博士論文
Doctoral Dissertation

Antibiotics shapes population-level diversity in the human gut microbiome

(抗生物質はヒト腸内細菌叢の集団レベルでの多様性を形成する)

西嶋 傑
Nishijima Suguru

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Prof. Masahira Hattori, whose expertise, knowledge and continuous encouragement throughout my research.

My sincere thanks also go to Assoc. Prof. Kenshiro Oshima (The University of Tokyo), Dr. Wataru Suda and Dr. Seok-Won Kim for their motivation, immense support and encouragement throughout my work.

I am also grateful to all my collaborators, Prof. Hidetoshi Morita (Okayama University) for his fecal sample collection, DNA isolation and sincere encouragement, Prof. Kenya Honda and Dr. Koji Atarashi (Keio University) for mice experiments, Assoc. Prof. Masahiro Umezaki (the University of Tokyo) for support for dietary data analysis, Dr. Todd D. Taylor (RIKEN) for support for writing manuscript and Dr. Yuu Hirose (Toyohashi University of technology) for DNA sequencing.

I also would like to thank all past and present members of our laboratory, Erica Iioka, Misa Takagi, Emi Omori, Hiromi Kuroyamagi, Naoko Yamashita, Keiko Komiya, Rina Kurokawa, Chie Shindo, Yukiko Takayama and Yasue Hattori for their great technical support and kind assistance.

Antibiotics shapes population-level diversity in the human gut microbiome

(抗生物質はヒト腸内細菌叢の集団レベルでの多様性を形成する)

Abstract

The human gut microbiome has profound influences on the host's physiology through its interference with various intestinal functions. The development of next-generation sequencing (NGS) technologies enabled us to comprehensively explore ecological and functional features of the gut microbiomes. Recent studies using the NGS-based metagenomic approaches have suggested high ecological diversity of the microbiome across countries. However, little is known about the structure and feature of the Japanese gut microbiome, and the factor that shapes the population-level diversity in the human gut microbiome. In this thesis, to address the above questions regarding the human gut microbiome, I analyzed metagenomic data of fecal DNA samples from healthy Japanese individuals and compared the data with that from individuals in other countries.

I obtained approximately 350 Gb of metagenomic sequences of the gut microbiome of 106 Japanese individuals in this study. By comparing the metagenomic data with that of 757 individuals from other 11 countries, I found that the Japanese gut microbiome showed more abundant in the phylum Actinobacteria, in particular in the genus *Bifidobacterium*, than that of the other 11 nations. Regarding the microbial functions, those of carbohydrate metabolism were overrepresented with a concurrent decrease in those for replication and repair and cell motility in the Japanese gut microbiome. The remarkable low prevalence of genes for methanogenesis with a significant depletion of the archaeon *Methanobrevibacter smithii* and significant enrichment of acetogenesis genes in the Japanese gut microbiome as compared to others

suggested a difference in the hydrogen metabolism pathway in the gut between them. These data suggested considerable uniqueness in the taxonomy and function of the Japanese gut microbiome (Nishijima S. *et al*., DNA res., in press).

To explore the factors that contribute to differences in the human gut microbiomes across the 12 countries, I further conducted an association study of the epidemiological data on dietary intake and antibiotic usage with metagenomic data of the 861 human gut microbiomes from the 12 countries. I found that the gut microbiome structure is significantly diverse across the 12 countries, which was strongly correlated with antibiotics as well as diet. Notably, the abundance of the major species *Bacteroides* showed a significant correlation with both antibiotic usage in humans and farm animals but not with diet; whereas, the abundance of another major species *Prevotella* showed a significant correlation with diet but not with antibiotic usage. Thus, the trade-off relation between these two major species appears to be a consequence of respective independent effects from dietary and antibiotic factors. The proliferation of antibiotic resistant genes, including the efflux pump, may underlie the positive correlation between *Bacteroides* and antibiotic usage. Collectively, these results suggest that antibiotics may have had a striking impact on the shaping of the gut microbiome structure of modern human populations (Nishijima S. *et al*., in preparation).

Table of Contents

1. Introduction

1.1. Human gut microbiomes and study background

A large number of microbes colonize the human body, and they have a profound influence on the host's physiology[1, 2]. The total number of these microbes is estimated to be similar to or more than that of human cells (Fig. 1.1)[1, 3, 4]. The majority of them reside in the intestinal tract, and the gut microbiome (the collective genomes of the microbes) comprises a complex microbial community with more than 100 microbial species. Because of the importance of the gut microbiome for the host's health and disease, it is called the "second genome" or "forgotten organ"[2, 5]. Therefore, the humans can be considered as "superorganism" united with their symbiotic microbiome[6].

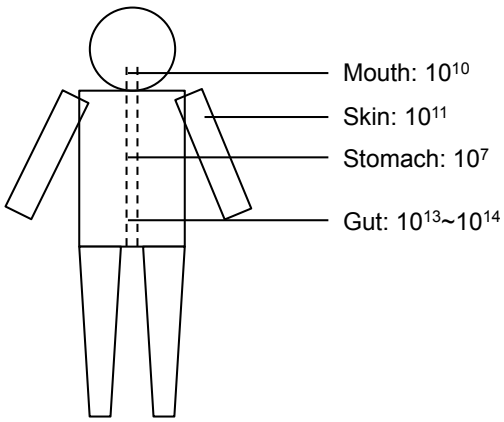


Fig. 1.1. The number of microbes on/in the human body. The number of microbial cells on/in whole human body. Data are from reference 1 and 3.

In the 17th century, Antonie van Leeuwenhoek observed indigenous microorganisms in feces using a microscope[7]. Thereafter, by the development of methods for culturing microorganisms, various commensal bacteria such as *Escherichia coli* and *Bifidobacterium* were isolated and identified from human feces around 1900,

which can be considered to be a starting point of the study of the human gut microbiome[7, 8]. In addition, *in vivo* analytical system using germ-free and gnotobiotic animals was established to evaluate the biological function of the microbes and revealed the indispensable association of the microbial community with the host's physiology[1, 9, 10]. Therefore, the culture-based method has been popularly used for investigating the gut microbial community structure. However, there were a large number of uncultivable species in the community[11]. As a result, knowledge about the gut microbial community obtained from the culture-based method has been limited and many unknown species have been left unanalyzed until recently.

1.2. Development of culture-independent method

To overcome the above mentioned problems encountered in the culture-based method, an alternative approach, a sequencing-based method was developed in 1980s, making it possible to directly acquire the DNA information of the collective genomes prepared from microbes in an environment[12-14]. Since a sequencing-based method did not require the culture step of microbes, these culture-independent methods enabled us to comprehensively elucidate the genomic and taxonomic information of the microbial community containing many uncultivable species.

There are two major culture-independent methods for the study of microbial communities (Fig. 1.2). One is a targeted sequencing of the ribosomal RNA (rRNA) gene in which the 16S rRNA gene regions are collectively amplified by PCR using the specific conserved primers and the amplified products are subsequently sequenced[14]. Variable regions in the 16S rRNA gene are useful for their taxonomic assignment and to know their phylogenetic relations. In the 16S rRNA gene analysis, the taxonomic assignment of microbes can be performed by similarity search of the 16S rRNA gene

sequences against the databases constructed from full-length 16S rRNA gene sequences of known individual microbes. Although this method can be applied only to taxonomic analysis of the microbial community, it can rapidly give us the information of an overview of the microbial content and abundance in the community with relatively low cost. In addition, the phylogenetic tree constructed from the 16S rRNA gene sequences can be used to quantitatively evaluate overall similarity or dissimilarity between different microbial communities[15].

Another is a metagenomic analysis, in which the collective genomes of microbes in an environment are randomly sequenced to collect whole genomic sequence data, from which both taxonomic and functional information can be obtained by employing several appropriate strategies. For example, mapping of the metagenomic reads to the reference genomes containing those of the isolated individual microbes can assign the reads to a particular taxonomy. The genes are computationally identified by using gene prediction softwares in the assembled data of the metagenomic reads or unassembled reads. The analysis of functional assignment of the genes identified in the metagenomic data can be performed by similarity search using KEGG and COG databases[12, 13].

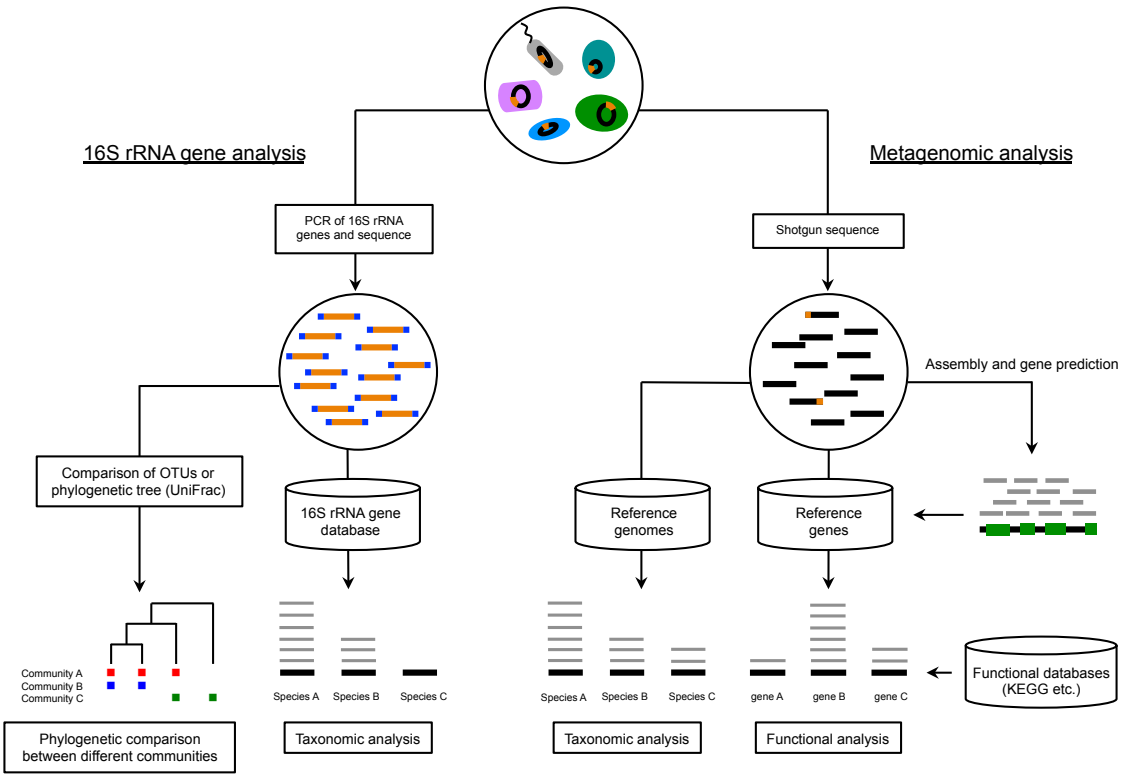


Fig. 1.2. Sequencing-based methods for the study of microbial communities.

1.3. International trends in the metagenomic study of the human microbiome

Up to date, many studies of human gut microbiomes using metagenomic and 16S rRNA gene analysis have been reported (Table 1.1). The first study of the human gut microbiome by metagenomic analysis was reported in 2006, in which the gut microbiomes of two Americans were sequenced[13], and the second study was published in 2007, in which the gut microbiomes of the 13 Japanese individuals were analyzed[16]. These two studies characterized the functional and microbial features of the community by analyzing the genes identified in the metagenomic data produced by the Sanger sequencing. After these two studies, in 2008, two large-scale international projects were

launched to study the human microbiome based on metagenomic analysis using next-generation sequencing (NGS) technologies. One is Metagenomics of the Human Intestinal Tract (Meta-HIT) project by the European Union and China. Another project is the Human Microbiome Project (HMP) by the United States. These two projects aimed to comprehensively analyze the human microbiome with a large dataset of microbiomes from numbers of individuals. In addition, the International Human Microbiome Consortium (IHMC) was also established by scientists from more than 10 countries including Japan, USA, several European countries and China to share technologies and exchange information about the human microbiome researches between the countries. Since then, many of the published projects have been carried out by metagenomic and 16S rRNA gene analysis using NGS.

The Meta-HIT and China groups conducted a large-scale metagenomic analysis of the human gut microbiome and identified 3.3 million (M) genes from the gut microbiomes of 124 European individuals[17], which was about five times more than that identified in the 13 Japanese individuals. The HMP aimed to collect microbiome data from the whole body including the gut, skin, and oral cavity[18, 19]. The HMP is also making an effort to construct the reference genome database comprising sequenced genomes of individual strains isolated from humans in collaboration with other IHMC members[20] and to develop bioinformatic tools and analytical pipelines for metagenomic data[21-23]. In addition to these large-scale projects, several studies targeting more than 100 individuals have been also conducted worldwide[16-18, 24-32].

Table 1.1. Metagenomic studies of human microbiomes.

Year	Sequencer	Research	Reference
2006	Sanger	Metagenomic and 16S rRNA gene analysis of 2 American gut microbiomes	13
2007	Sanger	Metagenomic analysis of 13 Japanese gut microbiomes (healthy and infant)	16
2009	454	Metagenomic and 16S rRNA gene analysis of 18 American gut microbiomes (twin and obesity)	24
2010	Illumina	Metagenomic analysis of 124 Danish and Spanish gut microbiomes (healthy and IBD)	17
2011	Sanger	Comparative metagenomics of 39 individuals and proposal of the concept of Enterotypes	25
2012	Illumina	Metagenomic analysis of 345 Chinese gut microbiomes (healthy and type II diabetes)	26
	Illumina	Metagenomic and 16S rRNA gene analysis of 139 American microbiomes	18
2013	SOLiD	Metagenomic analysis of 96 Russian gut microbiomes	28
	Illumina	Metagenomic analysis of 145 Swedish gut microbiomes (healthy and type II diabetes)	27
2014	Illumina	Metagenomic and 16S rRNA gene analysis of 196 French and German gut microbiomes (healthy and colorectal cancer)	33
	Illumina	Metagenomic analysis of 263 samples of American skin microbiomes	34
	Illumina	Metagenomic analysis of 237 Chinese gut microbiomes (healthy and liver cirrhosis)	29
2015	Illumina	Metagenomic analysis of 156 Austrian gut microbiomes (healthy and colorectal cancer)	30
	Illumina	Metagenomic analysis of 200 Swedish gut microbiomes (Mother and infant)	31
	Illumina	Metagenomic analysis of 212 Chinese oral and gut microbiomes (healthy and rheumatoid arthritis)	32

1.4. Structure of the human microbiome

Current studies based on a large-scale metagenomic and 16S rRNA gene analysis have more or less clarified details of the structure and function of the human microbiome. The human gut microbiome comprises mainly four phyla: Firmicutes, Bacteroidetes, Actinobacteria and Proteobacteria (Table 1.2)[14, 18]. Additionally, other phyla such as Euryarchaeota, Fusobacteria and Verrucomicrobia are also detected as minor phyla. At

the genus level, about 10 to 20 genera represent the majority of the microbial community (Table 1.2). The structure of the community in an individual is relatively stable over at least a few years despite dietary and life style variations in the individuals[35-37]. However, the relative abundances of each phylum, genus and species in the communities are significantly varied between individuals even for twins and within a family[18, 24]. In addition, microbiomes from the skin[34, 38-40], oral[18, 41], vaginal[42, 43], nasal[44], gastric[45], esophageal[46] and placental[47] have been also studied, which revealed significant variations and diversity in the microbial community structures among different body sites.

Table 1.2. Major phyla and genera comprising the human gut microbiome.

Phylum	Genus
Firmicutes	*Clostridium*
	Eubacterium
	Lactobacillus
	Ruminococcus
	Roseburia
	Blautia
	Dorea
	Enterococcus
	Faecalibacterium
	Streptococcus
	Dialister
	Anaerostipes
	Coprococcus
Bacteroidetes	*Bacteroides*
	Prevotella
	Parabacteroides
	Porphyromonas
	Alistipes
Actinobacteria	*Bifidobacterium*
	Eggerthella
	Collinsella
Proteobacteria	*Escherichia*
	Klebsiella
	Bilophila
Fusobacteria	*Fusobacterium*
Verrucomicrobia	*Akkermansia*
Euryarchaeotae	*Methanobrevibacter*

The human gut microbiome encodes an enormous number of and functionally diverse genes particularly involved in carbohydrate metabolism, which mainly metabolize dietary fibers that cannot be digested by the host[16, 48]. The over-representation of the genes is one of the characteristic features of the human gut microbiomes as compared with other environments[16]. Specifically, porphyranase, which degrades the cell walls of aquatic plants, was identified in many of Japanese gut microbiomes but not in those of Americans[49]. The prevalence of the porphyranase gene in the Japanese can be explained by the functional adaptation of the gut microbiome to the Japanese traditional dietary style. In contrast, genes for cell motility such as flagella and chemotaxis are relatively underrepresented in human gut microbiomes[16]. This feature may be due to the unnecessary movement of microbes in the gut because the stool content is stirred by peristalsis, and/or the host's immune system eliminating flagellated microbes that are highly immunogenic.

The formation of the gut microbiome begins after birth by colonizing environmental microbes mainly from mother's skin and the vaginal microbiome. The structure of the microbiome of newborn infants is largely influenced by several factors, such as ways of childbirth (natural childbirth or caesarean section), and breast or formula feeding[31, 50]. Also, the gut microbiome of infants is significantly different from that of children or adults[16, 51], but its structure becomes stable and similar to the adult gut microbiome around three years old[31, 51, 52]. The more long-term variation in the gut microbiome of individuals, such as from birth to old age, has not been studied yet.

The human gut microbiome structure is significantly different from those of microbial communities in other environments such as the sea and soil[53, 54]. Additionally, the gut microbiomes of various mammalians in the zoo and in the wild showed similarities in the taxonomic and functional components to that of humans, but those

possessed their own gut microbiome structure distinct from each other and humans[54-56], suggesting the influence of both host phylogeny and diet, and co-evolution between host and its microbiome while maintaining their symbiotic relation[57, 58].

1.5. Association with health and diseases

The human gut microbiome is profoundly associated with the host's health and diseases[59-61]. For example, the gut microbiome produces short chain fatty acids (SCFA) such as butyrate, acetate and propionate, which are known to be nutrients and have various biological activities to host cells[59, 60, 62]. In total, about 10% of the total calories are estimated to be derived from the gut microbiome in the form of these nutrients[63]. Butyrate induces colonic regulatory T cells, which play a central role in the suppression of autoimmune diseases[64, 65]. Acetate protects the host from infection by the pathogenic bacteria *Escherichia coli* O157:H7[66]. Furthermore, the gut microbiome produces several vitamins (vitamins B and K and others), some of which are essential for human health because the human genome lacks the biosynthesis genes of these vitamins[60, 62]. These bacterial metabolites are absorbed by epithelial cells and delivered throughout the body, which accounts for about 10% of total metabolites in the blood[67].

The aberrant structure of the gut microbiome is associated with various diseases such as obesity[24, 68], inflammatory bowel disease (IBD)[69, 70], colorectal cancer[30, 33, 71] and type 2 diabetes[26, 27] (Table 1.3). Typically, gut microbiomes in these patients commonly show an imbalance in the community (called dysbiosis) and a low microbial diversity[24, 68] (e.g., a lower number of species in samples with disease). A depletion of butyrate-producing species such as *Faecalibacterium* and other *Clostridiales* is often observed in the gut microbiomes of many of these patients[26, 72, 73]. In addition, several recent studies have indicated an association of the altered gut microbiomes with

neurologic diseases such as autism and multiple sclerosis[74-76], and normal gut microbiome with brain development and host's behavior, suggesting a tight interaction between the gut microbiome and the host's nervous system through the gut-brain axis[77, 78]. Thus, the human gut microbiome has a systemic impact on the host's physiological states (Fig. 1.3).

In 2013, van Nood *et al.* reported that fecal microbiota transplantation (FMT) showed an extremely higher therapeutic effect on recurrent *Clostridium difficile* infection than conventional antibiotic treatment alone[79]. FMT involves transplanting fecal microbiota from a healthy donor into the gut of a patient, suggesting the gut microbiome of healthy individuals has a strong biological activity for the host's physiologies. For therapeutics using FMT, many other studies have also been performed to better understand the recovery process of transplanted microbial community in patients[80], develop safer and more efficient FMT manipulations[81-83], and apply its use for other diseases associated with the gut microbiome such as inflammatory bowel disease[84, 85]. In addition, a variety of clinical applications using the gut microbiome have also been studied, which include exploration of beneficial microbes, development of drugs targeting or using the microbiome, and identification of microbial biomarkers specific to particular diseases[86].

Table 1.3. Diseases associated with the human gut microbiome.

Disease	Reference
Obesity	10, 24, 68
Inflammatory bowel disease	69, 70, 87
Colorectal cancer	30, 33, 71
Type II diabetes	26, 27
Type I diabetes	88, 89
Rheumatoid arthritis	32, 90
Atherosclerosis	91, 92
Irritable bowel syndrome	93, 94
Malnutrition	95, 96
Multiple sclerosis	72
Liver cirrhosis	29
Allergy	97
Eczema	98
Liver cancer	99
Autism	74-76

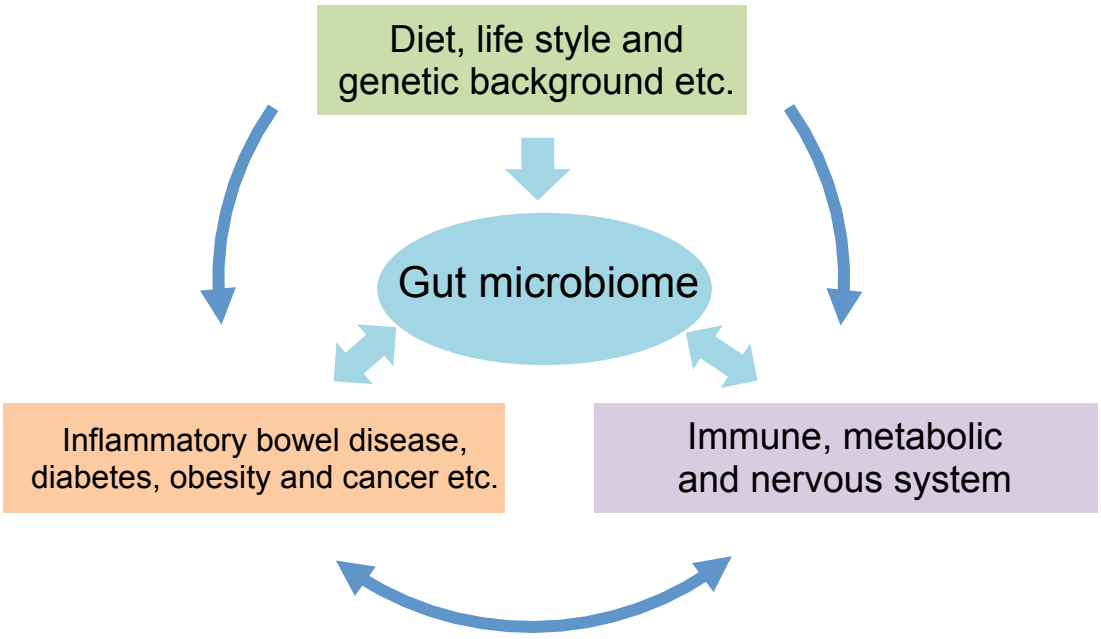


Fig. 1.3. Relationship between the host's physiology and the gut microbiome.

2. The gut microbiome of healthy Japanese and its microbial and functional uniqueness

2.1. Introduction

Various cohort studies of the human gut microbiome based on a metagenomic approach using NGS have been reported[17, 18, 24, 26-33]. These studies included patients with diseases such as obese[24], inflammatory bowel disease[17], type 2 diabetes[26, 27], colon cancer[30, 33], liver cirrhosis[29], and rheumatoid arthritis[32] patients, as well as numbers of healthy individuals in various countries including the United States, several European countries, and China. In addition, several studies have reported on gut microbiomes of Asian children and natives from rural areas[51, 100-103]. These studies suggested that the human gut microbiome is more or less affected by various factors such as diet and host's genetic background[104], and the altered microbiome is associated with diseases[105]. In addition, several comparative analyses have suggested a great diversity in human gut microbiomes at the population level[27, 28, 51, 100-103, 106], and even in those of patients with diseases[27, 107]. More basically, human gut microbiomes can be classified into 'enterotypes' on the basis of differences in the abundance of a few major species largely linked with dietary habits[25, 108].

Japanese have unique dietary culture and habits as compared with Western people, being reflected in the finding that their gut microbiomes have more genes for aquatic plant-derived polysaccharide-degrading enzymes with higher frequency than those of Americans[49]. In addition, Japanese exhibit the highest average life span and very low body mass index (BMI)[109]. A study on the gut microbiomes of 13 Japanese individuals has been previously published[16]. However, the dataset size was too small to allow comparison with other large datasets to precisely evaluate distinct features of the

Japanese gut microbiome (JPGM). Therefore, in this study, I collected and analyzed the metagenomic data from gut microbiomes of 106 Japanese individuals by sequencing of fecal DNA samples using NGS, and I further explored the unique microbial and functional features of the JPGM by comparing the microbiomes of a total of 861 healthy individuals selected from Japan and 11 other countries. The results of this study are shown in this chapter.

2.2. Methods

2.2.1. Subjects and fecal sample collection

One hundred and six healthy Japanese volunteers (age: 32 ± 11, BMI: 22 ± 2.7 [mean ± s.d.]; Appendix 1) were recruited by Azabu University (Japan). This study was approved by the Human Research Ethics Committee of Azabu University and the Research Ethics Committee of the University of Tokyo, and written consent was obtained from all subjects. No subjects were treated with antibiotics during fecal sample collection. Among them, fecal samples were longitudinally collected twice from 26 individuals every eight weeks and five times from nine individuals every two weeks, of which 16 individuals were shared with the previous study[110]. In total, 168 fecal samples were collected from the 106 individuals. The collected fresh feces were stored under anaerobic conditions using an AneroPack™ (Mitsubishi Gas Chemical Co. Inc., Tokyo, Japan) at 4 ºC. Within 36 hours after sampling, the feces were frozen in 20% glycerol (Wako Pure Chemical Industries, Osaka, Japan)/phosphate buffer saline (PBS) solution (Life Technologies, Tokyo, Japan) by liquid nitrogen and stored at −80 ºC until use.

2.2.2. Recovery of bacteria and archaea from fecal samples

Each frozen fecal sample (1.0 g) was thawed on ice and suspended vigorously in a 50 mL tube. The suspension was filtered with a 100 μm-mesh nylon filter (Becton Dickinson, Tokyo, Japan) to separate the bacterial cells from most of eukaryotic cells and other debris. When I compared microbial compositions at the genus level from seven samples with and without the filter, the Pearson's correlations between them were from 0.90 to 0.99, suggesting almost no significant difference in microbial compositions

between with and without filtration of feces. The debris on the filter was washed off twice using a glass or plastic bar with 10 mL PBS buffer. The filtrate was centrifuged at 5,000 X g for 10 min at 4 ºC. The bacterial pellet was rinsed twice with PBS, and finally with TE10 buffer (10 mM Tris-HCl, 10 mM EDTA, pH 8.0).

2.2.3. DNA isolation and purification

Each fecal DNA sample was isolated and purified according to the literature[110-112] with minor modifications. The bacterial pellet was suspended in 10 mL of TE10 buffer and incubated with 15 mg/mL lysozyme (Sigma-Aldrich Co. LCC., Tokyo, Japan) at 37 ºC for 1 h. Purified achromopeptidase (Wako) was added at a final concentration of 2,000 units/mL and further incubated at 37 ºC for 30 min. The suspension was treated with 1% (wt/vol) sodium dodecyl sulfate (SDS) and 1 mg/mL proteinase K (Merck, Tokyo, Japan), and incubated at 55 ºC for 1 h. The lysate was mixed with equal volumes of phenol/chloroform/isoamyl alcohol (Life Technologies) and centrifuged at 5,000 g for 10 min. DNA was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 4.5) and 2 volumes of ethanol, and pelleted by centrifugation at 5,000 g at 4 ºC for 15 min. The DNA pellet was rinsed with 75% ethanol, dried and dissolved in TE. DNA samples were treated with 1 mg/mL RNase A (Wako) at 37 ºC for 30 min, precipitated by adding equal volumes of 20% PEG solution (PEG6000-2.5 M NaCl), and kept on ice for 10 min. DNA was pelleted by centrifugation at 20,000 g at 4 ºC for 10 min, rinsed twice with 75% ethanol and dissolved in TE buffer.

2.2.4. Metagenomic sequencing of fecal DNA

The sequencing of each fecal DNA sample was performed by the 454 (Roche), Ion PGM/Proton (Life Technologies) and MiSeq (Illumina) platforms according to the

suppliers' protocols, respectively. For 454, 5 μg of fecal DNA was sheared to obtain fragments ranging from 300 to 700 bp for the FLX Titanium platform and 500 to 1,000 bp for the FLX+ platform. The libraries were prepared using the GS FLX Titanium Rapid Library MID Adaptors Kit. For Ion PGM/Proton, 100 ng of fecal DNA was sheared to obtain fragments ranging from 350 to 470 bp and the library was prepared using the Ion Xpress Plus Fragment Library Kit. For the 454 and Ion PGM/Proton reads, artificially redundant reads were removed using a replicate filter if any sequences had ≥ 95% identity to other sequences with exactly the same starting point[113]. Reads that mapped to the human genome (HG19) with Newbler (version 2.7) were also removed. Finally, reads with an average Quality Value (QV) less than 20 or less than 75 bp in length were removed. For 150 bp paired-end sequencing of MiSeq, 20 ng of fecal DNA was sheared to obtain fragments ranging from 300 to 400 bp and the library was prepared using the TruSeq DNA Sample Prep Kit. For 300 bp paired-end sequencing by MiSeq, fecal DNA library was prepared using the Nextera DNA Sample Prep Kit. Any 5' end low quality (< 20 QV) bases in MiSeq reads were trimmed off. Reads having bases less than 20 QV for more than half of the read length and reads whose length was less than 50 bp, were also filtered out. These procedures were done using the FASTX-Tool kit (http://hannonlab.cshl.edu/fastx_toolkit/). The filter-passed reads were then mapped to the human and phiX genomes using Bowtie2[114] (version 2.2.1) and any mapped reads were removed. Sequencing statistics are summarized in Appendix 1.

2.2.5. Assembly and gene prediction for metagenomic sequences

For each individual sample, the filter-passed reads were assembled by Newbler assembler for 454 and Ion PGM/Proton, and MiSeq separately. The contigs (≥ 500bp)

generated from assembly of the reads of 454 and IonPGM/Proton, and MiSeq were further assembled with Minimus2[115] with default settings for each individual. Un-assembled reads (singletons) of all individuals were merged and re-assembled again for each sequencer. MetageneAnnotator[116] was employed to predict protein-coding genes (≥ 100 bp) in the contigs (≥ 500 bp) and singletons from 454 longer than 300 bp. The genes were clustered using CD-HIT[117] with a 95% identity and 90% length coverage thresholds, in which a longest gene in the cluster was selected as the representative gene.

2.2.6. Publically available metagenomic data

I collected publically available metagenomic data of individuals from Denmark (DK), Spain (ES), the United States (US), China (CN), Sweden (SE), Russia (RU), Venezuela (VE), Malawi (MW), Austria (AT), France (FR) and Peru (PE). Metagenomic reads from DK[17] and ES[17] were downloaded from http://public.genomics.org.cn. Filter-passed reads and non-redundant genes from US[18] were downloaded from the HMP DACC (http://www.hmpdacc.org). Raw reads from DK[68], ES[106], CN[26, 29], SE[27], RU[28], AT[30], FR[33], PE[102], Yanomami (VE)[118] and US[102] were also downloaded from public databases and quality control steps were conducted using the same methods described above. The SOLiD reads from RU were subjected to error correction using the SOLiD Accuracy Enhancement Tool (SAET), and the quality control steps were performed as described in the previous study[28]. Raw reads for Amerindian (VE)[51], MW[51] and US[51] were downloaded from MG-RAST (http://metagenomics.anl.gov) and reads mapped to the human genome were excluded.

2.2.7. Country-specific metagenomic datasets of healthy individuals

To construct metagenomic datasets consisting of healthy individuals from each country, the data for individuals with BMI ≥ 30, those with diseases such as inflammatory bowl disease, type 2 diabetes, liver cirrhosis or colorectal cancer, and infants age < 3 years old were excluded from the data collected from a total of 1,734 individuals. I then combined the data from remaining healthy individuals per each country to construct datasets for each country. Although I could not access the metadata for the individuals in US[18], I used all the data with an average BMI of 24 ± 4 (s.d.) for this cohort. In total, 861 individuals from the 12 countries were selected and used for the analysis (Table 3.2).

2.2.8. Construction of microbial reference genomes

For the microbial composition analysis, I used in-house reference genome database comprised of a total of 6,107 genomes representing 2,373 clusters at the species level of Bacteria and Archaea, which were selected from 2,788 complete and 22,317 draft genomes available in GenBank/DDBJ/EBI (as of July 2014), 20 genomes from the study published by Atarashi *et al.*[119] and two unpublished genomes in my laboratory. The reference genomes are listed in Appendix 2.

The reference genome database was constructed by the following procedures. First, genomes matching with either of the following criterion were selected as references; (i) genomes mapped with ≥ 10 metagenomic reads when 60 M metagenomic reads from six countries (Japan, China, Denmark, Spain, Sweden and the United States) were mapped to the 25,085 genomes in GenBank/DDBJ/EBI (BLASTN; 95% identity and 90% length coverage). (ii) genomes for which the 16S rRNA gene, identified with

RNAmmer[120], were mapped with ≥ 10 in-house 16S rRNA V1-V2 region sequences of human microbes, which comprised of about 600 thousand reads (BLASTN; 85% identity and 90% length coverage). Second, 25 typical known pathogenic species such as *Bacillus anthracis*, *Bordetella pertussis, Burkholderia pseudomallei, Campylobacter coli, Campylobacter jejuni, Clostridium botulinum*, *Clostridium chauvoei*, *Clostridium tetani, Corynebacterium diphtheria, Francisella tularensis, Leptospira interrogans, Listeria monocytogenes, Mycobacterium abscessus, Mycobacterium tuberculosis*, *Salmonella enterica, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Vibrio cholera, Vibrio vulnificus* and *Yersinia pestis*, and four genera including *Borrelia*, *Chlamydia*, *Mycoplasma*, and *Rickettsia* were excluded from the reference dataset. Third, to reduce complexity and excess load in computing, for the species with ≥ 50 sequenced genomes at the strain level, some of the genomes were excluded to the extent that the genomes still covered ≥ 99% of the total reads mapped to the species. Those included *Acinetobacter baumannii*, *Bacillus cereus*, *Bacteroides fragilis*, *Enterococcus faecalis*, *Enterococcus faecium*, *Escherichia coli*, *Helicobacter pylori*, *Klebsiella pneumonia*, *Peptoclostridium difficile*, *Propionibacterium acnes*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pneumonia*, *Streptococcus pyogenes* and *Streptococcus suis*.

