Determining fecal bacterial profiles of a human-habituated wild chimpanzee population in Mahale Mountains National Park, Tanzania.

Brian Adam Szekely

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Biomedical and Veterinary Sciences

Dr. Taranjit Kaur
Dr. Jatinder Singh
Dr. William Huckle
Dr. Charles Hagedorn

(May 4, 2009)
Blacksburg, Virginia

Keywords: Gut flora, fecal flora, T-RFLP, chimpanzee, endangered species

Determining fecal bacterial profiles of a human-habituated wild chimpanzee population in Mahale Mountains National Park, Tanzania.

Brian Adam Szekely

ABSTRACT

Intestinal flora of wild chimpanzee has not been studied. Fecal flora analyses currently give insight to this environment. We collected 1 fecal sample from each of 12 human-habituated wild chimpanzees from Mahale Mountains National Park, 4 individuals in each of 3 age groups: juveniles, sub-adults, and adults. We analyzed samples for bacterial diversity using Terminal-Restriction Fragment Length Polymorphism (T-RFLP) of amplified 16S rDNA to determine bacterial diversity present. Between 1 and 14 terminal-restriction fragments (T-RFs) were observed in each sample. A total of 26 unique T-RFs were produced from the samples and ranged in size from 92 to 837 base pairs (bps). Twenty-four of these T-RFs corresponded to 5 bacterial phyla: *Actinobacteria, Bacteroidetes, Firmicutes, Mollicutes*, and *Proteobacteria*, as well as uncultured and unidentified bacteria. The remaining T-RFs corresponded solely to uncultured or unidentified bacteria. *Firmicutes* was the most common phylum, observed in 11 of the samples. *Bacteroidetes* and *Mollicutes* were the second-most common phyla, detected in 8 of the samples. Principal Components Analysis (PCA) revealed clustering of 10 of the 12 samples for two components, and 11 of the 12 samples for a third component, which accounted for 72.5% of the variation. Morisita indices were computed to compare T-RF profiles of two samples at a time, and were between 0 and 0.886. We conclude that there is more variation in the fecal flora of wild chimpanzees living in the same social group than previously expected. This suggests that factors other than diet and environment influence wild chimpanzee fecal bacterial profiles.

Acknowledgements

I would like to thank the following individuals for their time, effort, and collaboration with my research.

Dr. Taranjit Kaur for her mentorship and guidance throughout every aspect of the study, and especially for her endless effort to allow my Masters experience of a lifetime come to fruition.

Dr. Jatinder Singh for his vast knowledge, constant assistance, and relentless enthusiasm of which I've relied on heavily during my time in the program.

Dr. Bill Huckle for being a member of my committee and always having both ears ready to listen to my progress, trials, and tribulations, and for continually offering sound advice when I needed it.

Dr. Charles Hagedorn for being a member of my committee and for his unique point of view and motivational questions that helped shape my research project.

This work was supported by the National Science Foundation's (NSF) #0238069 Award to TJS/TK and #0244916 Sub-Award to TK. Additionally, I am very grateful to the Virginia-Maryland Regional College of Veterinary Medicine Research and Graduate Studies Department for their financial support.

Table of Contents

Figures

Tables

List of Abbreviations and Definitions

16S- component of the small prokaryotic ribsomal subunit
ARDRA- Amplified rDNA restriction analysis
BLAST- Basic Local Alignment Search Tool
bp(s)-Base pair(s)
DNA- Deoxyribonucleic acid
E. coli- Escherichia coli
FAM ™-Most commonly used carboxyfluoresceine label
FU- Fluorescence Units
G-Gravity
GI- Gastrointestinal
IUCN- International Union for the Conservation of Nature
kb-Kilobase
M-Molarity (Moles/liter)
min-Minutes
mL-Milliliter
mv-Milivolts
NCBI- National Center for Biotechnology Information
ng-Nanograms
PCA- Principal Components Analysis
PCR- Polymerase chain reaction
Prin1,2,3- Principal Components 1, 2 and 3
qPCR-Quantitative real time Polymerase Chain Reaction
Rep-PCR- Repetitive extragenic palindromic-polymerase chain reaction
RFLP- Restriction Fragment Length Polymorphism
RNA- Ribonucleic acid
ROX-Fluorescein dye X-Rhodamine
rRNA-Ribosomal ribonucleic acid
rDNA- Ribosomoal deoxyribonucleic acid
s-Seconds
SSCP- Single Strand Conformation Polymorphism
T-RFLP- Terminal-Restriction Fragment Length Polymorphism
µL- microliter
VBI- Virginia Bioinformatics Institute

Chapter 1-Introduction and Literature Review

Chapter 1.1

Research Objective

To generate the most complete view of fecal bacterial diversity within the chimpanzee study population using molecular-based techniques.

Hypothesis

Chimpanzees of the M-group (one of the last human-habituated chimpanzee social groups in Mahale Mountains National Park) will have similar fecal bacterial profiles, as diet, social interactions, and exposure to the same environment may play a large role in shaping such profiles.

Introduction

Microbial ecology is an emerging field that aims to describe how microorganisms interact with each other, their environment, as well as all other forms of life. One of the densest microbial habitats recorded is located in the human colon, with numbers reaching 10^{11}-10^{12} cells/mL [Whitman et al., 1998]. There are actually more bacterial cells located in a person's GI tract than human cells in the entire body [Gilmore & Ferretti, 2003]. Describing relationships between microorganisms that live within the intestinal tract of their hosts is of great importance. Beneficial intestinal flora perform critical functions for their human hosts and are collectively considered an "essential organ" [Eckburg et al., 2005]. It receives this "organ" status for several reasons. Some intestinal flora provide nourishment to the host through synthesizing essential amino acids and vitamins, while others can promote epithelial cell develpoment and thereby aid innate immunity in humans [Hooper & Gordon, 2001]. Epithelial cell development is critical for creating a barrier between pathogens in the external environment and the host gastrointestinal (GI) tract [Kagnoff & Eckmann, 1997]. Pathogens in the GI tract are also limited in part due to intestinal flora promoting host epithelial cell production of fucosylated glycans, on which many GI tract bacteria feed [Gilmore & Ferretti, 2003; Hooper & Gordon, 2001; Hooper et al., 1999]. Therefore, certain bacteria actually create additional food sources for other beneficial bacteria which could help to out-compete possible pathogens.

Human and non-human primate endogenous intestinal flora are not fully understood. Culture-based studies have been able to conclude that initial colonization of

microbiota within the gut of humans is through the vagina and feces of mothers to infants during childbirth [Mandar & Mikelsaar, 1996]. It has also been suggested that vertical transmission of important commensal GI bacteria occurs in wild ape populations between mother and offspring [Uenishi et al., 2007].

Describing relationships between intestinal flora and pathogenic bacteria are of great interest to microbiologists and epidemiologists alike. One can use information based on intestinal flora to possibly help combat or prevent pathogen transmission between humans and wild non-human primates. Pathogens threaten the continued existence of wild and captive chimpanzee populations. Describing relationships between microorganisms that live within the intestinal tract of chimpanzees is of great importance and is currently understudied. Chimpanzees are listed as *endangered* by the International Union for Conservation of Nature (IUCN), corresponding to being at a "very high risk" for extinction. Chimpanzees are also genetically closely related to humans [Demuth et al., 2006]. Characterization of the intestinal flora of this species is of primary interest for its health implications for the endangered chimpanzee as well as possible human health applications. The more health-related information that can be ascertained about the chimpanzee, the better chance we have of conserving the few populations that are still left in the wild, as well as those in captivity.

Intestinal flora characterization can potentially offer pertinent information regarding bacterial, and possibly pathogen transmission events (zoonotic, anthropozoonotic, and between chimpanzees as well). Nearly 50-80% of GI flora cannot be cultivated outside of the human gut [Suau et al., 1999; Wilson & Blitchington, 1996]. It is for this reason that molecular techniques have been used to study this complex ecosystem both in humans and non-human primates [Eckburg et al., 2005; Frey et al., 2006; Fujita et al., 2007; Goldberg et al., 2007; Rwego et al., 2008; Uenishi et al., 2007]. By understanding bacterial transmission dynamics between humans and chimpanzees as well as among the chimpanzee population, this will allow researchers to deduce sustainable, sound observational practices. Before transmission dynamics can be determined however, characterization of intestinal flora must first be undertaken. To this end, fecal bacterial profiles of a small group of wild, human-habituated chimpanzees were examined using a molecular-based approach.

Chapter 1.2 Background: Studying Microbial Ecology

Microbes were first classified in terms of phenotypic features (growth temperature, motility, nutrient requirements) using microscopes and traditional culturing techniques from the 17th century up until the 20th century. Ancestoral relationships could be inferred by similarities in such phenotypic features. Additionally, thousands of culturing studies have helped researchers to isolate bacterial pathogens. However, it is very difficult to establish growth conditions for intestinal flora outside of the gut due to unexplained complex interactions that occur in the host [Gilmore & Ferretti, 2003]. Cultureable bacterial species may be analyzed using a number of tools such as antibiotic sensitivity and resitance testing as well as biochemical characterization. There are several disadvantages to performing culturing experiments however. Culturing methods can be extremely time consuming and therefore possibly expensive. Also, a complete survey of gut flora cannot be performed using culturing methods alone, due to the lack of ability to culture many intestinal bacterial species. A molecular-based taxonomic approach is needed to more accurately characterize intestinal and fecal bacterial profiles.

Molecular-based intestinal flora researchers readily use the 16S ribosomal RNA (rRNA) gene to determine bacterial species present in a given sample. The 16S rRNA gene encodes for the small ribosomal subunit in prokaryotes, and is of great value for phylogenetic studies because it is extremely variable between bacterial species, yet highly conserved throughout evolutionary time [Olsen & Woese, 1993]. The first person to decipher the role that 16S rRNA would play in phylogenetic studies was Carl Woese. He and his colleagues proved that specific sequences of the 16S rRNA were highly conserved and could be used to infer phylogeny throughout the bacterial domain using T1 ribonucleases [Woese et al., 1975]. Several years later, 16S rRNA was found to be useful for identification and phylogenetic determinations because of the presence of two specific, separate portions of the sequence. These two portions consist of eight highly conserved regions U1-U8 which exist across the entire bacterial domain, and variable regions V1-V9, in between "U" regions, which are confined to a few well defined areas in the molecule [Gray et al., 1984]. The "V" regions are presumably portions of less importance for ribosomal function [Jonasson et al., 2002]. These "V" regions are readily

used to identify bacterial species due to their inherent large number of different nucleotide substitutions that occur in different bacterial species [Van de Peer et al., 1996].

One technique that showed much promise for studying 16S rRNA for the purpose of bacterial characterization starting in the 1980s, was Restriction Fragment Length Polymorphism (RFLP). RFLP makes use of different DNA sequences within the "V" regions of the 16S rRNA gene that in essence acts as a "fingerprint." The first assignment of a random RFLP was reported in 1980 when a 5 kb single-copy DNA fragment was cloned into a pAW101 vector and then restriction fragments of at least eight different lengths were detected [Wyman & White, 1980]. In R-RFLP, DNA is first extracted and purified, and the 16S rRNA gene is amplified using PCR. After PCR is performed, the products are then restricted with specific endonucleases that allow different fragments to be formed based on variable DNA sequences. The enzymes may recognize four, six, or eight base pairs in length. A small recognition sequence has the ability to be more sensitive as it can recognize few nucleotide changes in the "V" regions that distinguish bacterial phyla and at times class from one another. The resulting fragments are separated according to size using gel electrophoresis. Based on the differing sequences between phyla and classes of bacteria, fragments of different lengths are generated by the endonucleases, and phylogeny.

RFLP analysis of 16S rDNA, otherwise known as amplified rDNA restriction analysis or simply ARDRA, is of limited use for demonstrating the presence of specific phylogenetic groups or for estimating species richness or evenness [Liu et al., 1997]. Terminal-Restriction Fragment Length Polymorphism (T-RFLP) is able to generate data in this regard and is the culmination of comparative genomics, RFLP, PCR, and nucleic acid electrophoresis. T-RFLP is different from RFLP in that one (or both) of the primers used in the PCR reaction is labeled with a fluorescent dye so that when the restricted fragments are analyzed with an automated DNA sequencer, the sizes of only the terminal restriction fragment can be determined and are quantified in order to measure relative abundance and bacterial diversity within and between samples. When both primers are fluorescently labeled, they are labeled with different fluorophores and generate two profiles for each sample to allow simultaneous detection of fragments in one T-RFLP run. T-RFLP electropherograms are generated which relate fragment length (number of

nucleotides) to fluorescence intensity, and are used to determine presence *and* relative abundance of fragments located in the sample. The resulting fragment peaks correspond to different bacterial phyla. It is a sensitive and relatively rapid means of assessing community diversity [Liu et al., 1997]. It allows a distinctive fingerprint of a microbial community to be produced using fragment length and its associated relative abundance as the basis for comparison. It should be stressed that T-RFLP is used as a comparative method only and cannot distinguish all of the phyla that may be located within a particular community. It is used to find the major phyla within the sample that are at or above a specified threshold intensity set by the researcher. The threshold used for this study is discussed later in Materials and Methods.

Chapter 1.3: Terminal-Restriction Fragment Length Polymorphism: Comparisons, Advantages, Disadvantages

The use of 16S rDNA

The 16S gene is heterogeneous between multiple copies within one species and therefore can confuse the interpretation of microbial diversity studies [Dollhopf et al., 2001]. It has also been shown that studies using the 16S rRNA gene can lack resolution at the species level, specifically in the genus *Bacillus* [Ash et al., 1991]. However, T-RFLP analysis is one of the most frequently used high-throughput fingerprinting methods and therefore an large amounts of microbial profiling data is widely available [Schutte et al., 2008]. This is one reason why 16S rDNA and T-RFLP was used to characterize microbial profiles for this study, as compared to other molecular techniques that use different regions of bacterial DNA.

Automated Ribosomal Intergenic Spacer Analysis

ARISA is based on the intergenic region between the 16S and 23S rRNA genes in the rRNA operon, which has very high variability in nucleotide sequence among different bacterial species [Daffonchio et al., 2003; Danovaro et al., 2006]. ARISA can also have higher specificity than T-RFLP patterns when dealing within the genus *Pseudomonas*, and *Propionibacterium* in particular, as well as being able to determine species richness and evenness with more accuracy [Danovaro et al., 2006; Tilsala-Timisjarvi & Alatossava, 2001]. However, it should be noted that this is true only for particular species, and is not necessarily true for all bacterial species. Also, according to Tilsala-

Timisjarvi and Alatossava, 2001, the σ factor for gene *rpoB* appears to be in much less copy number than the 16S gene and may therefore be a better candidate to show higher discrimination in certain species. For the present study, we are looking at the bacterial profiles at the phylum level, and are not concerned with deciphering organisms down to the species level. Due to the existing extensive 16S rDNA database when compared to intergenic spacer region data, T-RFLP is still the better candidate to assess overall bacterial diversity in our samples.

Sequencing from cloned PCR and qPCR products

There is a rapidly growing rRNA gene sequence database online. It can determine genus, species, and sometimes strains of bacteria without the need for culturing. Also, many other genes can be targeted for analysis. Quantitative-PCR can be used to determine the number of copies of a gene in question which can be very useful in distinguishing species as well determining if a gene is a good candidate for generating microbial community profile data. Sequencing every isolated fragment from such a diverse sample as feces would be tremendously time consuming and expensive, and is therefore not recommended for this Masters project.

Microarrays

DNA microarrays have the potential to map diversity [Dollhopf et al., 2001] However, they do not reveal unexpected species; the method can only search for the specific phyla and species requested by the researcher. This method is more useful for finding specific differences in metabolic functions between known species by determining what genes are being expressed at a given time, and therefore is not recommended for generating entire bacterial profiles.

Denaturing or Temperature Gradient Gel Electrophoresis (D/TGGE)

DGGE or TGGE takes advantage of the fact that DNA can be chemically or thermodynamically degraded, and electrophoretic patterns can be analyzed based upon the varying sequence of nucleotides between phyla. Intestinal microbial community studies of chimpanzees have been performed using this technique [Uenishi et al., 2007]. However species richness or evenness cannot be analyzed using this method. Only presence or absence of an electrophoretic band that indicates specific phyla can be determined. Additionally, DGGE is considered to be less sensitive than T-RFLP

[Moeseneder et al., 1999; Nunan et al., 2005]. It is for these two reasons that these techniques were not used in the current study.

Single Strand Conformation Polymorphism (SSCP)

This method takes into account that single stranded DNA fragments travel at different speeds due to their conformational folding even if the lengths are the same. SSCP patterns can characterize differentiating bands by cloning and sequencing, whereas T-RFLP seems to be more advisable for routine analysis because difficulties resulting from gel-to-gel variation do not exist [Smalla et al., 2007]. It should be noted that like T-RFLP, this method can only be used for comparative purposes. As with the other methods, less data is available for SSCP when compared to T-RFLP. Many of the primers used in this method are specific for certain bacterial species and are not universal as in T-RFLP studies. It appears that a better use for SSCP would be to show diversity within a particular species of bacteria.

Advantages with T-RFLP

T-RFLP is a popular high-throughput, economical method used to assess microbial diversity. Using T-RFLP to study fecal flora is a comprehensive technique that determine the lengths of DNA fragments that reflect the dominant phyla in the community [Schutte et al., 2008]. The T-RFLP approach is a reliable approach to detect the relative changes in microbial community structures with high quantitative precision, which is also a prerequisite for monitoring effects on microbial communities [Hartmann & Widmer, 2008].

Disadvantages with T-RFLP

Currently, there is no solution to correct for migration discrepancies that arise during the T-RFLP procedure due to the use of different fluorophores (FAM ™ label for PCR, and ROX label for Sequencing Gel), and therefore researchers should take this into account when using T-RFLP data to determine the community composition. Preferential, unspecific, and inhibited PCR reactions as well as residual polymerase activity during restriction may also play a role in determining microbial profiles using T-RFLP, therefore careful optimization of PCR as well as digestion protocols can reduce but not completely exclude the impact of these inconsistencies [Hartmann & Widmer, 2008]. Next, the number of phylotypes that exist as DNA sequences is tiny compared to the amount that

actually is in the environmental sample [Abdo et al., 2006]. Therefore, T-RFLP does not detect all phyla represented in the sample. Additionally, it should be noted that with any PCR-based method, the relative abundance of products from mixed-template reactions are prone to several biases, including lower diversity revealed with more cycles [Bonnet et al., 2002].

As we knew when starting the procedure, T-RFLP is not an all encompassing method to analyze fecal bacterial profiles. It offers possible profiles that may exist in the sample, and only detects the major phyla. Additionally, it is hard to distinguish between actual peak heights and noise data that are generated from the DNA analyzer in the T-RFLP method. Computer programs that analyze Terminal-Restriction Fragments (T-RFs) can determine where a peak starts and ends, its height and area, but the true baseline must be determined by the researcher and can be difficult, as this can change from sample to sample. One simple approach to distinguish signal from noise is to impose a fixed detection threshold that is some arbitrarily chosen value, e.g., 50 or 100 fluorescence units (FU). Employing a high detection threshold such as 100 FU insures that the number of peaks that could be considered noise is very low, but risks excluding small reproducible peaks [Dunbar et al., 2001].

Chapter 1.4: Literature Review: Gut and Fecal Flora Analytical Studies

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. 2005. Diversity of the human intestinal microbial flora. Science 308(5728):1635-8.

Eckburg et al. reported the bacterial profiles of human subjects based on mucosal tissue and fecal sample analysis. These samples were obtained from three healthy adults to detect differences in bacterial isolates found in mucosal sites and feces within and between individuals. The three adults were selected at random from the control group of a large population-based case-control inflammatory bowel disease study at the University of Manitoba. Two 50 year olds and one 43 year old constituted the three subjects. Mucosal site samples were taken during colonoscopies from the cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and the rectum. The fecal samples were taken one month later during a follow-up exam. Both mucosal and fecal samples

were immediately flash-frozen with liquid nitrogen and stored at -70˚C until shipment on dry ice for analysis. This is an excellent way to preserve the samples.

The 16S rRNA gene was used to determine the presence of bacterial and archaeal species and was amplified using PCR with broad range bacterial primers. Products were cloned and sequenced. Of the 395 bacterial phylotypes discovered, 244 were novel and 80% represented species that hadn't been cultured previously. Most of the inferred organisms were members of the *Firmicutes-Clostridia* class and the *Bacteroidetes* phylum. Large variations among the 65 *Bacteroidetes* phylotypes were noted between subjects.

