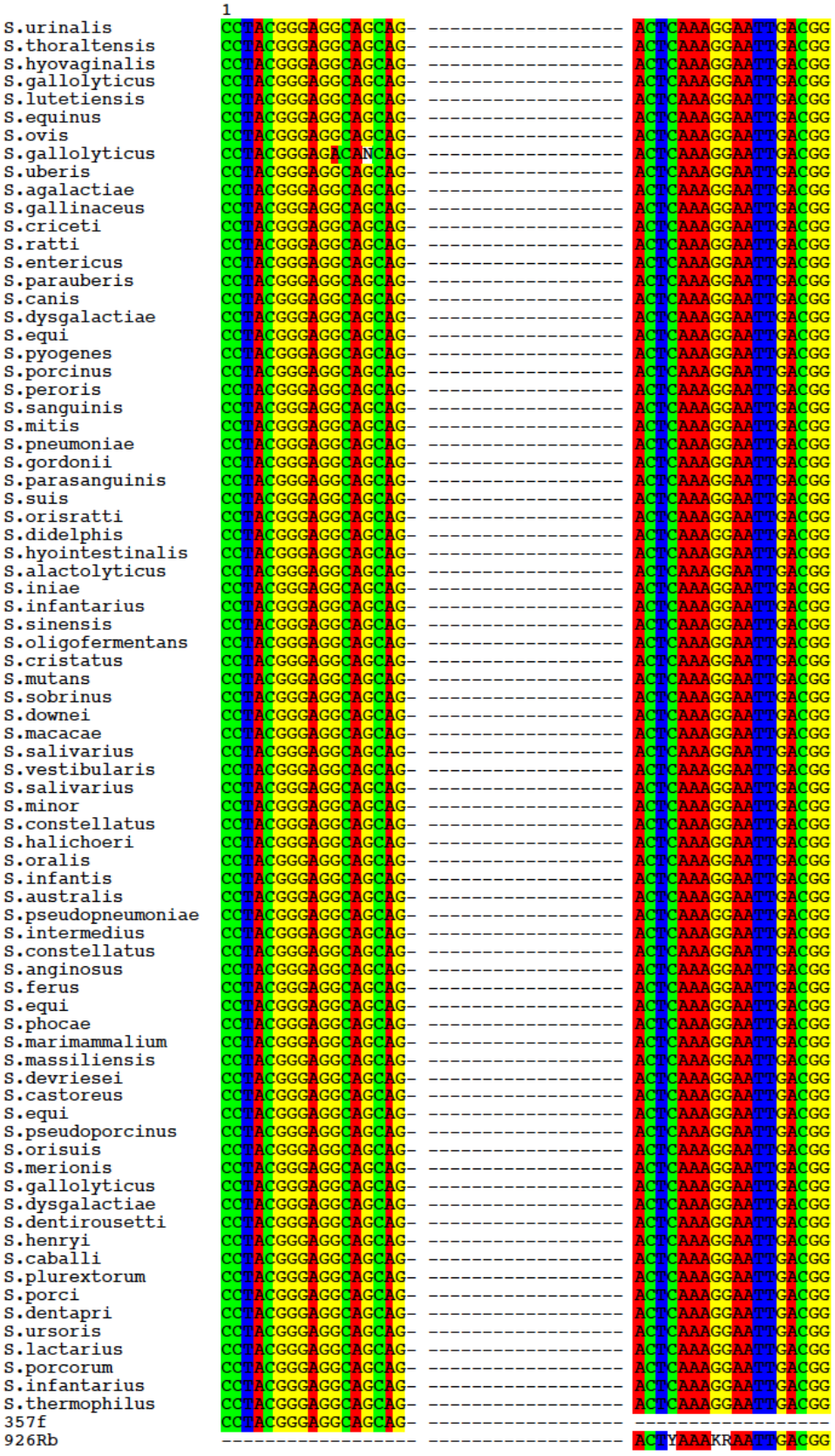


Figure J1 – Multiple Sequence Alignment of *Streptococcus* Type Strains and *Bifidobacterium* Optimised Primers

The *Bifidobacterium* optimised primers used in this study (357F - CCTACGGGAGGCAGCAG, 926Rb – CCGTCAATTYMTTTRAGT) and described by Sim et al (44) would not have inhibited detection of *Streptococcus* spp. There is one miss-match in *S. gallolyticus*. Importantly the primers will anneal to *S. thermophilus,* one of the species present in the probiotic formula.