The reference genomes selected above were further clustered to reduce the complexity at the species level. The 16S rRNA genes of the reference genomes were clustered with a 98.8% identity cut-off, and the obtained clusters were defined as single species. Reads that mapped to the genomes in the same cluster were merged and assigned to the representative species of the cluster. For a few clusters that were composed of obviously different species, such as *Streptococcus salivarius* and

Streptococcus thermophiles, both of which have 16S rRNA genes of > 98.8% identity, I manually separated these clusters into different species. Of a few species, such as *Fusobacteroium nucleatum* and certain *E. coli* strains, that formed distinct clusters, even when the species' names were identical, the species/clusters were merged when a sufficient number of multi-hit reads were commonly shared among them. Several draft genomes lacking 16S rRNA genes were assigned to the most similar species or clusters when the species' names were related and multi-hit reads were commonly shared among the genomes.

2.2.9. Microbial composition by mapping of metagenomic reads to the reference genomes

One M metagenomic reads per individual were mapped to the reference genomes using Bowtie2 with a 95% identity threshold. For the SOLiD reads from RU, Bowtie[121] (version 0.12.7) was employed with the same threshold. For several samples of which the number of metagenomic reads was less than one M, all of the reads for the individual were mapped to the reference genomes. In the 861 individuals selected, the minimum number of reads per individual was about 60,000, but Pearson's correlation coefficients (PCCs) between microbial compositions obtained from the mapping of 1 M reads and 60,000 reads from several same Japanese individuals was 0.99996, indicating that the number of reads per individual between 60,000 and 1 M did not significantly affect the results obtained from the mapping analysis in this study. The number of multi-hit reads that mapped to several different genomes with equal scores were divided among those genomes in proportion to the number of reads uniquely mapped to each genome. A similar normalization was also used for the quantification of transcripts in the RNA-Seq analysis[122]. For genome g, I defined the abundance π_g as follows,

$$\pi_g = \frac{U_g + \sum_{r \in GtoR(g)} P_{r,g}}{l_g}$$

where U_g is the number of reads that are uniquely mapped to genome g, *GtoR(g)* is the set of reads that are equally mapped to several genomes including genome g, and l_g is the length of g. $P_{r,g}$ is the probability that a read r is assigned to genome g, and is calculated as follows,

$$P_{r,g} = \frac{U_g}{\sum_{g' \epsilon RtoG(r)} U_{g'}}$$

where *RtoG(r)* is a set of genomes to which a read *r* mapped. The relative abundance of each genome was calculated by normalizing the number of reads mapped to the genome by the total number of reads. NCBI taxonomy information was used for taxonomic assignment of genus and species for each genome. Genomes that could not be accurately assigned to a particular genus were assigned to their higher rank classification and designated as "Unclassified a higher rank".

2.2.10. Comparison of microbial compositions among the countries

A multi-dimensional scaling (MDS) plot was constructed using the Jensen-Shannon divergence between the microbial compositions at the genus level of the 861 individuals. Hierarchical clustering of the countries based on the average microbial composition at the genus level was performed using the Ward method and the Bray-Curtis distances. Construction of a predictive model for each country based on microbial composition at the genus level was performed using the randomforest package in R. Evaluation of the predictive power of the model was conducted by 10-fold cross-validation with 90% of the training data and 10% of the prediction data. The number of trees was set to 500 and

the sample size option was set to the minimum number of individuals among the countries. The receiver operating characteristic (ROC) and area under the ROC curves (AUCs) of the predictive model were calculated by a one vs. all approach and plotted with the smooth function using the ROCR and pROC packages. Pearson's correlation coefficients were calculated between individuals within and between countries and the statistical differences between them were evaluated by permutation test with 10,000 random samplings.

2.2.11. Assessment and comparison of different methodologies

To evaluate the effect of different methodologies on the metagenomic analysis, the same fecal samples were subjected to sequencing with different sequencers, different DNA extraction methods, and different fecal sample storage conditions (Table 3.3). The microbial compositions at the genus level were calculated with the method described above. Similarity of the microbial compositions was evaluated using Pearson's correlation coefficient. Permutation testing with 10,000 times randomization was conducted to test the statistical significance of the similarity between the data obtained by different methodologies and between individuals within and between countries.

2.2.12. PCR detection of *Methanobrevibacter smithii* in the Japanese individuals

M. smithii was detected by PCR using *M. smithii* 16S rRNA gene-specific primers 5′-ATGCACCTCCTCTCAGCTAGTC-3′ and 5′-AGAGGTACTCCCAGGGTAGAGG-3′, of which sequences were designed using Primer3[123]. PCR was conducted in 10.0 μL PCR solution containing 0.2 μL of template

DNA, 0.02 μL of each primer, 1.0 μL of 10 × PCR buffer, 1.0 μL of dNTP mixture, 0.04 μL of Ex Taq polymerase (Takara Bio Inc., Shiga, Japan) and 7.52 μL of ddH_2O using GeneAmp PCR System 9700 (Applied Biosystems, Tokyo, Japan) with 40 cycles of denaturation (30 sec at 96°C), annealing (20 sec at 60°C), and elongation (3 min at 70°C). The PCR products were separated on 1.5% of agarose gels with a positive control from genomic DNA of *M. smithii* JCM 30028^T. PCR without DNA was also performed as negative control. Genomic DNA of *M. smithii* JCM 30028^T was obtained from Japan Collection of Microorganisms, RIKEN BRC.

2.2.13. Generation of a merged reference gene set of Japanese and integrated gene catalog

I constructed the merged reference gene set by clustering the JP non-redundant genes (4.9 M) and the non-redundant genes (9.9 M) in the integrated gene catalogue (IGC)[106], which was constructed from metagenomic data of more than 1,000 individuals from DK, ES, US and CN using CD-HIT with a 95% nucleotide identity and 90% length coverage cut-off. The IGC genes were downloaded from http://meta.genomics.cn/metagene/meta/home.

2.2.14. Functional assignment of non-redundant genes

Functional assignment of the non-redundant genes was performed using BLASTP searches (e-value ≤ 1.0e-5) against the KEGG (Kyoto Encyclopedia of Genes and Genomes) database (release 63) to obtain the KEGG orthologies (KOs). The genes with a besthit to eukaryotic genes were excluded from further analysis. Additionally, The non-redundant genes were searched using BLASTP (e-value ≤ 1.0e-5) against the eggNOG database (version 4.5)[124] to assign them to non-supervised orthologous groups

(NOGs). The phylum-level taxonomic assignment of the genes assigned to NOGs was conducted by BLASTN searches to the reference genomes with a ≥ 65% identity and ≥ 85% length coverage cut-off.

2.2.15. Quantification of the annotated genes in human gut microbiomes

One M metagenomic reads per individual were mapped to the JP and IGC merged reference gene set using Bowtie2 with a 95% identity cut-off. To adjust the mapping conditions for long reads (*e.g.,* 454 FLX+ reads with an average read length of 730 bp) to the short reference genes with an average length of 620 bp, reads in the JP dataset > 100 bp were split into 100 bp fragments, which were used independently for similarity searches, while fragments < 50 bp were discarded. The number of reads that mapped equally to more than one gene was normalized by the number of reads uniquely mapped to the genes as was conducted for the reference genome mapping. The quantities of KOs were calculated from the number of reads mapped to them. I first compared the relative abundance of KOs between the 104 JP individuals and the 757 individuals in the other 11 countries using Student's *t*-test to enumerate the KOs enriched and depleted with statistical significance (FDR adjusted p-value < 0.01) in the JPGM. Of them, those having the highest and lowest relative abundance among the 12 countries were identified. P-values were adjusted for multiple testing using p.adjust(p, "BH") in R language, which is based on the Benjamini-Hogberg approach[125].

2.3. Results

2.3.1. Metagenomic sequencing of the Japanese gut microbiome

My colleagues collected 168 fecal samples from 106 healthy JP individuals (age: 32 ± 11, BMI: 22±2.7 (mean ± s.d.); Appendix 1), extracted DNA from them with enzymatic lysis method and sequenced those using 454, Ion PGM/Proton and MiSeq sequencers. After the quality filtering, in total, I obtained about 350 gigabases (Gb) of metagenomic reads with an average of 3.4 ± 1.9 Gb (mean ± s.d.) for each individual (Appendix 1).

I performed *de novo* assembly of the metagenomic reads of the JPGM using Newbler for each individual to generate 13 Gb of contigs (≥ 500 bp) and 0.6 Gb of singletons (≥ 300 bp) (Fig. 2.1). In the sequences, approximately 23 M genes (the length ≥ 100bp) were predicted with MetaGeneAnnotator. I further clustered them with CD-HIT with a 95% nucleotide identity and 90% length coverage cut off. This procedure excluded many redundant genes shared among the individuals. Finally, I obtained 4,854,719 non-redundant genes as the reference gene set of the JPGM (Fig. 2.1).

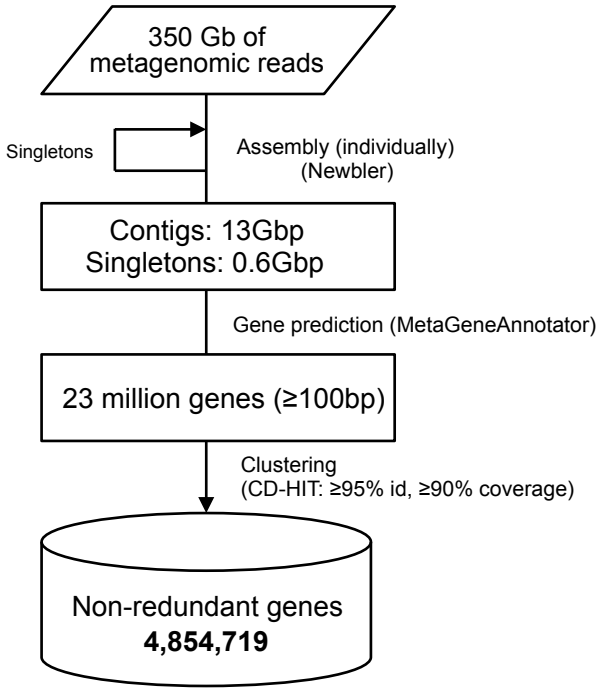


Fig. 2.1. Assembly and gene prediction of the metagenomic reads of the JPGM.

2.3.2. Analysis of the non-redundant genes of the Japanese gut microbiome

Rarefaction analysis of the number of non-redundant genes against the number of individuals sequenced showed that the genes shared by ≥ 1.9% (2/106) in the JP individuals were almost constant with approximately 100 individuals, while unique genes detected only in an individual were increased even beyond 100 individuals (Fig. 2.2a). The number of the genes shared by ≥ 1.9% in the JP individuals accounted for about 3.8 M genes, which covered 98.8% of the total mapped reads to the JP gene set. This result suggested that the reference gene set covers most of the genes encoded by the JPGM. The number of the genes shared by ≥ 50% in the JP individuals was

approximately 0.21 M, which accounted for about only 4% of the total JP gene set, indicating that only the small fraction constitutes the core gene set of the JPGM (Fig. 2.2b).

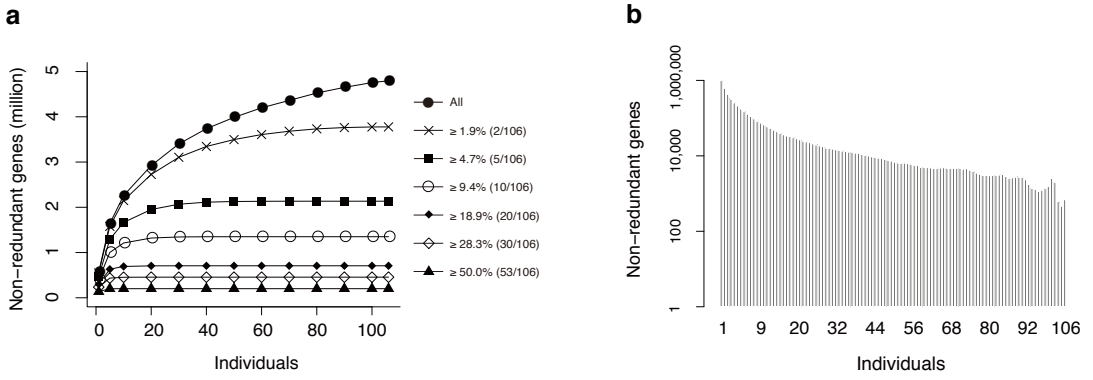


Fig. 2.2. Non-redundant genes in the JPGM. (a) The numbers of detected non-redundant genes plotted against the numbers of the JP individuals. Different symbols show the frequency of the genes in the JP individuals. (b) The number of the non-redundant genes shared by each number of the individuals plotted against the number of individuals.

Functional analysis of the non-redundant genes annotated with the KEGG database showed that the individual-specific genes contained significantly more functions involved in restriction-modification system such as type I restriction enzyme (K01154 and K03427), DNA methyltransferase (K00588) and site-specific DNA-methyltransferase (K00571) than the core genes shared by ≥ 50% in the JP individuals (Table 2.1a). Methyltransferases have the ability to frequently exchange their target recognition domains between them to increase the sequence diversity in these proteins[126], suggesting that restriction-modification system is a driving force for specifying individual gut microbiomes. On the other hand, the core genes were significantly rich in functions related to horizontal gene transfer (HGT) such as conjugal transfer ATP-binding protein TraC (K12063), site-specific DNA recombinase (K06400)

and integrase (K14059) in addition to many of essential genes such as ribosomal proteins (K02935, K02886, K02950 and K02967) as compared with individual-specific genes (Table 2.1b). From the finding that the genes involved in HGT are highly conserved, I hypothesize that HGT is crucial more than considered previously for functional adaptation of the gut microbial community to various ecological changes in the gut because novel or additional genes are acquired and spread in the microbial community by HGT. Similar results were also obtained in the analysis of gut microbiomes of China, Denmark, Spain and the United States[106], suggesting that the accelerated evolution of the individual-specific restriction-modification system and the highly conserved genes involved in HGT is a common feature of the human gut microbiome.

Table 2.1. Functions enriched in individual-specific genes (a) and core genes (b).

a

KEGG	Function	FDR adjusted p-value
K01154	type I restriction enzyme, S subunit [EC:3.1.21.3]	2.40E-180
K00754	not assigned	4.19E-97
K00558	DNA (cytosine-5-)-methyltransferase [EC:2.1.1.37]	1.80E-79
K01153	type I restriction enzyme, R subunit [EC:3.1.21.3]	8.75E-79
K00571	site-specific DNA-methyltransferase (adenine-specific) [EC:2.1.1.72]	8.66E-67
K03427	type I restriction enzyme M protein [EC:2.1.1.72]	2.10E-57
K07316	adenine-specific DNA-methyltransferase [EC:2.1.1.72]	5.36E-39
K00012	UDPglucose 6-dehydrogenase [EC:1.1.1.22]	3.08E-37
K07741	anti-repressor protein	4.10E-34
K01156	type III restriction enzyme [EC:3.1.21.5]	5.50E-32
K01791	UDP-N-acetylglucosamine 2-epimerase [EC:5.1.3.14]	4.61E-23
K08679	UDP-glucuronate 4-epimerase [EC:5.1.3.6]	3.19E-21
K06909	phage terminase large subunit	1.21E-19
K03328	polysaccharide transporter, PST family	4.22E-19
K06223	DNA adenine methylase [EC:2.1.1.72]	6.76E-19
K06915	not assigned	1.42E-18
K07474	phage terminase small subunit	5.60E-18
K01155	type II restriction enzyme [EC:3.1.21.4]	5.84E-18
K00786	not assigned	5.20E-16
K07459	putative ATP-dependent endonuclease of the OLD family	1.01E-15
K02334	DNA polymerase bacteriophage-type [EC:2.7.7.7]	1.56E-15
K02337	DNA polymerase III subunit alpha [EC:2.7.7.7]	1.71E-15
K06904	not assigned	4.79E-15
K15125	filamentous hemagglutinin	7.06E-15
K09952	hypothetical protein	2.29E-14
K01599	uroporphyrinogen decarboxylase [EC:4.1.1.37]	2.89E-14
K15914	N-acetyl-D-galactosamine transferase [EC:2.4.1.-]	1.24E-13
K03657	DNA helicase II / ATP-dependent DNA helicase PcrA [EC:3.6.4.12]	1.24E-13
K15342	CRISP-associated protein Cas1	1.58E-13
K06877	DEAD/DEAH box helicase domain-containing protein	2.88E-13

b

KEGG	Function	FDR adjusted p-value
K00936	not assigned	1.52E-69
K12063	conjugal transfer ATP-binding protein TraC	4.08E-67
K06400	site-specific DNA recombinase	1.05E-41
K00599	not assigned	6.44E-38
K03205	type IV secretion system protein VirD4	4.24E-37
K03088	RNA polymerase sigma-70 factor, ECF subfamily	7.01E-35
K07467	phage replication initiation protein	3.49E-33
K01144	exodeoxyribonuclease V [EC:3.1.11.5]	3.02E-21
K00689	dextransucrase [EC:2.4.1.5]	2.66E-19
K02574	ferredoxin-type protein NapH	2.66E-19
K02358	elongation factor Tu	4.17E-18
K07706	two-component system, AgrA family, sensor histidine kinase AgrC [EC:2.7.13.-]	3.05E-17
K01420	CRP/FNR family transcriptional regulator, anaerobic regulatory protein	1.56E-15
K14059	integrase	4.35E-13
K02855	AraC family transcriptional regulator, L-rhamnose operon regulatory protein RhaS	8.84E-13
K02982	small subunit ribosomal protein S3	1.34E-12
K07496	putative transposase	1.34E-12
K11527	two-component system, unclassified family, sensor histidine kinase and response regulator	3.16E-12
K07487	transposase	1.16E-11
K07814	putative two-component system response regulator	5.27E-11
K02935	large subunit ribosomal protein L7/L12	1.48E-10
K02886	large subunit ribosomal protein L2	1.55E-10
K02030	polar amino acid transport system substrate-binding protein	2.61E-10
K07052	not assigned	4.00E-10
K10117	multiple sugar transport system substrate-binding protein	6.43E-10
K02950	small subunit ribosomal protein S12	9.24E-10
K02967	small subunit ribosomal protein S2	1.28E-09
K02099	AraC family transcriptional regulator, arabinose operon regulatory protein	1.28E-09
K03327	multidrug resistance protein, MATE family	1.74E-09
K07775	two-component system, OmpR family, response regulator ResD	5.07E-09

P-values were calculated using the Fisher's exact test and adjusted for multiple testing.

2.3.3. Structure of the Japanese gut microbiomes

I analyzed the gut microbiomes from the 106 JP individuals by mapping the metagenomic reads to the reference genomes and genes. I identified 14 phyla and 178 genera of Bacteria and Archaea in the gut microbiomes of the 106 JP individuals (average relative abundance ≥ 0.001%). The obtained microbial compositions were mainly composed of four phyla such as Firmicutes (59.7 ± 14.7% (mean ± s.d.)), Actinobacteria (21.8 ± 16.3%), Bacteroidetes (16.7 ± 10.4%) and Proteobacteria (1.4 ± 1.6%) (Fig. 2.3a)[14, 18]. At the genus level, *Bifidobacteirum* (18 ± 15%), *Blautia* (17 ± 8.7%), *Bacteroides* (11 ± 8.5%), *Eubacterium* (6.5 ± 6.0%), *Faecalibacterium* (5.7 ± 4.7%) and *Ruminococcus* (5.6 ± 5.4%) dominated (mean relative abundance ≥ 5%) (Fig. 2.3b and 2.4), and at the species level, *Bifidobacterium adolescentis* (6.0 ± 8.9%), *Bifidobacterium pseudocatenulatum* (5.4 ± 7.5%), *Bifidobacterium longum* (4.6 ± 4.0%), *Blautia wexlerae* (4.3 ± 3.6%) and *Blautia* sp. CAG:37 (4.0 ± 5.1%) (mean relative abundance ≥ 3%) dominated in the JPGM (Fig. 2.3c and 2.4). Consistent with the previous studies, I observed high taxonomic diversity in the microbial compositions between individuals. The factors that determine the diversity in the microbial compositions of individuals have been poorly understood. On the other hand, the functional categories of the genes showed a relatively low diversity between individuals as compared with that in the microbial composition (Fig. 2.3). Similar results were also observed in the previous study of American gut microbiomes, implying the strong selective pressure for functions in shaping of the human gut microbiome[18, 24].

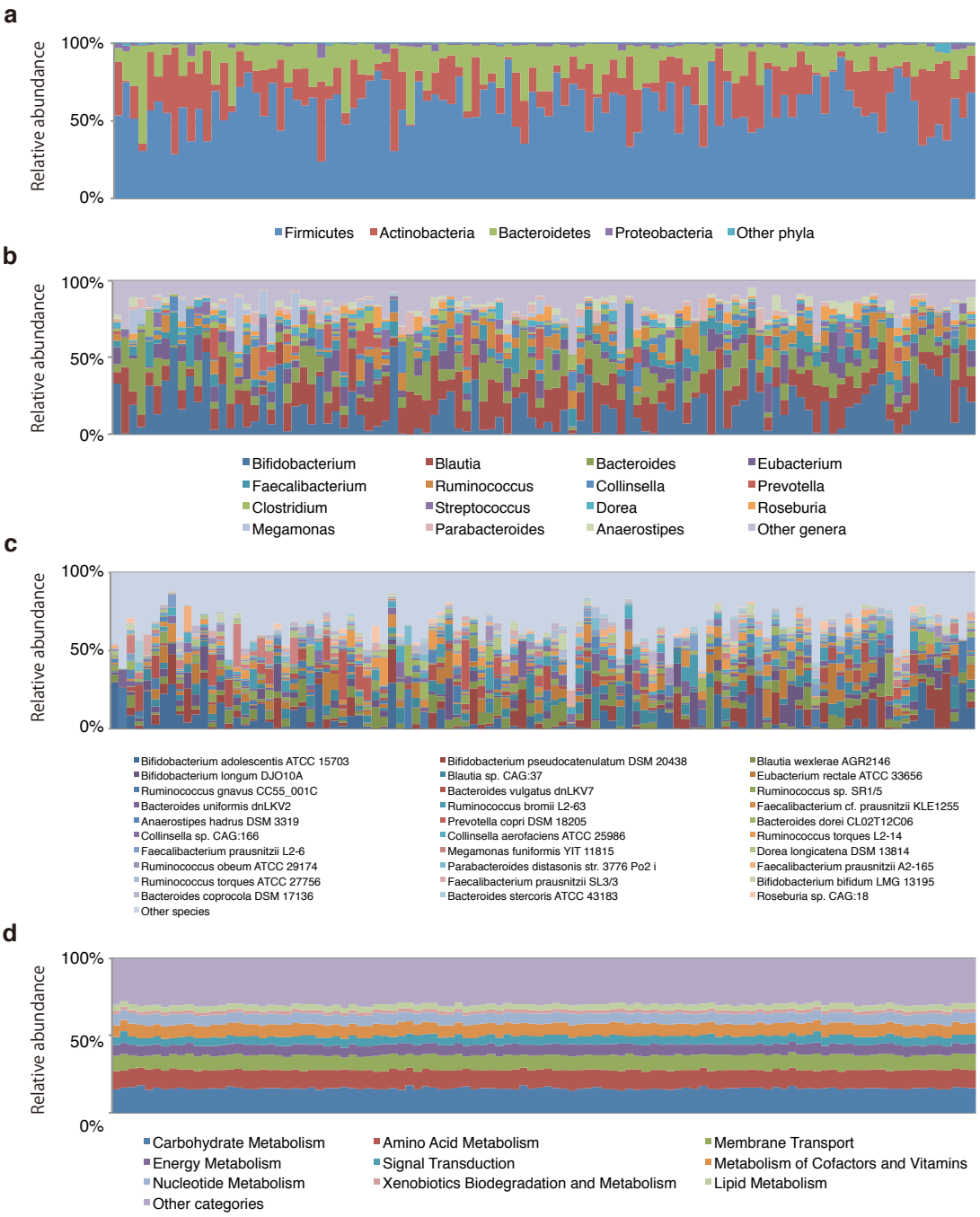


Fig. 2.3. Microbial and functional compositions of the 106 JP individuals. Metagenomic reads were analyzed with the pipeline and the microbial compositions at the phylum (a), the genus (b), the species level (c) and the functional profiles (d) were shown.

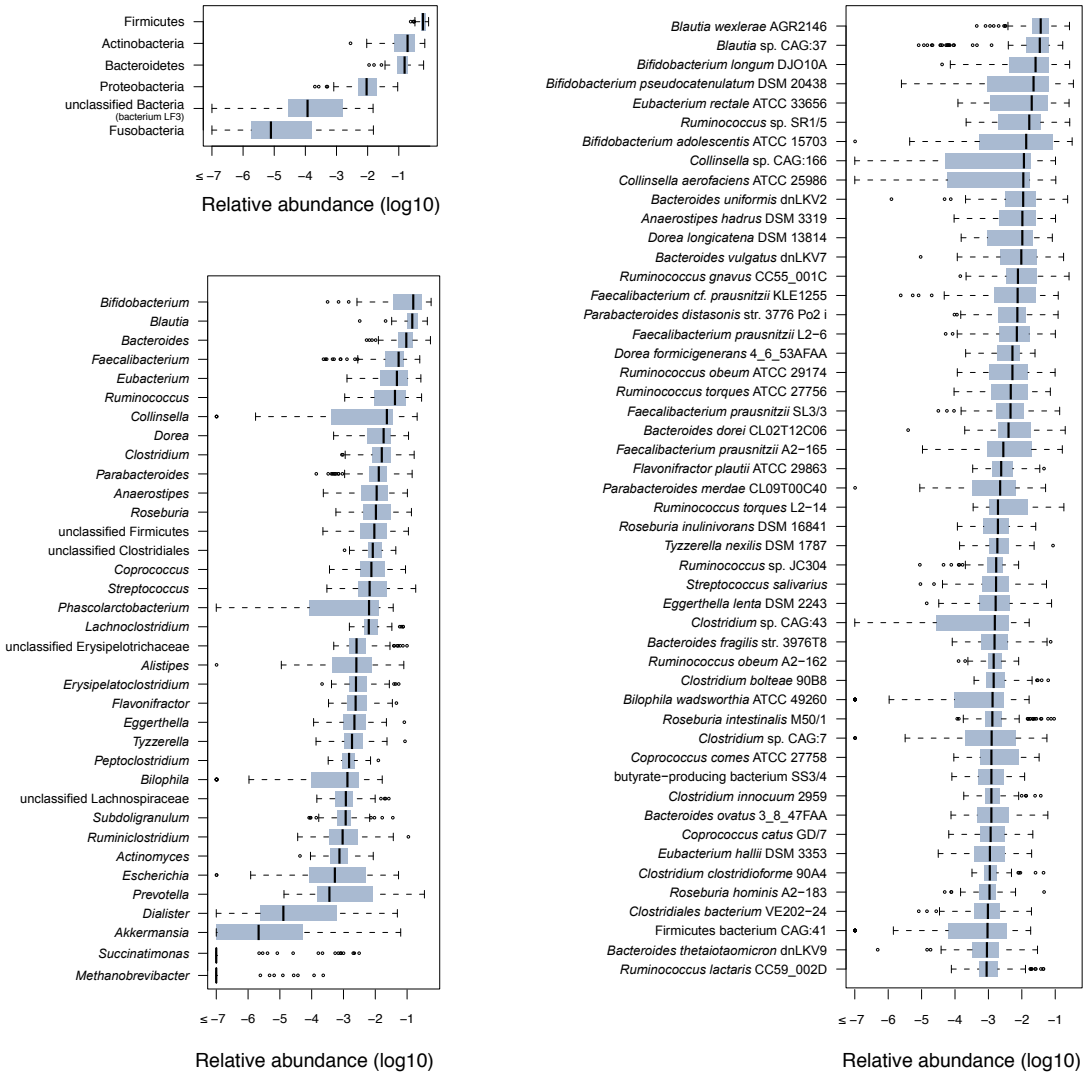


Fig. 2.4. Boxplot of the relative abundance of the taxonomy in the JPGM. Relative abundances of each genus (a) and species (b) in the 106 JP individuals were shown. Boxes represent the interquartile range (IQR) and the lines inside show the median. Whiskers denote the lowest and highest values within 1.5 times the IQR.

2.3.4. Population-level diversity in the human gut microbiome

In addition to metagenomic data of the JPGM, I collected metagenomic data of publically available samples from 11 other countries: i.e., Denmark (DK)[17, 68], Spain

(ES)[17, 106], the United States (US)[18, 51, 102], China (CN)[26, 29], Sweden (SE)[27], Russia (RU)[28], Venezuela (VE)[51, 118], Malawi (MW)[51], Austria (AT)[30], France (FR)[33] and Peru (PE)[102]. I combined the independent cohort data per country to construct country-specific metagenomic datasets, comprising a total of 861 healthy individuals in which the data for individuals with BMI ≥ 30, those designates with the following diseases based on the literature: IBD, type 2 diabetes, colorectal cancer, liver cirrhosis, and infants < 3 years old were excluded (Table. 2.2).

Table 2.2. The number of individuals from the 12 countries used in this study

Country	Total number of individuals in each country	The number of individuals used in this study*	Average age	Average BMI	References
Austria (AT)	156	41	66.6	26.0	30
China (CN)	382	187	39.1	22.0	25, 28
Denmark (DK)	290	121	55.2	24.2	5, 67
France (FR)	156	55	60.6	23.9	33
Japan (JP)	106	104	32.0	21.9	This study
Malawi (MW)	23	5	20.2	20.8	51
Peru (PE)	36	31	20.9	20.8	102
Russia (RU)	96	83	35.7	22.6	28
Spain (ES)	141	54	40.4	24.5	17, 106
Sweden (SE)	145	36	70.4	24.9	27
The United States (US)	174	126	26.5	23.6	17, 50, 100
Venezuela (VE)	29	18	20.2	20.8	50, 103
Total	1,734	861			

*: Patients with obesity (BMI ≥ 30), IBD, type 2 diabetes, colorectal cancer and liver cirrhosis and infants (< 3 years) were excluded (see text).

To investigate population-level variations in human gut microbiome structures among the 12 countries, I evaluated the microbial composition at the genus level by mapping the metagenomic reads to the reference genomes. The MDS plot of the microbial compositions showed that each country had a tendency to form distinct clusters (Fig. 2.5a). A permutation test confirmed significantly higher similarity of the microbial composition between individuals within a country than those between different countries (Fig. 2.5b). To test whether the microbial composition can predict an

individual's country of origin, I employed randomforest analysis[27, 51] to construct a predictive model for the 10 countries except VE and MW, for which sample numbers were too small to analyze. The results showed that AUCs ranged from 0.82 for US to near 1.00 for PE (Fig. 2.5c), demonstrating the high predictive accuracy of the model. Taken together, these results strongly suggested that the gut microbiome structure is significantly diverse across the 12 countries.

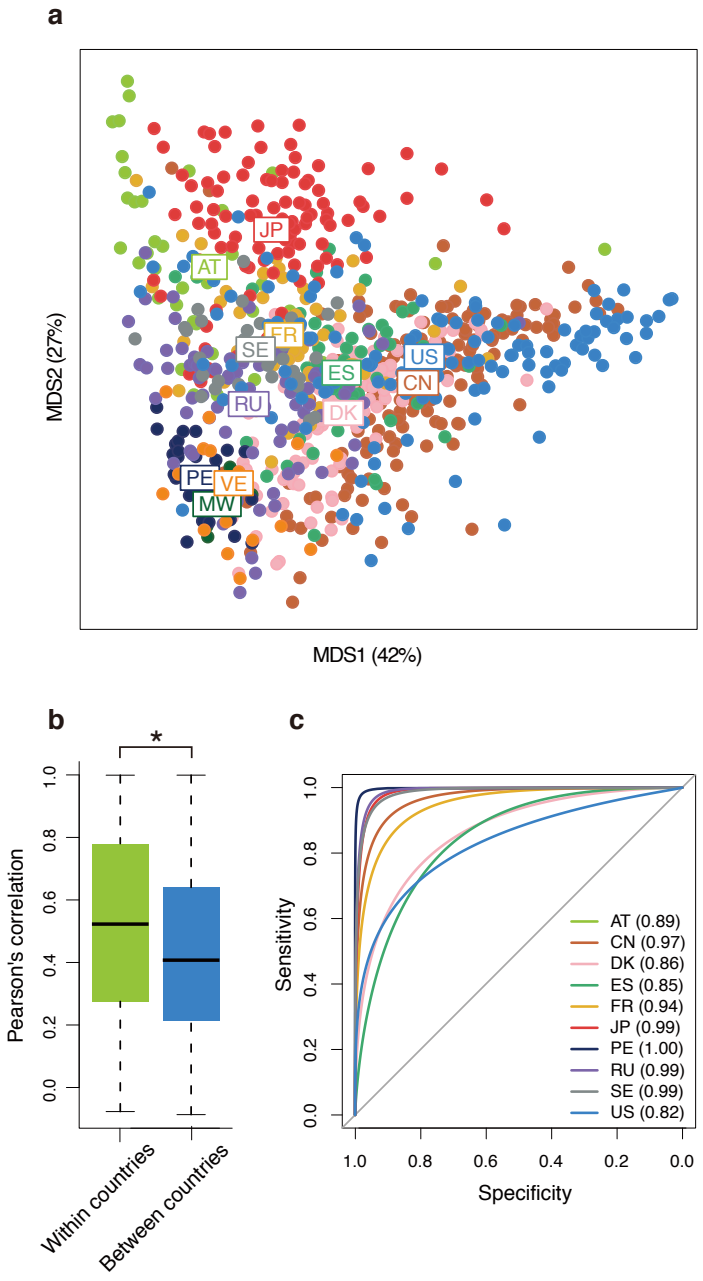


Fig. 2.5. Population-level diversity in human gut microbiomes from the 12 countries. a, MDS plot of microbial compositions at the genus level for the 861 individuals. Each circle represents an individual microbial composition and colors represent the country of origin. The position based on the average microbial composition for each country is displayed by abbreviations of the country name. b, Comparison of Pearson's correlation coefficients of microbial compositions in individuals within a country and those between different countries. Boxes represent the interquartile range (IQR) and the lines inside show the median. Whiskers denote the lowest and highest values within 1.5 times the IQR. Asterisk represents *P*-value < 0.05. c, ROC curves and AUC values from the randomforest model. The number in parenthesis represents the AUC values of the 10 countries.