Each mucosal library from one subject was a subset of their stool library, while the other two subjects' mucosal libraries were significantly different from their corresponding stool libraries. The author mentions the inconsistency of when the mucosal and fecal samples were taken. Ultimately the patchiness found between mucosal site and fecal analyses within an individual cannot be explained fully. Fecal samples were speculated then to consist of mucosal-specific bacterial populations that have been shed during defecation, as a well as a nonadherent luminal population. Bacterial flora in specific mucosal sites from each patient were similar, but fecal samples differed greatly between individuals. Therefore, most of the intersubject differences in bacterial phylotypes were between mucosal sites of one patient and fecal samples from another, showing that unknown factors are contributing to the constitution of bacteria found in feces.

Frey JC, Rothman JM, Pell AN, Nizeyi JB, Cranfield MR, Angert ER. 2006. Fecal bacterial diversity in a wild gorilla. Appl Environ Microbiol 72(5):3788-92.

Frey et al.'s rationale of the study was that normal intestinal flora of the endangered gorilla remains unexplored. However, nowhere in the paper does it mention why intestinal flora is important to the health of the host, of which I have already mentioned. The lack of purpose of the study other than characterization is evident.

In the study, four fecal samples were taken (once a month for four months between September and December of 2002) from one male silverback gorilla. The samples were collected from night nests less than 12 hours after defecation and identified by size, location, and the presence of silver hairs. Therefore, feces could have been

sitting in the nest anywhere up to 12 hours after defecation, which makes the collection procedure not standardized. Due to variable collection times, the bacterial profile could have changed to preferentially grow aerobic organisms for example. Additionally, fecal flora that was present in small concentrations at the time of defecation could have been out-competed by other organisms and subsequently become erased from the bacterial profile. These biases can be avoided or decreased by watching an individual animal defecate and collecting the sample immediately, which was done during our collection process (see Materials and Methods).

Results based on T-RFLP and sequencing analysis indicated that fecal bacterial profiles did not differ significantly over the course of the four month sampling period. *Firmicutes* was the most abundant phylum in the fecal samples. *Verucomicrobia, Actinobacteria, Lentisphaerae, Bacteroidetes, Spriochetes* and *Planctomycetes* were also found in the feces of the wild gorilla. Frey et al. states that due to the high amount of bacteria falling within the phylum *Verrucomicrobia,* it is obvious that this phylum plays a large role in the intestine of gorillas. Using only one individual, and only four samples, it appears premature to make such a conclusion. There could be an aggregation of this phylum in this individual due to a number of different factors including, diet, presence of disease, age, gender, or other possible unforeseen factors.

Goldberg TL, Gillespie TR, Rwego IB, Wheeler E, Estoff EL, Chapman CA. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. Biological Conservation 135:511-517.

In this study, fecal samples were collected from 25 humans (13 high-risk-overlap with chimpanzee habitat, 12 low-risk-no overlap with chimpanzee habitat) and 23 human-habituated chimpanzees from several different locations in and around Kibale National Park between May and June of 2004. Only fecal samples were collected from the chimpanzees while fecal swabs were analyzed from humans. Additionally, these fecal swabs were self-administered by the 25 humans, indicating possible inconsistencies during the collection of these samples. The preferential collection of bacteria from the lowest portions of the colon from humans should not have been compared to fecal samples collected from chimpanzees. Mucosal sites can vary largely from fecal samples

even from the same individual [Eckburg et al., 2005]. The high-contact humans and the chimpanzees lived in the same area and had access to the same environmental reservoirs of possible microbes. The data showed through antibiotic sensitivity and resistance testing, as well as Rep-PCR (collectively a polyphasic taxonomic approach), that the chimpanzees had more fecal flora in common with the humans living within overlapping areas than with humans living in a different non-overlapping location.

Only *Escherichia coli (E. coli)* strains were isolated to perform this study. Rep-PCR was used, as it has been shown to have a high power for discriminating among closely related *E. coli* isolates [Goldberg et al., 2007; Johnson & O'Bryan, 2000; Woods et al., 1993]. The title of the study is misleading as patterns of bacterial exchange were not deduced. It is still unsure as to how these *E. coli* isolates are being exchanged, either directly or indirectly through a common environmental reservoir. Goldberg et al. even say that the modes of transmission are undetermined. Goldberg et al. have shown that culture-based and molecular-based techniques each play an important role in characterizing fecal flora. Additionally, the study has indicated that wild, human-habituated chimpanzees and humans are actively sharing enteric bacteria due to ecological overlap.

Uenishi G, Fujita S, Ohashi G, Kato A, Yamauchi S, Matsuzawa T, Ushida K. 2007. Molecular analyses of the intestinal microbiota of chimpanzees in the wild and in captivity. Am J Primatol 69(4):367-76.

This research article is the first and only solely molecular analysis of wild chimpanzee fecal bacterial profiles to date. TGGE and ARDRA along with sequencing analyses were used to compare wild and captive chimpanzee populations. Fecal samples were taken from ten captive chimpanzees from the Primate Research Institute of Kyoto University in April 2004, while thirteen chimpanzees from Bossou, Guinea were used as the wild population during the months of July and August of 2004. Uenishi et al. aliquotted fecal samples in >90% ethanol within five minutes after defecation for both captive and wild chimpanzees. Caution was taken to avoid soil contamination by collecting the portion of feces that did not contact the forest floor for the wild chimpanzee samples.

PCR amplification for TGGE analysis was carried out using U968-GC and L1401 primers. These primers are specific to the V6-V8 regions of the bacterial 16S rRNA gene. This does not account for all variable regions within the gene and therefore can be considered less sensitive than the T-RFLP method that makes use of all nine variable regions of the 16S rRNA gene used in the current study. TGGE band profiles were described using hierarchical clustering analysis. Hierarchical clustering analysis is a statistical method for finding relatively homogenous clusters of cases based on measured characteristics. This type of statistical analysis was prudent for the question at hand as Uenishi et al. were trying to determine differences between wild and captive chimpanzee fecal flora. Six TGGE bands common to all individuals tested were excised from the gel and DNA was extracted. A portion of the eluted DNA was amplified for sequencing and results were compared via the BLAST program at National Center for Biotechnology Information (NCBI) with those registered in the databases to suggest possible taxonomic names.

DNA extracted for the TGGE protocol was also used for the ARDRA analysis. The V3-V8 region was amplified using primers 338f and L1401 and amplified products were then ligated into a pGEM-T vector. Clones were randomly picked from a plate raised from the feces of each chimpanzee. Of all the clone groups, the ones commonly shared by at least two individuals were subjected to further sequence analysis and compared using BLAST.

A cluster analysis of TGGE profiles indicated a difference in the composition of the fecal flora between wild and captive chimpanzees. The ARDRA results showed the predominant presence of plant polymer and sugar-fermenting bacteria in chimpanzees in the wild. There were also several bacteria common to both groups, namely *Clostridium, Lactobacillus,* and *Bifidobacterium*. Uenishi et al. conclude that these bacteria may be so essential that they have not been "washed out" by the modification of environmental conditions.

The main shortcomings of this study are with the methods. TGGE and ARDRA coupled with sequence analysis can only describe what species are located in the sample, but they do not reveal their relative abundance in the samples. T-RFLP coupled with robust data analysis is capable of this task, and that is why I have chosen this technique.

Saying whether a certain phylum of bacteria is located within a sample is important, however explaining its relative abundance in the sample can be more meaningful.

Fujita S, Kageyama T. 2007. Polymerase chain reaction detection of Clostridium perfringens in feces from captive and wild chimpanzees, Pan troglodytes. J Med Primatol 36(1):25-32.

The prevalence of *Clostridium perfringens (C. perfringens)* in wild human-habituated and captive chimpanzee fecal samples was examined. A total of 81 fecal samples were collected from at least 40 different chimpanzees from Mahale Mountains National Park (the current study site) during two wet seasons in 2001 and 2005 and one dry season in 2001. A total of 53 fecal samples from at least 15 chimpanzees were collected from Bossou, Guinea in 2001, and comprised the wild population samples. Samples were collected immediately after defecation and care was taken to only use the portion of the sample not in contact with the forest floor. A total of 15 fecal samples collected from five individuals from the Primate Research Institute in Inuyama, Japan were used for the captive population samples. Samples were collected immediately after defecation from indoor experimental booths and sleeping rooms. All samples were analyzed for the presence of *C. perfringens* using two sets of primer pairs for the *plc* gene using nested-PCR. Using chi-square testing, wild chimpanzee samples had a significantly lower prevalence of *C. perfringens* than their captive counterparts. These findings suggest that differing diets and environmental conditions, including stress levels, may affect the presence of an opportunistic pathogen found in the feces of chimpanzees.

Chapter 1.5: Study Rationale-Why study fecal flora?

No information is available regarding the intestinal microbiota of wild apes. However, several studies outlined above have used molecular techniques to characterize wild ape fecal flora, which currently give insight to gut flora. By the eventual determination of fecal flora found in humans and chimpanzees that have ecological overlap, it could enable researchers to determine pathogen transmission dynamics between the two species. This data can have an impact on shaping guidelines and procedures for humans working with, living near, or visiting areas harboring wild chimpanzees. Additionally, evaluating differences in bacterial profiles from members of

the same social group could help to explain if and how collective fluctuations in fecal flora are occurring.

This study can offer a reasonably sound approximation of fecal flora for the chimpanzees of Mahale Mountains National Park. Feces cannot offer a complete characterization of gut flora for any individual, as only the flora that are shed during defacation can be analyzed. Characterizing microbial diversity is the first step in deciphering its role in baseline physiological health status and human and animal disease. An imbalance in the gut flora of humans is hypothesized to be a major factor in some human intestinal disorders [Mazmanian et al., 2008]. It is possible then that an imbalance in gut flora of chimpanzees could also be a factor in their health and well-being.

Currently little research has been conducted detailing the normal intestinal and fecal flora of chimpanzees. Studies have not been conducted to determine variations in the diversity and relative abundance of bacteria in the feces of different individuals in the same social group. In our study, we use T-RFLP to characterize and compare fecal bacterial profiles and relative abundance of phyla and at times class between wild, human-habituated chimpanzees of the M-group in Mahale Mountains National Park in western Tanzania.

Chapter 2-Materials and Methods

Study Site: Designated in 1985, Mahale Mountains National Park is 623 square miles and located on the eastern shore of Lake Tanganyika in the western region of Tanzania. It is home to some of the last remaining chimpanzees in the world. The study site was located in the northern portion of the park near the village of Kasiha where chimpanzee behavioral research has been conducted since 1968.

Animal Population: The M group of chimpanzees, which number approximately 64 individuals at the present time, has been studied since 1968 by Japanese researchers. Due to their extensive efforts, the M group has been habituated to human presence and tolerates long periods of observations by researchers and tourists. The group has therefore become a valuable population for research purposes. Juveniles (ages 4-8 years), sub-adults (9-14 years) and adults (≥15 years) of each gender are found in the population.

Table 1: Sample animals

Sample MM07-	Age	Age Category	Gender	Date of Collection
01	7	Juvenile	M	11/25/2007
02	6	Juvenile	M	10/18/2007
03	5	Juvenile	F	7/23/2007
04	7	Juvenile	F	7/26/2007
05	12	Sub-Adult	M	7/16/2007
06	9	Sub-Adult	M	12/3/2007
07	9	Sub-Adult	F	7/5/2007
08	14	Sub-Adult	F	7/1/2007
09	29	Adult	M	7/11/2007
10	16	Adult	M	7/7/2007
11	26	Adult	F	7/15/2007
12	20	Adult	F	10/29/2007

The identity of each chimpanzee was known at the time of defecation. Twelve fecal samples were collected between June and December of 2007 from twelve chimpanzees, and are listed in Table 1. Three parent-offspring relationships exist within the sample animals. The animal associated with MM07-11 is the mother of MM07-07, MM07-12 is the mother of MM07-03, and MM07-09 is the father of MM07-02. The

fecal samples were collected immediately after defecation using wooden applicator sticks to remove only the superficial layer of fecal material and avoiding collection of any feces that may have directly contacted the forest floor to eliminate the risk of bacterial contamination from environmental sources. At the time of fecal sample collection, all chimpanzees were apparently healthy. All fecal samples were collected and transferred by personnel wearing face masks and gloves. Fecal samples were transferred into a sterile tube containing RNAlater (Applied Biosystems Inc. Austin, TX) in a ratio of 1:1 and stored at -20°C within 6 hours of collection. Samples placed in shatter-proof plastic tubes were shipped wrapped in absorbent material and placed in leak-proof plastic bags with ice packs to the Virginia-Maryland Regional College of Veterinary Medicine, where they were stored at -20°C upon arrival five days later. Feces can be stored at ambient temperature in RNAlater up to one month.

Terminal-Restriction Fragment Length Polymorphism

DNA Extraction- A Qiagen DNA Mini Fecal DNA Isolation Kit (Qiagen, Inc. Valencia, CA) was used to isolate total DNA from the fecal samples. Four extractions were performed for each sample. All centrifugation steps were performed at maximum speed (16,100 x G). Fecal **s**amples were first placed at 4˚C for approximately 45 minutes or until thawed depending upon the volume of sample. The samples were vortexed continuously for 1 minute to achieve homogenization. Pipette tips were cut with sterile scissors to allow a hole of approximately 0.5 cm in diameter to be formed in order to draw up fecal material without hindrance. Next, 200 µL of fecal sample in RNAlater was placed into a clean 2 mL tube. Then, 1.6mL of ASL buffer was added to each tube and vortexed for 30 seconds or until sample was homogenized. Samples incubated at room temperature overnight with slight agitation at 550 rotations per minute. The following morning, the isolation continued as directed by the manufacturer's instructions beginning with an incubation period of 20 min at 70˚C. The manufacturer's instructions were followed through the final step with an elution of 50 µl of DNA. All four replicates of the same sample were pooled together and using 200 µL of total DNA, 50 µL 1M NaCl and 500 µL 96-98% ethanol was added and then tubes incubated at -20˚C overnight to precipitate the DNA further. The following morning, the tubes were centrifuged at 4˚C for 30 minutes. The supernatant was discarded and 50 µL of 70% ethanol was added and

then tubes were centrifuged for 15 minutes. The supernatant was discarded and samples were left to dry under the hood for 30 minutes or until no ethanol droplets were left over. The DNA was resuspended in 30 µL of AE buffer and was then ready for a clean-up procedure.

DNA Clean-Up- A PowerClean DNA Clean-Up Kit (MoBio Laboratories, Carlsbad, CA) was used to purify the DNA, as several of the samples were not able to produce amplified PCR products due to impurities. PCR inhibitors such as humic and phenolic substances readily found in plants eaten by chimpanzees could have passed through the GI tract and therefore needed to be removed to successfully amplify the extracted DNA. The resulting clean DNA was stored at -20˚C until further analysis. To determine the concentration and purity of total DNA in the sample, a ND-1000 Full spectrum Spectrophotometer was used (NanoDrop Technologies, Wilmington, DE, USA). All DNA samples' 260nm/280nm and 260nm/230nm ratios must have been between 1.7 and 2.3 in order to be used for PCR.

Polymerase Chain Reaction- Amplification of the 16S rDNA gene was carried out using forward primer Bact-27F 5'-AGAGTTTGATCCTGGCTCAG-3' and reverse primer Bact-1387R 5'-GACGGGCGGTGTGTRCA-3' (Eurofins MWG Operon Inc., Huntsville, AL). These forward and reverse primers were chosen because of their ability to make full 16s rRNA gene copies of approximately 1550 base pairs (bps), found in nearly all bacterial species. PCR conditions were as follows in a final volume of 25 mL, with final concentrations listed for each reagent: 8.6 µL PCR grade water, 2.5 µL (1X) PCR buffer (Applied Biosystems, Foster City, CA), 1 µL (1X) Bovine Serum Albumin, 2.5 µL (0.25mM) of each deoxyribonucleoside triphosphate, 0.5 µL (0.2µM) forward and 0.5 µL (0.2µM)reverse primer, 2.5 Units AmpliTaq® DNA polymerase (Applied Biosystems, Foster City, CA) and 100 ng of total DNA. The following cycling parameters were used: 5 min of initial denaturation at 94ºC followed by 30 cycles of denaturation (30s 94ºC), annealing (30s 55ºC), and elongation (110s 72ºC), with a final extension at 72ºC for 7 min.

Pilot reactions were completed to establish optimum conditions before using fluorescently labeled forward primers. Only 25 µl reaction volumes worked. Large-scale PCR using 100 µl was never successful. PCR reactions with FAM™ fluorescently

labeled forward primers were used in order to visualize peaks on the chromatogram generated from a DNA analyzer. A PCR clean-up kit was used to rid the PCR products of unwanted PCR reagents including buffers, dNTPs, and polymerase using a MoBio UltraClean PCR Clean-Up Kit (MoBio Laboratories, Carlsbad, CA) according to the manufacturer's instructions.

Gel Electrophoresis- After each PCR, a 1.5% agarose gel was run at 80 millivolts (mv) for 65 minutes using 0.02 mL/L of ethidium bromide in the gel solution. PCR product (10 µl) from each reaction was mixed with 2 µl of 5X loading dye (Spectrum Chemicals & Laboratory Products, Inc. Gardena, CA) and 10 µl of this mixture was loaded into each lane on the gel. HyperLadder II (BioLine USA Inc.) was used for the standard ladder, which gives a range of 50-2000 base pairs. Water was used as a negative control for each gel run. The gel was checked for the presence, intensity, and location of bands using an encased camera box fixed with an ultraviolet light to visualize the bands. A digital camera kept a record of all gels processed. Figure 1 is an example of a typical gel.

Restriction Enzyme Reactions- The enzyme *HhaI* was used to restrict the PCR products as it recognizes and cuts frequently (4 bp cutter). *RsaI* was also experimented with, but had more unresolved phyla associated with each fragment. Some of the T-RFs corresponded to more than 3 different phyla, and therefore *HhaI*, and not *RsaI* was used for the study. Each enzyme reaction produced two replicates to be submitted for DNA analysis, therefore two reactions were performed for each sample for a total of four replicates submitted per sample. Each 20 µL reaction consisted of 150 ng of PCR product, 1µL of restriction enzyme (20 Units), 2 µL of supplied buffer, 0.2 µL of supplied BSA, and enough water to equal 20 µL. After all components were added together, the reaction incubated at 37^{o}C for 2 hours. The reaction tubes were then placed in a heating block at 65^{o}C for 20 min to inactivate the enzyme. Fifty 50 µL of 100% isopropanol and 5 µL (0.2M) NaCl were added to the reaction tubes and mixed to further precipitate the DNA, and were placed at-20^{o}C overnight. After the overnight precipitation step, the samples were spun at 16,100G for 30 minutes at 4^{o}C. The supernatant was discarded, and the pellet washed with 50 µL 70% ethanol and subsequently spun at 16,100G for 15 minutes at room temperature. The ethanol was

discarded and samples were left to dry under a BL-2 hood at room temperature. Each of two restriction reactions was resuspended in 14 µl of sterile water and was aliquotted to two tubes with 6 µl of sample in each. A total of four replicates were submitted for fragment analysis for every fecal sample, each with 8 µl of ROX®-labeled ladder (BioVentures, Inc., Murfresboro, TN).

The products were run on an ABI 3130xl DNA analyzer to measure the fluorescently labeled terminal fragments length and associated abundance.

Standardization- Before experiments were conducted on all samples, one sample was used to standardize the entire procedure to ensure repeatability. MM07-07, a sample from an adult female, was used for the standardization experiments. Many samples were collected from her and therefore offered several opportunities to experiment with the procedure. After optimizing the extraction protocol, using 100 ng of DNA in each PCR reaction, and using a total volume of 20 µL for the enzyme reactions, samples were ready to be submitted to Virginia Bioinformatics Institute (VBI). Figure 2 illustrates the repeatability of the protocol by having four of the same electropherograms generated by T-RFLP from each of the replicates.

GeneMapper Software version 4.0- Electropherograms were analyzed using GeneMapper software v4.0 (Applied Biosystems, Inc. Foster City, CA). Each peak may correspond to a specific phylum, and sometimes several peaks may correspond to the same phylum.