To examine the effects of different protocols used in the present and other studies on the observed differences in the microbial composition at the genus level, I compared and assessed variations in the microbial composition estimated from three different NGS sequencers, four different DNA extraction methods including two enzymatic lysis methods and two commercially available kits based on mechanical disruption of cells, and four different fecal sample storage conditions (Table 2.3). For the fecal sample storage conditions, I focused on assessment of differences in the storage time from defecation until freezing of fecal samples because it was considered to be varied among the studies (Table 2.4). The results revealed that PCCs between the microbial compositions from different protocols were high (from 0.88 to 1.00) in any pair of comparisons between them (Fig. 2.6), and significantly higher than those observed for the individuals within and between countries (Fig. 2.7). Although several samples showed relatively low similarities in microbial composition between different DNA extraction methods, the lowered similarities observed were not caused by a particular protocol, rather due to differences in the individual samples used. These data suggested that the methodological differences had no significant effects on the observed variations among the human gut microbiomes.

Table 2.3. Methodologies and protocols used to assess their effects on microbial composition

Sequencer	Number of samples used
Roche 454	20
Ion PGM	20
Illumina MiSeq	20

DNA extraction method	Number of samples used
Enzymatic method with lysozyme and achromopeptidase	8
Enzymatic method with lysozyme only	8
FastDNA Spin Kit for soil (mechanical method)	8
PowerSoil DNA Isolation Kit (mechanical method)	8

Fecal sample storage conditions	Number of samples used
Stored for one day at room temperature under aerobic conditions (1d-air)	3
Stored for one day at room temperature under anaerobic conditions (1d-ane)	3
Stored for three days at room temperature under aerobic conditions (3d-air)	3
Stored for three days at room temperature under anaerobic conditions (3d-ane)	3

Table 2.4. Sequencers, DNA extraction methods and fecal sample storage conditions used in the present and the other studies.

Country	Sequencers	DNA extraction methods	Fecal sample storage conditions	References
Austria	Illumina	No information	Frozen at −20 ºC after defecation, then stored at −80 ºC within 48 hours	30
China	Illumina	Enzymatic lysis	Frozen at −20 ºC within one day after defecation, then stored at −80 ºC	26
China	Illumina	Mechanical lysis	Immediately transferred in an ice bag to laboratory after defecation, then stored at −80 ºC	29
Denmark	Illumina	Enzymatic lysis	Immediately frozen at −20 ºC after defecation, then stored at −80 ºC	17, 68
France	Illumina	GNOME DNA Isolation Kit (mechanical lysis)	Stored at −20 ºC within 4 hours after defecation	33
Japan	454, Illumina, Ion PGM/Proton	Enzymatic lysis	Transferred at 4 ºC to laboratory within 36 hours after defecation, then frozen in liquid nitrogen and stored at −80 ºC	This study
Malawi	454	No information	Stored at −80 ºC within 30 minites after defecation	51
Peru	Illumina	PowerSoil DNA Isolation Kit (mechanical lysis)	Stored on ice for at most 4 days after defecation, then frozen	102
Russia	SOLiD	Mechanical lysis	Immediately frozen at −20 ºC after defecation	28
Spain	Illumina	Enzymatic lysis	Immediately frozen at −20 ºC after defecation, then stored at −80 ºC	17, 106
Sweden	Illumina	Mechanical lysis	Stored at −80 ºC after defecation	
The United States	Illumina	PowerSoil DNA Isolation Kit (mechanical lysis)	Transferred to laboratory on ice within 24 hours after defecation, then stored at −80 ºC	18
The United States	454	No information	Stored at −80 ºC within 30 minites after defecation	51
The United States	Illumina	PowerMicrobiome RNA Isolation Kit (mechanical lysis)	Stored on ice after defecation, then frozen within 24 hours	102
Venezuela	454	No information	Stored at −80 ºC within 30 minites after defecation	51
Venezuela	Illumina	PowerSoil DNA Isolation Kit (mechanical lysis)	Immediately frozen in liquid nitrogen after defecation	118

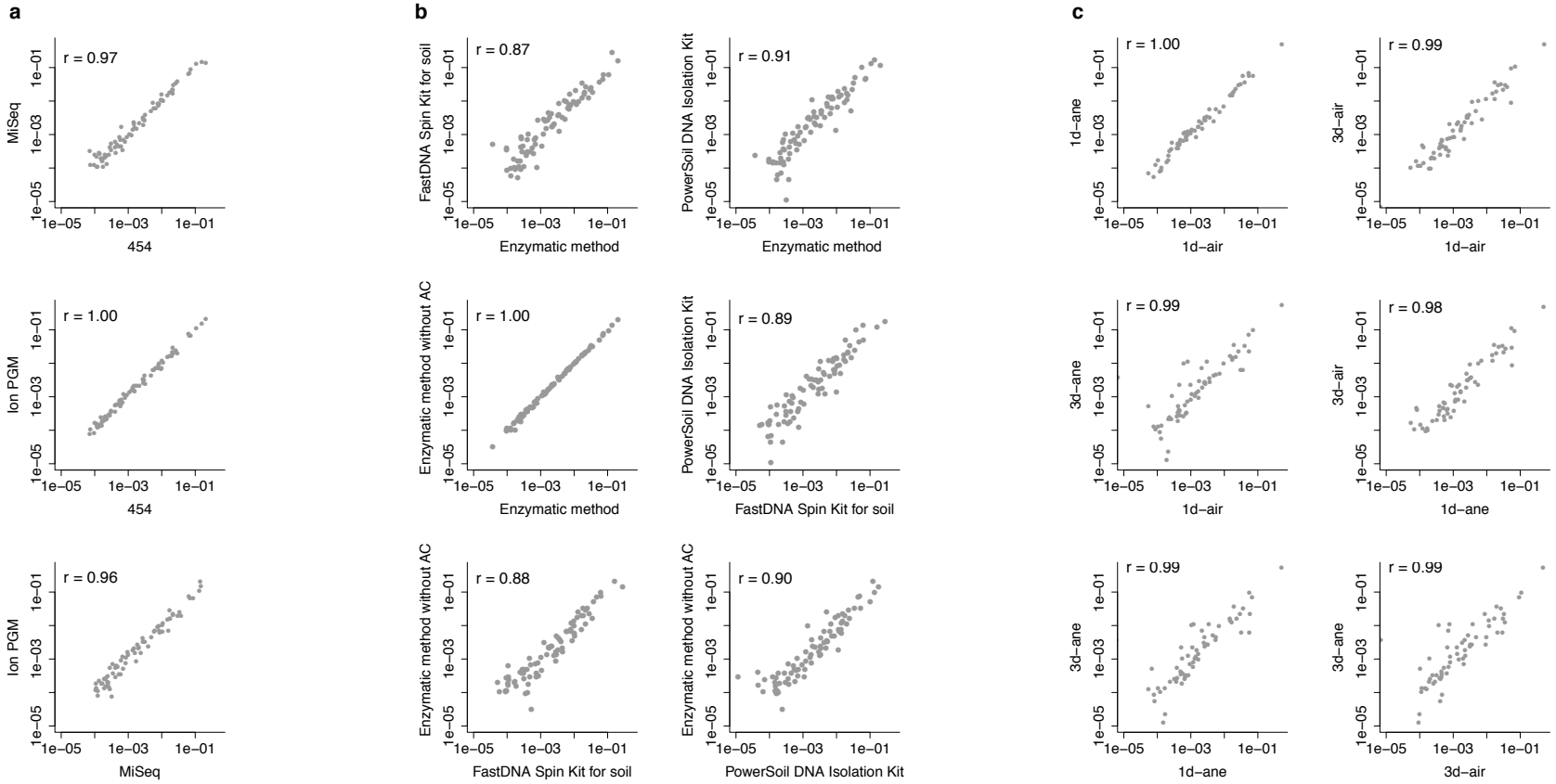


Fig. 2.6. **Comparison of PCCs of microbial compositions at the genus level among three different methodologies and those of individuals within a country and between different countries.** a, Twenty fecal DNA samples were subjected to sequencing with three different sequencers, Roche 454, Illumina MiSeq, and Ion PGM. b, Eight fecal samples were subjected to isolation of fecal DNA by five different DNA extraction methods. c, Three fecal samples were stored under four different storage conditions. Circles represent the genera with average relative abundance ≥ 0.01%. Vertical and horizontal axes indicate the average relative abundance of each genus. Abbreviations for fecal sample storage conditions are summarized in Table 3.3.

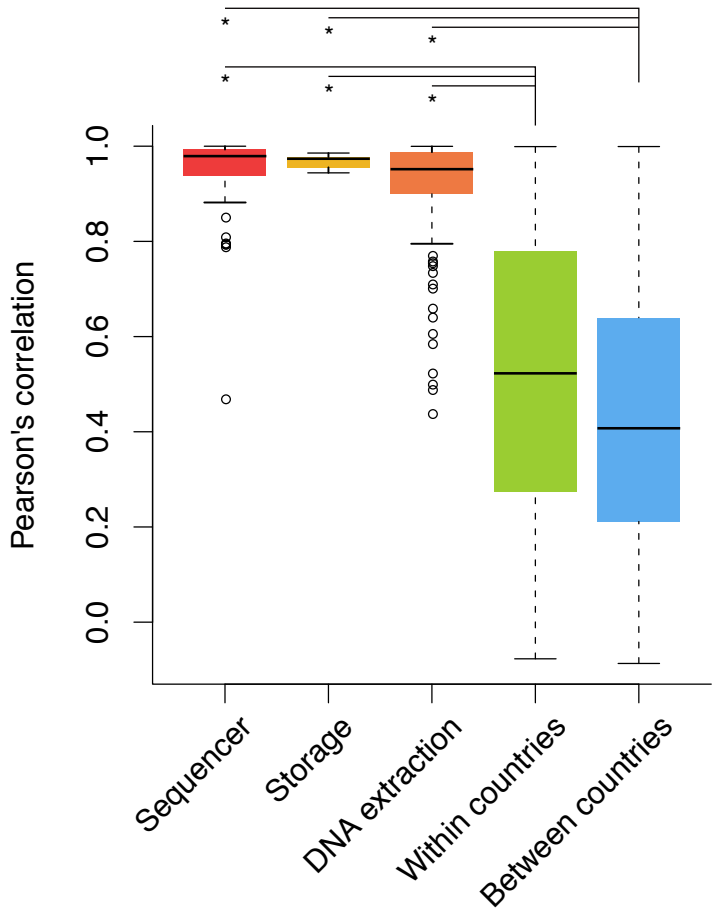


Fig. 2.7. Comparison of PCC obtained from different methodologies and individuals. PCCs of the microbial composition at the genus level in an individual obtained by different methodologies are shown in the left three boxes (red, yellow, and orange), and compared with those of individuals of within-country and between-countries shown in the right two boxes (green and blue). *P*-values were calculated by permutation tests with 10,000 random samplings. Asterisks represent *P*-values < 0.05.

Hierarchical clustering of the 12 countries based on the average microbial composition at the genus level showed that the JPGM was more similar to microbiomes of AT and SE than that of CN, while CN and US were most closely related, but far from the JP among the 12 countries (Fig. 2.8). These results strongly suggested that host ethnic and geographical closeness have no large influence on shaping of the overall

microbial composition of the human gut microbiome. We also assessed the contribution of variations in age and BMI to differences in the microbial abundance by using PERMANOVA. The coefficient of determination for the variation (R^2) in age and BMI was 0.16 (P-value = 0.07) and 0.2 (P-value = 0.14), suggesting that both factors had no significant influence on the observed results as well.

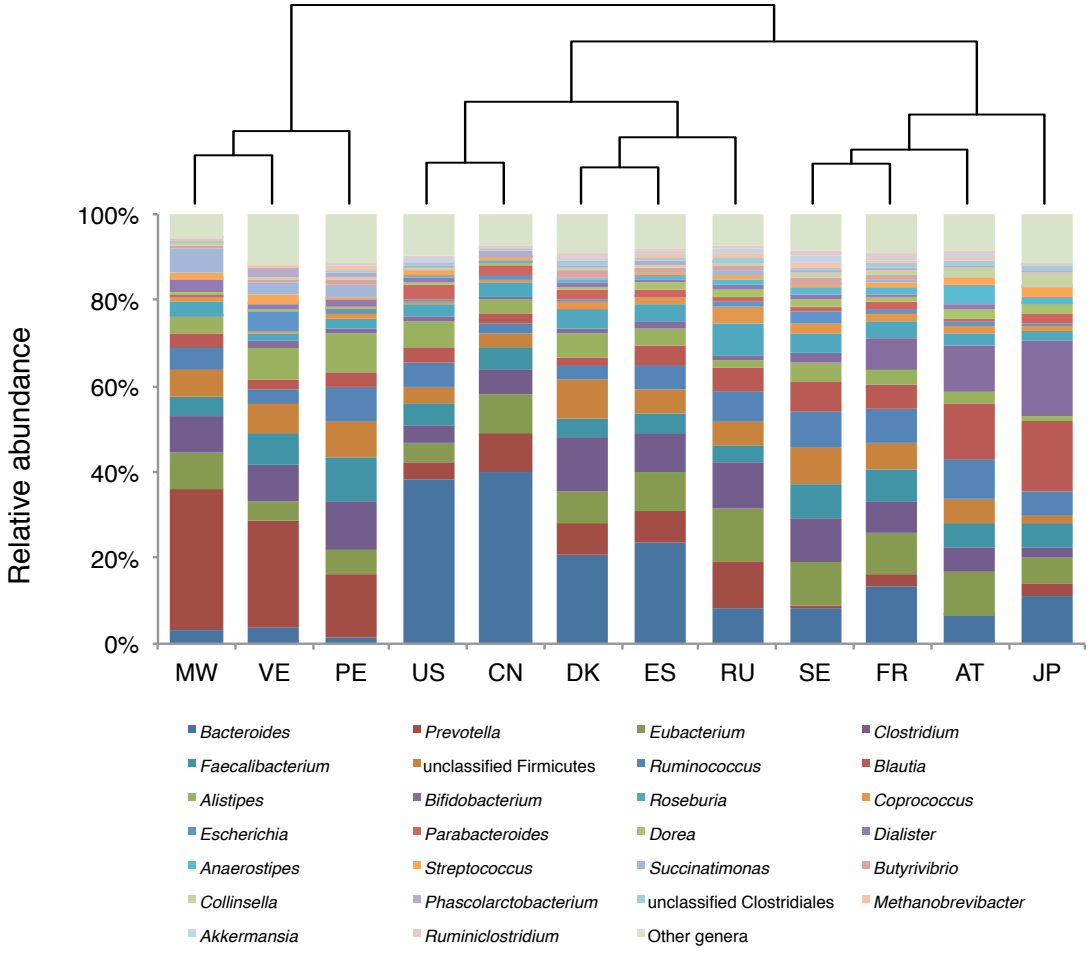


Fig. 2.8. **Hierarchical clustering of the 12 countries based on average microbial composition at the genus level.** Cluster dendrogram was generated with the Ward method using the Bray-Curtis distances. The top 26 genera with average relative abundance ≥ 0.5% are shown.

2.3.5. Characterization of the Japanese gut microbiome

When comparing the abundance of the bacterial phyla, the JPGM showed the highest abundance of Actinobacteria. In contrast, the abundance of Bacteroidetes and Proteobacteria in the JPGM was significantly lower than in the microbiomes of various other countries (Fig. 2.9a). Regarding the bacterial genera, the JPGM was characterized by the highest abundance of *Bidfidobacterium*, *Blautia*, *Collinsella*, *Streptococcus* and unclassified Clostridiales, but the lowest abundance of *Clostridium*, *Alistipes*, unclassified Firmicutes, *Dialister* and *Butyrivibrio* among the 12 countries (Fig. 2.9b).

Another characteristic feature of the JPGM was that it has the lowest frequency of *Methanobrevibacter smithii*, a methanogenic archaeon, among the 12 countries (Fig. 2.9c). Metagenomic mapping analysis showed that this species was detected only in eight (7.7%) JP individuals, while it was detected in a proportion of 39 – 100% of the individuals in other countries (relative abundance ≥ 0.0001%, Fig. 2.10a). The lowest prevalence of this archaeon in the JP cohort was also validated by PCR using species-specific 16S rRNA gene primers. The data showed that *M. smithii* was undetected in 97 (92%) of the 106 JP individuals both in the metagenomic mapping and PCR analysis, where five were positive in both analyses, three were positive only in the mapping analysis and one was positive only in the PCR analysis (Fig. 2.10b).

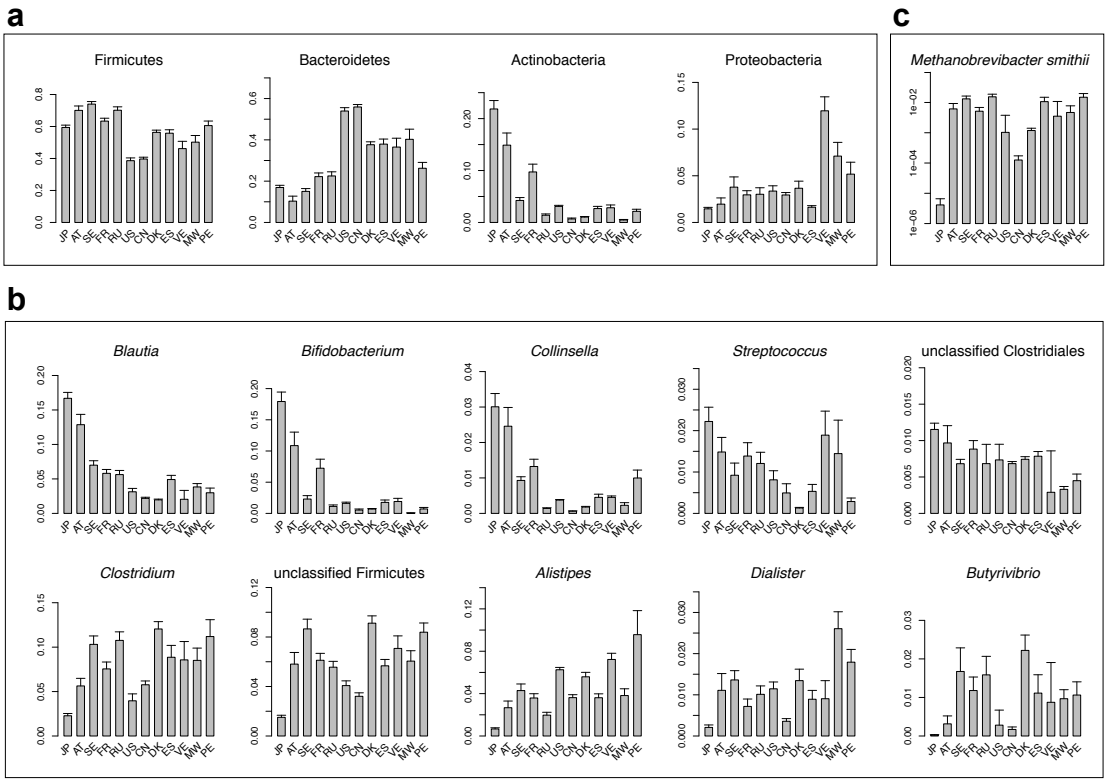


Fig. 2.9. **Taxonomic comparison of gut microbiomes of populations from the 12 countries.** Relative abundances of the four dominant phyla (a), the five genera with the highest and lowest abundance in the JPGM (b), and *M. smithii* (c) in the 12 countries are shown. Vertical axes represent the relative abundance of the species calculated from the number of mapped reads to the reference genomes.

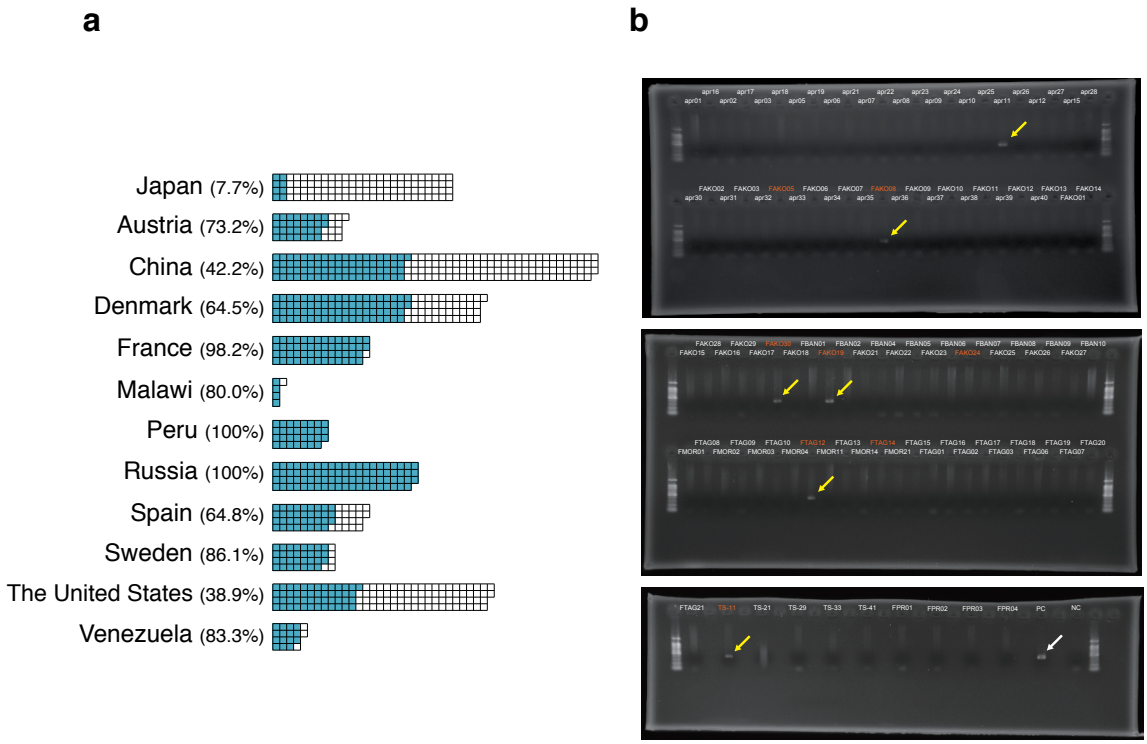


Fig. 2.10. **Detection of *M. smithii* in the human gut microbiome.** (a) Open and blue boxes indicate the individuals for which *M. smithii* was detected (relative abundance ≥ 0.0001%) and undetected, respectively, by mapping of metagenomic reads to the reference genomes. (b) PCR detection of *M. smithii* in the 106 JP individuals. Individual's IDs are represented at each lane, and the ones indicated in orange were *M. smithii*-positive in the mapping analysis. Yellow arrows indicate the bands for the PCR product of *M. smithii,* of which the positive control (PC) is shown by a white arrow. NC, negative control.

2.3.6. Functional comparison of the Japanese gut microbiome with the others

I compared the JP gene set with the IGC gene set[106]. The clustering of the JP (4.9 M) and the IGC genes (9.9 M) generated 11,929,034 non-redundant genes in total, of which about 2.0 M genes were shared by both gene sets, and 2.3 M and 7.7 M genes were unique to JP and IGC, respectively (Fig. 2.11a). This limited overlap between the JP and IGC gene sets was supported by the mapping analysis of metagenomic reads, in

which 45.6% of the JP metagenomic reads were mapped to the IGC gene set, while 80.0% were mapped to the JP gene set (Fig. 2.11b). In this clustering, 585,856 genes and 202,410 genes decreased from the original JP and IGC gene sets, respectively. This can be explained by the fact that these genes were fragmented and merged to longer authentic genes in either of the gene sets. As a result, the JP and IGC gene sets are composed of 4,268,863 and 9,676,237 non-redundant genes, respectively.

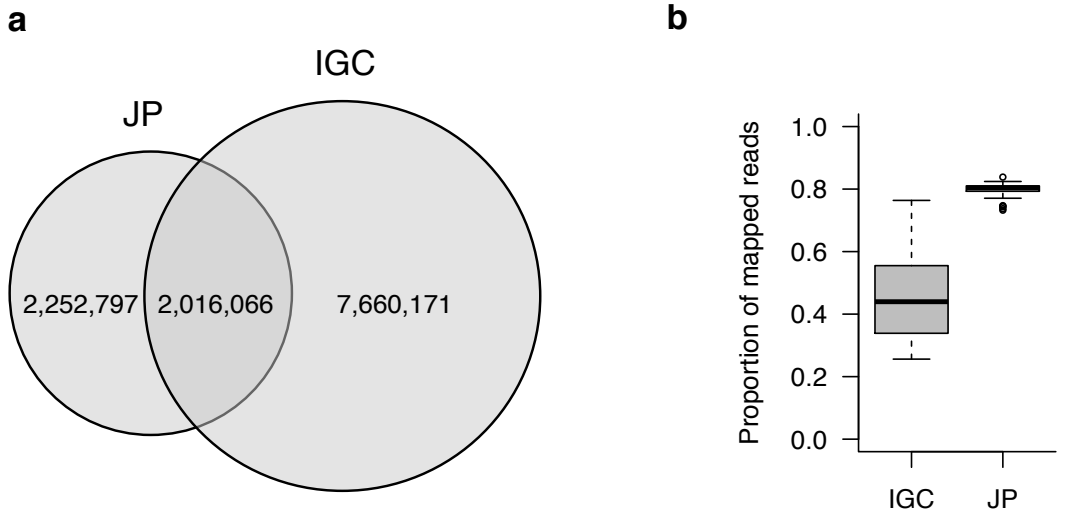


Fig. 2.11. **Comparison of JPGM and IGC non-redundant gene sets.** (a) Venn diagram of the number of genes in both gene sets. (b) Ratio of the mapped JP metagenomic reads to the JP and IGC gene sets.

Next, I annotated the gene sets with functions based the KEGG database. The analysis identified 5,789 KOs in the JPGM and a total of 6,205 KOs from both gene sets, in which 5,613 KOs (90%) were shared between both gene sets, demonstrating a significantly high similarity in functional profiles across the populations despite the small overlap in the gene sequences, which is concordant with the previous finding of a

high interindividual similarity of the functional profiles[24]. It was noted that the IGC-unique 416 KOs included multiple genes related to archaeal methane metabolism, while the JP-unique 176 KOs included more genes for spore formation than the IGC-unique KOs (Appendix 4 and 5).

To explore functions that are enriched or depleted in the JPGM as compared with microbiomes from the 11 other countries, I mapped the metagenomic reads of all individuals to the JP and IGC merged gene set. By comparing the numbers of mapped reads, we identified 563 and 521 KOs having the highest and lowest abundances in the JPGM among the 12 countries with statistical significance (Fig. 2.12, and Appendix 6 and7). The overrepresented KOs included functions for carbohydrate metabolism such as glucan 1,3-β-glucosidase (K01210), 6-phospho-β-galactosidase (K01220) and gluconokinase (K00851), and for membrane transport such as the phosphotransferase system of simple sugars including mannose, lactose, and *N*-acetylgalactosamine (K02796, K02787 and K02746). Thus, metabolic pathways for simple sugars such as mono- and oligosaccharides were significantly enriched in the JPGM as compared with the others. On the other hand, the depleted KOs included functions such as cell motility including chemotaxis protein CheX (K03409) and flagellar protein FliO/FilZ (K02418), replication and repair including DNA mismatch repair protein MutL (K03572) and DNA adenine methylase (K06223), suggesting a depletion of functions related to host immunity and DNA damage in the JPGM.

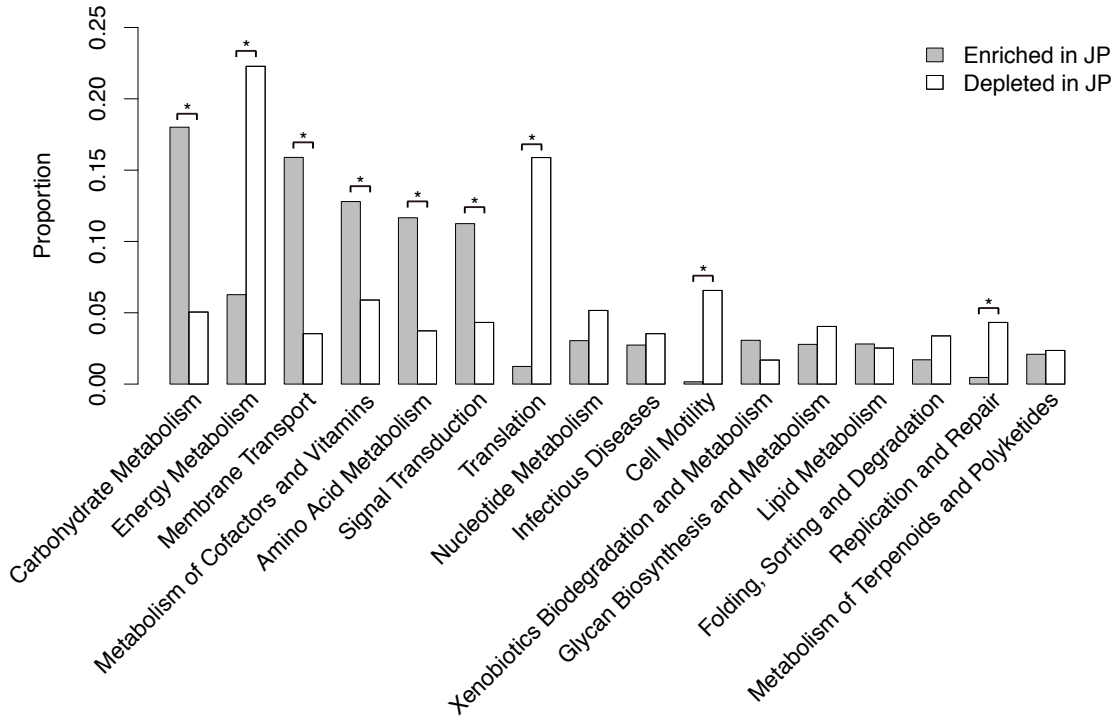


Fig. 2.12. **Enriched and depleted functions in the JP gut microbiome.** Functional categories of the KOs most enriched and depleted in the JPGM as compared with those of the other 11 countries are shown. The vertical axis represents the proportion of KOs assigned to the category. Asterisks indicate adjusted FDRs < 0.01 (Fisher's exact test).

In agreement with the lowest prevalence of *M. smithii* in the JP cohort, many of the KOs involved in methanogenesis were depleted in the JPGM. Of the 25 known KOs involved in methanogenesis, 18 were significantly depleted in the JPGM, with the lowest abundance among the 12 countries (Fig. 2.13 and 2.14). Conversely, I found a significant enrichment for multiple KOs involved in acetogenesis (the Wood–Ljungdahl pathway) in the JPGM. Of the 17 known KOs involved in acetogenesis, 13 were significantly enriched in the in the JPGM as compared with the other 11 countries, and five of them had the highest abundance among the 12 countries (Fig. 2.13 and 2.14). These two pathways utilize hydrogen to generate methane and acetate[127]. Furthermore the abundance of five known KOs involved in dissimilatory sulfate reduction (DSR),

which is the third pathway for hydrogen metabolism, was similar between the JP and other gut microbiomes (Fig. 2.13 and 2.14). These results indicated that the JPGM had a clear inverse pattern in the abundance of the KOs between both metabolic pathways as compared to all other microbiomes, suggesting a prominent difference in the pathways for hydrogen utilization in the gut between Japanese and other populations.

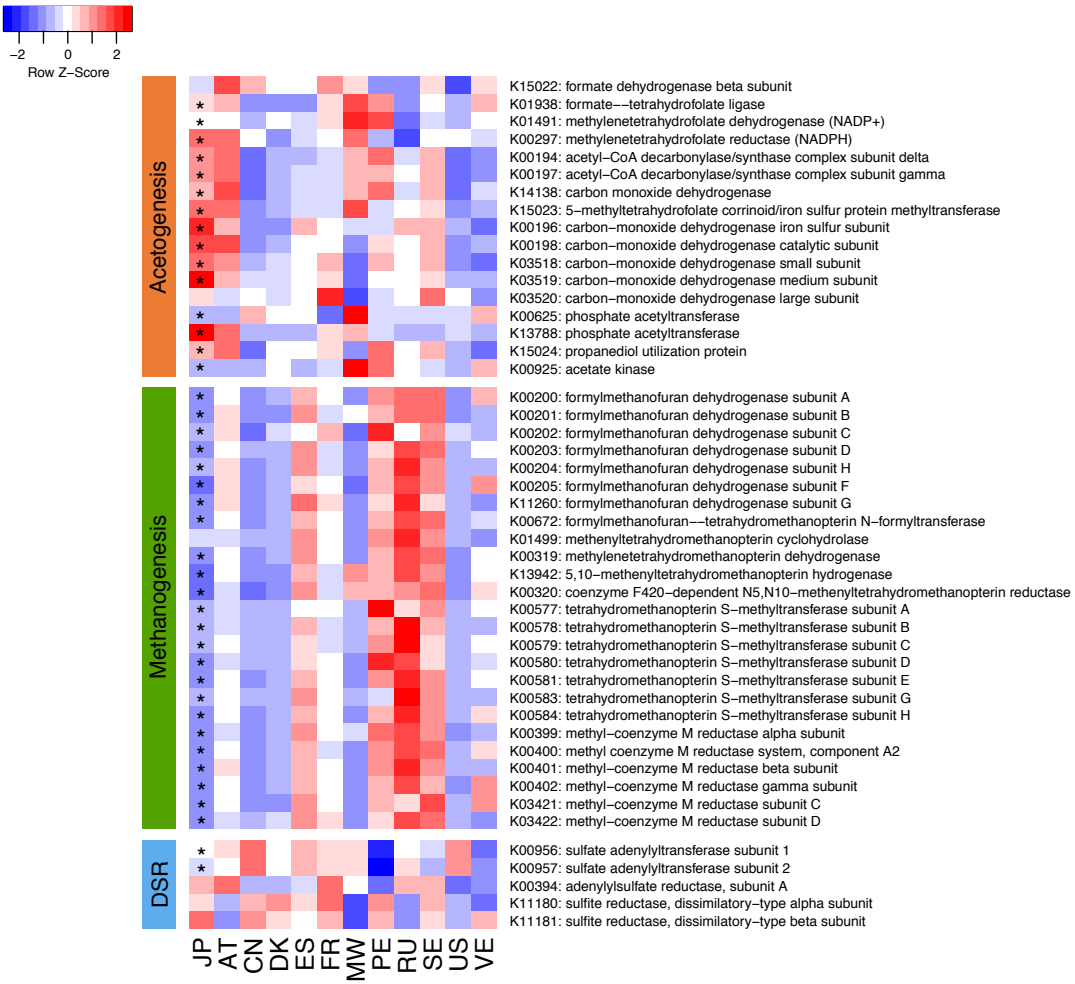


Fig. 2.13. **Enriched and depleted genes in the acetogenesis, methanogenesis, and dissimilatory sulfate reduction in the JPGM.** Relative abundances of KOs involved in the pathways for hydrogen metabolism in acetogenesis, methanogenesis and dissimilatory sulfate reduction among the 12 countries are shown. Red and blue boxes denote statistically high and low abundances as compared with the average abundance of the other 11 countries, respectively. Asterisks indicate adjusted FDRs < 0.01 (Student's *t*-test between the 104 JP individuals and the 757 individuals in the other countries).

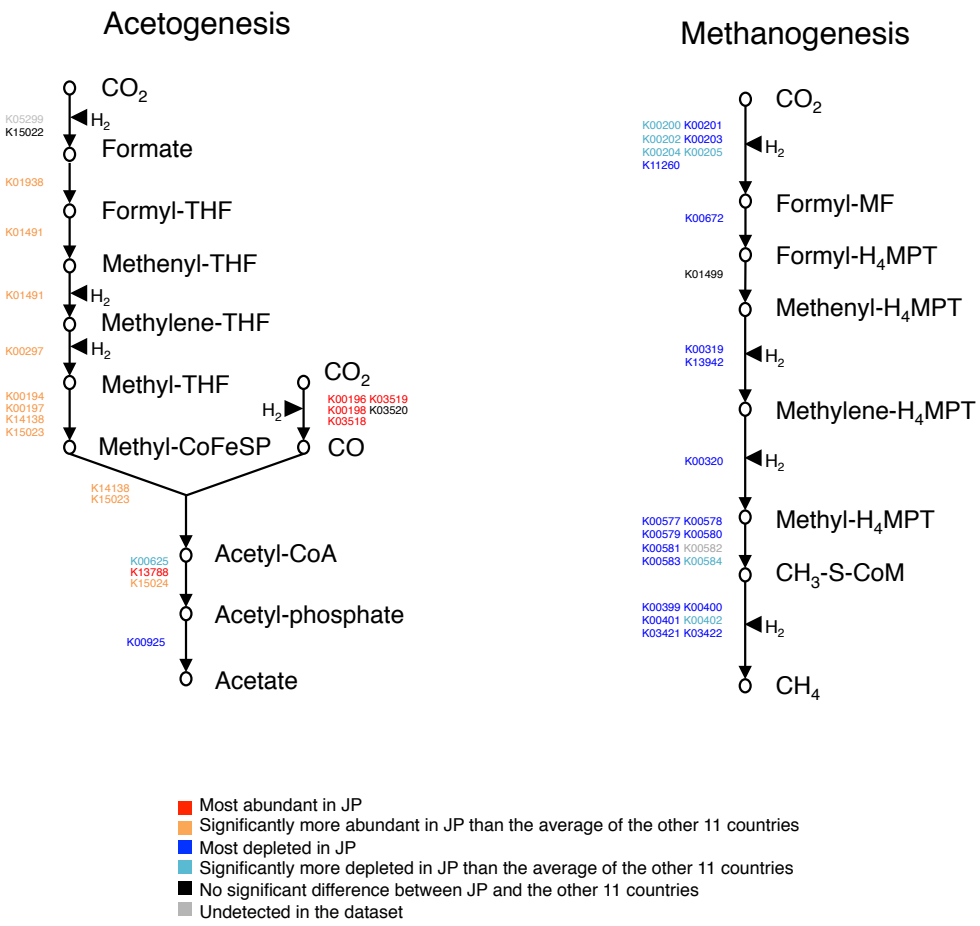


Fig. 2.14. **Metabolic pathways of acetogenesis and methanogenesis.** Metabolic pathways for acetogenesis and methanogenesis and KOs involved are shown. Colors indicate differences in the abundance of the KOs shown in the figure.

2.3.7. Gene families enriched in the Japanese population

I comprehensively surveyed gene families that are frequently present in the JP cohort by using the eggNOG database, which includes more compiled gene families than the KEGG database. The annotation of the merged JP and IGC non-redundant genes yielded 51,250 NOGs. In this analysis, I used 10 M metagenomic reads per individual

to detect a low content of NOGs, so that 60 individuals having < 10 M reads, including all individuals from MW and VE, were excluded from this analysis.

I mapped 10 M reads from the 801 individuals to the merged non-redundant genes to detect the NOGs present in the individual. For these NOGs, I compared the proportion of individuals possessing them in the JP cohort and with the average proportion of the individuals proportion in other the nine other countries, (Fig. 2.15). The results revealed 52 NOGs comprising a total of 1,114 genes that were detected in significantly higher proportions in the JP cohort than in the nine other countries using a threshold of a proportion of > 70% in JP, an average proportion of < 30% in other countries, and the a ratio of JP/others of ≥ 3. Of the 1,114 genes, 63% were taxonomically assigned to the known phyla. Among them, 30 (58%), eight (15%), and five (10%) NOGs were assigned to only Actinobacteria, Bacteroidetes, and Firmicutes, respectively, and nine other NOGs were distributed over more than two phyla (Fig. 2.16a). The high fraction of Actinobacteria may reflect the highest abundance of this phylum in the JP cohort among the 12 countries. Among the eight NOGs assigned to the Bacteroidetes, three NOGs (ENOG4108ZIS, ENOG4108MQB and ENOG4105WVE), that were detected in approximately 90% of the JP individuals cohort and in ~15% of other populations with the highest ratio of JP/others, were represented by the genes for aquatic plant-derived polysaccharide-degrading enzymes such as β-porphyranase (hydrolase family 16) and β-agarase published previously[49]. The functional distribution of the 52 NOGs revealed that 35% were of unknown function and no particular function was enriched (Fig. 2.16b).

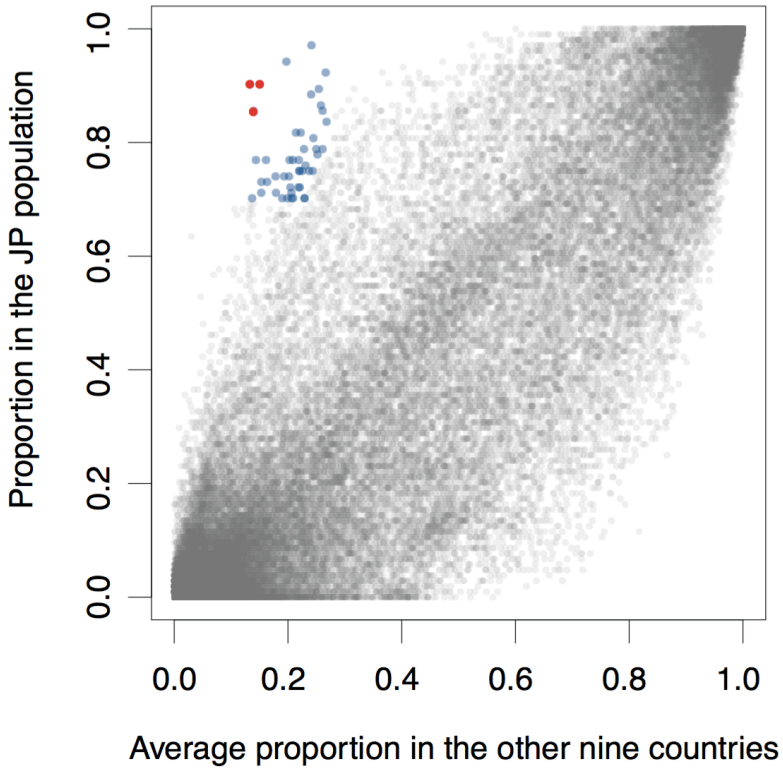


Fig. 2.15. Comparison of the prevalence of NOG gene families between the JP and the nine populations. The frequency of NOGs in the JP individuals plotted against those in the other nine countries. Each circle represents a NOG. The vertical axis represents the frequency of NOGs detected in the JP individuals. The horizontal axis represents the average frequency of NOGs detected in the individuals of the nine countries. Fifty-two NOGs significantly highly prevalent in the JP cohort as compared with the others (JP > 0.7 and the others < 0.3) are colored with blue. Three NOGs (ENOG4108MQB, ENOG4108ZIS, and ENOG4105WVE) were depicted in red.

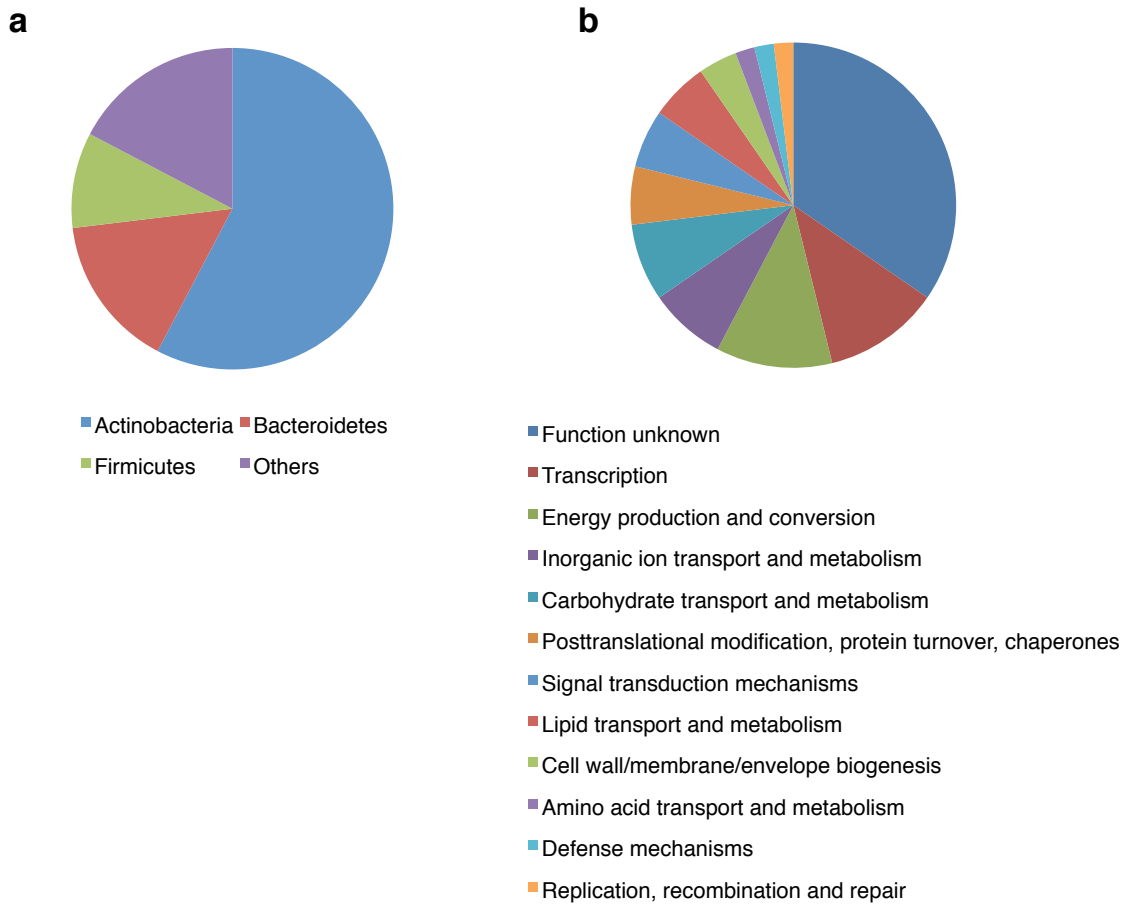


Fig. 2.16. Taxonomic and functional assignment of the 52 NOGs having a higher abundance in the JP cohort than in the other populations.
(a) Distribution of the 52 NOGs per phylum is shown. “Others” indicates more than two phyla.
(b) Distribution of the 52 NOGs per functions is shown.

2.1. Discussion

In this study, I conducted a metagenomic analysis of the JPGM from 106 individuals. By assembling the 350 Gb metagenomic sequences, predicting genes and clustering, I constructed a gene set of the JPGM comprising approximately 4.9 M non-redundant genes. The number of genes in this gene set is comparable to those (from 3.3 M to 6.0 M genes) reported in previous studies analyzing the gut microbiomes of individuals in other countries including DK, ES, CN, US and SE. The rarefaction analysis revealed that the JPGM gene set covered most of the genes shared by at least 1.9% in the JP individuals, indicating that the present gene set can be used as a reference for the genes in the JPGM.

The comparative analysis of the metagenomic datasets between the 104 JP individuals and the 757 individuals from other 11 other countries revealed a significant population-level diversity in the human gut microbiome across the 12 countries. The accuracy and reliability shown in this study is supported by the use of a larger dataset including more populations than that used in the previous studies. Additionally, the statistical assessment indicated no large effects of differences in experimental protocols such as DNA extraction method, fecal storage conditions and sequencer, BMI and age, on the observed results, which also supported the present findings. Thus, I provided the evidence for large variations in the structure and function of the human gut microbiomes of healthy adults at the population-level.

The present study also revealed various features specific to the JPGM. The JPGM showed the highest abundance of the phylum Actinobacteria among the microbiomes of the 12 countries, mainly because of the highest abundance of the genus *Bifidobacterium*. The high abundance of *Bifidobacterium* has also been observed in the gut microbiome

of Japanese children based on the 16S rRNA gene analysis[100], indicating that it is highly prevalent throughout the Japanese population. *Bifidobacterium* is thought to be a beneficial microbe having more glycoside hydrolases for degrading starch than other intestinal microbes[128]. Therefore, the high abundance of *Bifidobacterium* can be considered to be the consequence of the intake of various saccharides in traditional and unique Japanese foods. However, at present, it is unknown exactly which foods or nutrients unique to the Japanese diet contribute to the high abundance of *Bifidobacterium*.

Additionally, the JPGM is characterized by various unique functional features. For example, the high abundance of carbohydrate metabolism was observed in the JPGM, which leads to the production of high levels of short chain fatty acids and hydrogen as end products, both of which seems to be clinically beneficial[129, 130]. Concurrently, I found a depletion of deleterious functions such as cell motility, and replication and repair, suggesting a low abundance of the flagellated microbes leading to reduced proinflammatory responses by host cells and less DNA damage to be repaired in the gut of the Japanese individuals. Together, I suppose that such a gut ecosystem containing these beneficial functions might be globally associated with the highest average life span of Japanese in the world.

A remarkable depletion of the archaeon *M. smithii* is also characteristic of the JPGM, resulting in an overall depletion of genes for methanogenesis. In contrast, genes for acetogenesis, which are exclusively encoded by anaerobic acetogens such as the major species *Blautia*[131], were enriched in the JPGM as compared with other gut microbiomes. Both methanogenesis and acetogenesis are considered to be critical pathways for hydrogen consumption in the gut because these pathways are tightly linked with anaerobic fermentation of carbohydrates producing hydrogen[125]. My

findings suggest that acetogenesis is the preferable pathway for hydrogen metabolism in the JPGM, while methanogenesis is more actively utilized for hydrogen metabolism in many of the other gut microbiomes. Additionally, since the abundance of intestinal *M. smithii* is positively associated with the level of breath methane[132], the present data strongly suggest that *M. smithii* is the primary factor for ethnic differences in the level of methane in human breath reported previously[133, 134].

Many of microbial and functional uniqueness I found in the JPGM may be more or less influenced by various internal and external factors, contributing to the population-level diversity in the human gut microbiome. Therefore, to more deeply understand the diversity in human gut microbiomes, elucidation of such factors is further required.

3. Antibiotics shapes population level diversity in the human gut microbiome

3.1. Introduction

Human gut microbiome structure is affected by various factors such as diet[108, 135], antibiotics[136, 137], host's physiology[26, 29] and genetics[104]. For example, long-term dietary habit rich in protein and animal fat was correlated with a high abundance of *Bacteroides* in human gut, while that rich in carbohydrate was correlated with an enrichment of *Prevotella*[108]. Abundances of several taxonomies, particularly *Christensenellaceae*, were more largely associated between monozygotic twin pairs than between dizygotic twin pairs, suggesting the influence of host's genetic background[104]. In addition, antibiotic intake had a strong impact on the gut microbiome structure, resulting in an incomplete recovery of the structure to the initial state[138]. However, little is known about the factors and extent of their contribution to the population-level variability in the microbial abundance observed in chapter 1. To explore such factor, I conducted a large-scale association study of the epidemiological data of several external factors including dietary intake and antibiotic usage with metagenomic data of the 861 individuals from the 12 countries.

3.2. Methods

3.2.1. Collection of dietary intake data

Dietary intake information of 119 food items for the 12 countries was downloaded from the Food and Agriculture Organization Corporate Statistical database (FAOSTAT) (http://faostat3.fao.org/home/, as of June 2015). The averaged dietary intakes (g/capita/day) from 2002 to 2011 in the 12 countries were used for correlation analysis. According to the Standard Tables of Food Composition in Japan, 2010[139], three major nutrient compositions (carbohydrates, lipids and proteins) were calculated from the averaged dietary intakes of 119 food items, and the 119 food items were grouped into nine food categories based on their nutrient similarities. Nutrient quantities were transformed to z-scores before clustering, and dendrograms were generated using the Ward method and Spearman's correlation as dissimilarity. The amount of the nine food categories was normalized by the total dietary intake of each country.

3.2.2. Collection of antibiotic usage data

The data for antibiotic usage in humans, the defined daily dose (DDD) per 1,000 inhabitants, were collected from Hogberg LD *et al*[140], and the European Surveillance of Antimicrobial Consumption (ESAC) yearbook in 2009[141]. The data for China were obtained from Wang X *et al*[142]. The data of antibiotic usage in farm animals in kilograms were obtained from Van Boeckel TP *et al*[143]. The antibiotic usage in farm animals normalized by counting population number of the country was used for the correlation analysis. The details about the antibiotics are summarized in Tables 3.5 and 3.6.

3.2.3. Statistical analysis

Correlations between microbes and epidemiological factors (dietary and antibiotic usage data) were evaluated using Pearson's correlation coefficient (PCC) and Spearman's correlation coefficient (SCC). P-values were adjusted for multiple testing using p.adjust(p, "BH") in R language, which is based on the Benjamini-Hogberg approach[125]. Permutational Multivariate Analysis of Variance (PERMANOVA) was used to assess the association of these factors with variation of the overall gut microbiome structure using the adonis function in the Vegan package in R with Euclidian distances as dissimilarity.

3.2.4. Analysis of gut microbiomes of Asian children

The 16S rRNA gene sequence data of gut microbiomes of children from five Asian countries (Japan, China, Taiwan, Indonesia and Thailand)[100] were publically available and were obtained from GenBank/DDBJ/EMBL. Low quality bases (< 20 QV) at the 5′ ends were removed and reads with an average quality ≤ 25 were discarded. For each sample, up to 3,000 reads were used for similarity searches against the reference genomes using BLASTN with a 94% identity and a 90% length coverage cut-off. Antibiotic usage data of the five countries were obtained from the literature by Hogberg LD *et al*[144]. Correlations of between the relative abundance of *Bacteroides* and antibiotic usage were evaluated with PCC and SCC.

3.2.5. Antibiotic resistance genes analysis

To identify antibiotic resistance genes (ARs) in the gut microbiome, the Resfams database[144] was employed. The merged reference genes were searched against the Resfam database using the hmmscan function of HMMER3[145] with gathering thresholds.

The genes assigned to ‘transcriptional factor’ were excluded from further analysis. Up to one M metagenomic reads of each individual were mapped to the ARs using Bowtie2 with a 95% identity cut-off. The number of mapped reads to the ARs was normalized by the number of the total reads used to the mapping. One third of the genes assigned to ‘ABC transporter’ were used for the analysis since two thirds of them in this class were estimated to be false positives in the original paper[144].

ARs in the reference genomes were also examined by the same methods. For this analysis, we used 126 genomes of the species with > 0.05% abundance in average among the 12 countries. The genomes generated from assembly of only metagenomic reads[146] were not used in this analysis because they possessed significantly fewer ARs than the other genomes generated by sequencing of the cultured strains. For example, for *Bacteroides*, 8.2 ARs per genome were annotated for cultured strains, while 3.8 ARs per genome were found in the genomes generated from assembly only of the metagenomic reads. These data suggested the difficulty of sufficient reconstruction of genomes only from short metagenomic reads, particularly for the genomes containing ARs showing high sequence similarity due to horizontal gene transfer. Therefore, metagenomic reads containing ARs may not be efficiently incorporated into contigs in the assembly step to avoid misassembly, and contigs containing ARs may not be accurately assigned to particular species as well.

3.2.6. Analysis of the microbial composition in mice treated with beta-lactam antibiotics

The 16S rRNA gene sequence data of gut microbiomes of mice treated with and without beta-lactam antibiotics[119, 147] were obtained from GenBank/DDBJ/EMBL. For the data of reference 119, 16S rRNA gene sequence data were analyzed using the method

described previously[110]. In brief, the reads were quality-checked, and those lacking their PCR primer sequences at either sequence termini or the average quality value < 25 were discarded. The 3,000 filter-passed 16S reads for each sample were assigned to the genus using BLASTN with a 94% identity and a 90% length coverage cut off against the full-length 16S rRNA gene sequences database constructed from the Ribosomal Database Project[148]. The relative abundance of the top seven predominant genera was compared between untreated and treated mice (untreated: n=4, treated: n=5). Statistical differences were evaluated with Student's *t*-test. For the data of the reference 147, low quality bases (< 20 QV) at the 5' end were removed and the reads with an average quality ≤ 25 were discarded. For each sample, up to 3,000 reads were used to calculate microbial compositions at the genus level, which were compared between untreated and treated mice (untreated: n=36, treated: n=39) as described above.

3.3. Results

3.3.1. Dietary intake data of the 12 countries

I explored factors that are associated with population-level diversity in the gut microbiomes across the 12 countries. Since diet is considered to be a major factor influencing microbial composition[108, 135], I accessed the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT) to collect the dietary intake data (g/capita/day) for 119 food items in the 12 countries. I calculated the average intakes of three main nutrients (carbohydrates, proteins and lipids) for the 10-year period from 2002 to 2011, and grouped the food items in nine food categories by clustering them by the compositional similarity of nutrients (Fig. 3.1). Cluster analysis based on these dietary data roughly segregated most of the Western countries from other countries in Asia, South America, and Africa (Fig. 3.2). This indicated that the FAOSTAT data properly represents the current diversity in dietary habits of the 12 countries, allowing for its use in the correlation analysis with the top 26 genera with an average abundance of > 0.5%, accounting for 91% of the total abundance (Table 3.1).

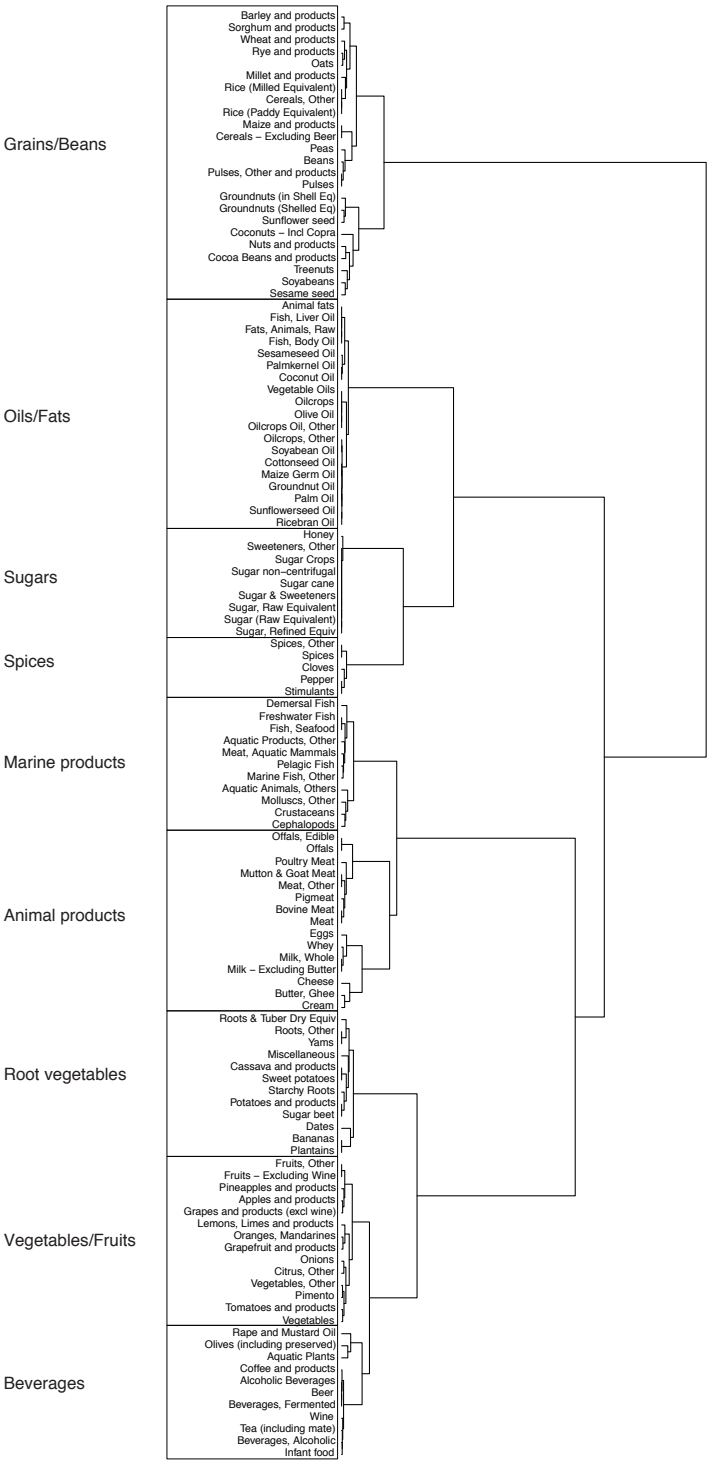


Fig. 3.1. Grouping of 119 food items into nine food categories. The 119 food items in the FAOSTAT database were clustered based on the compositional similarity of nutrients of which levels were calculated according to a book of the Standard Tables of Food Composition in Japan, 2010 (Methods). Quantity of the nutrients was transformed to z-scores before clustering and the dendrogram was generated using the Ward method and Spearman's correlation as dissimilarity.

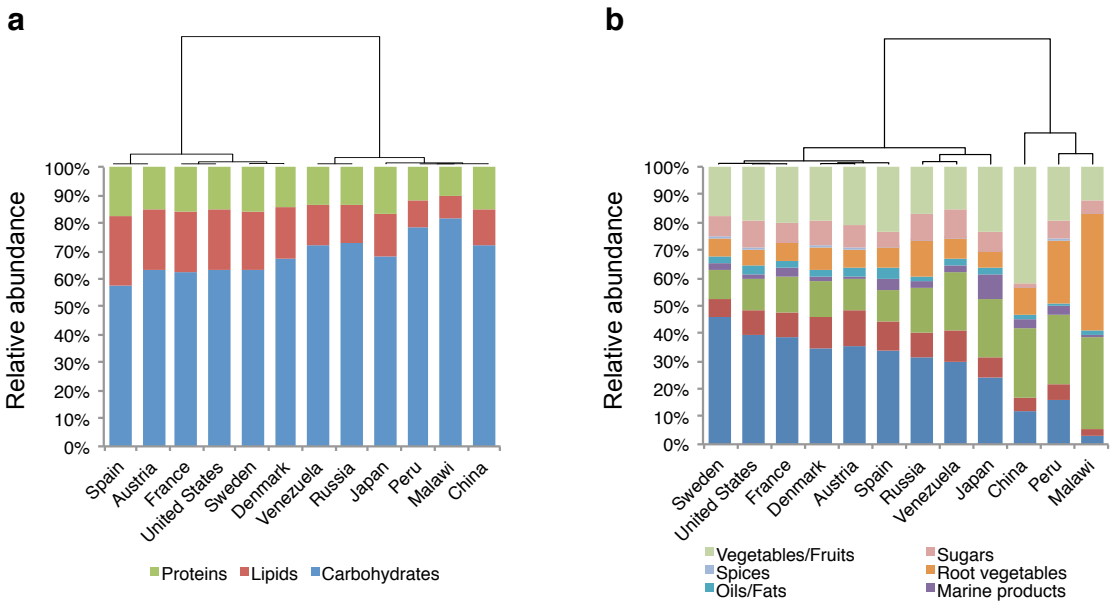


Fig. 3.2. Hierarchical clustering of the 12 countries based on average dietary intake data in the 10 years from 2002 to 2011. a, The dendrogarm of the 12 countries based on the ratio of three main nutrients (carbohydrates, proteins, and lipids) is shown. b, The dendrogarm of the 12 countries based on the ratio of nine food categories is shown. The dendrograms were generated using the Ward method and Pearson's correlation as dissimilarity.

Table 3.1. The average abundance of the top 26 genera used in this study

Genera	Average relative abundance
Bacteroides	14.89%
Prevotella	9.78%
Eubacterium	8.18%
Clostridium	7.95%
Faecalibacterium	5.96%
Unclassified Firmicutes	5.94%
Ruminococcus	5.86%
Blautia	5.75%
Alistipes	4.42%
Bifidobacterium	3.90%
Roseburia	3.61%
Coprococcus	1.53%
Escherichia	1.28%
Parabacteroides	1.27%
Dorea	1.24%
Dialister	1.12%
Anaerostipes	1.09%
Streptococcus	1.07%
Succinatimonas	1.05%
Butyrivibrio	0.96%
Collinsella	0.89%
Phascolarctobacterium	0.81%
Unclassified Clostridiales	0.70%
Methanobrevibacter	0.63%
Akkermansia	0.62%
Ruminiclostridium	0.53%
Total	91.02%

3.3.2. Correlation analysis of the microbiomes with dietary data

Correlation analysis with the three main nutrients revealed inverse relations between carbohydrate and protein/lipid levels for many of the genera tested (Fig. 3.3)[108]. Among them, the abundance of *Prevotella* and *Succinatimonas* positively correlated with carbohydrate and negatively with lipid and protein levels ($P < 0.05$ for all; Fig. 3.3a). At a finer level using nine food categories, the abundance of *Prevotella* and *Succinatimonas* positively correlated with "Grains/beans" and "Root vegetables", and negatively with "Animal products" ($P < 0.05$ for all; Fig. 3.3b). The abundance of *Ruminiclostridium* and *Akkermansia* showed inverse relations to those of *Prevotella* and

Succinatimonas (Fig. 3.3 and 3.4). The positive and negative associations of *Prevotella* with carbohydrate-rich and protein/lipid-rich diets respectively were also reported in several studies[51,101,108]. Unexpectedly, none of dietary factors showed a significant association with the major species *Bacteroides*, which is positively associated with Western diets rich in animal products and has a prominent trade-off relation with *Prevotella* in dietary association[51, 108, 135]

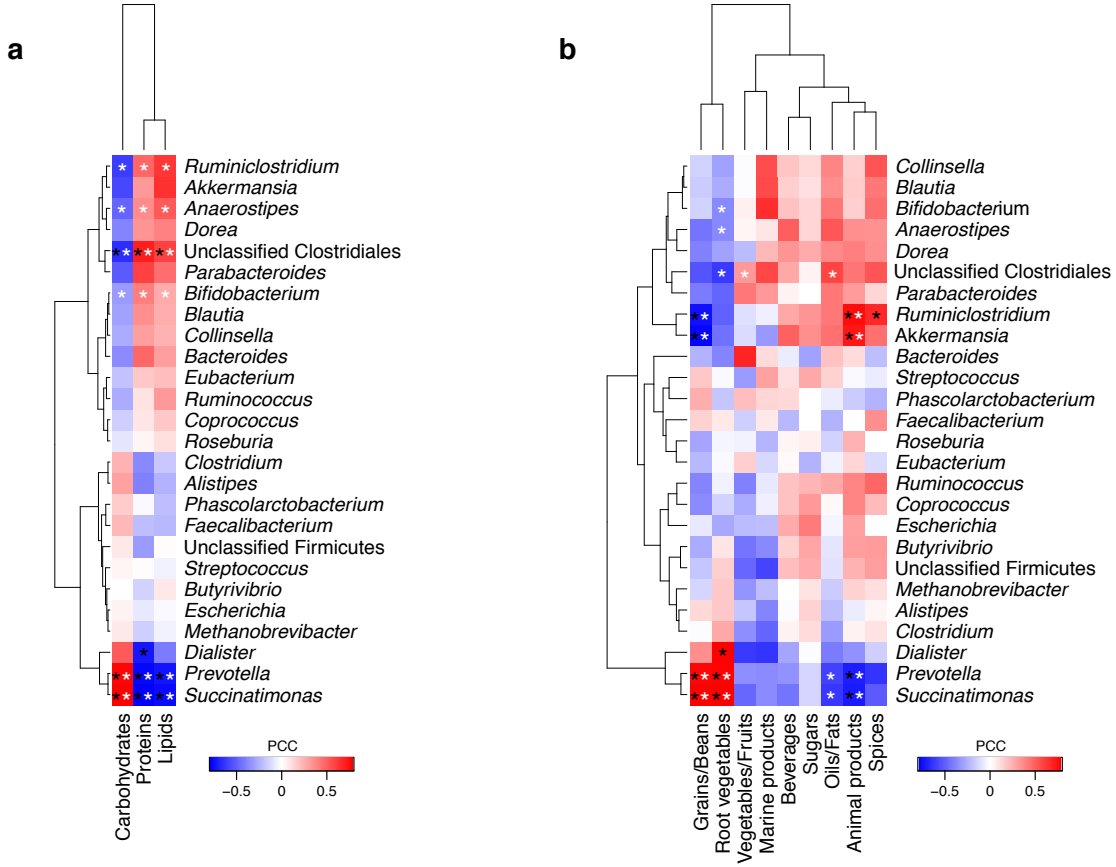


Fig. 3.3. Heat maps based on PCCs between the FAOSTAT dietary intake data and the abundance of the top 26 genera. Boxes depicted in red and blue indicate positive and negative correlations between microbial abundance and dietary intake data, for three main nutrients (a) and nine food categories (b). P-values were adjusted for multiple testing for each genus. Closed and open asterisks represent FDR adjusted p-values < 0.05 for PCC and SCC, respectively.