A procedure called "binning" was used for analysis. All fragments that exist within 1 bp of each other are considered to be the same size. For example, fragment sizes of 56.5 and 57.5 will be "binned" together for a fragment size of 57. One reason why some fragment sizes have decimal portions is because of the small inconsistencies in the speeds at which the fragments travel down the gel when two different fluorophores are used during the T-RFLP procedure (FAM for the PCR, and ROX for the gel). Only peaks at or above 100 FUs found in at least three of the four replicates were considered positive peaks. The area under the curve of each fragment peak was used to calculate the abundance measurement.

Microbial Community Analysis (MiCA)- The program used to convert fragment length and abundance data generated from GeneMapper to possible microbial

profiles is known as the Microbial Community Analysis III online tool (mica.ibest.uidaho.edu). MiCA enables researchers to perform the following tasks:

(a). *in silico* PCR amplification and restriction of 16S rRNA gene sequences found in public databases;

(b). automatic retrieval of data and archival storage in a relational database;

(c). comparison of multiple T-RFLP profiles obtained from a single sample using different primer-enzyme combinations;

(d). statistical analysis of T-RFLP data and clustering of samples based on similarities and differences.

The primary database contains a large number of 16S rRNA genes retrieved from the Michigan State University Ribosome Database Project II (rdp8.cme.msu.edu/html/). Ideally, investigators want to determine what known and unknown organisms could produce the observed community T-RF profiles. To do so, one would need to search a database of DNA sequences from known organisms and perform *in silico* PCR amplification and restriction digests of the genes to determine which ones could produce the fragments seen in the profile. The MiCA algorithm compares the data from one or more T-RFLP profiles to the outcomes(s) of in silico analyses of sequences in the database done using the same primers and restriction enzymes. The outcome is one or more plausible communities.

For this study, the T-RFLP Analysis (PAT+) tool was used, as this is the tool needed when only one primer is fluorescently labeled. Several parameters may be changed within the MiCA program when analyzing T-RFLP profiles. The sensitivity parameters were as follows: digest allows at most "0" mismatches within "0" bases from the 5' end of the primer, which accounts for the primer not annealing to the target sequence exactly. The window bin size forward match was set to the lowest setting of +/- "1", allowing profiles to be generated based on actual measured T-RF sizes as well as T-RF sizes above and below the actual size. For example, if a T-RF of size 100 bps is found in the sample, the profile generated is then based on T-RF sizes 99, 100, and 101 bps.

The relative abundance of each T-RF in a sample was calculated based on the formula below.

Average area of a specific T-RF (in FUs)

total FUs of all T-RFs of the sample

Principal Components Analysis (PCA)- PCA is an ordination statistical tool used to analyze multidimensional datasets by transforming multiple correlated data points into a smaller number of uncorrelated variables called components. Both the testing for a normal distribution in the dataset as well as PCA was performed using JMP v7.0 (JMP for Windows, SAS Institute, Cary, NC). The Shapiro-Wilk W Test was used to determine if the dataset had a normal distribution. PCA has the ability to reveal the internal structure of the data by representing similar measurements closer to each other in data space, and dissimilar measurements further away. PCA was used in this study because it is a tested method that graphically represents the clustering and the variability of T-RFLP (and other multidimensional data) profiles.

Community Analysis- The vegan package [Oksanen et al., 2008] was used to calculate the Morisita index of community similarity in R software [R: Team, 2008]. Although MiCA has the ability to perform statistical analyses of T-RFLP data, Morisita indices have been used previously and are a documented method to calculate similarities between different T-RFLP samples [Dollhopf et al., 2001] The formulas below were used to calculate the Morisita indices, where n_i is the abundance of species *i*, *N* is the total abundance of all species in a community, *l* is Simpson's dominance index for each community, and *s* is the total number of species in the community [Brower et al., 1990]. These equations were first modified by Dollhopf et al., 2001 to use for T-RFLP data by using each T-RF as a different species and peak height as a measure of species abundance. In the present study, we considered each T-RF to be a different species, and peak area to be the measure of species abundance. A zero indicates no similarity and one indicates complete identity. Complete identity takes into account presence of T-RF and its abundance measured in peak area.

$$I_M = \frac{2\sum n_{1i}n_{2i}}{(l_1 + l_2)N_1N_2} \qquad l = \frac{\sum_{i-1}^{s}(n_i(n_i - 1))}{N(N-1)}$$

T-RFLP Flowchart

Sample Collection-Stored in RNAlater, and kept at -20°C.

↓

DNA Isolation-QIAamp DNA Stool Mini Kit (Qiagen Inc. USA)

↓

PowerClean DNA- DNA Clean-Up kit (MoBio Laboratories Inc. USA)

↓

PCR-Using fluorescently FAM ™-labeled 16S rRNA forward primer *Bact-27F* and an unlabelled reverse primer *Bact-1387* for broad-range bacterial amplification.

↓

Gel Electrophoresis-1.5% agarose gel. A fragment length of approximately 1550 bp was the target length.

↓

PCR Cleanup-UltraClean PCR Clean-Up Kit was used to remove dNTPs and other PCR reagents from products. (MoBio Laboratories Inc. USA)

↓

Restriction Digest-Using enzyme *HhaI*, PCR products were restricted to generate Terminal-Restriction Fragments.

↓

DNA Analyzer-Samples were run at the Virginia Bioinformatics Institute. The Terminal-Restriction Fragments were separated with an automated ABI 3130xl DNA analyzer.

↓

GeneMapper Software v4.0 (Applied Biosystems, Inc. Foster City, CA)-Fluorescently labeled fragments were compared to conclude fragment length and relative abundance of T-RFs within the sample.

MiCA III-This online tool is used to analyze data from the step above. MiCA converts fragment length and abundance measurement to microbial profiles that list genus, and sometimes species located in the sample based on *in silico* digestions.

Statistical Analyses- PCA, an ordination technique, was used to show the variability of fecal microbial files between the samples. JMP v7.0 (JMP for Windows, SAS Institute, Cary, NC) was used to perform this analysis. The Morisita indices were calculated using the Vegan package [Oksanen et al., 2008] in R software [R: Team, 2008] in order to measure community similarities.

Chapter 3-Results

Chapter 3.1: Standardization

Figure 1, located in the Appendix, is an example of a successful PCR of 3 different samples with 4 replicates for each sample. Figure 2 depicts electropherograms from our standardization experiments and is also found in the Appendix. The reproducibility of our T-RFLP protocol is shown in these figures.

Chapter 3.2: T-RF Analysis

Twelve different T-RFLP electropherograms were generated and are found in Figure 3, located in the Appendix. The data outlined in Tables 4 through 15 shows each sample's species profile, however only known species are listed. Many uncultured and unknown species appear in every sample's possible profile, as each T-RF corresponds to one or multiple phyla and several hundreds of uncultured or unidentified species. MiCA was used to determine the phyla represented in each sample. The phylum, class, and where organisms have been isolated from previously was determined using the NCBI accession number generated from MiCA for each possible species.

Number of T-RFs per sample ranged from 1 to 14, with 1, 2, 4, 6, 13 and 14 T-RFs in 1, 3, 5, 1, 1, and 1 sample, respectively. A total of 26 different T-RFs were produced that ranged in size from 92 to 837 bps. Twenty-four T-RFs represented bacterial species contained in five bacterial phyla: *Actinobacteria, Bacteroidetes, Firmicutes, Mollicutes*, and *Proteobacteria*, as well as bacterial species that were uncultured or unidentified. The two remaining T-RFs represented only uncultured or unidentified species. *Firmicutes* was the most common phylum, represented in 11 samples. *Bacteroidetes* and *Mollicutes* were both considered the second-most common phylum, and were detected in 8 of the samples. Samples from MM07-07, a female sub-adult, and MM07-09, a male adult had the most diverse profiles, with 14 and 13 different T-RFs respectively. Additionally, these two samples were the only ones that included all five phyla represented in the sample population. The fecal sample with the least bacterial diversity was from MM07-02, a male juvenile, which only had one fragment in their profile. Average number of T-RFs in each age group is as follows: juvenile, 2.25; sub-adult, 6.5; adult, 6.25. The T-RF lengths found in the most samples were of sizes 100,

188, and 837 bps, corresponding to *Bacteroidetes*, *Firmicutes-Clostridia*, and *Mollicutes*. Each of these T-RFs was found in seven of the samples. There were also 14 unique T-RFs generated (92, 96, 98, 99, 174, 189, 193, 219, 222, 226, 365, 379, 592, and 635). Each of these unique T-RFs was found in one of six samples (MM07-02, MM07-03, MM07-04, MM07-07, MM07-09, and MM07-11).

The T-RF with the highest abundance in the sample population was 837, corresponding to the *Mollicutes* phylum and unidentified or uncultured organisms. This T-RF accounted for 22.5% of the abundance of all T-RFs in all samples. This T-RF also had the highest relative abundance in each of five profiles (MM07-03, MM07-05, MM07-07, MM07-08, and MM07-10).

Chapter 3.3: Statistical Analyses

Principal Components Analysis

Tests concluded that the data was not normally distributed. The data was skewed to the right, with more T-RFs having a low abundance and few having a high abundance. This does not affect the capabilities of PCA to explain the internal structure of the data. A PCA ordination plot of all samples is provided in in the Appendix as Figure 4. A PCA loading plot which descrbes what T-RFs comprise the variability within each of the components is provided in Figure 5. Principal Component 1 (Prin1), Prin2, and Prin3 accounted for 40.8%, 22% and 9.7% of the variation respectively within the data, for a total of 72.5%. According to the PCA method, Prin1 must account for the most variation in the data, and with each successive principal component, the highest remaining amount of variation between samples is calculated. T-RF 100 accounted for most of the variation in Prin1. T-RFs 92, 379, and 635 accounted for most of the variation in Prin2, and T-RF 365 accounted for most of the variation in Prin3 For Prin1 and 2, ten of the samples are clustered just left of center. Only samples from MM07-07 (7) and MM07-09 (9) are considered not part of the cluster. Prin3 has eleven of the samples clustered together, and only MM07-11 (11) is not considered part of the cluster.

Morisita Indices

Table 3 located in the Appendix shows the Morisita indices of all possible combinations of sample within the study population. A total of 66 combinations were

calculated and ranged from 0 to 0.886. Twelve of the combinations were above 0.5 and involved seven of the samples. The mean of the Morisita indices was low, at 0.256.

Chapter 4-Discussion

Chapter 4.1: Phylum Analysis

We used species data generated by MiCA III to assign phyla to T-RF sizes. Overall, most fragment sizes correspond to a specific phylum and class, but at times corresponded to multiple phyla. The average area of each T-RF was used for the relative abundance calculation. Height can also be used for the abundance calculation, however using only the height does not necessarily represent the true abundance of the peak. Several peaks may have the same height but have different areas, corresponding to different relative abundance and therefore this fact was taken into consideration when calculating abundance.

According to the RDP Release 9.60 database, 16 of the 26 T-RFs generated in our samples from the M-Group chimpanzees corresponded exclusively to one of 5 phyla: *Actinobacteria, Bacteroidetes, Firmicutes,* and *Proteobacteria,* and *Mollicutes* . Of the remaining 10 T-RFs, each corresponds to a combination of two or more of these five phyla. Each of these five phyla has been found previously in humans and chimpanzees [Eckburg et al., 2005; Uenishi et al., 2007]. These phyla perform unique functions in humans, and therefore may perform similar functions in chimpanzees. Further research is required to deduce such information for these non-human primates.

Actinobacteria

This phylum represents a very large and diverse taxonomic unit within bacteria. Some species within this phylum are human and animal pathogens, while others are antibiotic-producing soil inhabitants [Maiese et al., 1989]. Additionally, *Bifidobacterium* spp. are a major part of healthy human colonic flora [Moore & Holdeman, 1974]. *Bifidobacterium* spp. were found by partial 16S sequence analysis to also be in both captive and wild chimpanzee feces, suggesting the physiological importance of this bacteria for these endangered animals [Uenishi et al., 2007]. One would then expect to see this phylum in all or most of the samples. It is interesting to note that only 3 of the 12 samples were positive for this phylum in the current study: MM07-11, a sample from her daughter MM07-07, and MM07-09. There are several reasons why this phylum may not have been found in all samples. T-RFLP is mainly used to determine the major phyla present in a sample and therefore the other nine samples could have had this phylum in

their profiles but were simply at low concentrations during the time of collection. It is also possible that each of the nine samples has this phylum in their profile, but was simply part of the uncultured or unidentified portion. Because of the very diverse nature of this phylum, further analysis of the T-RFs must be performed before any conclusion can be stated regarding what functions these organisms might be performing in the host.

Bacteroidetes

The *Bacteroidetes* phylum is one of the largest constituents of gut flora in humans [Eckburg et al., 2005]. T-RFs corresponding to the *Bacteroidetes* phylum were found in eight of the twelve samples. Three of the samples that had lower than detectable abundance of T-RFs corresponding to *Bacteroidetes* within their profile were taken from juveniles, two male and one female. One would expect to find *Bacteroidetes* in every sample in the population, based on the fact that *Bacteroidetes* is one of the two major phyla found in human gut and fecal flora. This phylum was either at a very low abundance in the intestine at the time of collection for the three juveniles, or low amounts were shed for unknown reasons. The other sample that had lower than detectable levels of T-RFs corresponding to *Bacteroidetes* in their microbial profile was MM07-11, an adult female. MM07-11 did not have any positive T-RFs even within 5 bps of a T-RF corresponding to *Bacteroidetes*, suggesting that it was not due to an inconsistency within the DNA sequencing gel that accounted for MM07-11's lack of a T-RF corresponding to this phylum. There is a possibility that some of the uncultured or unidentified species found in MM07-11's profile may correspond to the *Bacteroidetes* phylum and simply have not been sequenced yet.

Firmicutes

Clostridia

Clostridia are a class of bacteria found in normal human fecal flora, and in one study, 95% of the *Firmicutes* bacteria found in the feces of healthy human subjects were part of this class [Eckburg et al., 2005]. Two of the T-RFs (188 and 189) correspond to the *Clostridia* class, and 8 of the 12 samples have 1 of these 2 fragments in their profile. *Clostridia spp.* are therefore one of the most common groups of bacteria in the present sample community. In a previous study, results have shown the predominant presence of *Clostridia* (a plant polymer and sugar-fermenting bacteria) in wild chimpanzees of

Bossou, Guinea [Uenishi et al., 2007]. Therefore it is not surprising to see this bacterial class in the feces of chimpanzees in Mahale, as their diet is mainly vegetarian. Another study examined the prevalence of *Clostridium perfringens* in a captive population from the Primate Research Institute, Kyoto University, and the wild population from the M-group in Mahale [Fujita & Kageyama, 2007]. The study showed a very low prevalence (1 out of 81 samples in 16 different individuals) of this bacterium in the M-group population during the three field seasons (February-March 2001 and 2005, as well as September-October 2001). Further research is necessary to determine if the prevalence of this specific bacterium in the study population has changed since this previous study.

Bacilli

Three T-RFs (226, 581 and 586) are associated solely with the *Bacilli* class, while three additional T-RFs (218, 221, and 225) are associated with *Bacilli, Mollicutes* or *Bacteroidetes.* Two male juveniles (MM07-01 and MM07-02), two female sub-adults (MM07-07 and MM07-08), and one male adult's (MM07-10) samples generated T-RFs corresponding solely to the *Bacilli* class. Large variations in the relative abundance of these T-RFs exist between samples. *Bacilli* can be considered ubiquitous organisms that have been readily found in many vertebrate guts, including primates, other omnivores, carnivores, and herbivores, as well as soil and cultures derived from freshwater and saline habitats [Ley et al., 2008]. It is therefore not unusual to find these organisms within the study population.

Lactobacilli

Two T-RFs correspond solely to this class (222 and 460), and are found in 2 females and 1 male (MM07-03, MM07-07, and MM07-10), each from the 3 different age groups. *Lactobacilli* spp. are of the first groups of bacteria to colonize the guts of humans, and there is evidence that this genus develops antimicrobial compounds that aid in the host's GI system of defense against pathogens [Servin, 2004]. It is unknown if these organisms perform the same functions in chimpanzees. *Lactobacilli* may be considered a common intestinal bacteria in chimpanzees, as this genus has been found in wild and captive populations [Uenishi et al., 2007].

Proteobacteria

Two T-RFs correspond solely to this phyla (97 and 99), while one additional T-RF (92) is associated with either *Proteobacteria* or *Bacteroidetes*. Four samples coming from animals of all age groups and genders (MM07-04, MM07-05, MM07-07, and MM07-09) had T-RFs 97 or 99 associated with their microbial profiles. Other than MM07-04, each animal's relative abundance of these T-RFs was small compared to other T-RFs in their profiles. This is not surprising as facultative species may represent only a very small portion of the bacteria in the strict anaerobic environment of the colon [Eckburg et al., 2005]. The relative abundance of this phylum within MM07-04 may be skewed because she only had 2 positive peaks associated with her profile. By having so few T-RFs generated, the sample's relative abundance may be artificially inflated.

Tenericutes-Mollicutes

Only two T-RFs (219 and 837) are associated solely with this class of bacteria in the samples. There are four other T-RFs (191, 193, 221, and 225) that are associated with *Mollicutes* and other groups including *Bacteroidetes* and *Firmicutes.* Sample MM07-03 is the only sample that had T-RF 219 in their profile, while seven samples had T-RF 837 in their profiles. Some of the species within this class may cause disease, while others are commensals, and the list of hosts that harbor these parasitic bacteria is continuing to grow and includes, humans, other mammals and fish [Razin, 2006]. Additionally, 10 of 13 wild chimpanzees analyzed in a previous study had *Anaeroplasma spp.* in their fecal bacterial profiles, suggesting the possible importance of this class of bacteria for chimpanzees [Uenishi et al., 2007]. Due to T-RF size 837 being widely distributed in the samples as well as being found in other wild chimpanzees, it is assumed that *Mollicutes* could be important commensal bacteria for chimpanzees. However, unless further studies are conducted, this statement is merely conjecture.

Uncultured or Unidentified Organisms

According to MiCA, the majority of possible bacteria in the profile of every sample were associated with uncultured or unidentified bacteria. Every T-RF that corresponds to a specific phylum in the animal samples also has uncultured species associated with it. Two T-RFs of sizes 631 and 635 are associated solely with uncultured or unidentified bacteria. Only two samples (MM07-07 and MM07-09) had T-RF 631 or 635 in their profile, and in both samples, they were present in low relative abundance.

Overall, between 47 - 95% of the possible species in the samples were uncultured or unidentified organisms. However, not all uncultured and unidentified species have been sequenced yet and some may not be included in the RDP II database. Thus, a T-RF may have more uncultured or unidentified species associated with it than another T-RF simply because only certain 16S rRNA genes have been sequenced. We cannot say that a specific sample's profile consists of 95% uncultured or unidentified species. We can say out of all the possible species listed for a certain sample, 95% of them correspond to uncultured or unidentified organisms, however, it is possible that all, some or none of them are located in the sample.

Chapter 4.2: Statistical Analyses

PCA

PCA has been used previously to analyze T-RFLP data based on fecal bacterial profiles of human infants, and is useful for smaller data sets and hypothesis testing [Dollhopf et al. 2001; Wang et al. 2004]. Our data set is small with only 26 different fragments, and PCA ordination demonstrated important aspects of the internal structure of the data. As shown in Figure 4, there is a clustering of the samples indicating that they are in some ways similar to each other. Specifically, 3 of the fragments (100, 188 and 837) were found in 7 of the 12 samples, although the abundance of T-RF 100 differs considerably between samples (115 to 1307 FUs). Further, 3 principal components accounted for 72.5% of the variation and can be explained in part by the presence of 14 unique fragments. When a T-RF is unique, it is found in only one sample and hence, its abundance is greater than zero, whereas the other 11 samples have a value of 0 and are considered to be "similar." These aspects of the data structure contributed to the clustering of 10 of 12 samples in the PCA plot for Prin1 and Prin2.

The two samples (MM07-07 and MM07-09) that were not part of the cluster for Prin1 and Prin2 differed from the other samples in that they had more diverse profiles with 14 and 13 T-RFs respectively, while the average number of T-RFs of the other 10 samples was 3.3, with a range of 1-6 (see Table 2*)*. It is unknown why MM07-07 and MM07-09 had more diverse profiles than the other samples, but we will discuss some of the variables that could account for their differences within the sample population. It is possible that more intestinal flora was present in fecal samples MM07-07 and MM07-09

due to differences in physiological status, and not necessarily because they had more diverse intestinal flora profiles. Additionally, MM07-07 and MM07-09 had the highest number of unique T-RFs with 4 and 3 respectively, and this may contribute to their location outside the cluster. T-RFLP is a PCR-based method, and therefore PCR inhibitors such as humic and phenolic substances found readily in plants can play a role in the success of the protocol based on the mainly vegetarian diet of chimpanzees. It is possible that animals associated with samples MM07-07 and MM07-09 may have a lower amount of PCR inhibitors in their guts brought on by unexplained dietary or physiological reasons. Lower amounts of PCR inhibitors within the samples could allow more DNA to be amplified and subsequently more possible T-RFs generated.

Morisita Indices

Morisita indices were calculated to measure similarity between 2 profiles at a time and a total of 66 combinations were calculated (see Table 3). The mean of all Morisita indices calculated was rather low (0.256), showing that variation exists between sample bacterial profiles. This is concurrent with Eckburg et al.'s study with human intestinal and fecal flora, even though only 3 subjects were used in that study. Twelve combinations were at or above 0.50 and involved 7 different samples. Of these 12 combinations, 8 of them were male-female, 1 of them was male-male, and 3 of them were female-female. Also, 6 of the combinations were adult-sub-adult, 3 of them were sub-adult-sub-adult, 2 of them were sub-adult-juvenile, and 1 was adult-juvenile.

MM07-02 only had 1 T-RF in its profile, and the T-RF happened to be unique in the sample population. Therefore, every index calculated with MM07-02 had to be 0. Of the 20 Morisita indices calculated as 0, 11 of them involved MM07-02. Three of the highest indices involve MM07-10. Even though only 4 T-RFs were in this sample profile, 2 of these were found in 7 of the other samples. Indices involving MM07-10 were therefore rather high due to the fact that half of its T-RFs were common in the population. The highest index calculated (0.886) was between a male adult and a male sub-adult, MM07-10 and MM07-05. Each sample only had 4 T-RFs in their profile, and only 2 T-RFs were common to both profiles. These T-RFs were of length 100 and 837, 2 of the most widespread T-RFs found in the sample population. MM07-10 and MM07-05 had nearly identical peak areas for T-RF 837. This was one of the major features that

accounted for the high index found between these samples. Seven of the samples also had T-RF 837 in their profile, but 6 of them were an order of magnitude lower than MM07-05 or MM07-10. Additionally, the two samples (MM07-07 and MM07-09) that were not part of the PCA clustering had a high Morisita index of 0.71. The sample size was not sufficient to draw conclusions about community similarities based on age group or gender.

The M-group has portions of physiology, environmental sources of bacteria, and most of their diet in common with each other. However, Morisita indices revealed that variations in fecal flora profiles exist within the population and overall, when one sample is compared to another, they are generally not similar. Variation in the deposition of specific intestinal bacteria may therefore exist in chimpanzees as it has been shown in humans.

Chapter 4.3: T-RFLP

The 16S rDNA sequencing databases are updated and released by MiCA periodically, as more and more scientists are using T-RFLP and sequencing analysis for phylogenetic studies. As more data are generated, the number of possible known profiles will continue to grow, and the number of unknown species will possibly diminish. T-RFLP only takes into account the location and number of restriction sites that are located within the 16S rRNA gene and does not give any indication as to what the sequence is *between* the restriction sites. There are a fixed number of possible fragment sizes that can be generated, and therefore many of the same fragment sizes can correspond to multiple phyla and subsequently multiple species. This is one of T-RFLP's shortcomings that have been discussed previously in chapter 1. One way to counteract this is to use combinations of 2 or more restriction enzymes along with different primer sets in order to produce more discrete fragment profiles that correspond only to specific phyla and even possibly just species. However, the selection of an appropriate restriction enzyme(s) when the sequences are unknown is a problem inherent in T-RFLP research. Even using known sequences of gene libraries, many species have shown the same fragment lengths even when an optimal enzyme was selected [Marsh, 1999]. It is for this reason and others that T-RFLP cannot take the place of sequencing analysis. It is simply a more

economical and rapid means of assessing diversity and changes within the major phyla found in different samples.

Chapter 5-Conclusions and Suggested Future Research

5.1 Conclusions

The hypothesis tested was whether or not wild chimpanzees belonging to the same social group would have similar microbial profiles. PCA and Community Analysis were used to test this hypothesis, as well as explain how the profiles differed in phyla diversity and phyla evenness. The results indicate that chimpanzee fecal microbial profiles from members of the same social group vary and are generally not similar, thereby rejecting the hypothesis. The research objective was to generate the most complete view of fecal bacterial diversity within the chimpanzee study population using molecular-based techniques. Given the time allotted for a Masters project, this research objective was met.

There are several possibly important intestinal bacteria that are found in the feces of nearly all the samples tested, namely members of the *Clostridia* class and the *Bacteroidetes* and *Mollicutes* phyla. PCA concluded that measurements of fragment length and abundance were similar within samples from the M-group. Less than 75% of the variation in the data could be explained by three components. Therefore, even though a clustering of 10 of the samples is seen in the data space, conclusions saying that these profiles are similar cannot be made. Much of the "similarity" seen in the PCA was due to the 14 unique T-RFs found. For each T-RF found in only one sample, all the other samples' measurements had the same value, "0." Morisita indices calculated based on pairwise comparisons showed 2 combinations to be very similar based on shared T-RFs and comparable abundance for those shared T-RFs. However, in general, samples were not similar when compared with each other.

T-RFLP has shown to be a rapid, reliable screening method for determinig the major bacterial phyla within fecal samples from wild human-habituated chimpanzees. The data suggests that factors other than similar diet and environment affect the fecal bacterial profiles of the chimpanzee population, and may fluctuate based on unknown various factors. Further research is necessary to determine what bacterial species are found in the samples, as well as the eventual determination of what functions these species are performing in the study population. Analyzing multiple samples from the

same individual over extended time periods as well as documenting dietary intake and behavior may aid in deciphering these roles.

5.2Suggested Future Research

Sequencing Analysis

The eventual determination of bacterial species located in the samples would be beneficial in characterizing species diversity and evenness within the population. Sequencing may also determine that samples that seemed similar may in fact have more variation than previously thought.

Zoonotic and Anthropozoonotic Bacterial Transmission

One method of examining whether or not humans and chimpanzees are exchanging or sharing intestinal flora is through antibiotic resistance testing. A future study with the M-group population could involve researchers and Mahale Mountains Park personnel who are in close contact with the chimpanzees on a regular basis throughout the year. By comparing the fecal flora between chimpanzees, and two human populations (close contact with chimpanzees and no contact with chimpanzees), we can decipher a level of bacterial exchange between and among the populations in close contact with each other. This could aid in certain park regulations for observation and research within the park in order to curb possible pathogen transmission to the endangered chimpanzee.

Age of Colonization and Temporal Dynamics

The intestinal tracts of mammals are highly dynamic with respect to the amounts and types of bacteria located therein at any give time. However, mammals are born with sterile guts; an age in which the intestinal microbial profile reaches maturity with respect to the major phyla should exist. It would be interesting to determine this age in chimpanzees, but it is very difficult to acquire fecal samples from infants. Infants are always very close to their mother and it is difficult to see when they defecate. If we were to acquire samples from multiple infants from multiple time periods, we may be able to determine the age of colonization. Additionally, profiles are constantly changing during this time, and it would be extremely interesting to determine how the profiles fluctuate before, during, and after the weaning period.

Additionally, if we could get multiple samples from any individual over the course of several wet and dry seasons, we might be able to determine the extent to which fecal bacterial profiles fluctuate in wild chimpanzees. Also, a distinction between adherent and nonadherent intestinal bacteria may be determined.

Other Vertebrates

It would be interesting to reveal how similar the chimpanzee microbial profiles are with other vertebrates in the park including several species of monkey. Vervet monkeys and baboons tolerate human presence in the park and would be great candidates for gut flora research.

References

Abdo Z, Schuette UM, Bent SJ, Williams CJ, Forney LJ, Joyce P. 2006. Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environ Microbiol 8(5):929-38.

Ash C, Farrow JA, Dorsch M, Stackebrandt E, Collins MD. 1991. Comparative analysis of Bacillus anthracis, Bacillus cereus, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. Int J Syst Bacteriol 41(3):343-6.

Bonnet R, Suau A, Dore J, Gibson GR, Collins MD. 2002. Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. Int J Syst Evol Microbiol 52(Pt 3):757-63.

Brower JE, Zar JH, von Ende CN. 1990. Field and Laboratory Methods for General Ecology: WC Brown. 237 p.

Daffonchio D, Cherif A, Brusetti L, Rizzi A, Mora D, Boudabous A, Borin S. 2003. Nature of polymorphisms in 16S-23S rRNA gene intergenic transcribed spacer fingerprinting of Bacillus and related genera. Appl Environ Microbiol 69(9):5128-37.

Danovaro R, Luna GM, Dell'anno A, Pietrangeli B. 2006. Comparison of two fingerprinting techniques, terminal restriction fragment length polymorphism and automated ribosomal intergenic spacer analysis, for determination of bacterial diversity in aquatic environments. Appl Environ Microbiol 72(9):5982-9.

Demuth JP, De Bie T, Stajich JE, Cristianini N, Hahn MW. 2006. The evolution of mammalian gene families. PLoS ONE 1:e85.

Dollhopf SL, Hashsham SA, Tiedje JM. 2001. Interpreting 16S rDNA T-RFLP Data: Application of Self-Organizing Maps and Principal Component Analysis to Describe Community Dynamics and Convergence. Microb Ecol 42(4):495-505.

Dunbar J, Ticknor LO, Kuske CR. 2001. Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. Appl Environ Microbiol 67(1):190-7.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. 2005. Diversity of the human intestinal microbial flora. Science 308(5728):1635-8.

Frey JC, Rothman JM, Pell AN, Nizeyi JB, Cranfield MR, Angert ER. 2006. Fecal bacterial diversity in a wild gorilla. Appl Environ Microbiol 72(5):3788-92.

Fujita S, Kageyama T. 2007. Polymerase chain reaction detection of Clostridium perfringens in feces from captive and wild chimpanzees, Pan troglodytes. J Med Primatol 36(1):25-32.

Gilmore MS, Ferretti JJ. 2003. Microbiology. The thin line between gut commensal and pathogen. Science 299(5615):1999-2002.

Goldberg TL, Gillespie TR, Rwego IB, Wheeler E, Estoff EL, Chapman CA. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. Biological Conservation 135:511-517.

Gray MW, Sankoff D, Cedergren RJ. 1984. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. Nucleic Acids Res 12(14):5837-52.

Hartmann M, Widmer F. 2008. Reliability for detecting composition and changes of microbial communities by T-RFLP genetic profiling. FEMS Microbial Ecology 63:249-260.

Hooper LV, Gordon JI. 2001. Commensal host-bacterial relationships in the gut. Science 292(5519):1115-8.

Hooper LV, Xu J, Falk PG, Midtvedt T, Gordon JI. 1999. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc Natl Acad Sci U S A 96(17):9833-8.

Johnson JR, O'Bryan TT. 2000. Improved repetitive-element PCR fingerprinting for resolving pathogenic and nonpathogenic phylogenetic groups within Escherichia coli. Clin Diagn Lab Immunol 7(2):265-73.

Jonasson J, Olofsson M, Monstein HJ. 2002. Classification, identification and subtyping of bacteria based on pyrosequencing and signature matching of 16S rDNA fragments. APMIS 110(3):263-72.

Kagnoff MF, Eckmann L. 1997. Epithelial cells as sensors for microbial infection. J Clin Invest 100(1):6-10.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol 6(10):776-88.

Liu WT, Marsh TL, Cheng H, Forney LJ. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63(11):4516-22.

Maiese WM, Lechevalier MP, Lechevalier HA, Korshalla J, Kuck N, Fantini A, Wildey MJ, Thomas J, Greenstein M. 1989. Calicheamicins, a novel family of antitumor antibiotics: taxonomy, fermentation and biological properties. J Antibiot (Tokyo) 42(4):558-63.

Mandar R, Mikelsaar M. 1996. Transmission of mother's microflora to the newborn at birth. Biol Neonate 69(1):30-5.

Marsh TL. 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Curr Opin Microbiol 2(3):323-7.

Mazmanian SK, Round JL, Kasper DL. 2008. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 453(7195):620-5.

Moeseneder MM, Arrieta JM, Muyzer G, Winter C, Herndl GJ. 1999. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. Appl Environ Microbiol 65(8):3518-25.

Moore WE, Holdeman LV. 1974. Special problems associated with the isolation and identification of intestinal bacteria in fecal flora studies. Am J Clin Nutr 27(12):1450-5.

Nunan N, Daniell TJ, Singh BK, Papert A, McNicol JW, Prosser JI. 2005. Links between plant and rhizoplane bacterial communities in grassland soils, characterized using molecular techniques. Appl Environ Microbiol 71(11):6784-92.

Oksanen J, Kindt R, Legendre P, O'Hara N, Simpson G, Solymos P, Stevens M, Wagner H. 2008. vegan: Community Ecology Package. R package version 1.15-1. http://cran.r-project.org/, http://vegan.r-forge.r-project.org/.

Olsen GJ, Woese CR. 1993. Ribosomal RNA: a key to phylogeny. FASEB J 7(1):113-23.

R Core Development Team.2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Razin S. 2006. The Genus *Mycoplasma* and Related Genera (Class Mollicutes). Prokaryotes 4:836-904.

Schutte UM, Abdo Z, Bent SJ, Shyu C, Williams CJ, Pierson JD, Forney LJ. 2008. Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. Appl Microbiol Biotechnol 80(3):365-80.

Servin AL. 2004. Antagonistic activities of lactobacilli and bifidobacteria against microbial pathogens. FEMS Microbiol Rev 28(4):405-40.

Smalla K, Oros-Sichler M, Milling A, Heuer H, Baumgarte S, Becker R, Neuber G, Kropf S, Ulrich A, Tebbe CC. 2007. Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: Do the different methods provide similar results? J Microbiol Methods.

Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD, Dore J. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl Environ Microbiol 65(11):4799-807.

Tilsala-Timisjarvi A, Alatossava T. 2001. Characterization of the 16S–23S and 23S–5S rRNA intergenic spacer regions of dairy propionibacteria and their identification with species-specific primers by PCR. International Journal of Food Microbiology 68:45-52.

Uenishi G, Fujita S, Ohashi G, Kato A, Yamauchi S, Matsuzawa T, Ushida K. 2007. Molecular analyses of the intestinal microbiota of chimpanzees in the wild and in captivity. Am J Primatol 69(4):367-76.

Van de Peer Y, Chapelle S, De Wachter R. 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res 24(17):3381-91.

Wang M, Ahrne S, Antonsson M, Molin G. 2004. T-RFLP combined with principal component analysis and 16S rRNA gene sequencing: an effective strategy for comparison of fecal microbiota in infants of different ages. J Microbiol Methods 59(1):53-69.

Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A 95(12):6578-83.

Wilson KH, Blitchington RB. 1996. Human colonic biota studied by ribosomal DNA sequence analysis. Appl Environ Microbiol 62(7):2273-8.

Woese CR, Fox GE, Zablen L, Uchida T, Bonen L, Pechman K, Lewis BJ, Stahl D. 1975. Conservation of primary structure in 16S ribosomal RNA. Nature 254(5495):83-6.

Woods CR, Versalovic J, Koeuth T, Lupski JR. 1993. Whole-cell repetitive element sequence-based polymerase chain reaction allows rapid assessment of clonal relationships of bacterial isolates. J Clin Microbiol 31(7):1927-31.
Wyman AR, White R. 1980. A highly polymorphic locus in human DNA. Proc Natl Acad Sci U S A 77(11):6754-8.

Appendix

Figure 1: All gels consisted of 1.5% agarose and were run at 80 mv for 65 minutes. Three samples are depicted, samples 1, 3 and 6. All samples show a band with the correct size at approximately 1550-1600 bps.

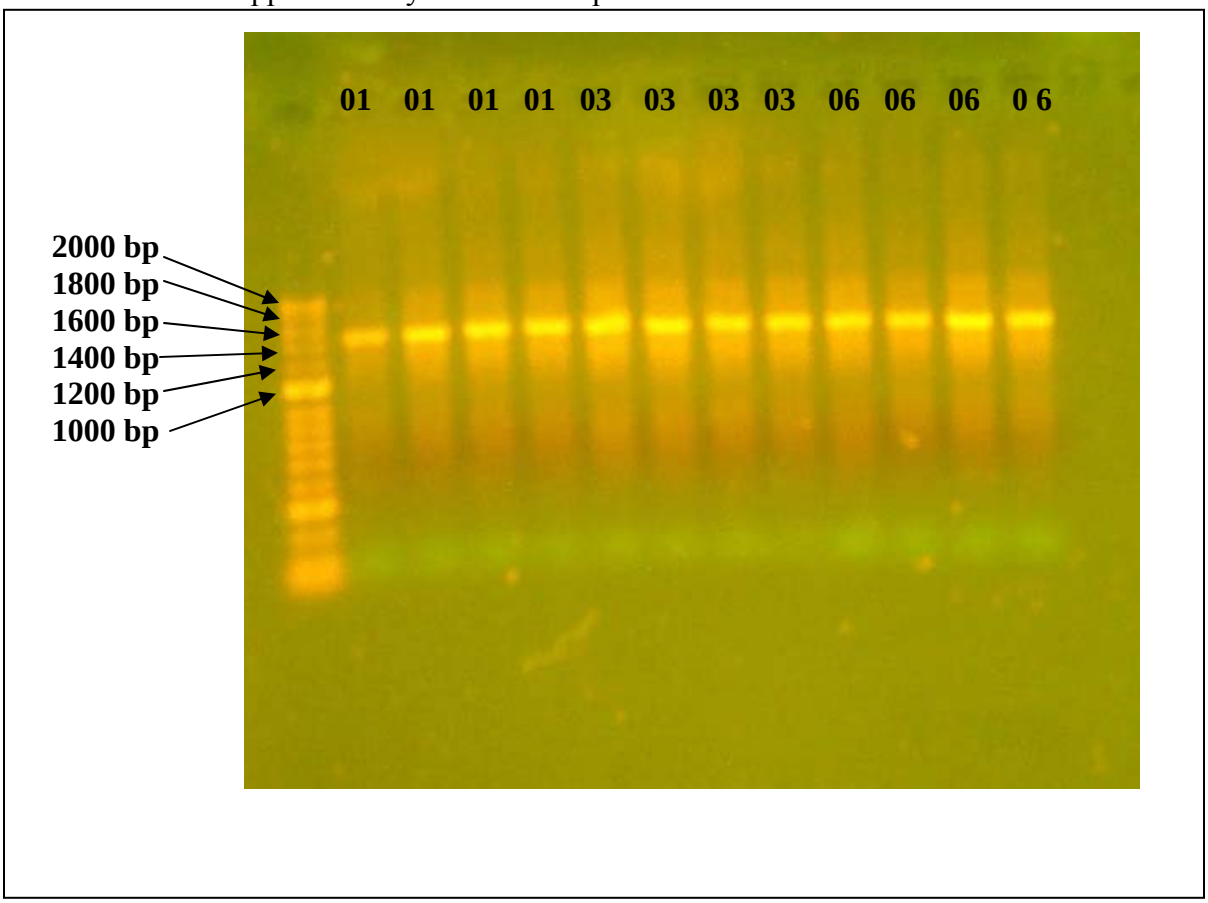
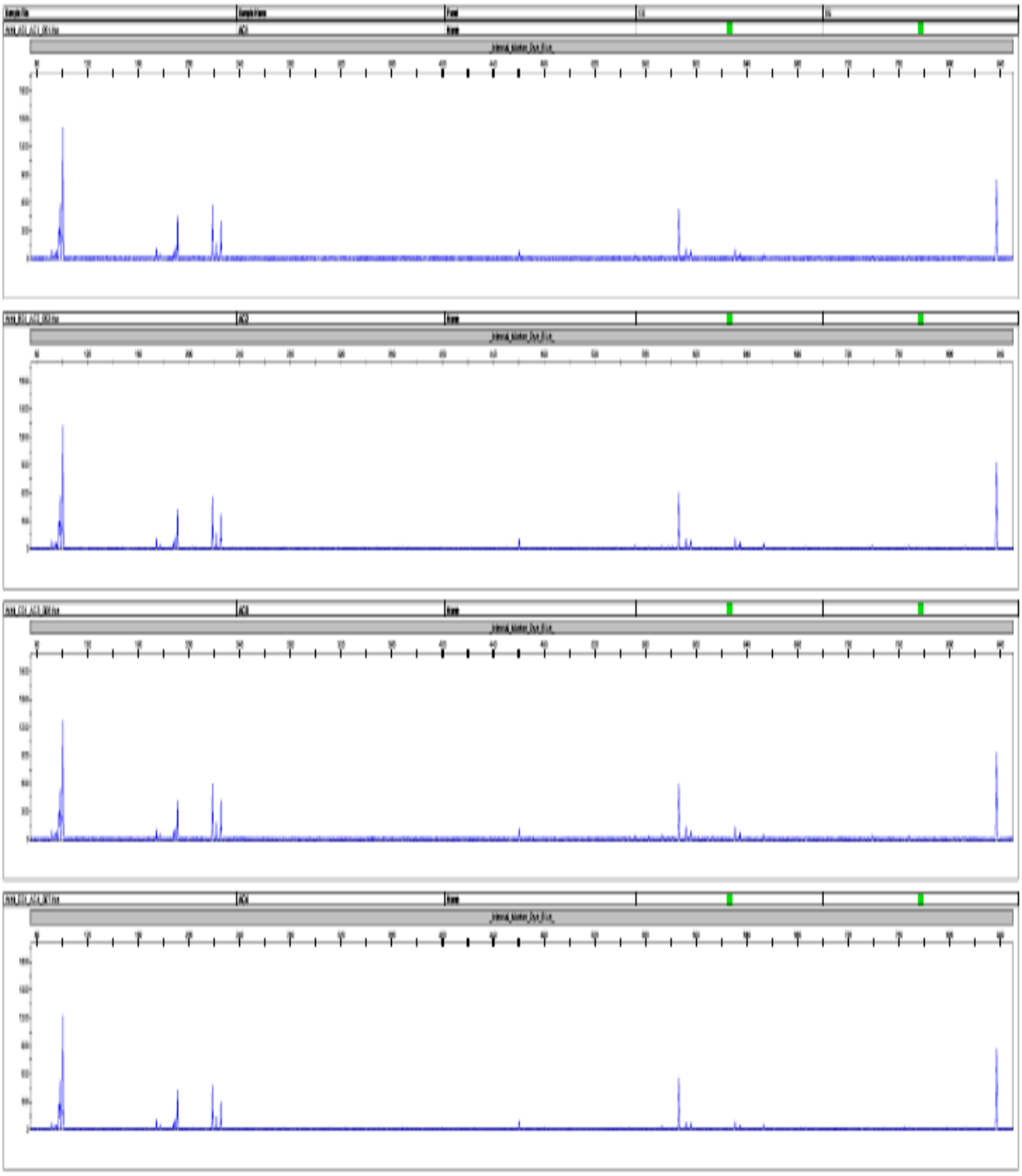


Figure 2: Electropherograms of four replicates from MM07-07 were used for standardization experiments. The x-axis is T-RF length measured in bps, while the y-axis is T-RF abundance measured in arbitrary FUs. A summary of MM07-07's data is outlined in Table 2.