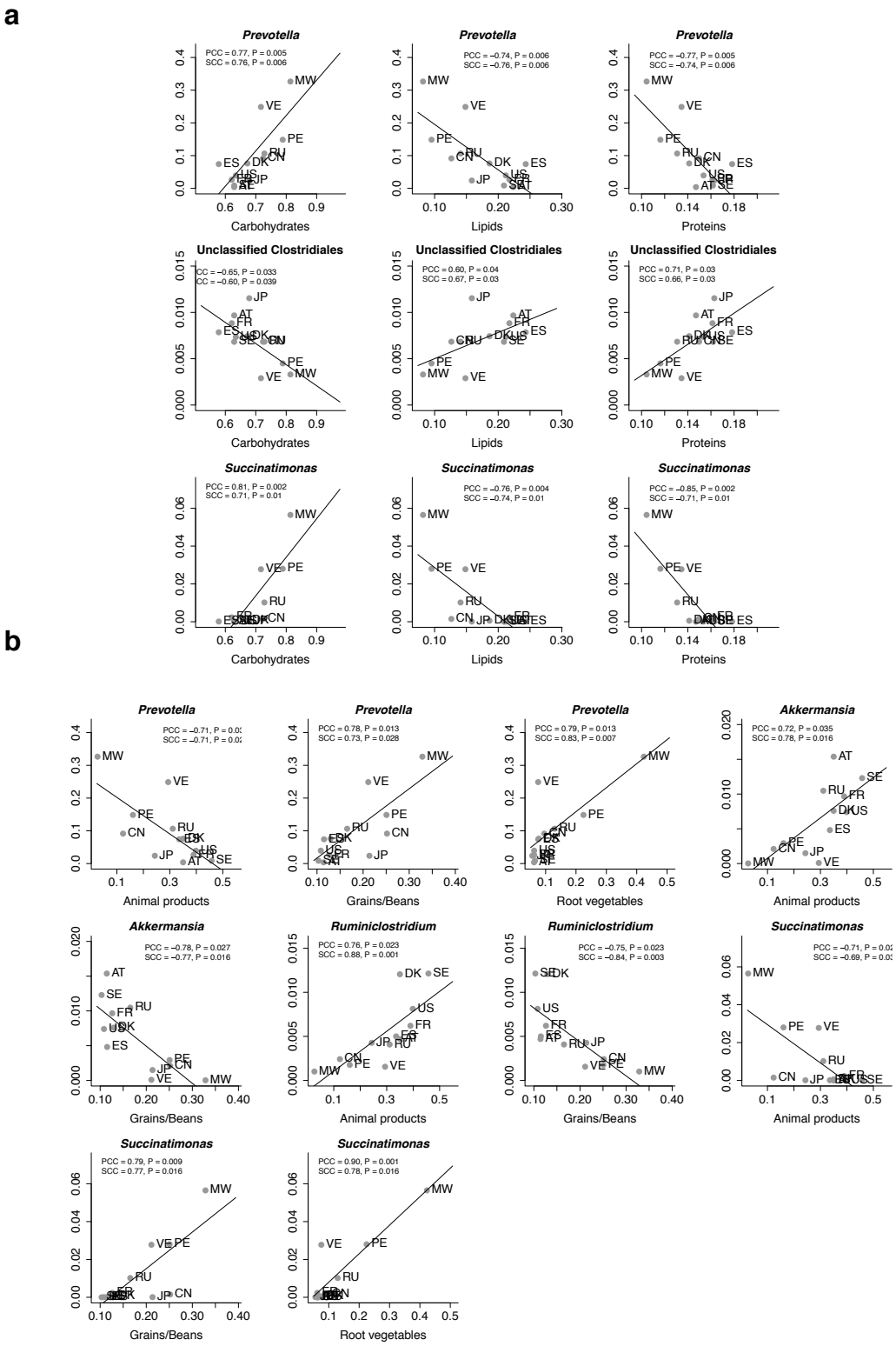


Fig. 3.4. Significant correlation between dietary data and genera. Correlations with statistical significance for both PCC and SCC between the abundance of genera and nutrients (a) and dietary categories (b) are shown**.**

3.3.3. Antibiotic usage in humans and farm animals in the countries

This shallow association between dietary intake and the abundance of *Bacteroides* suggested the existence of factors other than diet which might have a large influence on this major genus, as well as the population-level diversity in human gut microbiomes. In this context, I examined antibiotic usage because it can significantly alter the gut microbiome composition at the individual level[136-138]. Among the resources for antibiotic usage in humans, the collective data were available from two independent datasets: one mainly based on the IMS Health MIDAS database reported by Hogberg *et al*[140], and another based on the European Surveillance of Antimicrobial Consumption Network (ESAC)[141]. I found a significantly high correlation between the antibiotic usage of the countries common to both datasets (Fig. 3.5), suggesting the quantitative reliability and accuracy of both datasets. Since the Hogberg dataset included data for more countries than the ESAC dataset, I used the Hogberg dataset to collect the antibiotic usage information for 10 countries, since MW and CN were unavailable. The antibiotic usage for CN was obtained from other literature[142] (Tables 3.2). I also obtained the antibiotic usage in farm animals from a recent paper[143], for which data from seven countries were available (Table 3.3).

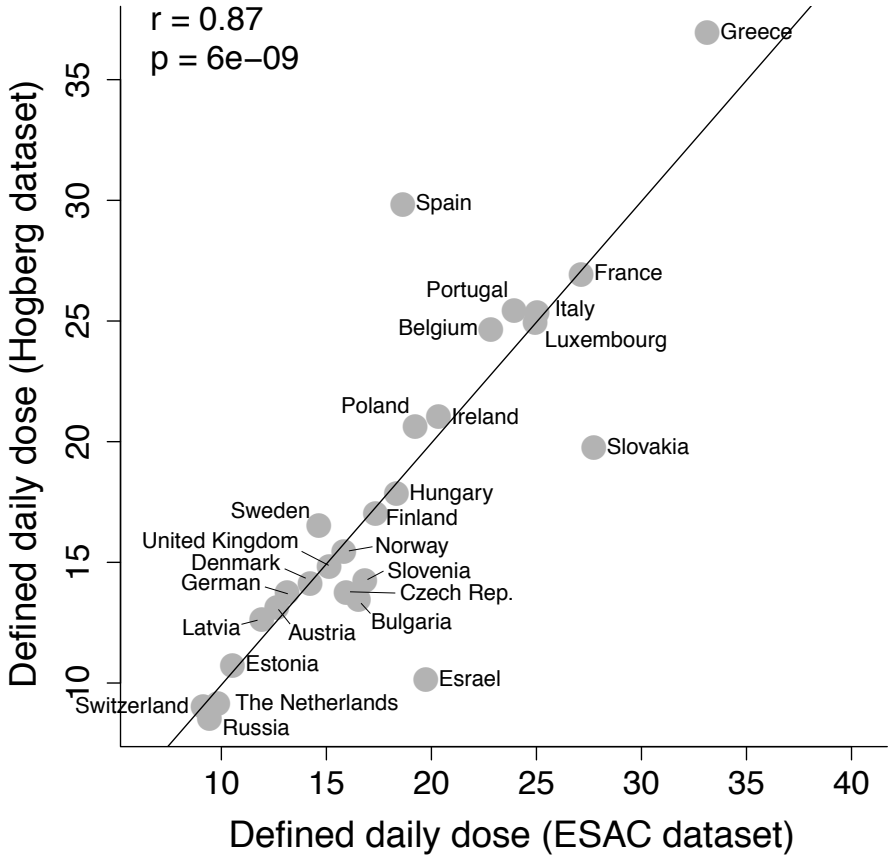


Fig. 3.5. Correlation between antibiotic usage data for countries common to both Hogberg and ESAC datasets. Two datasets of total antibiotic usages in humans, from the study by Hogberg *et al*[140] and the ESAC database[141], were used for the comparison. The antibiotic data of defined daily dose (DDD) in 2004 of the 26 countries common to both datasets were compared. PCCs and *P*-values are shown in the figure.

Table 3.2. Antibiotic usage in humans from 11 countries used in this study.

	AT (Austria)	CN (China)	DK (Denmark)	FR (France)	JP (Japan)	PE (Peru)	RU (Russia)	ES (Spain)	SE (Sweden)	US (The United States)	VE (Venezuela)
Beta-lactams	6.64	18.99	8.83	15.36	4.83	3.68	2.03	20.07	7.83	12.31	5.95
Macrolides/Liconsamides/Streptogramins	3.10	0.50	2.24	4.60	3.70	0.54	0.98	4.47	0.89	3.57	1.92
Quinolones	1.42	0.53	0.28	2.13	1.49	0.96	1.05	3.06	1.17	2.49	1.35
Tetracyclines	1.27	No data	1.17	3.50	0.71	0.42	1.29	1.19	3.65	4.65	0.55
Sufonamides/Trimethoprim	0.55	No data	0.77	0.43	0.03	0.59	2.01	0.54	0.79	1.34	1.13
Other antibiotics	0.16	No data	0.77	0.90	0.17	0.20	1.18	0.42	2.15	0.78	0.58
Total antibiotics	13.14	28.02	14.06	26.93	10.93	6.38	8.53	29.75	16.48	25.13	11.48

The antibiotic usage for all countries, except CN was obtained from Hogberg LD et al.

The antibiotic usage for CN was obtained from Wang X et al.

All the data was recorded in 2004.

Antibiotic usage was indicated by DDD (defined daily dose per 1,000 inhabitants).

No data for MW was available anywhere.

Table 3.2. Antibiotic usage in farm animals from seven countries used in this study

	AT (Austria)	DK (Denmark)	FR (France)	JP (Japan)	ES (Spain)	SE (Sweden)	US (The United States)
Total antibiotic usage	63,000	117,000	997,000	655,820	1,746,000	13,559	13,542,030
Total antibiotic usage per population	0.00739	0.0207	0.0154	0.00516	0.0371	0.00141	0.0420

The antibiotic usage was obtained from Boeckel TP et al.

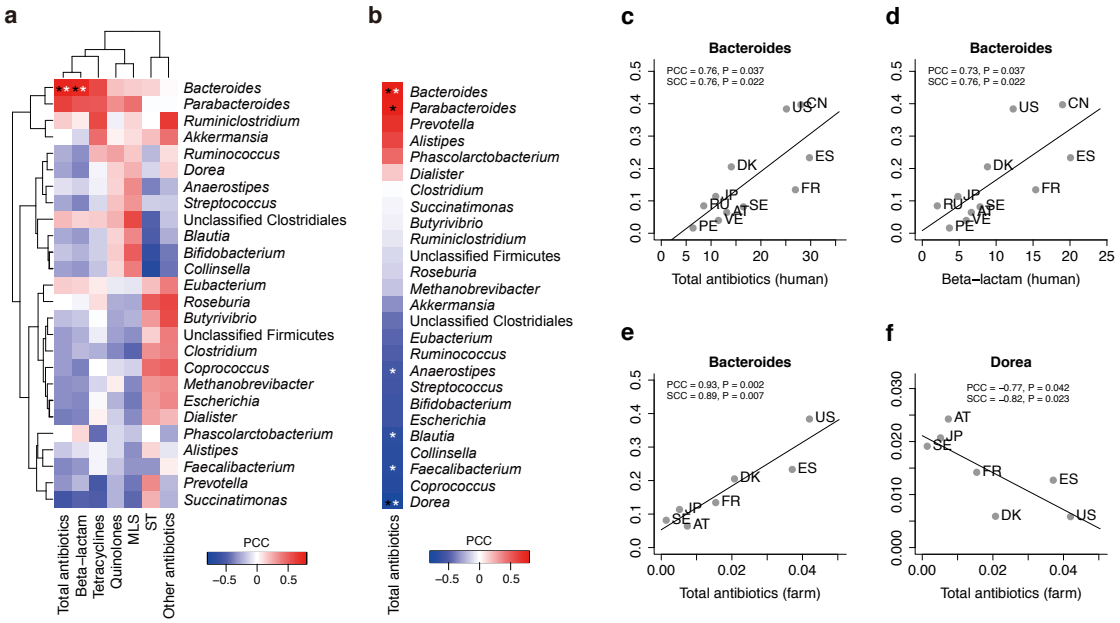
The population data was obtained from FAOSTAT.

All data was recorded in 2010, except for the US which was from 2011.

Antibiotic usage was indicated in kilograms

3.3.4. Correlation analysis of the microbiome with the antibiotic usage

Correlation analysis with the antibiotic usage data found significant positive correlations between the abundance of major species *Bacteroides* with both total antibiotic and beta-lactam usage in humans, and total antibiotic usage in farm animals ($P < 0.05$ for all; Fig. 3.6a-e). The abundance of minor genera, *Odoribacter*, *Parasutterella*, *Sutterella* and *Acetobacter,* also showed a significant positive correlation with total antibiotic usage in humans or farm animals ($P < 0.05$ for all; Fig. 3.7). In contrast, *Dorea* and *Eggerthella* showed significant negative correlation with total antibiotic usage in farm animals ($P < 0.05$ for both) and other five genera such as *Blautia, Collinsella, Coprococcus and Faecalibacterium* had the tendency of a high negative correlation with antibiotics (PCC and SCC < -0.60 for all; Fig. 3.6a and b). The strong association of antibiotics with the gut microbiome were further supported by PERMANOVA[108], where total antibiotic usage in humans and farm animals and beta-lactam usage in humans significantly contributed to the overall structure of the gut microbiome (coefficient of determination (R^2) = 0.34, 0.59 and 0.31 respectively) as well as plant-derived dietary factors such as “Root vegetables” and “Vegetables/Fruits” (R^2 = 0.26 and 0.26 respectively; Fig. 3.6g). These data are summarized in the correlation network of microbe-food-antibiotics (Fig. 3.8).



g

	Factor	R^2	P-value
Food categories	Animal products	0.15	0.151
	Beverages	0.04	0.731
	Grains/beans	0.21	0.063
	Marine products	0.11	0.312
	Oils/Fats	0.14	0.184
	Root vegetables	**0.26**	**0.023**
	Spices	0.17	0.134
	Sugars	0.05	0.681
	Vegetables/Fruits	**0.26**	**0.014**
Antibiotics uses (Human)	**Beta-lactam**	**0.31**	**0.018**
	MLS	0.09	0.440
	Quinolones	0.15	0.244
	Tetracyclines	0.21	0.107
	ST	0.11	0.407
	Total antibiotics	**0.34**	**0.009**
Antibiotics uses (Farm)	**Total antibiotics**	**0.59**	**0.004**
Host data	Age	0.16	0.140
	BMI	0.20	0.072

Fig. 3.6. Correlations between the abundance of the top 26 genera and antibiotic usage. a, b, Heat maps based on PCCs between the abundance of the top 26 genera and antibiotic usage in humans and farm animals are shown. Closed and open asterisks represent FDR adjusted p-values < 0.05 for PCC and SCC, respectively. c-f, Significant correlations for both PCC and SCC of the abundance of *genera* with antibiotic usage are shown. Y-axes represent relative abundance of the genus. g, PERMANOVA for dietary and antibiotic factors to gut microbiome variation are shown. R^2 indicates the coefficient of determination. Factors with *P*-values < 0.05 are shown in bold letters.

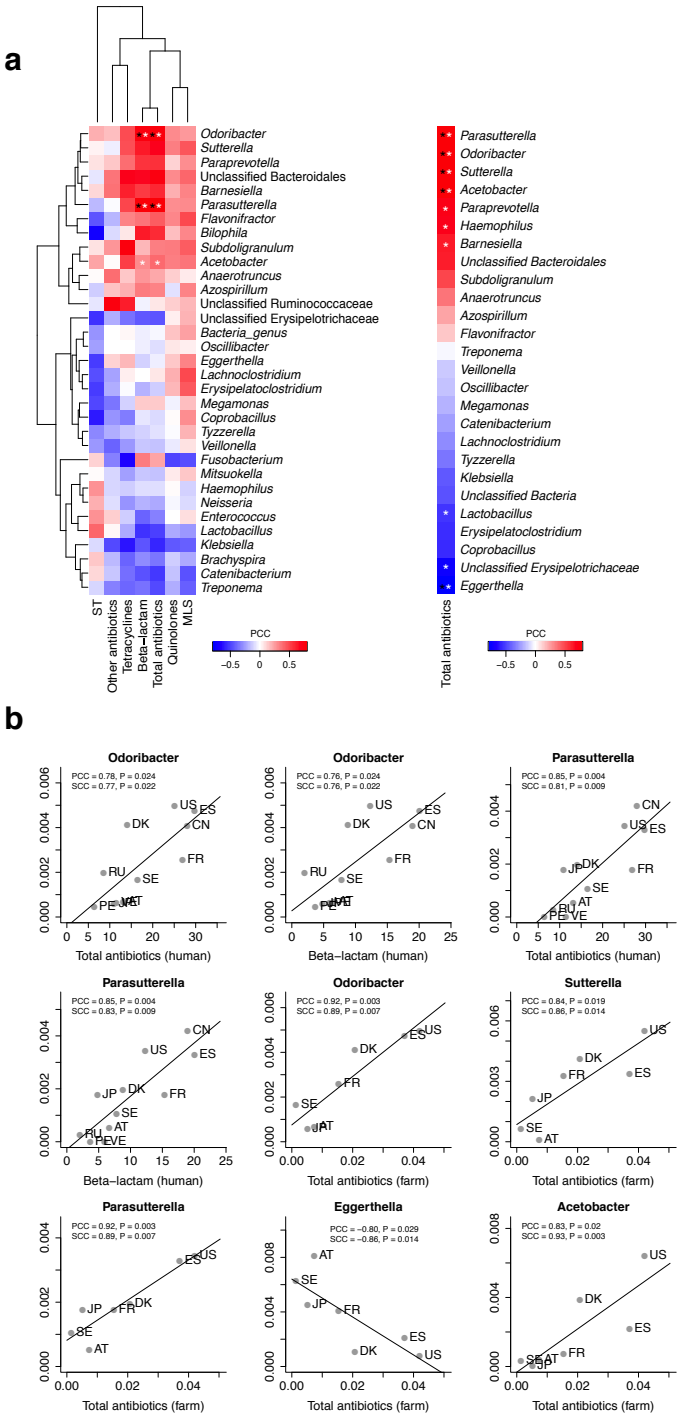


Fig. 3.7. Correlations between the abundance of minor genera and antibiotic usage. a, Heat maps based on PCCs that were calculated between 33 minor genera with the average abundance ≤ 0.05% and ≥ 0.01%, and antibiotic usage in humans and farm animals are shown. b, Significant correlations for both PCC and SCC of the abundance of the *genera* with antibiotic usage are shown. Y-axes represent relative abundance of the genus. P-values were adjusted for multiple testing for each genus. Closed and open asterisks represent FDR adjusted p-values < 0.05 for PCC and SCC, respectively. Significant correlations for both PCC and SCC of the abundance of several minor genera and antibiotic usage are shown in c.

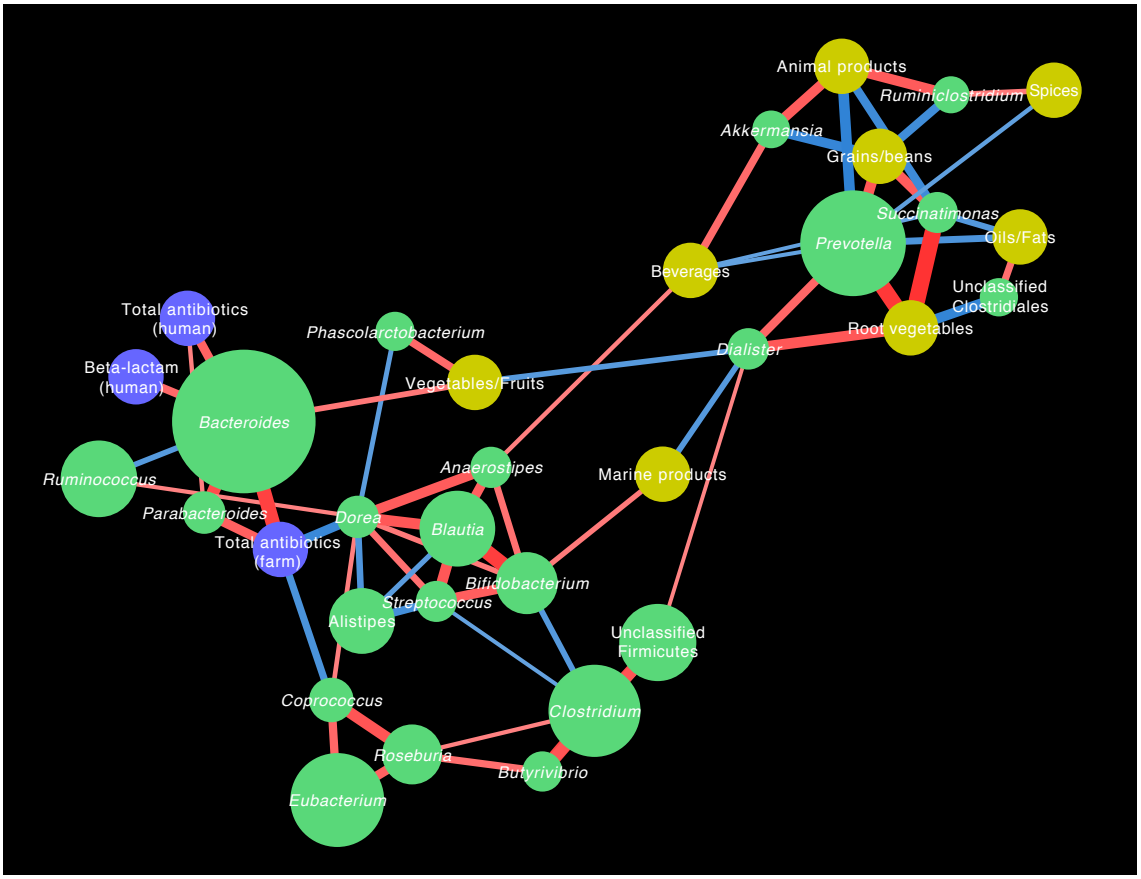


Fig. 3.8. Correlation network of microbes, antibiotics and diet. Circles in green, yellow and purple indicate microbes, dietary components, and antibiotics, respectively. Edges were drawn when the PCC is ≥ 0.6 or ≤ -0.6 between the abundance of microbes and dietary components, and antibiotic usage in humans, and when it is ≥ 0.7 or ≤ -0.7 between the abundance of microbes and antibiotic usage in farm animals. The red and blue colors of the edges indicate positive and negative correlations, respectively, and thickness indicates the degree of the PCC. The correlation network was drawn using Cytoscape 3.2.1.

Similarly, I also found the positive correlation between the abundance of *Bacteroides* and antibiotic usage in the gut microbiome of other independent cohort composed of 303 Asian children from five countries (China, Indonesia, Japan, Taiwan and Thailand)[97]. Since only 16S rRNA gene sequence data were used for the analysis of the microbial composition in this cohort, I used the 16S rRNA gene sequences publically available for estimation of the abundance of *Bacteroides* in the five countries.

Also, since the antibiotic usage data in humans for the five countries was available from the Hogberg dataset but that in farm animals for four countries was unavailable, I performed the correlation analysis between the abundance of *Bacteroides* and the antibiotic usage in humans among the five countries. The results showed the clear tendency of positive correlations between *Bacteroides* and antibiotic usage in humans in the five countries (Fig. 3.9). Thus, the positive correlation between *Bacteroides* and antibiotic usage seems to be also the case for the independent Asian's cohort although I did not statistically evaluate the correlation because of the number of countries was insufficient.

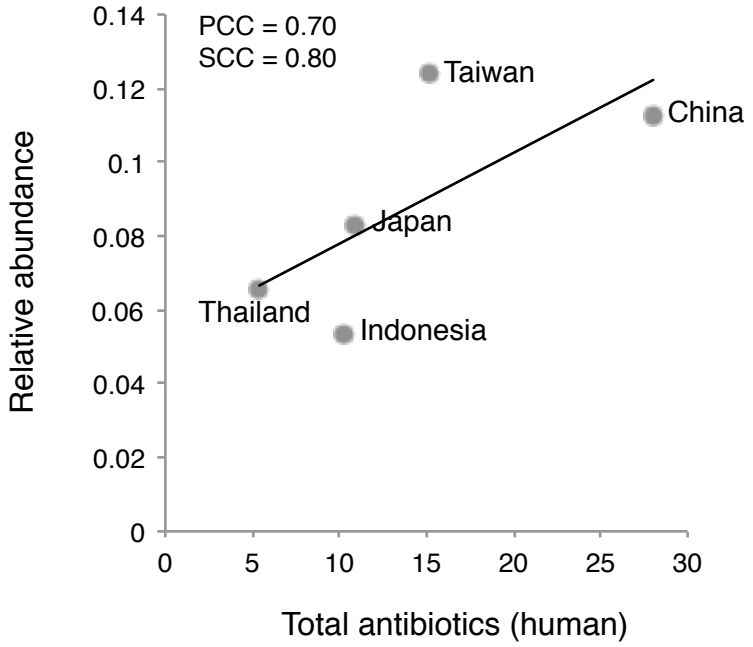


Fig. 3.9. Correlation between *Bacteroides* and antibiotic usage among the five Asian countries. Correlations between the average relative abundance of *Bacteroides* and the total antibiotic usage in humans are shown. PCC and SCC are indicated in the figure.

3.3.5. Antibiotic resistance genes

To explore the involvement of antibiotic resistance genes (ARs) in the association of antibiotics usage with gut microbiome structure, I compared the frequencies of ARs in the gut microbiome with antibiotic usage among the countries. The results showed that the total antibiotic usage and beta-lactam usage in humans showed a positive correlation with beta-lactam resistances, the resistance-nodulation-cell division (RND) efflux pump and total resistance genes ($P < 0.05$; Fig. 3.10a). Additionally, antibiotic usage in farm animals showed a positive correlation with RND efflux pomp among the countries (Fig. 3.10b). The positive correlations of the frequency of ARs in gut microbiomes with antibiotic usage in the country were also demonstrated previously[148, 149]. Collectively, these data suggested that the antibiotic usage tended to be associated with an increase of ARs in the individual gut microbiomes in the country.

To further investigate the contribution of ARs to association of the microbial abundance with antibiotic usage, I compared the frequencies of ARs annotated in genomes between *Bacteroides*, four minor genera positively associated with antibiotic usage, and other genera having little association with antibiotic usage. The results indicated that the positive-associated genera encoded more ARs than other genera, suggesting that the proliferation of ARs underlies the positive correlation between these genera and antibiotic usage (Fig. 3.11a). Among the ARs, RND efflux pump was significantly enriched in the positive-associated genera as compared with other genera (Fig. 3.11b).

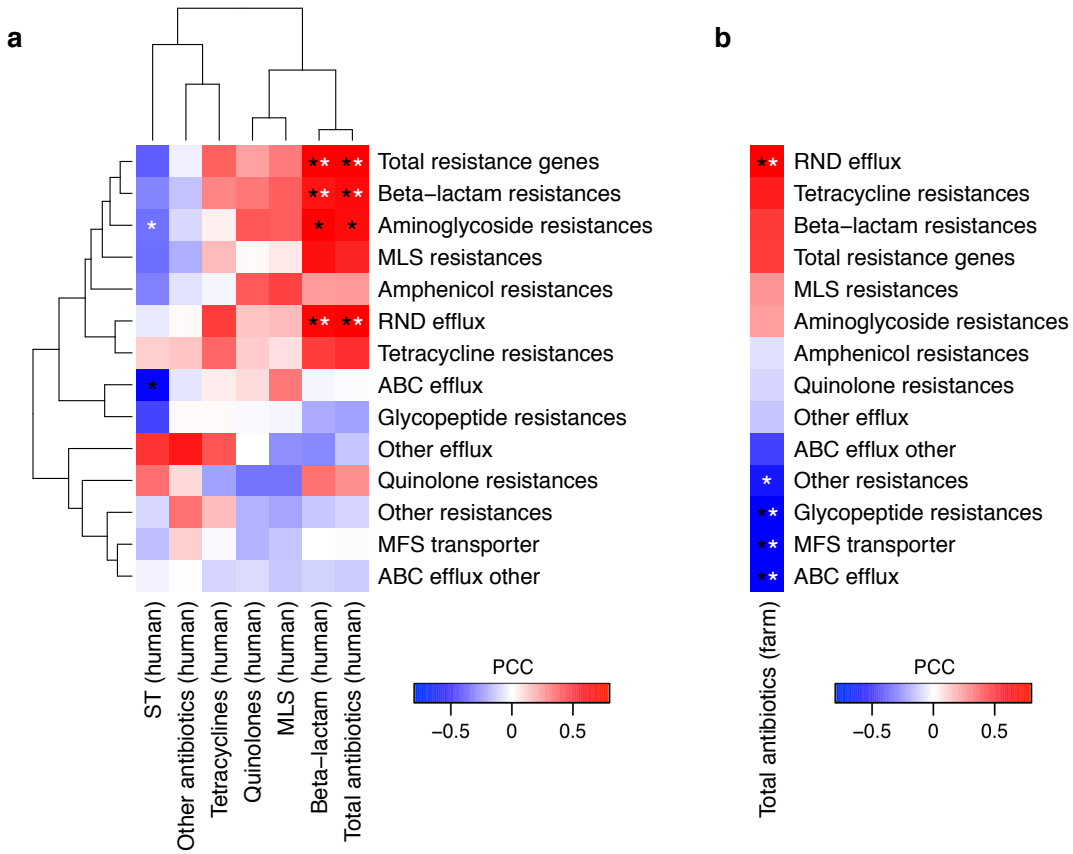


Fig. 3.10. Correlations between the abundance of ARs and antibiotic usage. Heat maps based on PCCs that were calculated between ARs and antibiotic usage in humans (a) and the total antibiotic usage in farm animals (b) are shown. P-values were adjusted for multiple testing for each AR. Closed and open asterisks represent FDR adjusted p-values < 0.05 for PCC and SCC, respectively.

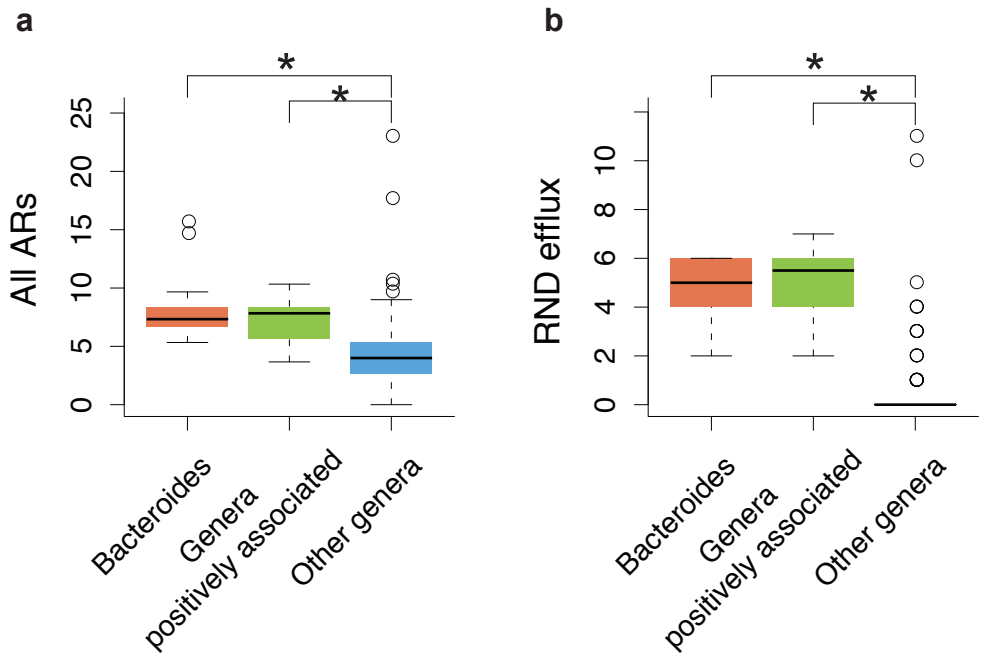


Fig. 3.11. Comparison of the frequency of ARs. a, b, The number of ARs annotated in genomes was compared among the genus *Bacteroides* (red) and four minor genera (*Parabacteroides*, *Parasutterella*, *Odoribacter* and *Sutterella*) having a significant positive correlation with antibiotic usage in humans or farm animals (green) and other genera having little association with antibiotic usage (blue). All ARs annotated and the RND efflux pump are shown in a and b, respectively. The vertical axis indicates the number of the corresponding ARs per genome. Asterisks represent *P*-values < 0.05 (Student's *t*-test).

3.3.6. Gut microbiomes of mice treated with antibiotics

I also experimentally validated the positive correlation between the abundance of *Bacteroides* and beta-lactam usage using mice treated and untreated with beta-lactam antibiotics. In two independent experiments[119, 147], the analysis of the 16S rRNA gene sequences of the gut microbiomes revealed that the antibiotic treatment increased the abundance of *Bacteroides* (Fig. 3.12). Similarly, the increase in *Bacteroides* abundance was also observed in a human intervention trial with beta-lactam[149].

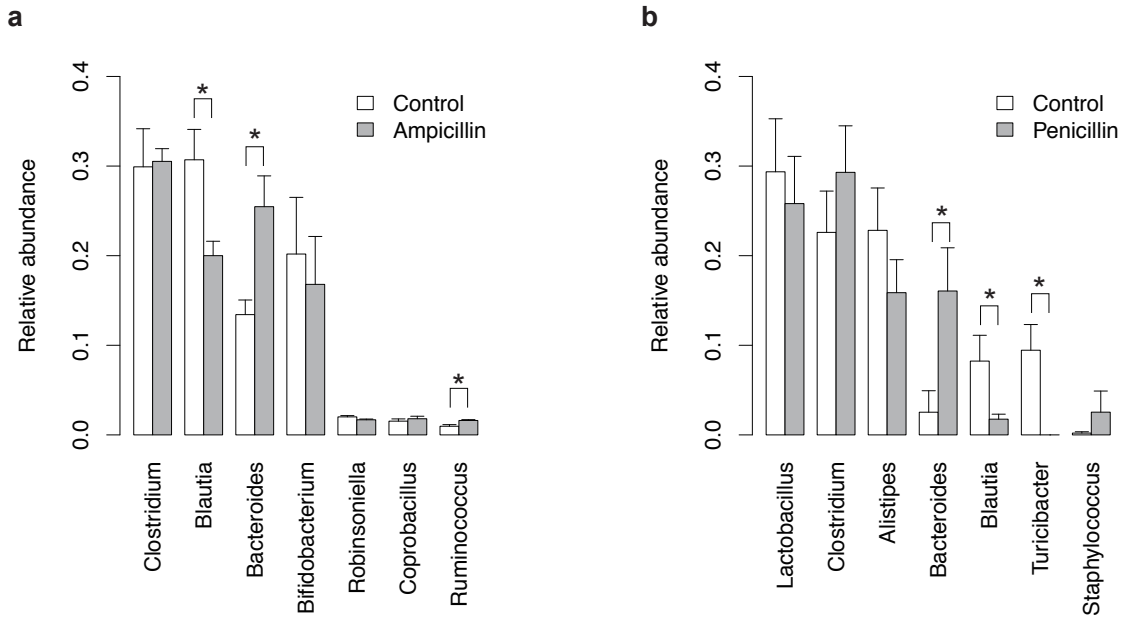


Fig. 3.12. Increase in the abundance of *Bacteroides* in mice treated with beta-lactams. a, b, The average microbial compositions at the genus level evaluated by the 16S rRNA gene sequences were compared between mice treated and untreated with beta-lactams. The results using the 16S rRNA sequenceing data obtained from reference 119 (control: n=4, ampicillin: n=5) and from reference 147 (control: n=36, penicillin: n=39) are shown in a and b, respectively. Top seven predominant genera including *Bacteroides* are shown in both analyses. Open and grey bars indicate control (untreated) and treated mice, respectively. Asterisks represent *P*-values < 0.05 (Student's *t*-test). Error bars indicate SEM.

3.4. Discussion

As presented here, a large-scale correlation analysis between gut microbes, diet and antibiotics provided evidence for the strong impact of antibiotics usage as well as diet on the gut microbiome structure. The present data also suggested that both antibiotics and diet more profoundly affect the microbial composition in the human gut microbiome than host's genetic and geographical closeness among the countries. Thus, both antibiotics and diet may be the primary factors that shape the human gut microbiome, resulting in the population-level diversity. It is also noted that antibiotics affects mainly *Bacteroides*, while diet affects mainly *Prevotella*, both of which are key species in determining the specificity and diversity in the human gut microbiome, such as enterotypes[25]. Therefore, the trade-off relation between these two major species, which was thought to be mainly due to differences in dietary habitat, may be a consequence of respective independent effects from dietary and antibiotic factors on the human gut microbiome. In addition, the present study also suggest that antibiotics has been involved in not only the emergence of the antibiotic resistance microbes[150, 151], but also in ecological changes to the human microbiome[152].

I showed that the RND efflux pomp is involved in the increase of the abundance of Bacteroides associated with antibiotic usage by conferring the antibiotic resistance property to this species. In addition, *Bacteroides* is also resistant to bile acids, which are excreted into the gut and act to form micelle with lipids in the diet. Interestingly, the RND efflux pump plays an important role in bacterial bile acid tolerance[150, 151]. Also, bile acid-tolerant species, including *Bacteroides*, were rapidly induced by animal-based dietary intervention[135]. Therefore, there may be a similar mechanism involving the RND efflux pump between the increase in *Bacteroides* due to antibiotic usage and that by short-term animal-based diets.

It has been demonstrated that antibiotics possess both rapid and long-term effects on the gut microbiome structure[136-138]. On the other hand, the dietary intervention studies for individuals indicated that diet-induced changes in the microbial abundance were smaller than inter-individual variations, and were restored rapidly when the intervention was eliminated[108, 135]. Thus, these observations support the present results showing that the association of antibiotics with gut microbiome structure is similar to or greater than that of the diet.

It should be also noted that antibiotics usage in farm animals showed a correlation with the human gut microbiome. Antibiotics usage in farm animals outnumbers that in humans in the United States[143]. As suggested for the high association between bacteria in the human gut and those in farm-associated environments[153, 154], it is also conceivable that the level of steady exposure of antibiotics from environments to humans may be stochastically greater than that of direct but occasional administration of antibiotics to humans[152]. The elucidation for plausible mechanisms proposed above remains as a future challenge.

Perturbations in the gut microbiome induced by antibiotic treatment have been proven to link to the etiology of several diseases in mice[147, 155, 156]. Additionally, for humans, several studies have suggested an association of antibiotic exposure with diseases such as IBD[157], obesity[158, 159] and asthma[160]. Thus, the present and other data imply a possible association of antibiotic usage with the prevalence of modern diseases in developed countries. However, further studies are required to address the influence of antibiotics on modern diseases through the gut microbiome.

4. Conclusion

In my study, I conducted a metagenomic analysis of fecal DNA samples from 106 JP individuals, compared the JP metagenomic data with those from other 11 countries, and performed an association study between 861 human gut microbiomes and the epidemiological data of dietary intake and antibiotic usage. Comparative analysis of the gut microbiomes among the 12 countries showed a great population-level diversity in the human gut microbiome. I found that the JPGM was characterized by a significant enrichment in the Actinobacteria phylum and *Bifidobacterium* genus and a remarkable depletion of the methanogenic archaeon, *Methanobrevibacter smithii*. These microbial differences in the abundance contributed to differences in functional features, such as the enrichment of carbohydrate metabolism genes with a concurrent depletion of genes involved in replication and repair. In addition, the high abundance of genes for acetogenesis in contrast to a depletion of genes for methanogenesis suggested a difference in hydrogen metabolism in the gut between the Japanese and other populations.

Additionally, I found that the population-level diversity in the human gut microbiome among the 12 countries was significantly associated with antibiotics usage in humans and farm animals as well as dietary intake. In particular, one of the major genus *Bacteroides* showed a strong positive correlation with total antibiotic usage in humans and farm animals and with beta-lactam antibiotic usage in humans. Another major genus *Prevotella* showed no association with antibiotic usages but a strong association with dietary intake. Comparative genome analysis revealed that the genera positively associated with antibiotics have more ARs than the genera that showed little association with antibiotics. Additionally, gut microbiomes of mice treated with a

beta-lactam antibiotic increased in the relative abundance of *Bacteroides*. These results suggested a strong contribution of antibiotics to the human gut microbiome.

Taken together, I revealed the characteristic features of the JPGM and found the strong association of antibiotics as well as diet with population-level diversity in the human gut microbiome. These results included many invaluable and novel findings, therefore, the present study provided new insights into the human gut microbiome research fields. In addition, I disclosed all resources which are of great use for metagenomics to public domains. I anticipate that these results will be helpful for future studies to promote human health and well-being.

5. References

1. Luckey, T.D. Introduction to intestinal microecology. *Am. J. Clin. Nultr.* **25**, 1292-1294 (1972).
2. Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. & Gordon, J.I. Host-bacterial mutualism in the human intestine. *Science* **307**, 1915-20 (2005).
3. Sender, R., Fuchs, S. & Milo R. Revised estimates for the number of human and bacteria cells in the body. *bioRxiv* (2016).
4. Bianconi, E. et al. An estimation of the number of cells in the human body. *Ann. Hum. Biol.* **40**, 463-71 (2013).
5. O'Hara, A.M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO reports* **7**, 688-93 (2006).
6. Lederberg, J. Infectious history. *Science* **288**, 287-293 (2000).
7. Tortora, G.J., Funke, B.R. & Case, C.L. *Microbiology: An Introduction, 11th edition*. (2012). San Fransisco: Benjamin Commings.
8. Leahy, S.C., Higgins, D.G., Fitzgerald, G.F. & van Sinderen, D. Getting better with bifidobacteria. *J. Appl. Microbiol.* **98**, 1303-15 (2005).
9. Hooper, L.V. & Gordon, J.I. Commensal host-bacterial relationships in the gut. *Science* **292**, 1115-1118 (2001).
10. Turnbaugh, P.J. et al. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-31 (2006).
11. Hayashi, H., Sakamoto, M. & Benno, Y. Phylogenetic analysis of human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbial Immunol.* **46**, 535-548 (2002).
12. Venter, J.C. et al. Environmental genome shotgun sequencing of the Sargasso

Sea. *Science* **304**, 66-74 (2004).

13. Gill, S. et al. Metagenomic analysis of the human distal gut microbiome. *Science* **312**, 1355-9 (2006).
14. Eckburg, P.B. et al. Diversity of the human intestinal microbial flora. *Science* **308**, 1635-8 (2005).
15. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228-35 (2005).
16. Kurokawa, K. et al. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res.* **14**, 169-81 (2007).
17. Qin, J. et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65 (2010).
18. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214 (2012).
19. The Human Microbiome Project Consortium. A framework for human microbiome research. *Nature* **486**, 215-21 (2012).
20. The Human Microbiome Jumpstart Reference Strains Consortium. A catalog of reference genomes from the human microbiome. *Science* **328**, 994-9 (2010).
21. Abubucker, S. et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* **8**, e1002358 (2012).
22. Langille, M.G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814-21 (2013).
23. Sunagawa, S. et al. Metagenomic species profiling using universal phylogenetic marker genes. *Nature methods* **10**, 1196-9 (2013).
24. Turnbaugh, P.J. et al. A core gut microbiome in obese and lean twins. *Nature* **457**, 480-4 (2009).

25. Arumugam, M. et al. Enterotypes of the human gut microbiome. *Nature* **473**, 174-180 (2011).

26. Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55-60 (2012).

27. Karlsson, F.H. et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99-103 (2013).

28. Tyakht, A.V. et al. Human gut microbiota community structures in urban and rural populations in Russia. *Nat. Commun.* **4**, 2469 (2013).

29. Qin, N. et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59-64 (2014).

30. Feng, Q. et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* **6**, 6528 (2015).

31. Backhed, F. et al. Dynamics and tabilization of the human gut microbiome during the first year of life. *Cell host & microbe* **17**, 690-703 (2015).

32. Zhang, X. et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat. Med.* **21**, 895-905 (2015).

33. Zeller, G. et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **10**, 766 (2014).

34. Oh, J. et al. Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59-64 (2014).

35. Faith, J.J. et al. The long-term stability of the human gut microbiota. *Science* **341**, 1237439 (2013).

36. Schloissnig, S. et al. Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45-50 (2013).

37. David, L.A. et al. Host lifestyle affects human microbiota on daily timescales.

Genome Biol. **15**, R89 (2014).

38. Gao, Z., Tseng, C.H., Pei, Z. & Blaser, M.J. Molecular analysis of human forearm superficial skin bacterial biota. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2927-32 (2007).

39. Fierer, N. et al. Forensic identification using skin bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6477-81 (2010).

40. Findley, K. et al. Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498**, 367-70 (2013).

41. Dewhirst, F.E. et al. The human oral microbiome. *J. Bacteriol.* **192**, 5002-17 (2010).

42. Ravel, J. et al. Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4680-4687 (2011).

43. Aagaard, K. et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* **7**, e36466 (2012).

44. Yan, M. et al. Nasal microenvironments and interspecific interactions influence nasal microbiota complexity and *S. aureus* carriage. *Cell host & microbe* **14**, 631-40 (2013).

45. Bik, E.M. et al. Molecular analysis of the bacterial microbiota in the human stomach. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 732-7 (2006).

46. Pei, Z. et al. Bacterial biota in the human distal esophagus. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4250-5 (2004).

47. Aagaard, K. et al. The placenta harbors a unique microbiome. *Sci. Transl. Med.* **6**, 237ra65 (2014).

48. El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D. & Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut

microbiota. *Nat. Rev. Microbiol.* **11**, 497-504 (2013).

49. Hehemann, J.H. et al. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908-12 (2010).

50. Dominguez-Bello, M.G. et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11971-5 (2010).

51. Yatsunenko, T. et al. Human gut microbiome viewed across age and geography. *Nature* **486**, 222-227 (2012).

52. Koenig, J.E. et al. Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4578-4585 (2011).

53. Dinsdale, E.A. et al. Functional metagenomic profiling of nine biomes. *Nature* **452**, 629-32 (2008).

54. Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R. & Gordon, J.I. Worlds within worlds- evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* **6**, 776-788 (2008).

55. Ley, R.E. et al. Evolution of Mammals and Their Gut Microbes. *Science* **320**, 1647-1651 (2008).

56. Muegge, B.D. et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**, 970-4 (2011).

57. Ley, R.E., Peterson, D.A. & Gordon, J.I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837-48 (2006).

58. Moeller, A.H. et al. Rapid changes in the gut microbiome during human evolution. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16431-5 (2014).

59. Kau, A.L., Ahern, P.P., Griffin, N.W., Goodman, A.L. & Gordon, J.I. Human nutrition, the gut microbiome and the immune system. *Nature* **474**, 327-36

(2011).

60. Cho, I. & Blaser, M.J. The human microbiome: at the interface of health and disease. *Nature reviews. Genetics* **13**, 260-70 (2012).

61. Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220-30 (2012).

62. Nicholson, J.K. et al. Host-Gut microbiota metabolic interactions. *Science* **336**, 1262-1267 (2012).

63. Bergman, E.N. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiol. Rev.* **70**, 567-590 (1990).

64. Furusawa, Y. et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446-50 (2013).

65. Atarashi, K. et al. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232-6 (2013).

66. Fukuda, S. et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-7 (2011).

67. Wikoff, W.R. et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3698-703 (2009).

68. Le Chatelier, E. et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541-546 (2013).

69. Morgan, X.C. et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, R79 (2012).

70. Gevers, D. et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell host & microbe* **15**, 382-92 (2014).

71. Zackular, J.P., Rogers, M.A., Ruffin, M.T.t. & Schloss, P.D. The human gut microbiome as a screening tool for colorectal cancer. *Cancer prevention research* **7**, 1112-1121 (2014).

72. Miyake, S. et al. Dysbiosis in the gut microbiota of patients with multiple sclerosis, with a striking depletion of species belonging to clostridia XIVa and IV clusters. *PLoS One* **10**, e0137429 (2015).

73. Sokol, H. et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16731-6 (2008).

74. Parracho, H.M., Bingham, M.O., Gibson, G.R. & McCartney, A.L. Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *J. Med. Microbiol.* **54**, 987-91 (2005).

75. Kang, D.W. et al. Reduced incidence of *Prevotella* and other fermenters in intestinal microflora of autistic children. *PLoS One* **8**, e68322 (2013).

76. Diaz Heijtz, R. et al. Normal gut microbiota modulates brain development and behavior. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3047-52 (2011).

77. Cryan, J.F. & Dinan, T.G. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* **13**, 701-12 (2012).

78. Hsiao, E.Y. et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451-63 (2013).

79. van Nood, E. et al. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *N. Engl. J. Med.* **368**, 407-15 (2013).

80. Weingarden, A. et al. Dynamic changes in short- and long-term bacterial composition following fecal microbiota transplantation for recurrent *Clostridium*

difficile infection. *Microbiome* **3**, 10 (2015).

81. Youngster, I. et al. Oral, capsulized, frozen fecal microbiota transplantation for relapsing *Clostridium difficile* infection. *JAMA* **312**, 1772-8 (2014).

82. Hamilton, M.J., Weingarden, A.R., Unno, T., Khoruts, A. & Sadowsky, M.J. High-throughput DNA sequence analysis reveals stable engraftment of gut microbiota following transplantation of previously frozen fecal bacteria. *Gut microbes* **4**, 125-35 (2013).

83. Cammarota, G. et al. Randomised clinical trial: faecal microbiota transplantation by colonoscopy vs. vancomycin for the treatment of recurrent *Clostridium difficile* infection. *Aliment. Pharmacol. Ther.* **41**, 835-43 (2015).

84. Angelberger, S. et al. Temporal bacterial community dynamics vary among ulcerative colitis patients after fecal microbiota transplantation. *Am. J. Gastroenterol.* **108**, 1620-30 (2013).

85. D., K., Hotte, N., Gillevet, P. & Madsen, K. Fecal microbiota transplantation inducing remission in Crohn's colitis and the associated changes in fecal microbiota profile. *J. Clin. Gastroenterol.* **48**, 625-628 (2014).

86. Olle, B. Medicines from microbiota. *Nat. Biotechnol.* **31**, 309-315 (2013).

87. Frank, D.N. et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13780-13785 (2007).

88. Giongo, A. et al. Toward defining the autoimmune microbiome for type 1 diabetes. *ISME J.* **5**, 82-91 (2011).

89. Wen, L. et al. Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* **455**, 1109-13 (2008).

90. Scher, J.U. et al. Expansion of intestinal *Prevotella copri* correlates with

enhanced susceptibility to arthritis. *eLIFE*, e01202 (2013).

91. Koren, O. et al. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4592-4598 (2011).

92. Karlsson, F.H. et al. Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nat. Commun.* **3**, 1245 (2012).

93. Saulnier, D.M. et al. Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome. *Gastroenterology* **141**, 1782-91 (2011).

94. Rigsbee, L. et al. Quantitative profiling of gut microbiota of children with diarrhea-predominant irritable bowel syndrome. *Am. J. Gastroenterol.* **107**, 1740-51 (2012).

95. Smith, M.I. et al. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**, 548-54 (2013).

96. Subramanian, S. et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417-21 (2014).

97. Bisgaard, H. et al. Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. *J. Allergy Clin. Immunol.* **128**, 646-52 e1-5 (2011).

98. Abrahamsson, T.R. et al. Low diversity of the gut microbiota in infants with atopic eczema. *J. Allergy Clin. Immunol.* **129**, 434-40, 440 e1-2 (2012).

99. Yoshimoto, S. et al. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97-101 (2013).

100. Nakayama, J. et al. Diversity in gut bacterial community of school-age children in Asia. *Sci. Rep.* **5**, 8397 (2015).

101. De Filippo, C. et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad.*

Sci. U.S.A. **107**, 14691-14696 (2010).

102. Obregon-Tito, A.J. et al. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* **6**, 6505 (2015).

103. Martinez, I. et al. The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. *Cell reports* **11**, 527-38 (2015).

104. Goodrich, J.K. et al. Human genetics shape the gut microbiome. *Cell* **159**, 789-99 (2014).

105. Clemente, J.C., Ursell, L.K., Parfrey, L.W. & Knight, R. The impact of the gut microbiota on human health: an integrative view. *Cell* **148**, 1258-70 (2012).

106. Li, J. et al. An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* **32**, 834-841 (2014).

107. Rehman, A. et al. Geographycal patterns of the standing and active human gut microbiome in health and IBD. *Gut* **65**, 238-248 (2015).

108. Wu, G.D. et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105-108 (2011).

109. World Health Organization. Global Health Observatory (GHO) data. (2015).

110. Kim, S.W. et al. Robustness of gut microbiota of healthy adults in response to probiotic intervention revealed by high-throughput pyrosequencing. *DNA Res.* **20**, 241-253 (2013).

111. Morita, H. et al. An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine. *Microbes Environ.* **22**, 214-222 (2007).

112. Ueno, M. et al. Assessment and improvement of methods for microbial DNA preparation from fecal samples. *Handbook of Molecular Microbial Ecology, Volume II: Metagenomics in Different Habitats, First Edition*, 191-198 (2011).

New Jersey: Wiley-well.

113. Gomez-Alvarez, V., Teal, T.K. & Schmidt, T.M. Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* **3**, 1314-1317 (2009).

114. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359 (2012).

115. Treangen, T.J., Sommer, D.D., Angly, F.E., Koren, S. & Pop, M. Next generation sequence assembly with AMOS. *Curr Protoc Bioinformatics* **30**, 11.8-11.8.18 (2011).

116. Noguchi, H., Taniguchi, T. & Itoh, T. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res.* **15**, 387-396 (2008).

117. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-9 (2006).

118. Clemente, J.C. et al. The microbiome of uncontacted Amerindians. *Sci. Adv.* **1**, e1500183 (2015).

119. Atarashi, K. et al. Th17 cell induction by adhesion of microbes to intestinal epithelial cells. *Cell* **163**, 367-380 (2015).

120. Lagesen, K. et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100-3108 (2007).

121. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

122. Mortazave A., Williams B.A., McCue K., Schaeffer L. & Wold B. Mapping and quantifying mammlian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621-8

(2008).

123. Rosen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**, 365-386 (2000).

124. Huerta-Cepas, J. et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286-93 (2016).

125. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **75**, 305-322 (1995).

126. Furuta, Y. & Kobayashi, I. Movement of DNA sequence recognition domains between non-orthologous proteins. *Nucleic Acids Res.* **40**, 9218-32 (2012).

127. Nakamura, N., Lin, H.C., McSweeney, C.S., Mackie, R.I. & Gaskins, H.R. Mechanisms of microbial hydrogen disposal in the human colon and implications for health and disease. *Annual review of food science and technology* **1**, 363-95 (2010).

128. Liu, S. et al. Starch and starch hydrolysates are favorable carbon sources for bifidobacteria in the human gut. *BMC Microbiol.* **15**, 54 (2015).

129. Tedelind, S., Westberg, F., Kjerruf, M. & Vidal, A. Anti-inflammatory properties of the short-chain fatty acids acetate and propionate: A study with relevance to inflammatory bowel disease. *World J. Gastroenterol.* **28**, 2826-2832 (2007).

130. Ohsawa, I. et al. Hydrogen acts as a therapeutic antioxidant by selectively reducing cytotoxic oxygen radicals. *Nat. Med.* **13**, 688-94 (2007).

131. Rey, F.E. et al. Dissecting the in vivo metabolic potential of two human gut acetogens. *J. Biol. Chem.* **285**, 22082-90 (2010).

132. Kim, G. et al. *Methanobrevibacter smithii* is the predominant methanogen in patients with constipation-predominant IBS and methane on breath. *Dig. Dis. Sci.* **57**, 3213-8 (2012).

133. Morii, H., Oda, K., Suenaga, Y. & Nakamura, T. Low Methane Concentration in the Breath of Japanese. *Journal of University of Occupational and Environmenatal Health* **25**, 397-407 (2003).

134. Levitt, M.D., Furne, J.K., Kuskowski, M. & Ruddy, J. Stability of human methanogenic flora over 35 years and a review of insights obtained from breath methane measurements. *Clin. Gastroenterol. Hepatol.* **4**, 123-9 (2006).

135. David, L.A. et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559-563 (2014).

136. Dethlefsen, L., Huse, S., Sogin, M.L. & Relman, D.A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280 (2008).

137. Jakobsson, H.E. et al. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* **5**, e9836 (2010).

138. Dethlefsen, L. & Relman, D. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4554-4561 (2011).

139. Ministry of education, sports, science and technology: Japan. Standard Tables of Food Compositionin Japan: 2010. (2011). (in Japanese).

140. Högberg, L.D., Muller, A., Zorzet, A., Monnet, D.L. & Cars, O. Antibiotic use worldwide. *The Lancet Infectious Diseases* **14**, 1179-1180 (2014).

141. European Surveillance of Antimicrobial Consumption. ESAC Yearbook 2009. (2009).

142. Wang, X., Zeng, L. & Li, Z. Kang sheng su lin chuang ying yong ji he li xing ping jia (Antibiotic clinical application and rationality evaluation). *Ningxia Med. J.* **27**, 416 (2005). (in Chinese).

143. Van Boeckel, T.P. et al. Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5649-5654 (2015).

144. Gibson, M.K., Forsberg, K.J. & Dantas, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* **9**, 207-216 (2015).

145. Finn, R.D., Clements, J. & Eddy, S.R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29-37 (2011).

146. Nielsen, H.B. et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32**, 822-828 (2014).

147. Cox, L.M. et al. Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell* **158**, 705-721 (2014).

148. Cole, J.R. et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141-5 (2009).

149. Panda, S. et al. Short-term effect of antibiotics on human gut microbiota. *PLoS One* **9**, e95476 (2014).

150. Levy, S.B. & Marshall, B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat. Med.* **10**, S122-9 (2004).

151. Goossens, H., Ferech, M., Vander Stichele, R. & Elseviers, M. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *The Lancet* **365**, 579-587 (2005).

152. Blaser, M.J. & Falkow, S. What are the consequences of the disappearing human microbiota? *Nat. Rev. Microbiol.* **7**, 887-894 (2009).

153. Smillie, C.S. et al. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480**, 241-244 (2011).

154. Harrison, E.M. et al. Whole genome sequencing identifies zoonotic transmission of MRSA isolates with the novel *mecA* homologue *mecC*. *EMBO Mol. Med.* **5**, 509-515 (2013).

155. Cho, I. et al. Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* **488**, 621-626 (2012).

156. Hill, D.A. et al. Commensal bacteria-derived signals regulate basophil hematopoiesis and allergic inflammation. *Nat. Med.* **18**, 538-46 (2012).

157. Hviid, A., Svanstrom, H. & Frisch, M. Antibiotic use and inflammatory bowel diseases in childhood. *Gut* **60**, 49-54 (2011).

158. Trasande, L. et al. Infant antibiotic exposures and early-life body mass. *Int. J. Obes. (Lond.)* **37**, 16-23 (2013).

159. Saari, A., Virta, L.J., Sankilampi, U., Dunkel, L. & Saxen, H. Antibiotic exposure in infancy and risk of being overweight in the first 24 months of life. *Pediatrics* **135**, 617-26 (2015).

160. Kozyrskyj, A.L., Ernst, P. & Becker, A.B. Increased risk of childhood asthma from antibiotic use in early life. *Chest* **131**, 1753-9 (2007).

6. Appendix

Appendix 1. Metadata and sequencing statistics of the 106 JP individuals

Subject ID	Age	Sex	BMI	Roche 454		Ion PGM		Illumina MiSeq		Ion Proton		Total reads	Total base pairs
				Reads	Base pairs	Reads	Base pairs	Reads	Base pairs	Reads	Base pairs		
apr01	21	female	18.8	3,616,600	1,511,135,838	1,468,274	311,169,667	17,779,732	2,758,183,854	-	-	22,864,606	4,580,489,359
apr02	23	female	18.6	2,628,746	1,026,436,447	3,606,299	797,289,881	27,347,713	4,051,358,283	-	-	33,582,758	5,875,084,611
apr03	21	female	19.9	2,184,162	784,607,990	5,475,448	1,270,080,643	25,110,532	3,724,556,684	-	-	32,770,142	5,779,245,317
apr05	20	female	23.2	1,223,255	599,941,068	-	-	17,783,033	5,246,127,659	-	-	19,006,288	5,846,068,727
apr06	22	male	22.7	1,129,655	416,913,103	1,639,524	332,816,324	-	-	2,916,108	480,434,648	5,685,287	1,230,164,075
apr07	22	male	20.8	1,168,088	402,651,421	1,799,059	379,201,967	-	-	2,624,583	439,838,921	5,591,730	1,221,692,309
apr08	19	female	22.3	1,233,049	413,431,585	1,219,601	231,746,517	-	-	2,193,773	366,678,069	4,646,423	1,011,856,171
apr09	20	female	18.5	3,065,678	1,314,648,501	5,112,619	1,231,269,412	24,747,208	3,664,126,874	-	-	32,925,505	6,210,044,787
apr10	21	female	18.0	1,182,504	555,221,170	-	-	17,839,740	5,264,780,845	-	-	19,022,244	5,820,002,015
apr11	23	female	19.1	1,083,003	602,939,544	-	-	16,336,733	4,829,913,153	-	-	17,419,736	5,432,852,697
apr12	25	male	23.7	1,717,716	694,917,938	4,375,130	1,061,610,762	26,744,611	3,962,600,100	-	-	32,837,457	5,719,128,800
apr15	22	female	24.4	1,116,609	526,876,648	-	-	15,526,356	4,563,985,308	-	-	16,642,965	5,090,861,956
apr16	20	female	18.0	3,154,885	1,303,790,446	4,780,685	1,151,892,044	24,281,833	3,595,782,961	-	-	32,217,403	6,051,465,451
apr17	20	male	19.4	2,673,276	1,075,288,659	3,591,991	849,547,914	26,844,485	3,974,770,852	-	-	33,109,752	5,899,607,425
apr18	21	male	21.5	1,101,523	579,832,038	-	-	14,979,707	4,429,641,376	-	-	16,081,230	5,009,473,414
apr19	20	male	19.4	913,549	443,724,457	-	-	18,653,403	5,528,566,347	-	-	19,566,952	5,972,290,804
apr21	21	female	19.4	1,032,314	394,480,338	-	-	18,569,721	5,487,716,771	-	-	19,602,035	5,882,197,109
apr22	22	female	19.5	1,244,906	579,326,092	-	-	18,518,105	5,473,333,837	-	-	19,763,011	6,052,659,929
apr23	21	female	20.5	1,337,946	661,645,265	-	-	16,182,750	4,800,728,933	-	-	17,520,696	5,462,374,198
apr24	19	male	25.1	1,182,834	483,346,472	-	-	16,891,735	4,985,795,937	-	-	18,074,569	5,469,142,409
apr25	19	male	18.7	1,176,431	534,267,378	-	-	18,357,333	5,435,814,962	-	-	19,533,764	5,970,082,340
apr26	21	male	22.2	1,198,205	566,296,215	-	-	18,667,816	5,528,105,513	-	-	19,866,021	6,094,401,728
apr27	20	female	22.9	1,044,138	449,345,695	-	-	18,269,627	5,381,759,953	-	-	19,313,765	5,831,105,648
apr28	20	female	24.8	977,520	481,731,086	-	-	17,767,462	5,242,715,348	-	-	18,744,982	5,724,446,434
apr30	21	female	19.2	1,113,198	581,361,486	-	-	17,085,867	5,049,283,198	-	-	18,199,065	5,630,644,684
apr31	33	male	28.0	1,256,095	553,728,889	-	-	17,651,579	5,219,215,525	-	-	18,907,674	5,772,944,414
apr32	19	male	21.8	1,255,099	540,951,434	-	-	18,636,135	5,508,702,888	-	-	19,891,234	6,049,654,322
apr33	23	male	23.8	935,257	541,854,736	-	-	18,597,849	5,505,404,196	-	-	19,533,106	6,047,258,932
apr34	20	female	21.9	1,041,477	679,049,047	-	-	17,913,411	5,260,094,469	-	-	18,954,888	5,939,143,516
apr35	19	female	17.3	865,540	296,790,287	-	-	16,560,787	4,873,049,793	-	-	17,426,327	5,169,840,080
apr36	19	female	19.6	1,081,364	702,479,247	-	-	19,104,747	5,640,395,843	-	-	20,186,111	6,342,875,090
apr37	21	female	21.2	1,298,728	526,349,155	-	-	18,644,441	5,502,003,973	-	-	19,943,169	6,028,353,128
apr38	19	male	22.6	903,832	552,619,941	-	-	19,238,817	5,704,934,670	-	-	20,142,649	6,257,554,611
apr39	23	male	20.0	3,497,728	1,482,279,204	4,463,050	1,047,847,428	23,129,859	3,428,793,868	-	-	31,090,637	5,958,920,500
apr40	19	male	20.1	2,210,816	908,122,778	3,067,139	716,408,531	22,315,744	3,307,903,412	-	-	27,593,699	4,932,434,721
FAKO01	36	male	24.2	1,237,744	869,451,621	4,748,024	1,046,263,996	-	-	-	-	5,985,768	1,915,715,617
FAKO02	47	male	20.5	1,212,255	774,334,821	4,977,020	1,191,843,698	20,601,066	3,171,189,959	-	-	26,790,341	5,137,368,478

Subject ID	Age	Sex	BMI	Roche 454		Ion PGM		Illumina MiSeq		Ion Proton		Total reads	Total base pairs
				Reads	Base pairs	Reads	Base pairs	Reads	Base pairs	Reads	Base pairs		
FAKO03	50	male	21.3	813,005	474,528,076	3,060,941	722,273,852	23,416,757	3,635,198,059	-	-	27,290,703	4,831,999,987
FAKO05	50	male	25.2	1,013,519	772,877,228	4,135,434	983,799,832	18,049,491	2,806,997,676	-	-	23,198,444	4,563,674,736
FAKO06	35	male	25.1	896,724	691,460,477	3,663,174	867,073,373	-	-	-	-	4,559,898	1,558,533,850
FAKO07	37	male	21.8	1,067,246	843,542,653	4,994,918	1,222,374,117	-	-	-	-	6,062,164	2,065,916,770
FAKO08	42	female	21.2	1,278,699	1,004,402,889	3,226,732	815,214,279	23,173,959	3,582,193,756	-	-	27,679,390	5,401,810,924
FAKO09	38	male	21.7	1,230,727	653,861,422	4,357,923	1,067,163,963	-	-	-	-	5,588,650	1,721,025,385
FAKO10	28	female	20.9	1,283,259	961,493,654	4,584,414	1,120,482,459	-	-	-	-	5,867,673	2,081,976,113
FAKO11	39	male	24.2	1,289,524	1,020,526,742	5,142,605	1,147,275,552	-	-	-	-	6,432,129	2,167,802,294
FAKO12	42	male	25.3	1,157,631	977,103,944	3,003,363	700,304,286	-	-	-	-	4,160,994	1,677,408,230
FAKO13	41	male	22.2	1,165,486	962,790,091	5,772,999	1,378,317,392	-	-	-	-	6,938,485	2,341,107,483
FAKO14	39	male	23.6	1,128,975	884,274,843	5,539,788	1,295,742,281	-	-	-	-	6,668,763	2,180,017,124
FAKO15	48	female	19.5	1,216,894	905,201,847	2,694,894	624,040,817	23,536,367	3,631,105,909	-	-	27,448,155	5,160,348,573
FAKO16	40	male	22.1	1,070,336	864,313,128	5,251,067	1,246,229,222	-	-	-	-	6,321,403	2,110,542,350
FAKO17	39	female	22.6	991,117	775,552,908	5,475,359	1,294,615,850	-	-	-	-	6,466,476	2,070,168,758
FAKO18	33	female	17.7	1,204,743	875,237,342	4,193,796	998,072,479	-	-	-	-	5,398,539	1,873,309,821
FAKO19	48	female	21.4	1,122,447	543,290,354	4,373,108	1,042,114,670	20,386,865	3,154,605,955	-	-	25,882,420	4,740,010,979
FAKO21	42	male	21.1	1,177,223	941,278,809	4,444,365	1,151,398,095	-	-	-	-	5,621,588	2,092,676,904
FAKO22	50	male	25.5	976,949	790,243,385	2,877,561	678,295,459	16,251,627	2,498,181,407	-	-	20,106,137	3,966,720,251
FAKO23	45	male	21.9	1,220,750	1,012,362,085	4,739,884	1,204,939,249	21,753,613	3,343,822,908	-	-	27,714,247	5,561,124,242
FAKO24	35	male	25.0	973,738	798,939,235	2,286,503	536,921,141	-	-	-	-	3,260,241	1,335,860,376
FAKO25	41	male	22.0	1,144,317	915,839,466	3,484,222	841,216,134	-	-	-	-	4,628,539	1,757,055,600
FAKO26	34	male	23.3	1,103,457	800,679,479	3,562,065	869,101,657	-	-	-	-	4,665,522	1,669,781,136
FAKO27	50	male	20.9	1,116,983	876,718,876	6,178,406	1,358,286,547	25,430,117	3,958,248,462	-	-	32,725,506	6,193,253,885
FAKO28	35	male	23.0	1,178,034	842,873,775	4,749,838	1,192,897,406	-	-	-	-	5,927,872	2,035,771,181
FAKO29	46	male	30.1	989,403	559,118,324	3,663,125	887,626,961	25,101,385	3,907,335,310	-	-	29,753,913	5,354,080,595
FAKO30	31	male	23.1	1,110,845	825,752,152	4,614,918	998,457,552	-	-	-	-	5,725,763	1,824,209,704
FBAN01	41	male	22.0	878,733	717,225,682	4,512,715	1,156,400,900	-	-	-	-	5,391,448	1,873,626,582
FBAN02	36	male	25.6	1,270,826	809,732,938	4,752,649	1,186,547,726	-	-	6,767,151	1,161,937,321	12,790,626	3,158,217,985
FBAN04	31	male	20.5	572,667	462,253,834	4,587,752	1,048,495,854	-	-	-	-	5,160,419	1,510,749,688
FBAN05	38	male	21.5	843,706	633,605,730	4,474,844	999,852,346	-	-	-	-	5,318,550	1,633,458,076
FBAN06	33	male	24.2	1,263,073	964,216,429	4,519,180	1,151,424,172	-	-	-	-	5,782,253	2,115,640,601
FBAN07	31	male	24.6	979,165	760,821,899	3,592,110	899,974,202	-	-	6,584,827	1,085,930,759	11,156,102	2,746,726,860
FBAN08	29	male	22.2	1,001,583	441,929,434	4,419,297	1,084,654,590	-	-	-	-	5,420,880	1,526,584,024
FBAN09	28	male	26.5	896,341	724,591,809	4,711,730	1,150,786,567	-	-	-	-	5,608,071	1,875,378,376
FBAN10	28	male	21.4	1,122,997	816,962,343	4,378,734	1,126,727,661	-	-	-	-	5,501,731	1,943,690,004
FMOR01	23	female	20.7	1,039,730	776,147,874	4,234,927	1,028,449,248	-	-	-	-	5,274,657	1,804,597,122
FMOR02	22	female	21.5	1,280,287	1,053,221,731	4,392,453	1,127,834,787	-	-	-	-	5,672,740	2,181,056,518
FMOR03	23	male	21.4	1,003,142	818,978,759	3,190,182	752,074,790	-	-	-	-	4,193,324	1,571,053,549
FMOR04	22	male	20.5	639,688	528,052,776	4,056,646	977,374,544	-	-	-	-	4,696,334	1,505,427,320
FMOR11	22	female	17.7	1,297,180	972,570,359	4,789,473	1,178,775,252	-	-	6,705,707	1,137,599,554	12,792,360	3,288,945,165
FMOR14	22	male	20.5	1,199,711	939,575,474	4,644,847	1,150,837,494	-	-	-	-	5,844,558	2,090,412,968
FMOR21	49	male	21.0	1,049,399	806,315,831	3,799,519	874,085,087	-	-	-	-	4,848,918	1,680,400,918
FPR01	40	female	22.3	-	-	-	-	-	-	11,750,594	1,956,925,399	11,750,594	1,956,925,399