Fluorescence Intensity (Arbitrary Fluorescent Units)

Terminal-Restriction Fragment Length (base pairs)

Figure 3 Sample Terminal-Restriction Fragment Length Polymorphism (T-RFLP) electropherograms for each of the three age groups studied. (A) Juveniles, (B) Sub-Adults, (C) Adults. Profiles were created using the restriction enzyme *HhaI*. The x-axis signifies T-RF length measured in base pairs, while the y-axis signifies fluorescence intensity measured in arbitrary units. Only peaks at or above 100 units are considered positive.

A

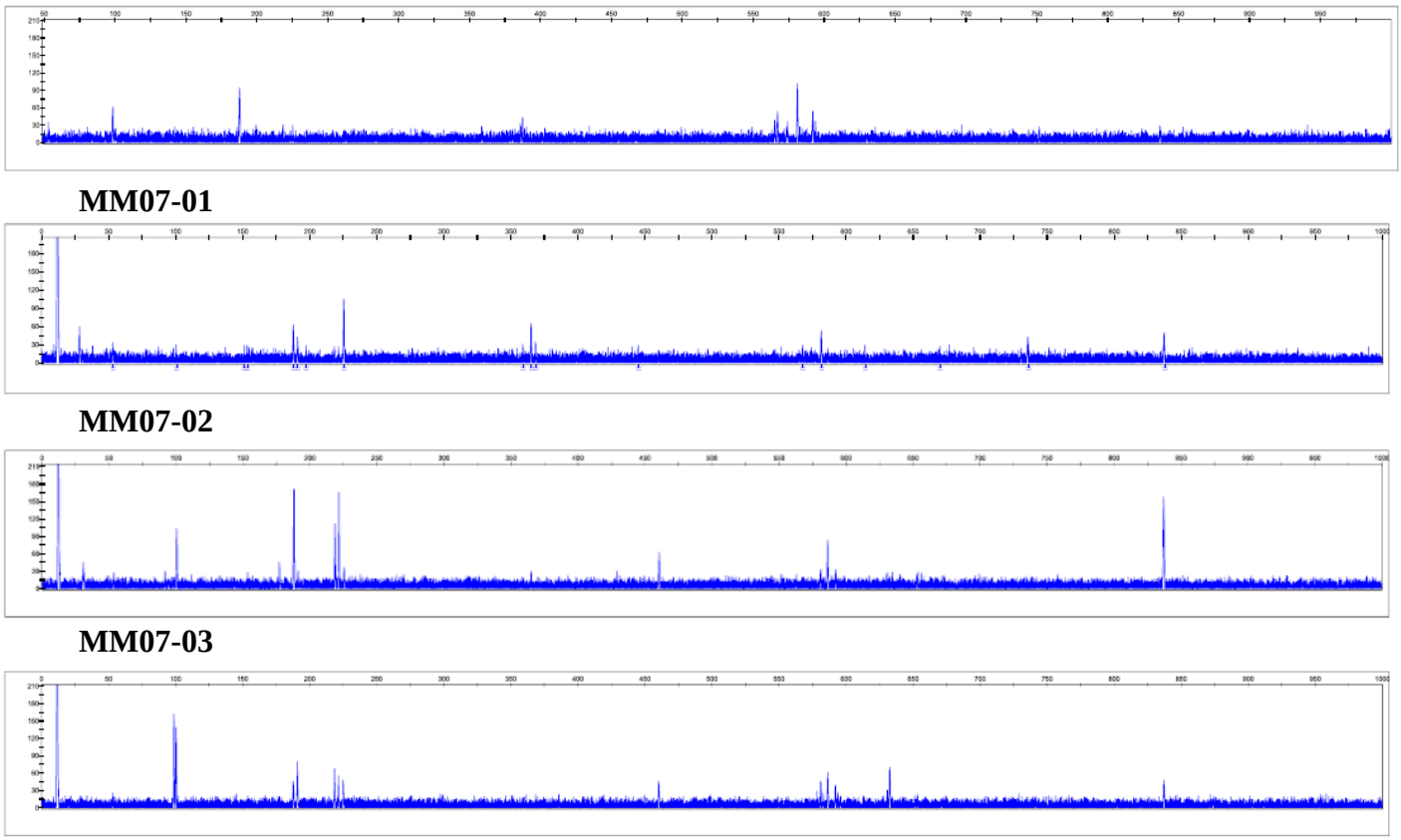


Figure 3

B

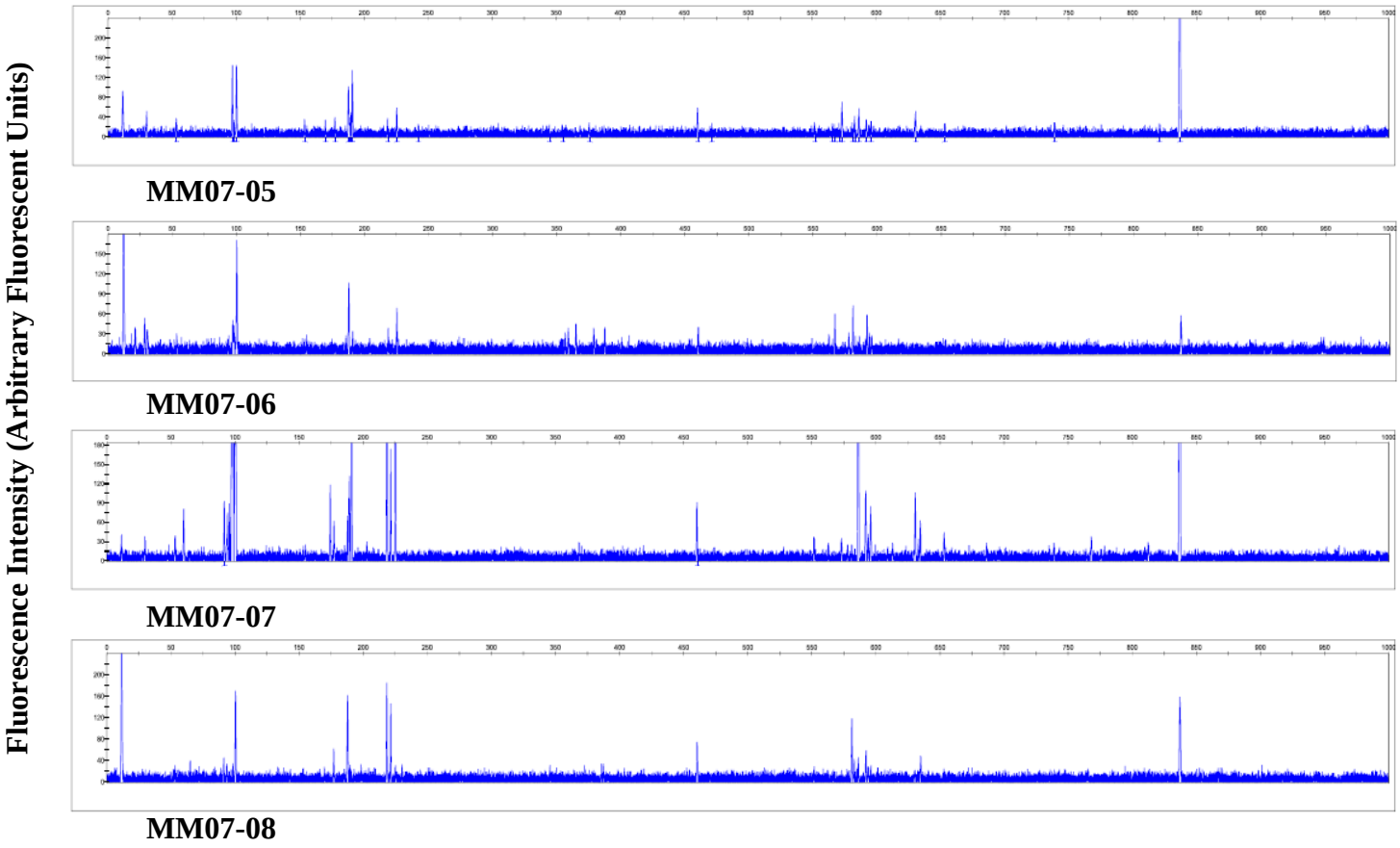

Fluorescence Intensity (Arbitrary Fluorescent Units)
MM07-05
MM07-06
MM07-07
MM07-08
Terminal-Restriction Fragment Length (base pairs)

Figure 3

C

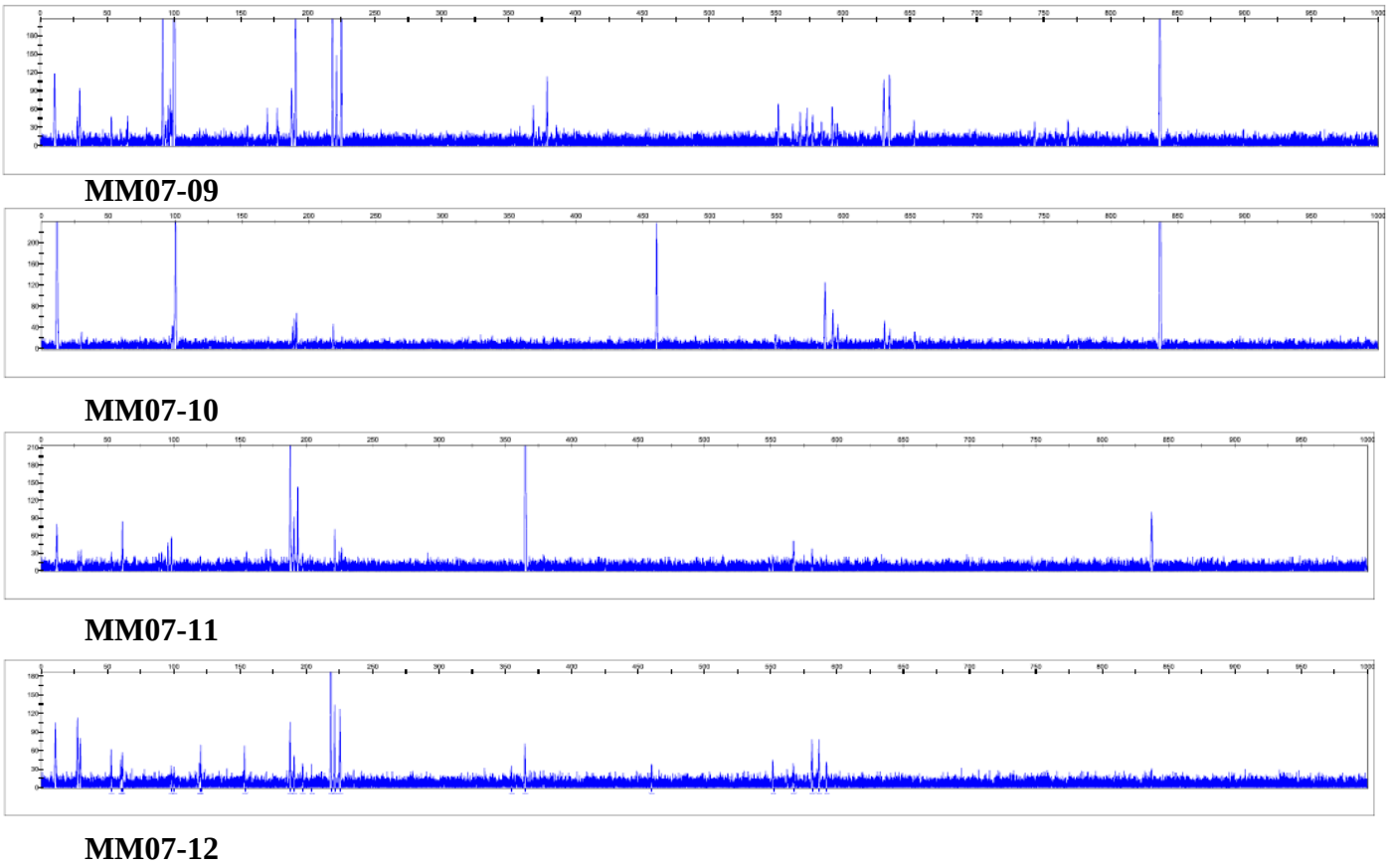
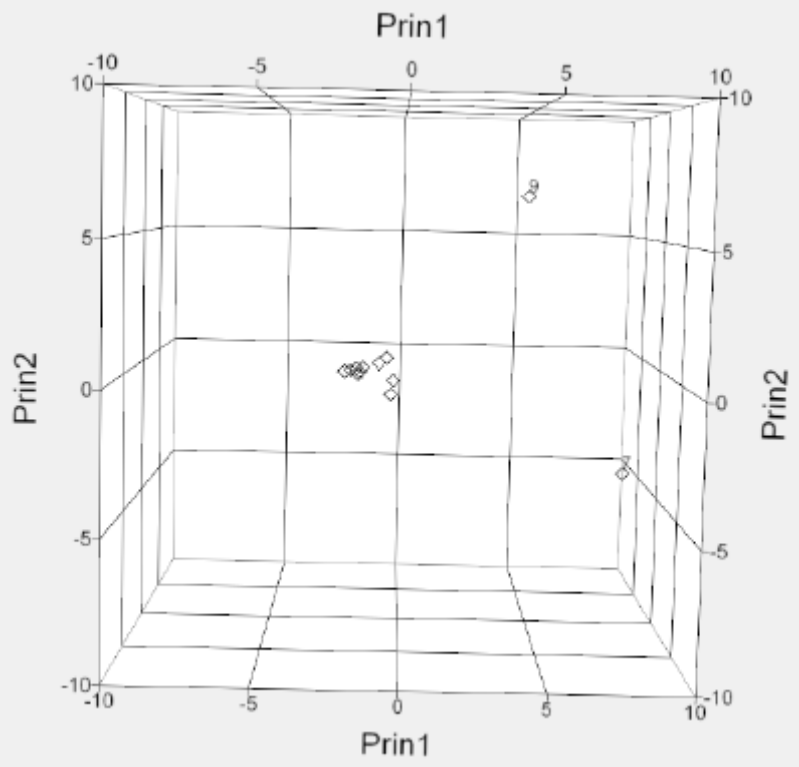

MM07-09
MM07-10
MM07-11
MM07-12
Fluorescence Intensity (Arbitrary Fluorescent Units)
Terminal-Restriction Fragment Length (base pairs)

Figure 4. Principle Component Analysis (PCA) ordination plots of bacterial T-RFLP peak area data for all 12 samples. Principle Component 1 (Prin1) accounted for 40.8% of the variation within the data, while Prin2 accounted for 22%, and Prin3 accounted for 9.7%, for a total of 72.5% of the variation within the data. (A) shows the clustering of 10 of the 12 samples along Prin1and Prin2, while (B) shows the clustering of 11 of the 12 samples along Prin3.

A



B

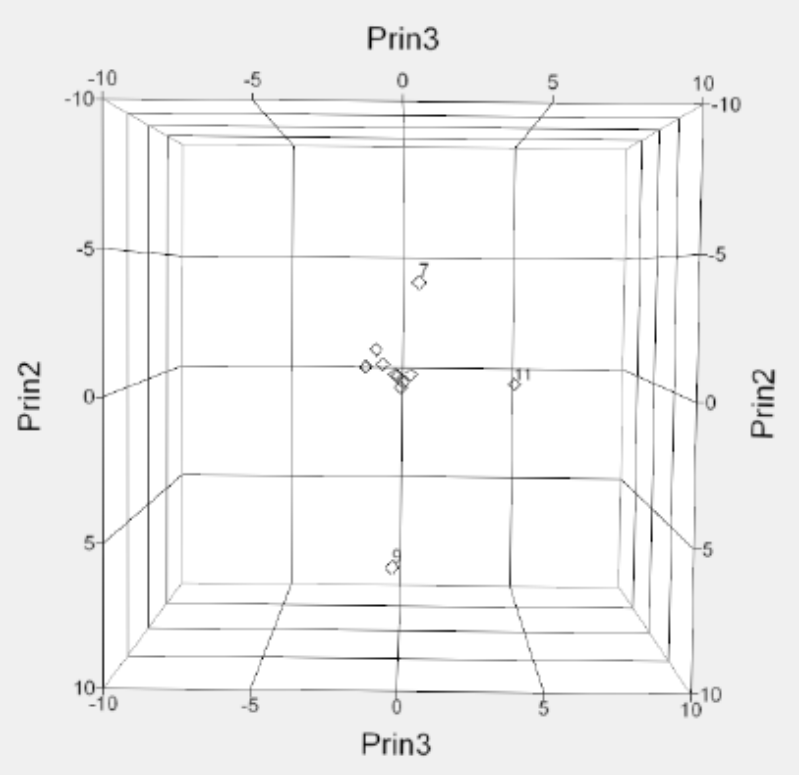


Figure 5- Loading Plot- The Loading Plot shows the relative importance of each T-RF in determining Morisita similarity indices. The circled cluster consists of T-RF sizes: 99, 193, 219, 222, 226, 365 and 581. These T-RFs are close to zero for both component 1 and 2, showing that they contributed very little to the variation in the data for these components. These highlighted T-RFs are either unique or only found in two individuals and therefore attributed to a large amount of "similarity" within the samples, as nearly all other samples had a similar value 0 for these T-RFs. For all components, the T-RFs that contributed more to the variation within the data are farther away from the zero axis of a particular component.

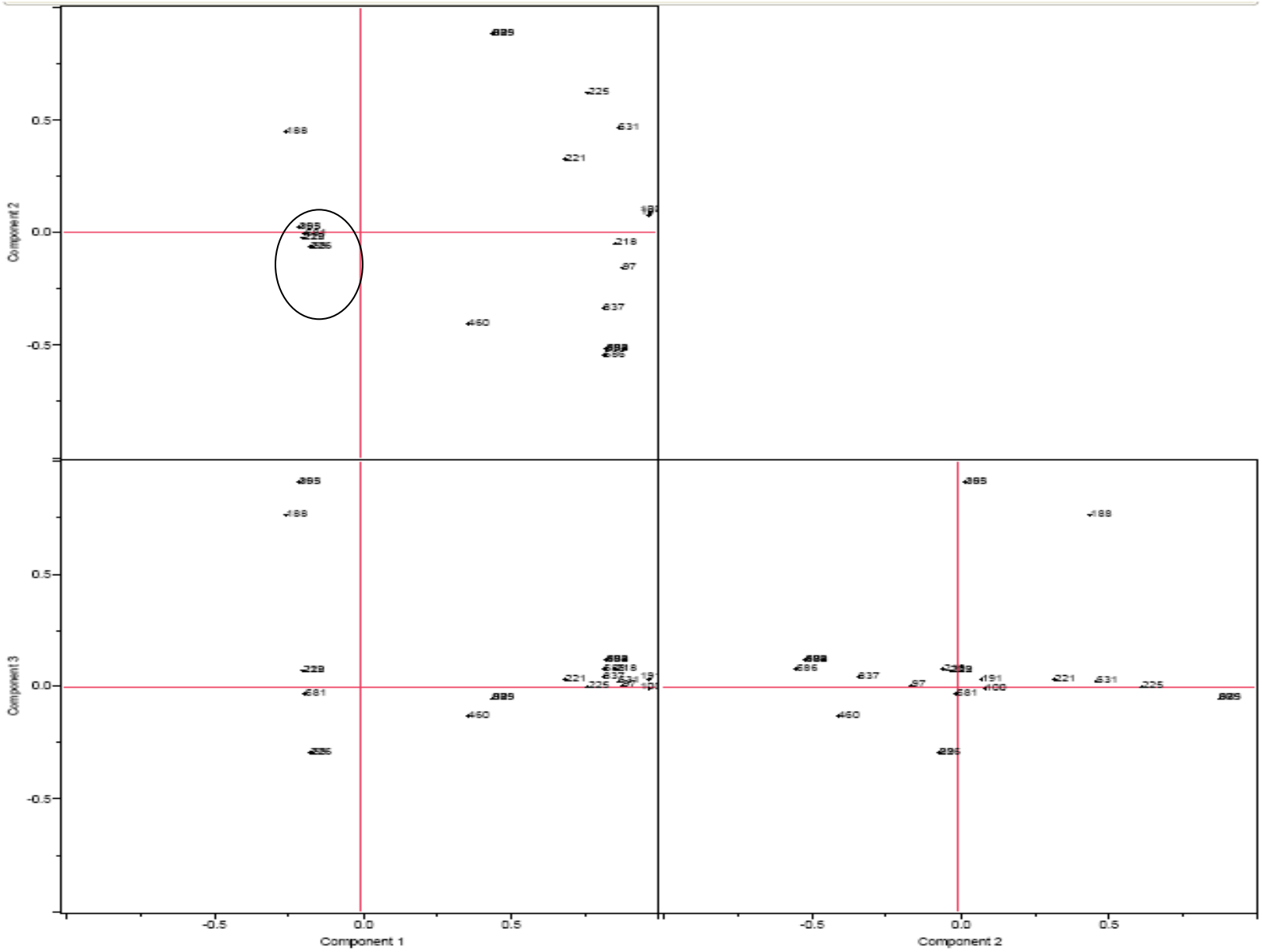


Table 2*:* Terminal-Restriction Fragment (T-RF) profile for 12 chimpanzee fecal samples

Age Group	Gender	MM #	Total T-RFs [a]	T-RF Length in Base Pairs (Relative Abundance)[b]	Representative Phyla for each T-RF
Juveniles	**Male**	**07-01**	2	**188** (0.47), **581** (0.53)	Firmicutes-Clostridia, Firmicutes-Bacilli, Uncultured
	Male	**07-02**	1	**226** (1.0)	Firmicutes-Bacillales, Uncultured
	Female	**07-03**	4	**188** (0.28), **219** (0.17), **222** (0.21), **837** (0.34)	Firmicutes-Clostridia, Mollicutes, Firmicutes-Lactobacilli, Uncultured
	Female	**07-04**	2	**99** (0.56), **100** (0.44)	Proteobacteria, Bacteroidetes, Uncultured
Sub-Adults	**Male**	**07-05**	4	**97** (0.13), **100** (0.12), **191** (0.10), **837** (0.65)	Proteobacteria, Bacteroidetes, Firmicutes-Clostridia, Mollicutes, Uncultured
	Male	**07-06**	2	**100** (0.58), **188** (0.42)	Bacteroidetes, Firmicutes-Clostridia, Uncultured
	Female	**07-07**	14	**97** (0.04), **98** (0.10), **100** (0.19), **174** (0.02), **189** (0.02), **191** (0.06), **218** (0.08), **221** (0.02), **225** (0.06), **460** (0.02), **586** (0.13), **592** (0.03), **631** (0.02), **837** (0.22)	Proteobacteria, Bacteroidetes, Actinobacteria, Firmicutes-Clostridia, Mollicutes, Firmicutes-Bacilli, Firmicutes-Lactobacilli, Firmicutes, Uncultured
	Female	**07-08**	6	**100** (0.15), **188** (0.15), **218** (0.17), **221** (0.12), **581** (0.15), **837** (0.25)	Bacteroidetes, Firmicutes-Clostridia, Firmicutes-Bacilli, Mollicutes, Uncultured
Adults	**Male**	**07-09**	13	**92** (0.06), **96** (0.02), **97** (0.03), **100** (0.22), **188** (0.04), **191** (0.07), **218** (0.06), **221** (0.04), **225** (0.19), **379** (0.04), **631** (0.05), **635** (0.06), **837** (0.11)	Bacteroidetes, Proteobacteria, Firmicutes-Clostridia, Mollicutes, Firmicutes-Bacilli, Actinobacteria, Uncultured
	Male	**07-10**	4	**100** (0.19), **460** (0.17), **586** (0.12), **837** (0.52)	Bacteroidetes, Firmicutes-Lactobacilli, Firmicutes-Bacilli, Uncultured, Mollicutes
	Female	**07-11**	4	**188** (0.23), **193** (0.13), **365** (0.49), **837** (0.14)	Firmicutes-Clostridia, Mollicutes, Actinobacteria, Uncultured, Mollicutes
	Female	**07-12**	4	**188** (0.17), **218** (0.39), **221** (0.20), **225** (0.24)	Firmicutes-Clostridia, Bacteroidetes, Firmicutes-Bacilli, Mollicutes, Uncultured

[a] T-RFs were considered positive when 3 out of the 4 replicates from a given sample are equal to greater than 100 Fluorescent Units (FU) in height.

[b] Average area of positive peaks for each T-RF divided by the total area of all positive peaks for all T-RFs in the sample.

Table 3. Morisita index of community similarity is shown for each combination of fecal samples. Calculations are based on comparison of the Simpson's Dominance Index and similarities in T-RF length and abundance between 2 samples.

Sample MM:	07-02	07-03	07-04	07-05	07-06	07-07	07-08	07-09	07-10	07-11	07-12
07-01	0[a]	0.35	0[a]	0[a]	0.39	0[a]	0.45	0.06	0[a]	0.26	0.21
07-02		0[a]	0[a]	0[a]	0[a]	0[a]	0[a]	0[a]	0[a]	0[a]	0[a]
07-03			0[a]	0.6	0.31	0.38	0.57	0.25	0.57	0.38	0.18
07-04				0.10	0.49	0.26	0.20	0.31	0.19	0[a]	0[a]
07-05					0.14	0.60	0.56	0.37	0.89	0.23	0[a]
07-06						0.34	0.45	0.45	0.25	0.23	0.19
07-07							0.66	0.71	0.71	0.14	0.25
07-08								0.56	0.60	0.27	0.52
07-09									0.42	0.11	0.43
07-10										0.22	0[a]
07-11											0.13
07-12											

[a] 0 indicates that no T-RF of the same length was common between the 2 samples, and 1 indicates that T-RF length and abundance for all fragments were identical in the 2 samples.

Table 4- MM07-01 Species Profile

Phylum/Class	Name	Accession	Isolation
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Firmicutes/Bacilli	Atopostipes suicloacalis (T) PPC79.	AF445248	swine manure
Firmicutes/Bacilli	Enterococcus faecium ML4.	AB232954	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Lactococcus garvieae JCM 8735.	AB012306	shrimp freshwater pathogen
Firmicutes/Bacilli	Paenibacillus peoriae (T) IFO15541T.	D78476	Soil
Firmicutes/Bacilli	Streptococcus agalactiae.	AF459432	human pathogen
Firmicutes/Bacilli	Streptococcus bovis.	M58835	human pathogen
Firmicutes/Bacilli	Streptococcus canis CG40.	AJ413205	human pathogen
Firmicutes/Bacilli	Streptococcus didelphis (T) W94-11374-1.	AF176103	opossum pathogen
Firmicutes/Bacilli	Streptococcus dysgalactiae subsp. dysgalactiae.	AB159678	Swine
Firmicutes/Bacilli	Streptococcus gallolyticus subsp. gallolyticus ATCC 43143.	AF104114	human and ruminal
Firmicutes/Bacilli	Streptococcus iniae (T) ATCC29178.	AF335572	fish pathogen
Firmicutes/Bacilli	Streptococcus parasanguinis mother C3 5C3.	AM157421	human breast milk
Firmicutes/Bacilli	Streptococcus pyogenes MGAS10394.	CP000003	human PATHOGEN
Firmicutes/Bacilli	Streptococcus salivarius.	AF459433	human PATHOGEN
Firmicutes/Bacilli	Streptococcus sanguinis ChDC B203.	AF543281	human oral
Firmicutes/Bacilli	Streptococcus thermophilus.	DQ176426	human GI
Firmicutes/Bacilli	Streptococcus vestibularis ATCC 49124.	AY188353	human oral
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core
Proteobacteria	Photobacterium damselae UCP4.	DQ530294	fish pathogen
Proteobacteria	Salinivibrio costicola subsp. costicola GSP12.	AY505534	Great Salt Plains

Table 5- MM07-02 Species Profile

Phylum/Class	Name	Accession	Isolation
Bacteroidetes	Chryseobacterium indoltheticum (T).	M58774	marine mud
Bacteroidetes	Chryseobacterium taiwanensis (T) BCRC 17412.	DQ318789	soil
Bacteroidetes	Flavobacterium sp. Smarlab BioMol-2300973.	AY230767	human tissue
Bacteroidetes	Riemerella anatipestifer 11693.	AY871829	duck/turkey PATHOGEN
Firmicutes	Coprobacillus cateniformis JCM 10603.	AB030218	human feces
Firmicutes/Bacillales	Alicyclobacillus disulfidooxidans DSM 12064.	AB089843	mixed fruit juice
Firmicutes/Bacillales	Paenibacillus ourofinensis AC13MSD.	EU257517	cerrado soil
Firmicutes/Clostridia	Anaerobacter polyendosporus.	AJ222546	soil
Firmicutes/Clostridia	Clostridium cellulovorans DSM 3052.	X73438	poplar wood
Firmicutes/Clostridia	Clostridium pasteurianum (T).	M23930	soil
Proteobacteria	Ancylobacter aquaticus (T).	M62790	creek
Proteobacteria	Ancylobacter rudongensis (T) AS1.1761.	AY056830	roots of Spartina anglica
Proteobacteria	Haliangium ochraceum GS1.	EF108312	coastal sands
Proteobacteria	Halomonas korlensis XK1.	EU085033	saline/alkaline soil
Proteobacteria	Rhodomicrobium vannielii.	M34127	hot spring
Tenericutes/Mollicutes	Alfalfa witches-broom phytoplasma AlfWB.	EF193360	alfalfa plants
Tenericutes/Mollicutes	Anaeroplasma abactoclasticum (T).	M25050	cattle rumen
Tenericutes/Mollicutes	Cactus phytoplasma Martinez-Soriano 1999.	AF200718	cactus
Tenericutes/Mollicutes	Candidatus Phytoplasma brasiliense (T) HibWB26.	AF147708	plants
Tenericutes/Mollicutes	Cotton phyllody phytoplasma CoP.	EF186827	cotton
Tenericutes/Mollicutes	Mycoplasma genitalium (T) G37.	X77334	human origin, parasitic PATHOGEN
Tenericutes/Mollicutes	Mycoplasma lipophilum (T).	M24581	human/rhesus monkey oral
Tenericutes/Mollicutes	Mycoplasma pneumoniae (T).	M29061	Human pathogen
Tenericutes/Mollicutes	Mycoplasma sphenisci UCMJ.	AY756171	Choana of an Aquarium-Reared Jackass Penguin PATHOGEN
Tenericutes/Mollicutes	Sesame phyllody phytoplasma SIL.	EU072505	sesame
Tenericutes/Mollicutes	Soybean phyllody phytoplasma SOYP.	EF193353	soybean
Tenericutes/Mollicutes	Tomato big bud phytoplasma TBB.	EF193359	tomato

Table 6- MM07-03 Species Profile

Phylum/Class	Name	Acession	Isolation
Actinobacteria	Amycolatopsis orientalis subsp. orientalis (T) ssp. orien...	X76958	human clinical isolated
Actinobacteria	Patulibacter minatonensis (T) KV-614.	AB193261	Soil
Actinobacteria	Streptomyces nodosus ATCC14899.	AF114033	River
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Firmicutes	Alicyclobacillus acidoterrestris TK25.	EF095718	hot springs
Firmicutes	Alicyclobacillus tolerans K1.	Z21979	Environment
Firmicutes	Bacillus mannanilyticus (T) AM-001.	AB043864	Environment
Firmicutes	Caldaterra yamamurae YMO722.	AB308475	Compost
Firmicutes	Clostridium spiroforme DSM 1552.	X73441	rabbit cecum
Firmicutes	Lactobacillus delbrueckii (T).	M58814	Naturally fermented foods
Firmicutes/Bacilli	Enterococcus asini (T) AS2.	Y11621	donkey cecum
Firmicutes/Bacilli	Enterococcus avium (T) CIP 103 019.	AF133535	rare human PATHOGEN
Firmicutes/Bacilli	Enterococcus durans D1.	DQ239692	human clinical sample
Firmicutes/Bacilli	Enterococcus faecalis 47/3.	EF653454	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus faecium C228.	AB246407	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus mundtii HDYM-33.	EF428252	human clinical sample-sinus and abscess
Firmicutes/Bacilli	Streptococcus mutans 669.	AF139603	human dental PATHOGEN
Firmicutes/Clostridia	Clostridium acetobutylicum ATCC 824.	AE001437	Soil
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	Grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	Human PATHOGEN
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Ruminococcus obeum 1-4.	AY169411	ice core
Proteobacteria	Candidatus Glomeribacter gigasporarum.	AJ251636	Fungi
Proteobacteria	Vibrio fluvialis CIFAMVIFL01.	DQ683079	human feces pathogen
Tenericutes/Mollicutes	Acholeplasma laidlawii.	M23932	sewage, manure, humus, soil, parasite in mammals

Tenericutes/Mollicutes	Acholeplasma oculi ISM1499.	U14906	conjunctivae of sheep with contagious ophthalmia
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine respiratory PATHOGEN
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma opalescens MHS408.	AF538961	canine pathogen
Tenericutes/Mollicutes	Mycoplasma zalophi 4296C.	AF493543	PATHOGEN California sea lions
Tenericutes/Mollicutes	Spiroplasma citri Fewa.	AM157768	Citrus sinensis leaves in Egypt

Table 7- MM07-04 Species Profile

Phylum/Class	Name	Accession	Isolation
Bacteroidetes	Alkaliflexus imshenetskii (T) type strain: Z-7010.	AJ784993	soda lake
Bacteroidetes	Bacteroides cf. forsythus oral clone BU063.	AY008308	subgingival plaque human
Bacteroidetes	Bacteroides intestinalis (T) JCM 13265 341.	AB214328	human feces
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides splanchnicus (T).	L16496	normal human colonic flora
Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Dokdonia donghaensis MED134.	DQ481462	sea water
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Gramella forsetii KT0803.	CU207366	concentrated seawater
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Bacteroidetes	Parabacteroides distasonis JCM 5825.	AB238922	symbiotic distal human intestine
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge

Table 8- MM07-05 Species Profile

Phylum/Class	Name	Accession	Isolation
Bacteroidetes	Alkaliflexus imshenetskii (T) type strain: Z-7010.	AJ784993	soda lake
Bacteroidetes	Bacteroides intestinalis (T) JCM 13265 341.	AB214328	human feces
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides splanchnicus (T).	L16496	normal human colonic flora
Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Cytophaga fermentans (T).	M58766	river epilithon
Bacteroidetes	Dokdonia donghaensis MED134.	DQ481462	marine
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Gramella forsetii KT0803.	CU207366	concentrated seawater
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Bacteroidetes	Microscilla arenaria.	M60455	sand
Bacteroidetes	Parabacteroides distasonis ATCC 8503.	CP000140	symbiotic distal human intestine
Bacteroidetes	Perexilibacter aurantiacus Shu-F-UV2-2.	AB276355	sediment
Bacteroidetes	Persicobacter diffluens.	M58765	marine mud
Bacteroidetes	Polaribacter dokdonensis MED152.	DQ481463	marine
Bacteroidetes	Psychroflexus torquis BSi20642.	DQ007442	arctic sea ice
Bacteroidetes	Psychroserpens burtonensis (T) ACAM188.	U62913	antarctic lacustrine and sea ice
Bacteroidetes	Tannerella forsythensis (T) 338.	L16495	human oral pathogen
Bacteroidetes	Tenacibaculum maritimum (T).	M64629	fish pathogen
Bacteroidetes	Zobellia uliginosa (T).	M62799	marine
Firmicutes/Clostridia	Butyrivibrio fibrisolvens 49.	EF427365	human/animal feces
Firmicutes/Clostridia	Clostridium algidixylanolyticum (T) SPL73.	AF092549	vacuum-packed temp abused lamb
Firmicutes/Clostridia	Clostridium bolteae (T) type strain: 16351.	AJ508452	human sources
Firmicutes/Clostridia	Clostridium clostridioforme (T).	M59089	human feces
Firmicutes/Clostridia	Clostridium fimetarium (T) Z-2189 DSM 9179.	AF126687	vacuum-packed meat
Firmicutes/Clostridia	Clostridium hathewayi.	EF408243	feral pig cecum

Firmicutes/Clostridia	Eubacterium oxidoreducens G2-2.	AF202259	cattle rumen
Firmicutes/Clostridia	Eubacterium rectale 1-82.	AY169428	human feces
Firmicutes/Clostridia	Lachnospira pectinoschiza 1-10.	AY169414	glacier ice core
Firmicutes/Clostridia	Pseudobutyrivibrio ruminis L4.	AY699285	GI of dromedary camel
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Firmicutes/Clostridia	Roseburia faecalis M88/1.	AY804150	human gut
Firmicutes/Clostridia	Ruminococcus gnavus A2.	EU139255	healthy human fecal
Proteobacteria	Desulfarculus baarsii DSM 2075.	AF418174	Bioreactors
Proteobacteria	Desulfatibacillum alkenivorans (T) PF2803.	AY493562	oil-polluted sediment
Proteobacteria	Desulfatimicrobium mahresensis SA1.	DQ006288	aerobic sludge
Proteobacteria	Desulfobacter curvatus DSM 3379.	AF418175	marine sediment
Proteobacteria	Desulfobacter latus DSM 3381.	AJ441315	marine sediment
Proteobacteria	Desulfobacter postgatei (T) DSM 2034.	AF418180	brackish water
Proteobacteria	Desulfobacula toluolica DSM 7467.	AJ441316	marine sediment
Proteobacteria	Desulfococcus multivorans DSM 2059.	AF418173	sewage digester
Proteobacteria	Desulfococcus oleovorans Hxd3.	CP000859	oil field
Proteobacteria	Desulfomonile limimaris DCB-F.	AF282177	marine sediment
Proteobacteria	Desulfosarcina variabilis (T).	M26632	marine sediment
Proteobacteria	Desulfotignum balticum DSM 7044.	AF418176	marine sediment
Proteobacteria	Desulfotignum toluolica H3 DSM 18732 4.	EF207158	oil reservoir
Proteobacteria	Hyphomicrobium sulfonivorans CT.	AY468372	human mouth
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine respiratory PATHOGEN
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma ovis (T).	AF338268	sheep/goat PATHOGEN