Subject ID	Age	Sex	BMI	Roche 454		Ion PGM		Illumina MiSeq		Ion Proton		Total reads	Total base pairs
				Reads	Base pairs	Reads	Base pairs	Reads	Base pairs	Reads	Base pairs		
FPR03	25	male	22.5	-	-	-	-	-	-	6,711,654	1,101,100,297	6,711,654	1,101,100,297
FPR04	35	male	19.7	-	-	-	-	-	-	7,276,806	1,069,850,871	7,276,806	1,069,850,871
FPR05	55	male	27.6	-	-	-	-	-	-	9,039,035	1,426,959,232	9,039,035	1,426,959,232
FTAG01	54	male	22.4	1,244,303	893,078,531	4,646,145	1,106,652,891	22,511,838	3,513,074,899	-	-	28,402,286	5,512,806,321
FTAG02	39	female	22.9	1,017,091	851,572,311	4,749,277	1,216,521,098	-	-	-	-	5,766,368	2,068,093,409
FTAG03	34	female	18.7	1,197,243	893,910,927	5,758,559	1,371,519,499	-	-	3,671,046	567,203,425	10,626,848	2,832,633,851
FTAG06	41	male	21.2	1,303,059	978,525,985	3,769,323	869,067,092	-	-	-	-	5,072,382	1,847,593,077
FTAG07	27	male	25.0	550,482	333,209,786	4,478,255	1,157,372,262	-	-	-	-	5,028,737	1,490,582,048
FTAG08	29	male	21.3	1,119,122	844,573,882	4,348,464	1,045,592,291	-	-	-	-	5,467,586	1,890,166,173
FTAG09	34	male	21.1	987,649	761,335,728	4,556,858	1,015,283,188	-	-	3,676,699	611,041,699	9,221,206	2,387,660,615
FTAG10	30	male	23.6	866,651	310,698,085	3,967,579	949,151,507	-	-	-	-	4,834,230	1,259,849,592
FTAG12	37	female	20.6	1,061,262	781,617,041	4,650,110	1,076,978,044	-	-	-	-	5,711,372	1,858,595,085
FTAG13	25	female	18.9	1,088,397	863,477,979	3,839,673	888,047,924	-	-	-	-	4,928,070	1,751,525,903
FTAG14	39	female	20.0	910,757	586,172,905	2,709,444	584,238,167	-	-	-	-	3,620,201	1,170,411,072
FTAG15	37	male	22.8	1,279,507	1,052,657,242	4,610,008	1,045,774,508	-	-	-	-	5,889,515	2,098,431,750
FTAG16	34	female	20.0	1,119,929	860,800,299	3,573,257	862,260,462	-	-	5,882,855	982,877,558	10,576,041	2,705,938,319
FTAG17	26	male	36.1	1,015,453	564,675,149	4,068,162	1,030,784,818	-	-	-	-	5,083,615	1,595,459,967
FTAG18	36	female	20.5	949,169	638,960,066	4,467,158	1,062,822,389	-	-	-	-	5,416,327	1,701,782,455
FTAG19	33	female	21.3	840,282	626,776,404	5,014,468	1,184,085,290	-	-	-	-	5,854,750	1,810,861,694
FTAG20	40	female	18.9	892,706	704,667,477	3,801,134	851,601,009	-	-	6,459,265	1,094,741,479	11,153,105	2,651,009,965
FTAG21	31	female	18.8	1,287,240	997,810,879	4,789,124	1,221,964,810	-	-	6,858,366	1,153,821,888	12,934,730	3,373,597,577
TS-11	60	male	24.8	-	-	-	-	-	-	8,836,531	1,549,259,449	8,836,531	1,549,259,449
TS-21	46	male	22.1	-	-	-	-	-	-	7,973,204	1,330,712,137	7,973,204	1,330,712,137
TS-29	48	female	18.7	-	-	-	-	-	-	7,752,836	1,374,864,110	7,752,836	1,374,864,110
TS-33	55	female	19.9	-	-	-	-	-	-	9,041,527	1,644,215,789	9,041,527	1,644,215,789
TS-41	48	male	23.0	-	-	-	-	-	-	7,996,333	1,396,681,648	7,996,333	1,396,681,648

Appendix 2. Reference genomes used in this study

a. Reference genomes collected from NCBI

Genome ID							
AABF00000000	AABM00000000	AAGP00000000	AAIQ00000000	AAJM00000000	AAJO00000000	AAJW00000000	AAKL00000000
AALD00000000	AALF00000000	AAMN00000000	AAMW00000000	AAMX00000000	AAMY00000000	AAMZ00000000	AANP00000000
AANX00000000	AAOB00000000	AAOJ00000000	AAOX00000000	AAPD00000000	AAPJ00000000	AAPS00000000	AAPZ00000000
AAQF00000000	AAQG00000000	AAQH00000000	AAQJ00000000	AARF00000000	AARG00000000	AASA00000000	AASB00000000
AAUA00000000	AAUJ00000000	AAUV00000000	AAVL00000000	AAVM00000000	AAVN00000000	AAVO00000000	AAVP00000000
AAVW00000000	AAWL00000000	AAXA00000000	AAXB00000000	AAXD00000000	AAXE00000000	AAXF00000000	AAXG00000000
AAXS00000000	AAXV00000000	AAXX00000000	AAYG00000000	AAYH00000000	AAYI00000000	AAYW00000000	AAZD00000000
AAZE00000000	AAZF00000000	AAZG00000000	AAZH00000000	AAZI00000000	AAZJ00000000	AAZP00000000	AAZZ00000000
ABAA00000000	ABAW00000000	ABAX00000000	ABBD00000000	ABBE00000000	ABBF00000000	ABBG00000000	ABBH00000000
ABBM00000000	ABBV00000000	ABBX00000000	ABBZ00000000	ABCA00000000	ABCB00000000	ABCC00000000	ABCF00000000
ABCM00000000	ABCO00000000	ABCQ00000000	ABCS00000000	ABDT00000000	ABDU00000000	ABDV00000000	ABDW00000000
ABDX00000000	ABDY00000000	ABED00000000	ABEE00000000	ABEQ00000000	ABEY00000000	ABEZ00000000	ABFK00000000
ABFX00000000	ABFY00000000	ABFZ00000000	ABGC00000000	ABGD00000000	ABHC00000000	ABHD00000000	ABHH00000000
ABID00000000	ABIK00000000	ABIL00000000	ABIX00000000	ABIY00000000	ABJD00000000	ABJK00000000	ABJL00000000
ABJO00000000	ABJP00000000	ABKL00000000	ABKW00000000	ABKX00000000	ABLC00000000	ABLK00000000	ABOO00000000
ABOT00000000	ABOU00000000	ABQA00000000	ABQC00000000	ABQK00000000	ABQL00000000	ABQM00000000	ABQN00000000
ABQP00000000	ABQQ00000000	ABQR00000000	ABQS00000000	ABQT00000000	ABQU00000000	ABRA00000000	ABRB00000000
ABRC00000000	ABRU00000000	ABSM00000000	ABVG00000000	ABVH00000000	ABVK00000000	ABVO00000000	ABVP00000000
ABVQ00000000	ABVR00000000	ABVX00000000	ABVY00000000	ABWG00000000	ABWH00000000	ABWJ00000000	ABWK00000000
ABWL00000000	ABWM00000000	ABWN00000000	ABWO00000000	ABWP00000000	ABWV00000000	ABWW00000000	ABWZ00000000
ABXA00000000	ABXB00000000	ABXG00000000	ABXH00000000	ABXI00000000	ABXJ00000000	ABXP00000000	ABXU00000000
ABXV00000000	ABXW00000000	ABXX00000000	ABXY00000000	ABYH00000000	ABYI00000000	ABYJ00000000	ABYK00000000
ABYL00000000	ABYN00000000	ABYO00000000	ABYP00000000	ABYQ00000000	ABYS00000000	ABYT00000000	ABYU00000000
ABYV00000000	ABYW00000000	ABZU00000000	ABZV00000000	ABZX00000000	ABZY00000000	ABZZ00000000	ACAA00000000
ACAB00000000	ACAC00000000	ACAH00000000	ACAI00000000	ACAJ00000000	ACAL00000000	ACAS00000000	ACBL00000000
ACBW00000000	ACBX00000000	ACBY00000000	ACBZ00000000	ACCC00000000	ACCF00000000	ACCG00000000	ACCH00000000
ACCI00000000	ACCJ00000000	ACCK00000000	ACCL00000000	ACCR00000000	ACDB00000000	ACDC00000000	ACDD00000000
ACDE00000000	ACDG00000000	ACDH00000000	ACDI00000000	ACDJ00000000	ACDK00000000	ACDL00000000	ACDM00000000
ACDN00000000	ACDO00000000	ACDP00000000	ACDQ00000000	ACDR00000000	ACDS00000000	ACDT00000000	ACDX00000000
ACDY00000000	ACDZ00000000	ACEA00000000	ACEB00000000	ACEC00000000	ACEN00000000	ACEO00000000	ACEP00000000
ACEQ00000000	ACET00000000	ACFB00000000	ACFD00000000	ACFE00000000	ACFG00000000	ACFH00000000	ACFR00000000
ACFT00000000	ACFU00000000	ACFX00000000	ACFY00000000	ACGA00000000	ACGB00000000	ACGC00000000	ACGD00000000
ACGE00000000	ACGG00000000	ACGH00000000	ACGI00000000	ACGK00000000	ACGO00000000	ACGP00000000	ACGQ00000000
ACGR00000000	ACGS00000000	ACGT00000000	ACGU00000000	ACGV00000000	ACGX00000000	ACGY00000000	ACGZ00000000
ACHA00000000	ACHB00000000	ACHD00000000	ACHE00000000	ACHF00000000	ACHG00000000	ACHI00000000	ACHJ00000000
ACHM00000000	ACHN00000000	ACIB00000000	ACIC00000000	ACID00000000	ACIE00000000	ACIF00000000	ACIH00000000
ACII00000000	ACIJ00000000	ACIK00000000	ACIL00000000	ACIM00000000	ACIN00000000	ACIO00000000	ACIP00000000
ACIQ00000000	ACIZ00000000	ACJB00000000	ACJM00000000	ACJW00000000	ACJX00000000	ACJY00000000	ACKO00000000
ACKP00000000	ACKQ00000000	ACKR00000000	ACKS00000000	ACKT00000000	ACKU00000000	ACKV00000000	ACKW00000000
ACKX00000000	ACKY00000000	ACKZ00000000	ACLA00000000	ACLB00000000	ACLE00000000	ACLF00000000	ACLH00000000
ACLI00000000	ACLJ00000000	ACLK00000000	ACLL00000000	ACLM00000000	ACLN00000000	ACLO00000000	ACLP00000000
ACLQ00000000	ACLR00000000	ACLU00000000	ACLY00000000	ACMH00000000	ACMI00000000	ACML00000000	ACNI00000000
ACNJ00000000	ACNN00000000	ACNO00000000	ACNV00000000	ACOB00000000	ACOF00000000	ACOG00000000	ACOH00000000
ACOI00000000	ACOK00000000	ACOM00000000	ACON00000000	ACOO00000000	ACOP00000000	ACOY00000000	ACOZ00000000
ACPK00000000	ACPL00000000	ACPM00000000	ACPN00000000	ACPO00000000	ACPP00000000	ACPQ00000000	ACPR00000000
ACPS00000000	ACPT00000000	ACPV00000000	ACPW00000000	ACPZ00000000	ACQC00000000	ACQD00000000	ACQE00000000
ACQF00000000	ACQG00000000	ACQH00000000	ACQL00000000	ACQN00000000	ACQV00000000	ACRB00000000	ACRC00000000
ACRE00000000	ACRF00000000	ACRG00000000	ACRH00000000	ACRI00000000	ACRJ00000000	ACRK00000000	ACRL00000000
ACRM00000000	ACRN00000000	ACRO00000000	ACRP00000000	ACRQ00000000	ACRS00000000	ACRT00000000	ACRX00000000
ACRY00000000	ACSB00000000	ACSE00000000	ACSF00000000	ACSH00000000	ACSK00000000	ACSL00000000	ACSM00000000
ACTB00000000	ACTC00000000	ACTG00000000	ACTI00000000	ACTJ00000000	ACTK00000000	ACTL00000000	ACTM00000000
ACTN00000000	ACTO00000000	ACTP00000000	ACTR00000000	ACTS00000000	ACTT00000000	ACTV00000000	ACTW00000000
ACTX00000000	ACUA00000000	ACUD00000000	ACUE00000000	ACUF00000000	ACUH00000000	ACUO00000000	ACUX00000000
ACUY00000000	ACUZ00000000	ACVA00000000	ACVB00000000	ACVE00000000	ACVI00000000	ACVN00000000	ACVO00000000
ACVP00000000	ACVQ00000000	ACVR00000000	ACVX00000000	ACWB00000000	ACWC00000000	ACWD00000000	ACWE00000000
ACWF00000000	ACWG00000000	ACWH00000000	ACWI00000000	ACWJ00000000	ACWK00000000	ACWL00000000	ACWM00000000
ACWN00000000	ACWO00000000	ACWP00000000	ACWQ00000000	ACWR00000000	ACWS00000000	ACWU00000000	ACWW00000000
ACWX00000000	ACWY00000000	ACXA00000000	ACXB00000000	ACXD00000000	ACXF00000000	ACXG00000000	ACXP00000000
ACXU00000000	ACXX00000000	ACXY00000000	ACYG00000000	ACYH00000000	ACYI00000000	ACYT00000000	ACYW00000000
ACYY00000000	ACZI00000000	ACZJ00000000	ACZK00000000	ACZL00000000	ACZM00000000	ACZR00000000	ACZS00000000
ACZW00000000	ADAD00000000	ADAM00000000	ADAN00000000	ADAV00000000	ADAX00000000	ADBA00000000	ADBE00000000
ADBF00000000	ADBG00000000	ADBR00000000	ADBT00000000	ADBY00000000	ADCD00000000	ADCH00000000	ADCJ00000000
ADCK00000000	ADCL00000000	ADCM00000000	ADCN00000000	ADCO00000000	ADCP00000000	ADCQ00000000	ADCR00000000
ADCS00000000	ADCT00000000	ADCU00000000	ADCV00000000	ADCW00000000	ADCX00000000	ADCY00000000	ADCZ00000000
ADDA00000000	ADDB00000000	ADDO00000000	ADDR00000000	ADDS00000000	ADDT00000000	ADDU00000000	ADDV00000000
ADDW00000000	ADDX00000000	ADDY00000000	ADEA00000000	ADEB00000000	ADEC00000000	ADEE00000000	ADEF00000000
ADEG00000000	ADEJ00000000	ADEK00000000	ADEL00000000	ADEM00000000	ADEN00000000	ADEO00000000	ADEP00000000
ADEQ00000000	ADER00000000	ADES00000000	ADET00000000	ADEU00000000	ADEV00000000	ADEW00000000	ADEY00000000
ADFC00000000	ADFO00000000	ADFP00000000	ADFQ00000000	ADFR00000000	ADFT00000000	ADFU00000000	ADFW00000000
ADGG00000000	ADGH00000000	ADGI00000000	ADGJ00000000	ADGK00000000	ADGP00000000	ADGQ00000000	ADGR00000000
ADHG00000000	ADHJ00000000	ADHN00000000	ADJL00000000	ADJN00000000	ADKM00000000	ADKO00000000	ADKP00000000
ADKX00000000	ADLB00000000	ADLD00000000	ADLE00000000	ADLF00000000	ADLG00000000	ADLJ00000000	ADLL00000000
ADLN00000000	ADLP00000000	ADLQ00000000	ADLR00000000	ADLS00000000	ADLT00000000	ADLU00000000	ADLV00000000
ADLW00000000	ADLY00000000	ADLZ00000000	ADMB00000000	ADMC00000000	ADMD00000000	ADME00000000	ADMF00000000
ADMG00000000	ADML00000000	ADMM00000000	ADMN00000000	ADMO00000000	ADMP00000000	ADMS00000000	ADMT00000000

Genome ID							
ADMV00000000	ADNA00000000	ADNS00000000	ADNT00000000	ADNU00000000	ADNV00000000	ADNW00000000	ADNX00000000
ADNY00000000	ADOM00000000	ADOP00000000	ADTK00000000	ADTL00000000	ADTN00000000	ADTP00000000	ADTR00000000
ADTV00000000	ADTZ00000000	ADUH00000000	ADUM00000000	ADUP00000000	ADVK00000000	ADVL00000000	ADVQ00000000
ADVR00000000	ADWO00000000	ADWW00000000	ADWZ00000000	ADXF00000000	ADXH00000000	ADXI00000000	ADXJ00000000
ADXM00000000	ADXN00000000	ADYQ00000000	ADYZ00000000	ADZU00000000	ADZV00000000	AEAA00000000	AEAT00000000
AEAW00000000	AEBA00000000	AEBG00000000	AEBI00000000	AEBU00000000	AEBV00000000	AEBY00000000	AECE00000000
AECM00000000	AECN00000000	AECQ00000000	AECS00000000	AECT00000000	AECU00000000	AECV00000000	AECW00000000
AECY00000000	AEDB00000000	AEDD00000000	AEDL00000000	AEDO00000000	AEDP00000000	AEDQ00000000	AEDR00000000
AEDS00000000	AEDT00000000	AEDU00000000	AEDV00000000	AEDW00000000	AEDX00000000	AEDY00000000	AEDZ00000000
AEEC00000000	AEED00000000	AEEE00000000	AEEF00000000	AEEG00000000	AEEH00000000	AEEI00000000	AEEJ00000000
AEEL00000000	AEEM00000000	AEEN00000000	AEEP00000000	AEEQ00000000	AEES00000000	AEET00000000	AEFH00000000
AEFL00000000	AEFQ00000000	AEGG00000000	AEGH00000000	AEGQ00000000	AEGV00000000	AEHJ00000000	AEHN00000000
AEHO00000000	AEHP00000000	AEHQ00000000	AEHR00000000	AEHT00000000	AEHW00000000	AEHX00000000	AEHY00000000
AEIA00000000	AEIH00000000	AEIQ00000000	AEJB00000000	AEJC00000000	AEJD00000000	AEJE00000000	AEJP00000000
AEJU00000000	AEJV00000000	AEJW00000000	AEJX00000000	AEJY00000000	AEJZ00000000	AEKA00000000	AEKH00000000
AEKI00000000	AEKJ00000000	AEKK00000000	AEKL00000000	AEKM00000000	AEKN00000000	AEKO00000000	AEKT00000000
AEKU00000000	AELK00000000	AELL00000000	AELM00000000	AELN00000000	AELO00000000	AELP00000000	AELQ00000000
AELR00000000	AELS00000000	AELT00000000	AELU00000000	AELV00000000	AELW00000000	AELX00000000	AELY00000000
AELZ00000000	AEMA00000000	AEMB00000000	AEMC00000000	AEMD00000000	AEME00000000	AEMF00000000	AEMI00000000
AEMJ00000000	AENH00000000	AENJ00000000	AENN00000000	AENO00000000	AENP00000000	AENQ00000000	AENS00000000
AENT00000000	AENU00000000	AENV00000000	AENW00000000	AENX00000000	AENY00000000	AENZ00000000	AEOA00000000
AEOF00000000	AEOH00000000	AEON00000000	AEOP00000000	AEOQ00000000	AEOR00000000	AEOS00000000	AEOU00000000
AEPB00000000	AEPC00000000	AEPD00000000	AEPE00000000	AEPF00000000	AEPG00000000	AEPH00000000	AEPM00000000
AEPO00000000	AEPP00000000	AEPS00000000	AEPT00000000	AEPU00000000	AEPV00000000	AEPW00000000	AEPX00000000
AEPY00000000	AEPZ00000000	AEQA00000000	AEQN00000000	AEQO00000000	AEQP00000000	AEQQ00000000	AEQR00000000
AERC00000000	AERE00000000	AERF00000000	AERG00000000	AERH00000000	AERI00000000	AERJ00000000	AERK00000000
AERU00000000	AEUD00000000	AEUE00000000	AEUH00000000	AEUN00000000	AEUU00000000	AEUY00000000	AEUZ00000000
AEVB00000000	AEVC00000000	AEVD00000000	AEVE00000000	AEVF00000000	AEVG00000000	AEVH00000000	AEVI00000000
AEVN00000000	AEVO00000000	AEVP00000000	AEVY00000000	AEWB00000000	AEWG00000000	AEWI00000000	AEWJ00000000
AEWT00000000	AEWU00000000	AEWV00000000	AEWX00000000	AEWY00000000	AEWZ00000000	AEXA00000000	AEXB00000000
AEXF00000000	AEXG00000000	AEXI00000000	AEXJ00000000	AEXK00000000	AEXM00000000	AEXN00000000	AEXO00000000
AEXP00000000	AEXQ00000000	AEXR00000000	AEXS00000000	AEXU00000000	AEXV00000000	AEXW00000000	AEXX00000000
AEXY00000000	AEXZ00000000	AEYC00000000	AEYL00000000	AEYM00000000	AEYS00000000	AEYT00000000	AEYY00000000
AEZI00000000	AEZJ00000000	AEZL00000000	AEZR00000000	AEZX00000000	AEZZ00000000	AFAM00000000	AFAY00000000
AFAZ00000000	AFBA00000000	AFBB00000000	AFBC00000000	AFBD00000000	AFBE00000000	AFBF00000000	AFBG00000000
AFBH00000000	AFBL00000000	AFBM00000000	AFBN00000000	AFBO00000000	AFBP00000000	AFBQ00000000	AFBR00000000
AFCC00000000	AFCE00000000	AFCH00000000	AFDH00000000	AFDI00000000	AFDP00000000	AFDW00000000	AFDX00000000
AFDZ00000000	AFEC00000000	AFER00000000	AFES00000000	AFFL00000000	AFFM00000000	AFFN00000000	AFFO00000000
AFFP00000000	AFFW00000000	AFFY00000000	AFGA00000000	AFGF00000000	AFGG00000000	AFHF00000000	AFHH00000000
AFHI00000000	AFHQ00000000	AFHR00000000	AFHS00000000	AFHT00000000	AFID00000000	AFIF00000000	AFIH00000000
AFII00000000	AFIJ00000000	AFIK00000000	AFIL00000000	AFIM00000000	AFIN00000000	AFJE00000000	AFMN00000000
AFMO00000000	AFMP00000000	AFNE00000000	AFNK00000000	AFNM00000000	AFNN00000000	AFNU00000000	AFOF00000000
AFOI00000000	AFOJ00000000	AFOX00000000	AFOY00000000	AFPO00000000	AFPR00000000	AFPT00000000	AFPW00000000
AFPX00000000	AFPY00000000	AFPZ00000000	AFQA00000000	AFQB00000000	AFQC00000000	AFQD00000000	AFQE00000000
AFQI00000000	AFQJ00000000	AFQN00000000	AFQO00000000	AFQP00000000	AFQQ00000000	AFQR00000000	AFQT00000000
AFQU00000000	AFQV00000000	AFQY00000000	AFRP00000000	AFRQ00000000	AFRW00000000	AFSF00000000	AFSL00000000
AFSU00000000	AFTH00000000	AFTL00000000	AFTX00000000	AFUB00000000	AFUC00000000	AFUD00000000	AFUE00000000
AFUF00000000	AFUG00000000	AFUH00000000	AFUI00000000	AFUJ00000000	AFUK00000000	AFUL00000000	AFUN00000000
AFUO00000000	AFUP00000000	AFUQ00000000	AFUR00000000	AFUS00000000	AFUT00000000	AFUU00000000	AFUV00000000
AFVQ00000000	AFVV00000000	AFWL00000000	AFWM00000000	AFWQ00000000	AFWR00000000	AFWT00000000	AFWX00000000
AFWZ00000000	AFXD00000000	AFXI00000000	AFXN00000000	AFXO00000000	AFXP00000000	AFXR00000000	AFXS00000000
AFXT00000000	AFXU00000000	AFXV00000000	AFXW00000000	AFXX00000000	AFYA00000000	AFYD00000000	AFYE00000000
AFYK00000000	AFYL00000000	AFYM00000000	AFYN00000000	AFYO00000000	AFYP00000000	AFYQ00000000	AFYR00000000
AFYS00000000	AFYT00000000	AFYU00000000	AFZC00000000	AFZD00000000	AFZE00000000	AFZF00000000	AFZG00000000
AFZX00000000	AFZY00000000	AFZZ00000000	AGAB00000000	AGAD00000000	AGAE00000000	AGAF00000000	AGAG00000000
AGAH00000000	AGAY00000000	AGAZ00000000	AGBA00000000	AGBB00000000	AGBD00000000	AGBU00000000	AGBV00000000
AGBX00000000	AGBY00000000	AGBZ00000000	AGCF00000000	AGCI00000000	AGCJ00000000	AGCK00000000	AGCL00000000
AGCM00000000	AGCN00000000	AGCQ00000000	AGCS00000000	AGDG00000000	AGDI00000000	AGDJ00000000	AGDL00000000
AGDM00000000	AGDP00000000	AGDQ00000000	AGDR00000000	AGDS00000000	AGDT00000000	AGDU00000000	AGDV00000000
AGDW00000000	AGDX00000000	AGDY00000000	AGDZ00000000	AGEA00000000	AGEB00000000	AGEC00000000	AGED00000000
AGEE00000000	AGEF00000000	AGEG00000000	AGEH00000000	AGEI00000000	AGEJ00000000	AGEK00000000	AGEL00000000
AGEM00000000	AGEN00000000	AGEO00000000	AGEQ00000000	AGER00000000	AGEU00000000	AGEV00000000	AGEW00000000
AGEX00000000	AGEY00000000	AGEZ00000000	AGFF00000000	AGFH00000000	AGFI00000000	AGFM00000000	AGFN00000000
AGFO00000000	AGFU00000000	AGFW00000000	AGFX00000000	AGGB00000000	AGGD00000000	AGGE00000000	AGGG00000000
AGGJ00000000	AGGL00000000	AGGM00000000	AGGO00000000	AGGQ00000000	AGGT00000000	AGGU00000000	AGGW00000000
AGGY00000000	AGHA00000000	AGHD00000000	AGHE00000000	AGHF00000000	AGHG00000000	AGHH00000000	AGHI00000000
AGHK00000000	AGHL00000000	AGHM00000000	AGHP00000000	AGHQ00000000	AGHR00000000	AGHS00000000	AGHT00000000
AGHU00000000	AGHW00000000	AGHX00000000	AGIP00000000	AGJI00000000	AGJJ00000000	AGKB00000000	AGKC00000000
AGKF00000000	AGKX00000000	AGLG00000000	AGMX00000000	AGNV00000000	AGNZ00000000	AGOA00000000	AGOE00000000
AGOL00000000	AGOT00000000	AGPG00000000	AGPH00000000	AGPJ00000000	AGQF00000000	AGQG00000000	AGQQ00000000
AGQV00000000	AGQX00000000	AGQY00000000	AGQZ00000000	AGRB00000000	AGRI00000000	AGRJ00000000	AGRK00000000
AGRL00000000	AGRU00000000	AGSO00000000	AGTD00000000	AGTS00000000	AGUG00000000	AGVO00000000	AGVP00000000
AGVT00000000	AGVU00000000	AGVV00000000	AGVW00000000	AGVZ00000000	AGWI00000000	AGWJ00000000	AGWK00000000
AGWL00000000	AGWM00000000	AGWN00000000	AGWP00000000	AGWQ00000000	AGWR00000000	AGWT00000000	AGWU00000000
AGWV00000000	AGWW00000000	AGWX00000000	AGWY00000000	AGWZ00000000	AGXA00000000	AGXC00000000	AGXE00000000
AGXF00000000	AGXG00000000	AGXH00000000	AGXI00000000	AGXJ00000000	AGXK00000000	AGXM00000000	AGXO00000000
AGXQ00000000	AGXR00000000	AGXS00000000	AGXT00000000	AGXU00000000	AGXV00000000	AGXW00000000	AGXX00000000
AGXY00000000	AGXZ00000000	AGYA00000000	AGYB00000000	AGYD00000000	AGYF00000000	AGYG00000000	AGYH00000000
AGYI00000000	AGYJ00000000	AGYK00000000	AGYL00000000	AGYM00000000	AGYN00000000	AGYO00000000	AGYP00000000
AGYQ00000000	AGYR00000000	AGYS00000000	AGYT00000000	AGYU00000000	AGYV00000000	AGYX00000000	AGYY00000000