Table 9- MM07-06 Species Profile

Phylum/Class	Name	Accession	Isolation
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides vulgatus (T).	M58762	Very common human/animal fecal
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface Antarctic marine water
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum aeronauticum cw-04.	AY703033	corroded Al alloy from an aircraft
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge

Table 10- MM07-07 Species Profile

Phylum/Class	Name	Accession	Isolation
Actinobacteria	Amycolatopsis lactamdurans LC411.	AJ243301	Soil
Actinobacteria	Kribbella sandramycini KACC 20249.	AY253864	potato tuber in soil
Actinobacteria	Mycobacterium bonickei (T) W5998 ATCC 49935.	AY012573	human soft tissue wounds PATHOGEN
Actinobacteria	Mycobacterium conceptionense (T) CIP 108544.	AY859684	human case osteitits PATHOGEN
Actinobacteria	Mycobacterium farcinogenes NCTC 10955T.	AY457084	human case PATHOGEN
Actinobacteria	Mycobacterium fortuitum CIP 104534T.	AY457066	human PATHOGEN found in water, sewage, soil
Actinobacteria	Mycobacterium mucogenicum D3.	DQ068744	human PATHOGEN
Actinobacteria	Mycobacterium peregrinum CIP 105382T.	AY457069	human PATHOGEN

Actinobacteria	Mycobacterium phocaicum (T) CIP 108542.	AY859682	human PATHOGEN
Actinobacteria	Mycobacterium porcinum CIP 105392T.	AY457077	swine PATHOGEN
Actinobacteria	Mycobacterium psychrotolerans (T) type strain: WA101.	AJ534886	pond water near Uranium mine
Actinobacteria	Mycobacterium senegalense CIP 104941T.	AY457081	human PATHOGEN
Actinobacteria	Mycobacterium septicum ATCC 700731.	AY457070	human PATHOGEN
Actinobacteria	Pseudonocardia saturnea IMSNU 20052T.	AJ252829	Air
Bacteroidetes	Alkaliflexus imshenetskii (T) type strain: Z-7010.	AJ784993	soda lake
Bacteroidetes	Bacteroides cf. forsythus oral clone BU063.	AY008308	subgingival plaque human
Bacteroidetes	Bacteroides intestinalis (T) JCM 13265 341.	AB214328	human feces
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides splanchnicus (T).		normal human colonic flora
Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bifissio spartinae AS1.1762.		Soil
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Capnocytophaga ochracea ChDC OS42.	AF543292	human oral
Bacteroidetes	Chryseobacterium indoltheticum (T).	M58774	rhizosphere of lettuce
Bacteroidetes	Chryseobacterium taiwanensis (T)	DQ318789	Soil
Bacteroidetes	Cytophaga fermentans (T).	M58766	river epilithon
Bacteroidetes	Dokdonia donghaensis MED134.	DQ481462	sea water
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Gramella forsetii KT0803.	CU207366	Concentrated seawater
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Bacteroidetes	Microscilla arenaria.	M60455	sea water
Bacteroidetes	Parabacteroides distasonis ATCC 8503.	CP000140	symbiotic distal human intestine
Bacteroidetes	Perexilibacter aurantiacus Shu-F-UV2-2.	AB276355	Sediment
Bacteroidetes	Persicobacter diffluens.	M58765	river epilithon
Bacteroidetes	Polaribacter dokdonensis MED152.	DQ481463	sea water
Bacteroidetes	Psychroserpens burtonensis (T) ACAM188.	U62913	antarctic sea ice
Bacteroidetes	Riemerella anatipestifer 11693.	AY871829	duck/turkey PATHOGEN

Bacteroidetes	Tannerella forsythensis TR6.	AB053947	human dental PATHOGEN
Bacteroidetes	Tenacibaculum maritimum (T).	M64629	fish PATHOGEN
Bacteroidetes	Zobellia uliginosa (T).	M62799	Marine
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Deferribacteres	Deferribacter thermophilus (T) BMA1.	U75602	oil reservoir
Firmicutes/Bacilli	Alicyclobacillus acidiphilus (T) TA-67.	AB076660	acidic beverage
Firmicutes/Bacilli	Alicyclobacillus acidoterrestris ATCC 49025T.	AB042057	herbal tea
Firmicutes/Bacilli	Alicyclobacillus cycloheptanicus DSM 4007.	AB059681	Soil
Firmicutes/Bacilli	Alicyclobacillus disulfidooxidans DSM 12064.	AB089843	mixed fruit juice
Firmicutes/Bacilli	Alicyclobacillus fastidiosus S-TAB.	AB264021	apple juice
Firmicutes/Bacilli	Alicyclobacillus ferripilum TC-71.	EU137837	hot spring
Firmicutes/Bacilli	Carnobacterium alterfunditum.	L08623	antarctic
Firmicutes/Bacilli	Carnobacterium divergens (T).	M58816	Vacuum-meat/seafood products poultry flora
Firmicutes/Bacilli	Carnobacterium gallinarum DSM 4847(T), NCFB 2766(T).	AJ387905	poultry flora, meat, cheese, seafood, fish intestines
Firmicutes/Bacilli	Enterococcus asini (T) AS2.	Y11621	donkey cecum
Firmicutes/Bacilli	Enterococcus avium (T) CIP 103 019.	AF133535	rare human PATHOGEN
Firmicutes/Bacilli	Enterococcus dispar LMG 13521.	AJ301829	human clinical sample
Firmicutes/Bacilli	Enterococcus durans D1.	DQ239692	human clinical sample
Firmicutes/Bacilli	Enterococcus faecalis 47/3.	EF653454	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus faecium C228.	AB246407	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus mundtii HDYM-33.	EF428252	human clinical sample-sinus and abscess
Firmicutes/Bacilli	Paenibacillus ourofinensis AC13MSD.	EU257517	Brazillian soil
Firmicutes/Bacilli	Paenibacillus ruminocola CA8.	DQ085278	cattle rumen
Firmicutes/Bacilli	Streptococcus downei ChDC YD1.	AY942561	human dental plaque
Firmicutes/Bacilli	Streptococcus mutans 669.	AF139603	human dental PATHOGEN
Firmicutes/Bacilli	Streptococcus orisuis Dentisuis.	AB212809	swine oral PATHOGEN
Firmicutes/Bacilli	Streptococcus sobrinus ChDC YS211 PD1164.	DQ677790	human dental plaque
Firmicutes/Clostridia	Butyrivibrio fibrisolvens 49.	EF427365	human/animal

			feces
Firmicutes/Clostridia	Clostridium algidixylanolyticum (T) SPL73.	AF092549	vacuum-packed temp abused lamb
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium bolteae (T) type strain: 16351.	AJ508452	human sources
Firmicutes/Clostridia	Clostridium clostridioforme (T).	M59089	human feces
Firmicutes/Clostridia	Clostridium fimetarium (T) Z-2189 DSM 9179.	EF408243	vacuum-packed meat
Firmicutes/Clostridia	Clostridium hathewayi.	M23930	feral pig cecum
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Clostridium pasteurianum (T).		soil
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum aeronauticum cw-04.	AY703033	corroded Al alloy from an aircraft
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	cmixed cow/pig manure
Firmicutes/Clostridia	Eubacterium oxidoreducens G2-2.	AF202259	cattle rumen
Firmicutes/Clostridia	Eubacterium rectale 1-82.	AY169428	human feces
Firmicutes/Clostridia	Lachnospira pectinoschiza 1-10.	AY169414	glacier ice core
Firmicutes/Clostridia	Megasphaera elsdenii 5T.	DQ146765	human feces
Firmicutes/Clostridia	Megasphaera hominis.	L79909	human colonic biota
Firmicutes/Clostridia	Moorella thermoacetica.	M59121	stagnant ponds
Firmicutes/Clostridia	Pseudobutyrivibrio ruminis pC-XS2.	AF202260	cow rumen
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	Sheep rumen
Firmicutes/Clostridia	Roseburia faecalis M88/1.	AY804150	human gut
Firmicutes/Clostridia	Ruminococcus gnavus A2.	EU139255	healthy human fecal
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core
Firmicutes/Clostridia	Selenomonas lacticifex (T) DSM20757.	AF373024	Breweries
Firmicutes/Clostridia	Selenomonas ruminantium DSM2872.	AF373022	Breweries
Firmicutes/Clostridia	Selenomonas sputigena (T) ATCC 35185.	AF287793	human subgingival plaque
Firmicutes/Clostridia	Zymophilus paucivorans (T) DSM20756.	AF373025	pitching yeast
Proteobacteria	Acetobacter pasteurianus.	EU034024	fruits/vegetables
Proteobacteria	Candidatus Glomeribacter gigasporarum.	AJ251636	endocellular of fungus
Proteobacteria	Desulfarculus baarsii DSM 2075.	AF418174	bioreactors
Proteobacteria	Desulfatibacillum alkenivorans (T) PF2803.	AY493562	oil-polluted sediment
Proteobacteria	Desulfatimicrobium mahresensis SA1.	DQ006288	aerobic sludge
Proteobacteria	Desulfobacter curvatus DSM 3379.	AF418175	marine sediment
Proteobacteria	Desulfobacter latus DSM 3381.	AJ441315	marine sediment
Proteobacteria	Desulfobacter postgatei (T) DSM	AF418180	brackish water

	2034.		
Proteobacteria	Desulfobacula toluolica DSM 7467.	AJ441316	marine sediment
Proteobacteria	Desulfococcus multivorans DSM 2059.	AF418173	sewage digester
Proteobacteria	Desulfococcus oleovorans Hxd3.	CP000859	oil field
Proteobacteria	Desulfomonile limimaris DCB-F.	AF282177	marine sediment
Proteobacteria	Desulfosarcina variabilis (T).	M26632	marine sediment
Proteobacteria	Desulfotignum balticum DSM 7044.	AF418176	marine sediment
Proteobacteria	Desulfotignum toluolica H3 DSM 18732 4.	EF207158	oil reservoir
Proteobacteria	Halomonas korlensis XK1.	EU085033	Saline soil, Human PATHOGEN from fish bite
Proteobacteria	Hyphomicrobium sulfonivorans CT.	AY468372	human mouth
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge
Proteobacteria	Vibrio fluvialis CIFAMVIFL01.	DQ683079	human feces PATHOGEN
Proteobacteria	Xanthomonas campestris EGS09.	EF101972	plant PATHOGEN
Tenericutes/Mollicutes	Acholeplasma laidlawii.	M23932	sewage, manure, humus, soil, parasite in mammals
Tenericutes/Mollicutes	Acholeplasma oculi ISM1499.	U14906	conjunctivae of sheep with contagious ophthalmia
Tenericutes/Mollicutes	Mycoplasma agalactiae PG2.	CU179680	Sheep/goat PATHOGEN
Tenericutes/Mollicutes	Mycoplasma bovis (T) Donetta (type strain) pMb16S.	U02968	Mammal PATHOGEN
Tenericutes/Mollicutes	Mycoplasma californicum (T).	M24582	Cattle PATHOGEN
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine respiratory PATHOGEN
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma genitalium (T) G37.	X77334	human origin, parasitic PATHOGEN
Tenericutes/Mollicutes	Mycoplasma lipophilum (T).	M24581	human/rhesus monkey oral
Tenericutes/Mollicutes	Mycoplasma ovis (T).	AF338268	Sheep/goat PATHOGEN
Tenericutes/Mollicutes	Mycoplasma pneumoniae (T).	M29061	human PATHOGEN

Tenericutes/Mollicutes	Mycoplasma sphenisci UCMJ.	AY756171	Choana of an Aquarium-Reared Jackass Penguin PATHOGEN
Tenericutes/Mollicutes	Mycoplasma zalophi 4296C.	AF493543	PATHOGEN California sea lions
Tenericutes/Mollicutes	Spiroplasma citri Qualubia.	AM157769	plant PATHOGEN

Table 11- MM07-08 Species Profile

Phylum/Class	Name	Accession	Isolation
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bifissio spartinae AS1.1762.	AY056829	soil
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Firmicutes/Bacilli	Atopostipes suicloacalis (T) PPC79.	AF445248	swine manure
Firmicutes/Bacilli	Enterococcus asini (T) AS2.	Y11621	donkey cecum
Firmicutes/Bacilli	Enterococcus avium (T) CIP 103 019.	AF133535	rare human PATHOGEN
Firmicutes/Bacilli	Enterococcus dispar LMG 13521.	AJ301829	human clinical sample
Firmicutes/Bacilli	Enterococcus durans D1.	DQ239692	human clinical sample
Firmicutes/Bacilli	Enterococcus faecalis 47/3.	EF653454	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus faecium C228.	AB246407	GI of humans/mammals PATHOGEN

Firmicutes/Bacilli	Enterococcus mundtii HDYM-33.	EF428252	human clinical sample-sinus and abscess
Firmicutes/Bacilli	Lactococcus garvieae JCM 8735.	AB012306	shrimp freshwater pathogen
Firmicutes/Bacilli	Lactococcus lactis subsp. cremoris (T).	M58836	Radish
Firmicutes/Bacilli	Paenibacillus peoriae (T) IFO15541T.	D78476	Soil
Firmicutes/Bacilli	Streptococcus agalactiae.	AF459432	human pathogen
Firmicutes/Bacilli	Streptococcus bovis.	M58835	human pathogen
Firmicutes/Bacilli	Streptococcus canis CG40.	AJ413205	human pathogen
Firmicutes/Bacilli	Streptococcus didelphis W95-4849.	AF176100	opposum pathogen
Firmicutes/Bacilli	Streptococcus dysgalactiae subsp. dysgalactiae (T) ATCC 4...	AB002485	Swine
Firmicutes/Bacilli	Streptococcus gallolyticus subsp. gallolyticus ATCC 43143.	AF104114	human and ruminal
Firmicutes/Bacilli	Streptococcus iniae (T) ATCC29178.	AF335572	fish pathogen
Firmicutes/Bacilli	Streptococcus mutans 669.	AF139603	human dental PATHOGEN
Firmicutes/Bacilli	Streptococcus parasanguinis mother C3 5C3.	AM157421	human breast milk
Firmicutes/Bacilli	Streptococcus pyogenes MGAS10394.	CP000003	human PATHOGEN
Firmicutes/Bacilli	Streptococcus salivarius.	AF459433	human PATHOGEN
Firmicutes/Bacilli	Streptococcus sanguinis ChDC B203.	AF543281	human oral
Firmicutes/Bacilli	Streptococcus thermophilus.	DQ176426	human GI
Firmicutes/Bacilli	Streptococcus vestibularis ATCC 49124.	AY188353	human oral
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum aeronauticum cw-02.	AY703031	corroded Al alloy from an aircraft
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core

Proteobacteria	Candidatus Glomeribacter gigasporarum.	AJ251636	endocellular of fungus
Proteobacteria	Photobacterium damselae subsp. damselae (T) ATCC33539.	AB032015	fish pathogen
Proteobacteria	Salinivibrio costicola subsp. costicola GSP12.	AY505534	Great Salt Plains
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge
Proteobacteria	Vibrio fluvialis CIFAMVIFL01.	DQ683079	human feces pathogen
Tenericutes/Mollicutes	Acholeplasma laidlawii.	M23932	sewage, manure, humus, soil, parasite in mammals
Tenericutes/Mollicutes	Acholeplasma oculi ISM1499.	U14906	conjunctivae of sheep with contagious ophthalmia
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine pathogen
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma zalophi 4296C.	AF493543	PATHOGEN California sea lions
Tenericutes/Mollicutes	Spiroplasma citri Fewa.	AM157768	Plant source

Table 12- MM07-09 Species Profile

Phylum/Class	Name	Accession	Isolation
Actinobacteria	Corynebacterium genitalium ATCC 33978.	U87824	human PATHOGEN
Actinobacteria	Kocuria halotolerans YIM 90716.	DQ979377	saline soil
Actinobacteria	Kocuria rosea CU24.	EF522129	Cuban sugarcane plant
Aquificae	Persephonella hydrogeniphila (T) 29W.	AB086419	deep sea hydrothermal vent
Bacteroidetes	Bacteroides cf. forsythus oral clone BU063.	AY008308	subgingival human plaque
Bacteroidetes	Bacteroides intestinalis (T) JCM 13265 341.	AB214328	human feces
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides splanchnicus (T).	L16496	normal human colonic flora

Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bifissio spartinae AS1.1762.	AY056829	soil
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Capnocytophaga sputigena.	L14636	human PATHOGEN
Bacteroidetes	Chryseobacterium hispanicum (T) type strain:VP48.	AM159183	drinking water, Spain
Bacteroidetes	Chryseobacterium indoltheticum (T).	M58774	rhizosphere of lettuce
Bacteroidetes	Chryseobacterium shigense (T) GUM-Kaji.	AB193101	lactic acid beverage
Bacteroidetes	Chryseobacterium taiwanensis (T) BCRC 17412.	DQ318789	Soil
Bacteroidetes	Cytophaga fermentans (T).	M58766	river epilithon
Bacteroidetes	Dokdonia donghaensis MED134.	DQ481462	Marine
Bacteroidetes	Elizabethkingia miricola L99 CGMCC 2295.	EU375848	human PATHOGEN
Bacteroidetes	Empedobacter brevis.	M59052	nosocomial human infections
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Gelidibacter algens IC147.	AF001367	Antarctic sea ice
Bacteroidetes	Gelidibacter gilvus (T) IC158.	AF001369	Antarctic sea ice
Bacteroidetes	Gramella forsetii KT0803.	CU207366	concentrated seawater
Bacteroidetes	Leeuwenhoekiella blandensis MED217.	DQ294291	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Bacteroidetes	Microscilla arenaria.	M60455	Sand
Bacteroidetes	Ornithobacterium rhinotracheale (T).	L19156	avian respiratory tract
Bacteroidetes	Parabacteroides distasonis ATCC 8503.	CP000140	symbiotic distal human intestine
Bacteroidetes	Perexilibacter aurantiacus Shu-F-UV2-2.	AB276355	Sediment
Bacteroidetes	Persicobacter diffluens.	M58765	marine mud
Bacteroidetes	Psychroflexus torquis BSi20642.	DQ007442	arctic sea ice
Bacteroidetes	Psychroserpens burtonensis (T) ACAM188.	U62913	antarctic lacustrine and sea ice
Bacteroidetes	Riemerella anatipestifer 11693.	AY871829	duck/turkey PATHOGEN
Bacteroidetes	Tannerella forsythensis (T) 338.	L16495	human oral pathogen
Bacteroidetes	Tenacibaculum maritimum (T).	M64629	fish pathogen
Bacteroidetes	Wautersiella falsenii subsp. genomovar 2 NF 622.	AM238670	human clinical isolates
Bacteroidetes	Zobellia uliginosa (T).	M62799	Marine
Chlorobi	Chlorobaculum tepidum (T).	M58468	hot spring
Chlorobi	Chlorobium limicola UdG 6042.	Y10644	Aquatic
Firmicutes/Bacillales	Bacillus arseniciselenatis E1H.	AF064705	Lake
Firmicutes/Bacillales	Paenibacillus ourofinensis AC13MSD.	EU257517	cerrado soil
Firmicutes/Bacilli	Alicyclobacillus acidiphilus (T) TA-67.	AB076660	acidic beverage

Firmicutes/Bacilli	Alicyclobacillus acidoterrestris ATCC 49025T.	AB042057	herbal tea
Firmicutes/Bacilli	Alicyclobacillus cycloheptanicus DSM 4007.	AB059681	Soil
Firmicutes/Bacilli	Alicyclobacillus disulfidooxidans DSM 12064.	AB089843	mixed fruit juice
Firmicutes/Bacilli	Alicyclobacillus fastidiosus S-TAB.	AB264021	apple juice
Firmicutes/Bacilli	Alicyclobacillus ferripilum TC-71.	EU137837	hot spring
Firmicutes/Bacilli	Enterococcus asini (T) AS2.	Y11621	donkey cecum
Firmicutes/Bacilli	Enterococcus avium (T) CIP 103 019.	AF133535	rare human PATHOGEN
Firmicutes/Bacilli	Enterococcus dispar LMG 13521.	AJ301829	human clinical sample
Firmicutes/Bacilli	Enterococcus durans D1.	DQ239692	human clinical sample
Firmicutes/Bacilli	Enterococcus faecalis 47/3.	EF653454	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus faecium C228.	AB246407	GI of humans/mammals PATHOGEN
Firmicutes/Bacilli	Enterococcus mundtii HDYM-33.	EF428252	human clinical sample-sinus and abscess
Firmicutes/Bacilli	Streptococcus mutans 669.	AF139603	human dental PATHOGEN
Firmicutes/Clostridia	Butyrivibrio fibrisolvens 49.	EF427365	human/animal feces
Firmicutes/Clostridia	Clostridium algidixylanolyticum (T) SPL73.	AF092549	vacuum-packed temp abused lamb
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium bolteae (T) type strain: 16351.	AJ508452	human sources
Firmicutes/Clostridia	Clostridium clostridioforme (T).	M59089	human feces
Firmicutes/Clostridia	Clostridium fimetarium (T) Z-2189 DSM 9179.	AF126687	vacuum-packed meat
Firmicutes/Clostridia	Clostridium hathewayi.	EF408243	feral pig cecum
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Clostridium pasteurianum (T).	M23930	soil
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum aeronauticum cw-02.	AY703031	corroded Al alloy from an aircraft
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Desulfotomaculum thermosapovorans MLF.	Z26315	rice hulls and peanut shells compost
Firmicutes/Clostridia	Eubacterium oxidoreducens G2-2.	AF202259	cattle rumen
Firmicutes/Clostridia	Eubacterium rectale 1-82.	AY169428	human feces
Firmicutes/Clostridia	Lachnospira pectinoschiza 1-10.	AY169414	glacier ice core
Firmicutes/Clostridia	Moorella thermoacetica.	M59121	stagnant ponds

Firmicutes/Clostridia	Natronoanaerobium halophilum G-M14CH-4.	AJ271451	soda lakes
Firmicutes/Clostridia	Pseudobutyrivibrio ruminis L4.	AY699285	GI of dromedary camel
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Firmicutes/Clostridia	Roseburia faecalis M88/1.	AY804150	human gut
Firmicutes/Clostridia	Ruminococcus gnavus A2.	EU139255	healthy human fecal
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core
Firmicutes/Lactobacilli	Aerococcus urinae (T).	M77819	human urinary tract infection
Proteobacteria	Anaeromyxobacter dehalogenans (T) 2CP-1 ATCC BAA-258.	AF382396	soils/sediment
Proteobacteria	Bacteriovorax stolpii (T) uki-2.	M34125	parasite of other GM(-) bacteria
Proteobacteria	Candidatus Glomeribacter gigasporarum.	AJ251636	Fungus
Proteobacteria	Desulfarculus baarsii DSM 2075.	AF418174	uncertain
Proteobacteria	Desulfatibacillum alkenivorans (T) PF2803.	AY493562	oil-polluted sediment
Proteobacteria	Desulfatibacillus olefinivorans LM2801.	DQ826724	industrial wastewater
Proteobacteria	Desulfatimicrobium mahresensis SA1.	DQ006288	aerobic sludge
Proteobacteria	Desulfobacter curvatus DSM 3379.	AF418175	marine sediment
Proteobacteria	Desulfobacter latus DSM 3381.	AJ441315	marine sediment
Proteobacteria	Desulfobacter postgatei (T) DSM 2034.	AF418180	brackish water
Proteobacteria	Desulfobacula toluolica DSM 7467.	AJ441316	marine sediment
Proteobacteria	Desulfococcus multivorans DSM 2059.	AF418173	sewage digester
Proteobacteria	Desulfococcus oleovorans Hxd3.	CP000859	oil field
Proteobacteria	Desulfodehalobacter spongiphilus AA1.	EF187256	marine sponge
Proteobacteria	Desulfofrigus fragile (T) LSv21.	AF099065	marine arctic sediment
Proteobacteria	Desulfofrigus oceanense (T) ASv26.	AF099064	marine arctic sediment
Proteobacteria	Desulfohalobium retbaense (T) DSM 5692.	U48244	oil pipeline
Proteobacteria	Desulfomicrobium baculatum (T) Type strain X VKM B-1378 D...	AF030438	Manganese ore
Proteobacteria	Desulfomicrobium thermophilum P6.2.	AY464939	terrestrial hot spring
Proteobacteria	Desulfomonile limimaris DCB-F.	AF282177	marine sediment
Proteobacteria	Desulfonatronum lacustre DSM 10312.	AF418171	mud from alkaline lake
Proteobacteria	Desulforhopalus vacuolatus (T) ltk10.	L42613	temperate estuary
Proteobacteria	Desulfosarcina variabilis.	M34407	marine black mud
Proteobacteria	Desulfotalea psychrophila (T) LSv54.	AF099062	arctic sediments
Proteobacteria	Desulfotignum balticum DSM 7044.	AF418176	marine mud
Proteobacteria	Desulfotignum toluolica H3 DSM 18732 4.	EF207158	oil reservoir
Proteobacteria	Desulfovibrio alaskensis HEB223.	DQ867001	Tunisian marine environment
Proteobacteria	Desulfovibrio alcoholovorans DSM 5433.	AF053751	African rice field

Proteobacteria	Desulfovibrio bizertensis MB3.	DQ422859	Tunisian marine environment
Proteobacteria	Desulfovibrio dechloracetivorans (T).	AF230530	marine sediment
Proteobacteria	Desulfovibrio desulfuricans EFX-DES.	DQ526427	Aquatic
Proteobacteria	Desulfovibrio giganteus DSM 4370.	AF418170	gut, soil-feeding termite *Cubitermes* sp.
Proteobacteria	Desulfovibrio oxamicus (T) DSM 1925.	DQ122124	Environment
Proteobacteria	Desulfovibrio piger (T) ATCC29098.	AF192152	human PATHOGEN
Proteobacteria	Desulfovibrio profundus DSM 11384.	AF418172	marine sediment
Proteobacteria	Desulfovibrio senezii (T) CVL DSM 8436.	AF050100	solar saltern
Proteobacteria	Geobacter metallireducens (T) GS-15.	L07834	aquatic sediments
Proteobacteria	Geobacter sulfurreducens PCA.	AE017180	Subsurface
Proteobacteria	Geobacter uraniireducens Rf4.	CP000698	Sediments of Ur field site
Proteobacteria	Geopsychrobacter electrodiphilus (T) A1.	AY187303	marine sediment
Proteobacteria	Halomonas korlensis XK1.	EU085033	saline/alkaline soil
Proteobacteria	Hyphomicrobium sulfonivorans CT.	AY468372	human mouth
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge
Proteobacteria	Syntrophus aciditrophicus SB.	CP000252	sludge/sewer treatment
Proteobacteria	Thalassospira profundimaris WP0211.	AY186195	deep sea sediment
Proteobacteria	Thalassospira xiamenensis M-5.	AY189753	diesel-polluted Bohai Gulf
Proteobacteria	Tistrella mobilis (T).	AB071665	Wastewater
Proteobacteria	Vibrio fluvialis CIFAMVIFL01.	DQ683079	human feces PATHOGEN
Proteobacteria	Xanthomonas campestris EGS09.	EF101972	plant pathogen
Tenericutes/Mollicutes	Acholeplasma laidlawii.	M23932	sewage, manure, humus, soil, parasite in mammals
Tenericutes/Mollicutes	Acholeplasma oculi ISM1499.	U14906	conjunctivae of sheep with contagious ophthalmia
Tenericutes/Mollicutes	Mycoplasma agalactiae PG2.	CU179680	sheep/goat PATHOGEN
Tenericutes/Mollicutes	Mycoplasma bovis (T) Donetta (type strain) pMb16S.	U02968	mammal PATHOGEN
Tenericutes/Mollicutes	Mycoplasma californicum (T).	M24582	Cattle PATHOGEN

Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine respiratory PATHOGEN
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma genitalium (T) G37.	X77334	human origin, parasitic PATHOGEN
Tenericutes/Mollicutes	Mycoplasma lipophilum (T).	M24581	human/rhesus monkey oral
Tenericutes/Mollicutes	Mycoplasma ovis (T).	AF338268	sheep/goat PATHOGEN
Tenericutes/Mollicutes	Mycoplasma pneumoniae (T).	M29061	human PATHOGEN
Tenericutes/Mollicutes	Mycoplasma sphenisci UCMJ.	AY756171	Choana of an Aquarium-Reared Jackass Penguin PATHOGEN
Tenericutes/Mollicutes	Mycoplasma zalophi 4296C.	AF493543	PATHOGEN California sea lions
Tenericutes/Mollicutes	Spiroplasma citri Fewa.	AM157768	Citrus sinensis leaves in Egypt
Tenericutes/Mollicutes	Spiroplasma citri Qualubia.	AM157769	Citrus sinensis leaves in Egypt

Table 13- MM07-10 Species Profile

Phylum/Class	Name	Accession	Isolation
Actinobacteria	Actinocorallia glomerata IMSNU 22179T.	AJ293704	plants/pond
Bacteroidetes	Bacteroides plebeius M35.	AB200219	human feces
Bacteroidetes	Bacteroides vulgatus (T).	M58762	very common human/animal fecal
Bacteroidetes	Bizionia argentinensis JUB59.	EU021217	surface antarctic marine water
Bacteroidetes	Flexibacter aggregans BSs20185.	DQ514301	marine sediment
Bacteroidetes	Leeuwenhoekiella blandensis MED217 (T).	DQ294290	marine
Bacteroidetes	Leeuwenhoekiella marinoflava (T).	M58770	sea water
Firmicutes	Lactobacillus vitulinus (T).	M23727	calf rumen
Firmicutes/Bacilli	Carnobacterium divergens (T).	M58816	vacuum-meat/seafood poultry flora

Firmicutes/Bacilli	Carnobacterium gallinarum DSM 4847(T), NCFB 2766(T).	AJ387905	poultry flora, meat, cheese, seafood, fish intestines
Firmicutes/Bacilli	Streptococcus downei ChDC YD1.	AY942561	human dental plaque
Firmicutes/Bacilli	Streptococcus sobrinus ChDC YS211 PD1164.	DQ677790	human dental plaque
Firmicutes/Clostridia	Desulfotomaculum aeronauticum cw-02.	AY703031	corroded Al alloy from an aircraft
Firmicutes/Clostridia	Quinella ovalis (T).	M62701	sheep rumen
Firmicutes/Clostridia	Selenomonas lacticifex (T) DSM20757.	AF373024	Breweries
Firmicutes/Clostridia	Selenomonas ruminantium DSM2872.	AF373022	Breweries
Firmicutes/Clostridia	Selenomonas sputigena (T) ATCC 35185.	AF287793	human subgingival plaque
Firmicutes/Clostridia	Zymophilus paucivorans (T) DSM20756.	AF373025	pitching yeast
Proteobacteria	Syntrophobacter fumaroxidans MPOB.	CP000478	anaerobic granular sludge
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine respiratory PATHOGEN
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY

Table 14- MM07-11 Species Profile

Phylum/Class	Name	Accession	Isolation
Actinobacteria	Amycolatopsis lurida DSM43187.	X81573	soil/ produces ristocetin
Actinobacteria	Amycolatopsis marina MS392A.	EU329845	marine sediment
Actinobacteria	Corynebacterium pseudodiphtheriticum CIP103420T, (ATCC107...	AJ439343	human PATHOGEN
Actinobacteria	Eggerthella lenta JCM9979.	AB011817	human feces
Actinobacteria	Microbacterium aurum (T) IFO 15204T.	D21340	corn
Actinobacteria	Mycobacterium abscessus CIP 104536T.	AY457071	water, soil, dust, Human Pathogen
Actinobacteria	Mycobacterium aichiense JS618.	AF498656	contaminated env.
Actinobacteria	Mycobacterium asiaticum (T) ATCC 25276.	X55604	human PATHOGEN
Actinobacteria	Mycobacterium aubagnense (T) CIP 108543.	AY859683	human PATHOGEN
Actinobacteria	Mycobacterium bolletii CIP 108541.	AY859681	human PATHOGEN

Actinobacteria	Mycobacterium brisbanense (T) W6743 ATCC 49938.	AY012577	water, soil, dust, Human Pathogen
Actinobacteria	Mycobacterium chelonae ATCC 19237.	AY457082	soil, human sputum, human pathogen
Actinobacteria	Mycobacterium chitae (T) ATCC 19627.	X55603	soil
Actinobacteria	Mycobacterium chlorophenolicum PCP-I.	X79094	soil
Actinobacteria	Mycobacterium chubuense (T) ATCC 27278.	X55596	soil
Actinobacteria	Mycobacterium cookii (T) ATCC 49103 (T) = NZ2..	X53896	moss
Actinobacteria	Mycobacterium immunogenum CIP 106684T.	AY457080	human PATHOGEN
Actinobacteria	Mycobacterium leprae.	U15186	human leprosy pathogen
Actinobacteria	Mycobacterium mageritense CIP 104973T.	AY457076	human PATHOGEN
Actinobacteria	Mycobacterium massiliense (T) CCUG 48898.	AY593980	human PATHOGEN
Actinobacteria	Mycobacterium obuense (T) ATCC 27023.	X55597	human PATHOGEN
Actinobacteria	Mycobacterium wolinskyi ATCC 700010.	AY457083	human PATHOGEN
Actinobacteria	Streptomyces bikiniensis (T) DSM40581.	X79851	human pathogen
Actinobacteria	Streptomyces espinosus NRRL 5729.	X80826	uncertain
Actinobacteria	Streptomyces glaucescens DSM40716.	X79322	uncertain
Actinobacteria	Streptomyces rimosus subsp. rimosus B4.	EF371440	uncertain
Actinobacteria	Streptomyces shandongensis 24#.	AY875718	uncertain
Actinobacteria	Williamsia maris (T) SJS0289-JS1 JoyO.	AB010909	deep sea
Bacteroidetes	Flavobacterium aquatile (T).	M62797	soil/freshwater
Chloroflexi	Chloroflexus aurantiacus J-10-fl.	CP000909	hot spring
Cyanobacteria/possibly Firmicutes	Microcystis elabens NIES42.	U40335	water
Firmicutes/Clostridia	Butyrivibrio fibrisolvens 49.	EF427365	human/animal feces
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Eubacterium rectale 1-82.	AY169428	human feces
Firmicutes/Clostridia	Lachnospira pectinoschiza 1-10.	AY169414	ice core
Firmicutes/Clostridia	Roseburia faecalis M88/1.	AY804150	human gut flora
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	ice core
Proteobacteria	Achromatium oxaliferum Rydal 8.	L79968	sediment

Proteobacteria	Aggregatibacter aphrophilus CIP 70.73.	EU083529	periodontal human pathogen
Proteobacteria	Haemophilus parainfluenzae CIP 102513.	EU083530	normal human oropharyngeal flora, can be pathogenic
Proteobacteria	Pasteurella multocida subsp. tigris.	AY057994	human pathogen from tiger bite
Proteobacteria	Pseudoalteromonas citrea L3.	DQ401135	seawater
Proteobacteria	Pseudoalteromonas piscicida L2.	DQ401136	seawater
Tenericutes/Mollicutes	Mycoplasma cynos H831.	AF538682	canine pathogen
Tenericutes/Mollicutes	Mycoplasma fermentans (T).	M24289	human lower genital tract uncertain PATHOGENICITY
Tenericutes/Mollicutes	Mycoplasma hominis 7488.	AJ002269	human lower genital tract
Tenericutes/Mollicutes	Mycoplasma penetrans HF-2.	BA000026	human PATHOGEN
Tenericutes/Mollicutes	Ureaplasma urealyticum (T).	M23935	normal human male/female genital flora

Table 15- MM07-12 Species Profile

Phylum/Class	Name	Accession	Isolation
Aquificae	Persephonella hydrogeniphila (T) 29W.	AB086419	deep sea hydrothermal vent
Bacteroidetes	Bifissio spartinae AS1.1762.	AY056829	soil
Bacteroidetes	Chryseobacterium indoltheticum (T).	M58774	rhizosphere of lettuce
Bacteroidetes	Chryseobacterium taiwanensis (T) BCRC 17412.	DQ318789	soil
Bacteroidetes	Riemerella anatipestifer 11693.	AY871829	duck/turkey PATHOGEN
Firmicutes/Bacillales	Paenibacillus ourofinensis AC13MSD.	EU257517	cerrado soil
Firmicutes/Bacilli	Alicyclobacillus acidiphilus (T) TA-67.	AB076660	acidic beverage
Firmicutes/Bacilli	Alicyclobacillus acidoterrestris ATCC 49025T.	AB042057	herbal tea
Firmicutes/Bacilli	Alicyclobacillus cycloheptanicus DSM 4007.	AB059681	soil
Firmicutes/Bacilli	Alicyclobacillus disulfidooxidans DSM 12064.	AB089843	mixed fruit juice
Firmicutes/Bacilli	Alicyclobacillus fastidiosus S-TAB.	AB264021	apple juice
Firmicutes/Bacilli	Alicyclobacillus ferripilum TC-71.	EU137837	hot spring

Firmicutes/Bacilli	Enterococcus asini (T) AS2.	Y11621	donkey cecum
Firmicutes/Bacilli	Enterococcus avium (T) CIP 103 019.	AF133535	rare human PATHOGEN
Firmicutes/Bacilli	Enterococcus dispar LMG 13521.	AJ301829	human clinical sample
Firmicutes/Bacilli	Enterococcus durans D1.	DQ239692	human clinical sample
Firmicutes/Bacilli	Enterococcus faecalis 47/3.	EF653454	GI of humans/mamm als PATHOGEN
Firmicutes/Bacilli	Enterococcus faecium C228.	AB246407	GI of humans/mamm als PATHOGEN
Firmicutes/Bacilli	Enterococcus mundtii HDYM-33.	EF428252	human clinical sample-sinus and abscess
Firmicutes/Bacilli	Streptococcus mutans 669.	AF139603	human dental PATHOGEN
Firmicutes/Clostridia	Clostridium aminophilum 152R-1b.	DQ278862	grazing ruminants
Firmicutes/Clostridia	Clostridium innocuum (T).	M23732	human PATHOGEN
Firmicutes/Clostridia	Clostridium pasteurianum (T).	M23930	soil
Firmicutes/Clostridia	Dendrosporobacter quercicolus (T).	M59110	oak leaves
Firmicutes/Clostridia	Desulfotomaculum alkaliphilum (T) S1.	AF097024	mixed cow/pig manure
Firmicutes/Clostridia	Moorella thermoacetica.	M59121	stagnant ponds
Firmicutes/Clostridia	Ruminococcus obeum 1-33.	AY169419	Greenland ice core
Proteobacteria	Candidatus Glomeribacter gigasporarum.	AJ251636	Fungus
Proteobacteria	Halomonas korlensis XK1.	EU085033	saline/alkaline soil
Proteobacteria	Vibrio fluvialis CIFAMVIFL01.	DQ683079	human feces PATHOGEN
Proteobacteria	Xanthomonas campestris EGS09.	EF101972	plant pathogen
Tenericutes/Mollicutes	Acholeplasma laidlawii.	M23932	sewage, manure, humus, soil, parasite in mammals
Tenericutes/Mollicutes	Acholeplasma oculi ISM1499.	U14906	conjunctivae of sheep with contagious ophthalmia

Tenericutes/Mollicutes	Mycoplasma agalactiae PG2.	CU179680	stagnant ponds
Tenericutes/Mollicutes	Mycoplasma bovis (T) Donetta (type strain) pMb16S.	U02968	sheep/goat PATHOGEN
Tenericutes/Mollicutes	Mycoplasma californicum (T).	M24582	mammal PATHOGEN
Tenericutes/Mollicutes	Mycoplasma genitalium (T) G37.	X77334	human origin, parasitic PATHOGEN
Tenericutes/Mollicutes	Mycoplasma lipophilum (T).	M24581	human/rhesus monkey oral
Tenericutes/Mollicutes	Mycoplasma pneumoniae (T).	M29061	human PATHOGEN
Tenericutes/Mollicutes	Mycoplasma sphenisci UCMJ.	AY756171	Choana of an Aquarium-Reared Jackass Penguin PATHOGEN
Tenericutes/Mollicutes	Mycoplasma zalophi 4296C.	AF493543	PATHOGEN California sea lions
Tenericutes/Mollicutes	Spiroplasma citri Fewa.	AM157768	Citrus sinensis leaves in Egypt
Tenericutes/Mollicutes	Spiroplasma citri Qualubia.	AM157769	Citrus sinensis leaves in Egypt