Genome ID							
AGZB00000000	AGZC00000000	AGZD00000000	AGZE00000000	AGZF00000000	AGZG00000000	AGZI00000000	AGZJ00000000
AGZK00000000	AGZL00000000	AGZM00000000	AGZN00000000	AGZO00000000	AGZP00000000	AGZQ00000000	AGZR00000000
AGZS00000000	AGZT00000000	AGZU00000000	AGZW00000000	AGZX00000000	AGZY00000000	AGZZ00000000	AHAA00000000
AHAB00000000	AHAC00000000	AHAD00000000	AHAE00000000	AHAF00000000	AHAZ00000000	AHBA00000000	AHBC00000000
AHBD00000000	AHBM00000000	AHBQ00000000	AHBV00000000	AHBW00000000	AHCA00000000	AHCB00000000	AHCD00000000
AHCF00000000	AHDB00000000	AHEM00000000	AHFB00000000	AHFH00000000	AHGS00000000	AHGW00000000	AHHG00000000
AHHJ00000000	AHHK00000000	AHIE00000000	AHIL00000000	AHIP00000000	AHIR00000000	AHIS00000000	AHIT00000000
AHIX00000000	AHIY00000000	AHJE00000000	AHKB00000000	AHKF00000000	AHKH00000000	AHKL00000000	AHKO00000000
AHKT00000000	AHKZ00000000	AHLK00000000	AHLV00000000	AHLW00000000	AHLZ00000000	AHMB00000000	AHPN00000000
AHPP00000000	AHRC00000000	AHRK00000000	AHRM00000000	AHRT00000000	AHRY00000000	AHSB00000000	AHSL00000000
AHSN00000000	AHSR00000000	AHSU00000000	AHTA00000000	AHTF00000000	AHWA00000000	AHWF00000000	AHWO00000000
AHWP00000000	AHXC00000000	AHXD00000000	AHXJ00000000	AHXS00000000	AHYH00000000	AHYJ00000000	AHYM00000000
AHYN00000000	AHYR00000000	AHYS00000000	AHYT00000000	AHYU00000000	AHYV00000000	AHYW00000000	AHYX00000000
AHYZ00000000	AHZA00000000	AHZV00000000	AIAI00000000	AIAM00000000	AIAR00000000	AIAV00000000	AIBF00000000
AIBH00000000	AIBM00000000	AICA00000000	AICL00000000	AICN00000000	AICP00000000	AICQ00000000	AICR00000000
AICS00000000	AICT00000000	AICU00000000	AIDW00000000	AIDX00000000	AIEA00000000	AIED00000000	AIEF00000000
AIEI00000000	AIEK00000000	AIEM00000000	AIEW00000000	AIFK00000000	AIFV00000000	AIGJ00000000	AIGK00000000
AIGX00000000	AIHM00000000	AIII00000000	AIIM00000000	AIJE00000000	AIJQ00000000	AIJV00000000	AIJY00000000
AIKI00000000	AIKP00000000	AIKQ00000000	AIKS00000000	AIKT00000000	AILB00000000	AILF00000000	AILI00000000
AILJ00000000	AILO00000000	AILP00000000	AILX00000000	AILY00000000	AIPP00000000	AISD00000000	AITA00000000
AITZ00000000	AIUV00000000	AIVD00000000	AIVF00000000	AJAB00000000	AJAC00000000	AJAH00000000	AJAI00000000
AJAK00000000	AJAL00000000	AJAM00000000	AJAN00000000	AJAP00000000	AJAQ00000000	AJBD00000000	AJBJ00000000
AJDG00000000	AJDL00000000	AJDQ00000000	AJEV00000000	AJFB00000000	AJFF00000000	AJFI00000000	AJFW00000000
AJGB00000000	AJGH00000000	AJGQ00000000	AJGR00000000	AJGS00000000	AJGT00000000	AJHI00000000	AJHK00000000
AJIK00000000	AJIN00000000	AJJL00000000	AJJN00000000	AJJO00000000	AJJQ00000000	AJKE00000000	AJKI00000000
AJKJ00000000	AJKN00000000	AJKO00000000	AJKP00000000	AJKQ00000000	AJKT00000000	AJKU00000000	AJKV00000000
AJKY00000000	AJLB00000000	AJLC00000000	AJLG00000000	AJLR00000000	AJLS00000000	AJML00000000	AJMM00000000
AJMT00000000	AJMU00000000	AJMV00000000	AJMW00000000	AJPQ00000000	AJPU00000000	AJQX00000000	AJQY00000000
AJQZ00000000	AJRA00000000	AJRB00000000	AJRC00000000	AJRD00000000	AJRE00000000	AJRF00000000	AJRG00000000
AJRH00000000	AJRZ00000000	AJSV00000000	AJSW00000000	AJSX00000000	AJSY00000000	AJSZ00000000	AJTA00000000
AJTB00000000	AJTC00000000	AJTF00000000	AJTG00000000	AJTH00000000	AJTI00000000	AJTJ00000000	AJTM00000000
AJTR00000000	AJTS00000000	AJTY00000000	AJUA00000000	AJUD00000000	AJUF00000000	AJUG00000000	AJUI00000000
AJUJ00000000	AJUK00000000	AJUM00000000	AJUP00000000	AJUY00000000	AJVA00000000	AJVC00000000	AJVN00000000
AJVO00000000	AJVQ00000000	AJVZ00000000	AJWA00000000	AJWL00000000	AJWQ00000000	AJWR00000000	AJWT00000000
AJWW00000000	AJXD00000000	AJXL00000000	AJXM00000000	AJXO00000000	AJXP00000000	AJXZ00000000	AJYX00000000
AJZR00000000	AJZS00000000	AJZT00000000	AKAO00000000	AKAU00000000	AKBD00000000	AKBT00000000	AKBU00000000
AKCB00000000	AKCC00000000	AKCF00000000	AKDO00000000	AKDP00000000	AKFD00000000	AKFF00000000	AKFH00000000
AKFK00000000	AKFO00000000	AKFP00000000	AKFQ00000000	AKFR00000000	AKFS00000000	AKFT00000000	AKFU00000000
AKFV00000000	AKGC00000000	AKGD00000000	AKGE00000000	AKGF00000000	AKGG00000000	AKHK00000000	AKHZ00000000
AKIA00000000	AKIB00000000	AKIH00000000	AKII00000000	AKIO00000000	AKIS00000000	AKIT00000000	AKIU00000000
AKIW00000000	AKIX00000000	AKJB00000000	AKJE00000000	AKJG00000000	AKJH00000000	AKJI00000000	AKJJ00000000
AKJK00000000	AKJM00000000	AKJN00000000	AKJQ00000000	AKJX00000000	AKJY00000000	AKJZ00000000	AKKB00000000
AKKE00000000	AKKF00000000	AKKK00000000	AKKR00000000	AKKS00000000	AKKT00000000	AKKV00000000	AKNH00000000
AKRA00000000	AKRF00000000	AKTT00000000	AKVJ00000000	AKVM00000000	AKVN00000000	AKVO00000000	AKVP00000000
AKWA00000000	AKXI00000000	AKXP00000000	AKXR00000000	AKXV00000000	AKYD00000000	AKYE00000000	AKYF00000000
AKYX00000000	AKZH00000000	AKZK00000000	AKZL00000000	AKZN00000000	AKZQ00000000	AKZR00000000	AKZS00000000
AKZT00000000	AKZU00000000	AKZW00000000	AKZX00000000	AKZY00000000	AKZZ00000000	ALAA00000000	ALAC00000000
ALAN00000000	ALAP00000000	ALAS00000000	ALAX00000000	ALBJ00000000	ALBQ00000000	ALBX00000000	ALCA00000000
ALCB00000000	ALCD00000000	ALCG00000000	ALCH00000000	ALCJ00000000	ALCL00000000	ALCT00000000	ALCV00000000
ALCX00000000	ALEG00000000	ALIF00000000	ALIH00000000	ALIJ00000000	ALIK00000000	ALIL00000000	ALIM00000000
ALIO00000000	ALIR00000000	ALIW00000000	ALIX00000000	ALIY00000000	ALIZ00000000	ALJB00000000	ALJG00000000
ALJH00000000	ALJJ00000000	ALJK00000000	ALJL00000000	ALJM00000000	ALJN00000000	ALJO00000000	ALJP00000000
ALJQ00000000	ALJR00000000	ALKE00000000	ALKF00000000	ALKG00000000	ALKH00000000	ALKI00000000	ALKJ00000000
ALKK00000000	ALKL00000000	ALKM00000000	ALKQ00000000	ALKT00000000	ALKX00000000	ALLB00000000	ALLL00000000
ALLM00000000	ALLO00000000	ALLR00000000	ALLT00000000	ALLW00000000	ALLY00000000	ALMB00000000	ALMI00000000
ALMJ00000000	ALMM00000000	ALMN00000000	ALMP00000000	ALND00000000	ALNE00000000	ALNG00000000	ALNI00000000
ALNJ00000000	ALNK00000000	ALNL00000000	ALNM00000000	ALNN00000000	ALNO00000000	ALNP00000000	ALNS00000000
ALOA00000000	ALOB00000000	ALOD00000000	ALOG00000000	ALOH00000000	ALOU00000000	ALOV00000000	ALOW00000000
ALOX00000000	ALOZ00000000	ALPF00000000	ALPT00000000	ALPU00000000	ALPV00000000	ALPY00000000	ALQA00000000
ALQS00000000	ALRK00000000	ALRL00000000	ALSA00000000	ALSC00000000	ALSG00000000	ALSW00000000	ALTJ00000000
ALTY00000000	ALUG00000000	ALVC00000000	ALVD00000000	ALVJ00000000	ALVM00000000	ALVT00000000	ALVU00000000
ALVW00000000	ALWD00000000	ALWJ00000000	ALWK00000000	ALWP00000000	ALWR00000000	ALXA00000000	ALXE00000000
ALXI00000000	ALXQ00000000	ALXS00000000	ALYD00000000	ALYK00000000	ALYL00000000	ALYM00000000	ALYN00000000
ALYO00000000	ALYP00000000	ALYS00000000	ALYV00000000	ALYW00000000	ALZX00000000	AMAB00000000	AMAC00000000
AMAE00000000	AMBL00000000	AMBP00000000	AMBQ00000000	AMBZ00000000	AMCA00000000	AMCE00000000	AMCM00000000
AMCP00000000	AMCX00000000	AMCZ00000000	AMDE00000000	AMDF00000000	AMDH00000000	AMEK00000000	AMEL00000000
AMEM00000000	AMEN00000000	AMEO00000000	AMEP00000000	AMEQ00000000	AMER00000000	AMES00000000	AMET00000000
AMEU00000000	AMEV00000000	AMEW00000000	AMEX00000000	AMEY00000000	AMEZ00000000	AMFA00000000	AMFD00000000
AMFE00000000	AMFQ00000000	AMGI00000000	AMGJ00000000	AMGM00000000	AMGT00000000	AMGU00000000	AMIC00000000
AMIH00000000	AMIW00000000	AMIX00000000	AMIY00000000	AMIZ00000000	AMJC00000000	AMJD00000000	AMJF00000000
AMJG00000000	AMJH00000000	AMJI00000000	AMJJ00000000	AMJM00000000	AMOO00000000	AMPE00000000	AMPG00000000
AMPI00000000	AMPJ00000000	AMPL00000000	AMPM00000000	AMPR00000000	AMPS00000000	AMPT00000000	AMQA00000000
AMQG00000000	AMQH00000000	AMQT00000000	AMQU00000000	AMQW00000000	AMQX00000000	AMRF00000000	AMRJ00000000
AMRL00000000	AMRN00000000	AMRP00000000	AMRQ00000000	AMSF00000000	AMSH00000000	AMSV00000000	AMSY00000000
AMUG00000000	AMUN00000000	AMUQ00000000	AMVC00000000	AMVM00000000	AMWD00000000	AMWL00000000	AMWQ00000000
AMXA00000000	AMXB00000000	AMXD00000000	AMXE00000000	AMXF00000000	AMXG00000000	AMXH00000000	AMXK00000000
AMXL00000000	AMXM00000000	AMXS00000000	AMXT00000000	AMYA00000000	AMYG00000000	AMYI00000000	AMYY00000000
AMZA00000000	AMZB00000000	AMZE00000000	AMZG00000000	AMZI00000000	AMZN00000000	AMZQ00000000	AMZX00000000
AMZY00000000	ANAJ00000000	ANAM00000000	ANAP00000000	ANAQ00000000	ANAU00000000	ANAV00000000	ANBT00000000
ANBV00000000	ANBY00000000	ANCA00000000	ANCE00000000	ANCT00000000	ANDJ00000000	ANEA00000000	ANET00000000

Genome ID							
ANFJ00000000	ANFK00000000	ANFS00000000	ANFT00000000	ANFX00000000	ANFZ00000000	ANGB00000000	ANGC00000000
ANGF00000000	ANGH00000000	ANHX00000000	ANHZ00000000	ANIA00000000	ANIC00000000	ANID00000000	ANIN00000000
ANIP00000000	ANIU00000000	ANJL00000000	ANJR00000000	ANJS00000000	ANJT00000000	ANJU00000000	ANJV00000000
ANJW00000000	ANJX00000000	ANJY00000000	ANJZ00000000	ANKA00000000	ANKB00000000	ANKC00000000	ANKD00000000
ANKE00000000	ANKF00000000	ANKG00000000	ANKH00000000	ANKI00000000	ANKJ00000000	ANKK00000000	ANKL00000000
ANKM00000000	ANKN00000000	ANKO00000000	ANKP00000000	ANKV00000000	ANKW00000000	ANKX00000000	ANKY00000000
ANLE00000000	ANLF00000000	ANLG00000000	ANMH00000000	ANMI00000000	ANMJ00000000	ANMK00000000	ANML00000000
ANMM00000000	ANMN00000000	ANMR00000000	ANNF00000000	ANNH00000000	ANNI00000000	ANNK00000000	ANNS00000000
ANNT00000000	ANNV00000000	ANNX00000000	ANOE00000000	ANOI00000000	ANOK00000000	ANPA00000000	ANPD00000000
ANPK00000000	ANQA00000000	ANQC00000000	ANQO00000000	ANQS00000000	ANRB00000000	ANRE00000000	ANRN00000000
ANRO00000000	ANSH00000000	ANSQ00000000	ANTA00000000	ANTB00000000	ANTW00000000	ANTX00000000	ANUO00000000
ANUR00000000	ANUS00000000	ANUX00000000	ANVO00000000	ANWL00000000	ANWM00000000	ANWO00000000	ANXH00000000
ANXP00000000	ANZB00000000	ANZF00000000	AOBN00000000	AOBO00000000	AOBQ00000000	AOBR00000000	AOBS00000000
AOBZ00000000	AOCB00000000	AOCD00000000	AOCE00000000	AOCJ00000000	AOCN00000000	AOCT00000000	AODD00000000
AODH00000000	AODI00000000	AODN00000000	AODQ00000000	AOFD00000000	AOGL00000000	AOGO00000000	AOGS00000000
AOHM00000000	AOHN00000000	AOHQ00000000	AOIF00000000	AOIG00000000	AOJA00000000	AOJP00000000	AOKO00000000
AOLV00000000	AOMS00000000	AOMT00000000	AONF00000000	AONQ00000000	AOOD00000000	AOOK00000000	AOOM00000000
AOON00000000	AOPY00000000	AORC00000000	AORL00000000	AORN00000000	AORU00000000	AORV00000000	AORX00000000
AORY00000000	AOSD00000000	AOSG00000000	AOSK00000000	AOSM00000000	AOSN00000000	AOSO00000000	AOSQ00000000
AOTF00000000	AOTH00000000	AOTK00000000	AOTY00000000	AOTZ00000000	AOUE00000000	AOUH00000000	AOUL00000000
AOUN00000000	AOUR00000000	APAS00000000	APAV00000000	APBN00000000	APBP00000000	APBV00000000	APCQ00000000
APCU00000000	APFF00000000	APFG00000000	APFI00000000	APGP00000000	APHP00000000	APHQ00000000	APHS00000000
APHX00000000	APHY00000000	APIS00000000	APIT00000000	APJE00000000	APJF00000000	APJG00000000	APJH00000000
APJI00000000	APJN00000000	APJS00000000	APJV00000000	APKC00000000	APKZ00000000	APLP00000000	APMB00000000
APMH00000000	APMP00000000	APMQ00000000	APMW00000000	APMX00000000	APMY00000000	APOE00000000	APOH00000000
APOI00000000	APOL00000000	APOM00000000	APON00000000	APOO00000000	APOS00000000	APOT00000000	APOY00000000
APOZ00000000	APPA00000000	APPC00000000	APPD00000000	APPE00000000	APPF00000000	APPG00000000	APPH00000000
APPI00000000	APPP00000000	APPR00000000	APPS00000000	APPU00000000	APPV00000000	APPW00000000	APPX00000000
APPY00000000	APPZ00000000	APQA00000000	APQB00000000	APQC00000000	APQE00000000	APQF00000000	APQG00000000
APQH00000000	APQI00000000	APQJ00000000	APQM00000000	APQN00000000	APQO00000000	APQP00000000	APQQ00000000
APQS00000000	APQT00000000	APQU00000000	APRI00000000	APRJ00000000	APRK00000000	APRL00000000	APRN00000000
APRO00000000	APRP00000000	APRQ00000000	APRS00000000	APRT00000000	APRU00000000	APRV00000000	APRW00000000
APRX00000000	APRY00000000	APRZ00000000	APSB00000000	APSC00000000	APSD00000000	APSE00000000	APSF00000000
APSH00000000	APSJ00000000	APSN00000000	APSR00000000	APTY00000000	APTZ00000000	APUE00000000	APUG00000000
APUI00000000	APVE00000000	APVQ00000000	APWE00000000	APWR00000000	APWT00000000	APWX00000000	APXE00000000
APXK00000000	APXW00000000	APXX00000000	APXY00000000	APYF00000000	APYJ00000000	APYS00000000	APYT00000000
APYV00000000	APYX00000000	APZD00000000	APZJ00000000	AQAF00000000	AQAW00000000	AQAY00000000	AQAZ00000000
AQBA00000000	AQBM00000000	AQBP00000000	AQBR00000000	AQBS00000000	AQBX00000000	AQCF00000000	AQCG00000000
AQCL00000000	AQCM00000000	AQCS00000000	AQCW00000000	AQCY00000000	AQCZ00000000	AQDD00000000	AQDG00000000
AQDJ00000000	AQDQ00000000	AQDR00000000	AQDS00000000	AQDT00000000	AQDU00000000	AQDV00000000	AQDX00000000
AQDY00000000	AQDZ00000000	AQEA00000000	AQEF00000000	AQEH00000000	AQEI00000000	AQEJ00000000	AQEV00000000
AQFO00000000	AQFP00000000	AQFQ00000000	AQFR00000000	AQFS00000000	AQFT00000000	AQFU00000000	AQFV00000000
AQGJ00000000	AQGK00000000	AQGL00000000	AQGN00000000	AQGO00000000	AQGR00000000	AQGT00000000	AQHM00000000
AQHX00000000	AQHY00000000	AQHZ00000000	AQOB00000000	AQOC00000000	AQOL00000000	AQOM00000000	AQON00000000
AQOO00000000	AQOP00000000	AQOS00000000	AQOZ00000000	AQPD00000000	AQPJ00000000	AQPL00000000	AQPP00000000
AQQB00000000	AQQF00000000	AQQG00000000	AQQH00000000	AQQO00000000	AQRI00000000	AQSD00000000	AQSG00000000
AQTP00000000	AQTQ00000000	AQTR00000000	AQTT00000000	AQTU00000000	AQUC00000000	AQUD00000000	AQUJ00000000
AQUM00000000	AQUN00000000	AQUO00000000	AQUS00000000	AQUV00000000	AQVA00000000	AQVC00000000	AQVS00000000
AQVT00000000	AQWE00000000	AQWG00000000	AQWI00000000	AQWJ00000000	AQWK00000000	AQWL00000000	AQWR00000000
AQWS00000000	AQWW00000000	AQWX00000000	AQWY00000000	AQWZ00000000	AQXC00000000	AQXH00000000	AQXR00000000
AQXU00000000	AQXZ00000000	AQYA00000000	AQYB00000000	AQYD00000000	AQYI00000000	AQYQ00000000	AQYR00000000
AQYY00000000	AQZD00000000	AQZF00000000	AQZG00000000	AQZM00000000	AQZR00000000	AQZT00000000	AQZV00000000
AQZX00000000	AQZZ00000000	ARAA00000000	ARAB00000000	ARAC00000000	ARAE00000000	ARAI00000000	ARAJ00000000
ARAN00000000	ARAY00000000	ARAZ00000000	ARBC00000000	ARBD00000000	ARBF00000000	ARBI00000000	ARBK00000000
ARBN00000000	ARBO00000000	ARBR00000000	ARBV00000000	ARBW00000000	ARBX00000000	ARBY00000000	ARBZ00000000
ARCA00000000	ARCC00000000	ARCD00000000	ARCE00000000	ARCF00000000	ARCG00000000	ARCH00000000	ARCI00000000
ARCJ00000000	ARCS00000000	ARCT00000000	ARCU00000000	ARDE00000000	ARDF00000000	ARDH00000000	ARDJ00000000
ARDK00000000	ARDL00000000	ARDO00000000	ARDP00000000	ARDR00000000	ARDT00000000	ARDU00000000	ARDW00000000
AREC00000000	AREF00000000	AREK00000000	AREM00000000	ARES00000000	ARET00000000	AREU00000000	ARFA00000000
ARFB00000000	ARFC00000000	ARFD00000000	ARFF00000000	ARFI00000000	ARFJ00000000	ARFK00000000	ARFL00000000
ARFP00000000	ARFQ00000000	ARFR00000000	ARFT00000000	ARFU00000000	ARFX00000000	ARFY00000000	ARGB00000000
ARGC00000000	ARGQ00000000	ARGS00000000	ARGU00000000	ARIP00000000	ARIR00000000	ARIU00000000	ARIX00000000
ARIY00000000	ARJN00000000	ARJO00000000	ARJQ00000000	ARJT00000000	ARJW00000000	ARJX00000000	ARKD00000000
ARKF00000000	ARKH00000000	ARKI00000000	ARKJ00000000	ARKX00000000	ARKY00000000	ARKZ00000000	ARLB00000000
ARLD00000000	ARLE00000000	ARLO00000000	ARLS00000000	ARMA00000000	ARMC00000000	ARMD00000000	ARME00000000
ARMF00000000	ARMM00000000	ARMW00000000	ARNA00000000	ARNB00000000	ARNC00000000	ARND00000000	ARNI00000000
ARNQ00000000	ARNR00000000	ARNS00000000	ARNT00000000	ARNV00000000	ARPV00000000	ARQD00000000	ARQJ00000000
ARQM00000000	ARQR00000000	ARRI00000000	ARRT00000000	ARSY00000000	ARTA00000000	ARTC00000000	ARTD00000000
ARTG00000000	ARTH00000000	ARTJ00000000	ARTK00000000	ARTM00000000	ARTV00000000	ARVS00000000	ARVT00000000
ARVV00000000	ARWA00000000	ARWD00000000	ARWR00000000	ARYB00000000	ARYO00000000	ARYY00000000	ARYZ00000000
ARZA00000000	ARZC00000000	ASAA00000000	ASAN00000000	ASDE00000000	ASDV00000000	ASEO00000000	ASEP00000000
ASGM00000000	ASHO00000000	ASHR00000000	ASHT00000000	ASHU00000000	ASHV00000000	ASHW00000000	ASJC00000000
ASJE00000000	ASJO00000000	ASJQ00000000	ASJR00000000	ASJW00000000	ASJX00000000	ASKA00000000	ASKB00000000
ASKC00000000	ASKD00000000	ASKE00000000	ASKF00000000	ASKG00000000	ASKH00000000	ASKI00000000	ASML00000000
ASNS00000000	ASNV00000000	ASOP00000000	ASOR00000000	ASPR00000000	ASPS00000000	ASPT00000000	ASPU00000000
ASPV00000000	ASPW00000000	ASPX00000000	ASPZ00000000	ASQG00000000	ASQH00000000	ASQJ00000000	ASQK00000000
ASQL00000000	ASQM00000000	ASQN00000000	ASQQ00000000	ASRG00000000	ASRU00000000	ASRV00000000	ASRY00000000
ASRZ00000000	ASSA00000000	ASSB00000000	ASSC00000000	ASSD00000000	ASSE00000000	ASSF00000000	ASSG00000000
ASSI00000000	ASSM00000000	ASSN00000000	ASSO00000000	ASSP00000000	ASSQ00000000	ASSR00000000	ASSS00000000
ASST00000000	ASSU00000000	ASSV00000000	ASSW00000000	ASSX00000000	ASSY00000000	ASSZ00000000	ASTA00000000

Genome ID							
ASTB00000000	ASTC00000000	ASTD00000000	ASTE00000000	ASTF00000000	ASTG00000000	ASTI00000000	ASTJ00000000
ASTQ00000000	ASTS00000000	ASTV00000000	ASUF00000000	ASUI00000000	ASUN00000000	ASUU00000000	ASUV00000000
ASUW00000000	ASVA00000000	ASVF00000000	ASVG00000000	ASVI00000000	ASVJ00000000	ASVL00000000	ASVN00000000
ASVR00000000	ASVU00000000	ASVX00000000	ASVZ00000000	ASWA00000000	ASWC00000000	ASWD00000000	ASWF00000000
ASWG00000000	ASWH00000000	ASWI00000000	ASWJ00000000	ASWK00000000	ASWL00000000	ASWM00000000	ASWO00000000
ASWT00000000	ASWZ00000000	ASXA00000000	ASXF00000000	ASXI00000000	ASXP00000000	ASXS00000000	ASXT00000000
ASXU00000000	ASYE00000000	ASYF00000000	ASYG00000000	ASYH00000000	ASYI00000000	ASYJ00000000	ASYK00000000
ASYL00000000	ASYX00000000	ASZP00000000	ASZT00000000	ASZU00000000	ASZV00000000	ASZW00000000	ASZX00000000
ASZZ00000000	ATAA00000000	ATAB00000000	ATAE00000000	ATAH00000000	ATAX00000000	ATAY00000000	ATAZ00000000
ATBA00000000	ATBB00000000	ATBC00000000	ATBD00000000	ATBE00000000	ATBF00000000	ATBI00000000	ATBQ00000000
ATBY00000000	ATBZ00000000	ATCA00000000	ATCB00000000	ATCC00000000	ATCD00000000	ATCE00000000	ATCF00000000
ATCG00000000	ATCH00000000	ATCI00000000	ATCK00000000	ATCL00000000	ATCO00000000	ATDC00000000	ATDD00000000
ATDE00000000	ATDJ00000000	ATDL00000000	ATDP00000000	ATDT00000000	ATEX00000000	ATFB00000000	ATFC00000000
ATFD00000000	ATFE00000000	ATFF00000000	ATFI00000000	ATFK00000000	ATFL00000000	ATFM00000000	ATFN00000000
ATFO00000000	ATFP00000000	ATGG00000000	ATGH00000000	ATGI00000000	ATGK00000000	ATHH00000000	ATHL00000000
ATHX00000000	ATIB00000000	ATIC00000000	ATIV00000000	ATJE00000000	ATJF00000000	ATJG00000000	ATJH00000000
ATJI00000000	ATJJ00000000	ATJK00000000	ATJL00000000	ATJM00000000	ATJN00000000	ATJO00000000	ATJP00000000
ATJQ00000000	ATJR00000000	ATJS00000000	ATJX00000000	ATKB00000000	ATKC00000000	ATKD00000000	ATKE00000000
ATKF00000000	ATKG00000000	ATKH00000000	ATKI00000000	ATKJ00000000	ATKK00000000	ATKN00000000	ATLK00000000
ATLQ00000000	ATMR00000000	ATMS00000000	ATMT00000000	ATNE00000000	ATNG00000000	ATNI00000000	ATNK00000000
ATNM00000000	ATOG00000000	ATOH00000000	ATOI00000000	ATOJ00000000	ATOK00000000	ATOL00000000	ATOM00000000
ATOV00000000	ATOX00000000	ATRO00000000	ATSX00000000	ATSY00000000	ATSZ00000000	ATTA00000000	ATTG00000000
ATTI00000000	ATTJ00000000	ATTK00000000	ATTN00000000	ATTO00000000	ATTP00000000	ATTS00000000	ATTU00000000
ATTZ00000000	ATUD00000000	ATUF00000000	ATUJ00000000	ATUR00000000	ATUT00000000	ATUX00000000	ATUZ00000000
ATVB00000000	ATVD00000000	ATVE00000000	ATVG00000000	ATVI00000000	ATVN00000000	ATVP00000000	ATVQ00000000
ATVR00000000	ATVS00000000	ATVT00000000	ATVU00000000	ATVV00000000	ATVW00000000	ATVX00000000	ATVY00000000
ATVZ00000000	ATWA00000000	ATWB00000000	ATWC00000000	ATWH00000000	ATWJ00000000	ATWK00000000	ATWL00000000
ATWP00000000	ATWS00000000	ATWX00000000	ATWY00000000	ATWZ00000000	ATXD00000000	ATXE00000000	ATXG00000000
ATXK00000000	ATXL00000000	ATXM00000000	ATXV00000000	ATXW00000000	ATXY00000000	ATXZ00000000	ATYA00000000
ATYG00000000	ATYJ00000000	ATYK00000000	ATYL00000000	ATYN00000000	ATYQ00000000	ATYZ00000000	ATZB00000000
ATZE00000000	ATZF00000000	ATZH00000000	ATZI00000000	ATZN00000000	ATZT00000000	ATZW00000000	ATZX00000000
ATZZ00000000	AUAA00000000	AUAB00000000	AUAD00000000	AUAE00000000	AUAK00000000	AUAN00000000	AUAO00000000
AUAR00000000	AUAS00000000	AUAT00000000	AUAW00000000	AUAX00000000	AUBA00000000	AUBD00000000	AUBF00000000
AUBI00000000	AUBJ00000000	AUBL00000000	AUBM00000000	AUBN00000000	AUBO00000000	AUBR00000000	AUBW00000000
AUCJ00000000	AUCL00000000	AUCM00000000	AUCN00000000	AUCO00000000	AUCY00000000	AUDA00000000	AUDD00000000
AUDF00000000	AUDH00000000	AUDN00000000	AUDO00000000	AUDP00000000	AUDR00000000	AUDS00000000	AUDW00000000
AUEE00000000	AUEF00000000	AUEK00000000	AUEM00000000	AUEP00000000	AUEU00000000	AUEX00000000	AUFA00000000
AUFB00000000	AUFC00000000	AUFD00000000	AUFF00000000	AUFG00000000	AUFH00000000	AUFI00000000	AUFJ00000000
AUFK00000000	AUFL00000000	AUFN00000000	AUFP00000000	AUFQ00000000	AUFR00000000	AUFS00000000	AUFT00000000
AUFX00000000	AUGD00000000	AUGJ00000000	AUGK00000000	AUGO00000000	AUGP00000000	AUGQ00000000	AUGT00000000
AUGX00000000	AUHB00000000	AUHC00000000	AUHE00000000	AUHH00000000	AUHM00000000	AUHO00000000	AUHQ00000000
AUHW00000000	AUIA00000000	AUIB00000000	AUIC00000000	AUID00000000	AUII00000000	AUIK00000000	AUIL00000000
AUIM00000000	AUIN00000000	AUIO00000000	AUIP00000000	AUIQ00000000	AUIV00000000	AUIX00000000	AUIY00000000
AUIZ00000000	AUJA00000000	AUJB00000000	AUJC00000000	AUJD00000000	AUJE00000000	AUJF00000000	AUJG00000000
AUJH00000000	AUJI00000000	AUJJ00000000	AUJK00000000	AUJL00000000	AUJM00000000	AUJN00000000	AUJO00000000
AUJP00000000	AUJQ00000000	AUJR00000000	AUJS00000000	AUJT00000000	AUJU00000000	AUJV00000000	AUJW00000000
AUJX00000000	AUJY00000000	AUJZ00000000	AUKA00000000	AUKB00000000	AUKC00000000	AUKD00000000	AUKE00000000
AUKF00000000	AUKN00000000	AUKQ00000000	AUKS00000000	AUKY00000000	AULA00000000	AULC00000000	AULD00000000
AULI00000000	AULL00000000	AULR00000000	AULW00000000	AUME00000000	AUMG00000000	AUML00000000	AUMP00000000
AUMS00000000	AUMT00000000	AUMU00000000	AUMV00000000	AUMW00000000	AUNB00000000	AUNM00000000	AUNN00000000
AUNO00000000	AUNR00000000	AUNS00000000	AUNW00000000	AUNY00000000	AUOB00000000	AUOC00000000	AUOE00000000
AUOK00000000	AUOP00000000	AUOR00000000	AUOS00000000	AUOT00000000	AUOU00000000	AUOV00000000	AUOX00000000
AUOZ00000000	AUPA00000000	AUPD00000000	AUPH00000000	AUPJ00000000	AUQW00000000	AURB00000000	AUSF00000000
AUSH00000000	AUSW00000000	AUTD00000000	AUTE00000000	AUTH00000000	AUTI00000000	AUTJ00000000	AUTK00000000
AUTL00000000	AUTM00000000	AUTN00000000	AUTO00000000	AUTP00000000	AUTQ00000000	AUTR00000000	AUTS00000000
AUTU00000000	AUTZ00000000	AUUB00000000	AUUC00000000	AUUD00000000	AUUE00000000	AUUR00000000	AUVG00000000
AUWB00000000	AUWR00000000	AUWS00000000	AUWT00000000	AUXA00000000	AUXF00000000	AUXP00000000	AUXQ00000000
AUYD00000000	AUYM00000000	AUYR00000000	AUZE00000000	AUZP00000000	AUZQ00000000	AVAA00000000	AVAB00000000
AVAE00000000	AVAI00000000	AVBI00000000	AVBJ00000000	AVBK00000000	AVBL00000000	AVBM00000000	AVCC00000000
AVCQ00000000	AVCR00000000	AVCS00000000	AVDX00000000	AVFB00000000	AVFE00000000	AVFL00000000	AVGK00000000
AVGO00000000	AVHE00000000	AVHN00000000	AVII00000000	AVIM00000000	AVIU00000000	AVJC00000000	AVJG00000000
AVJS00000000	AVJU00000000	AVKO00000000	AVKQ00000000	AVLA00000000	AVLC00000000	AVLD00000000	AVLH00000000
AVLI00000000	AVLJ00000000	AVLX00000000	AVMC00000000	AVMD00000000	AVMH00000000	AVMJ00000000	AVMO00000000
AVMP00000000	AVMQ00000000	AVNB00000000	AVNC00000000	AVND00000000	AVNG00000000	AVNK00000000	AVNM00000000
AVNN00000000	AVNS00000000	AVNT00000000	AVNU00000000	AVNV00000000	AVNW00000000	AVOE00000000	AVOP00000000
AVPC00000000	AVPD00000000	AVPM00000000	AVPP00000000	AVQL00000000	AVSU00000000	AVSV00000000	AVTC00000000
AVTP00000000	AVTY00000000	AVUA00000000	AVUJ00000000	AVUW00000000	AVVI00000000	AVVS00000000	AVVV00000000
AVXG00000000	AVYE00000000	AVYS00000000	AVZC00000000	AVZZ00000000	AWAE00000000	AWAM00000000	AWAO00000000
AWAR00000000	AWAX00000000	AWBD00000000	AWBF00000000	AWBI00000000	AWCI00000000	AWCR00000000	AWCV00000000
AWCW00000000	AWDE00000000	AWDK00000000	AWDO00000000	AWDV00000000	AWDW00000000	AWEA00000000	AWED00000000
AWEM00000000	AWEZ00000000	AWGF00000000	AWGV00000000	AWNJ00000000	AWNW00000000	AWOP00000000	AWOY00000000
AWPJ00000000	AWQR00000000	AWQW00000000	AWQX00000000	AWQY00000000	AWRZ00000000	AWSA00000000	AWSB00000000
AWSC00000000	AWSD00000000	AWSE00000000	AWSF00000000	AWSG00000000	AWSH00000000	AWSI00000000	AWSJ00000000
AWSK00000000	AWSV00000000	AWSW00000000	AWSX00000000	AWSY00000000	AWTB00000000	AWTC00000000	AWTD00000000
AWTF00000000	AWTI00000000	AWTS00000000	AWTX00000000	AWUC00000000	AWUF00000000	AWUG00000000	AWUJ00000000
AWUK00000000	AWUN00000000	AWUO00000000	AWUP00000000	AWUR00000000	AWVI00000000	AWVS00000000	AWVT00000000
AWVU00000000	AWVZ00000000	AWWC00000000	AWWG00000000	AWWH00000000	AWWJ00000000	AWWK00000000	AWXA00000000
AWXB00000000	AWXC00000000	AWXD00000000	AWXG00000000	AWXH00000000	AWXI00000000	AWXP00000000	AWXS00000000
AWXX00000000	AWXY00000000	AWYG00000000	AWZN00000000	AWZQ00000000	AWZR00000000	AWZU00000000	AXAA00000000
AXAD00000000	AXAF00000000	AXAG00000000	AXAI00000000	AXAP00000000	AXAQ00000000	AXAS00000000	AXAT00000000

Genome ID							
AXAU00000000	AXAV00000000	AXAZ00000000	AXBC00000000	AXBK00000000	AXBL00000000	AXBO00000000	AXBT00000000
AXBU00000000	AXDC00000000	AXDK00000000	AXDQ00000000	AXDT00000000	AXDW00000000	AXDX00000000	AXDY00000000
AXLJ00000000	AXLK00000000	AXLM00000000	AXLN00000000	AXLO00000000	AXLP00000000	AXLQ00000000	AXLR00000000
AXLS00000000	AXLT00000000	AXLU00000000	AXLV00000000	AXLW00000000	AXLX00000000	AXLY00000000	AXLZ00000000
AXMA00000000	AXMB00000000	AXMC00000000	AXMD00000000	AXME00000000	AXMF00000000	AXMG00000000	AXMH00000000
AXMI00000000	AXMJ00000000	AXMK00000000	AXML00000000	AXMM00000000	AXMN00000000	AXNO00000000	AXNU00000000
AXNV00000000	AXNW00000000	AXNX00000000	AXNY00000000	AXNZ00000000	AXOF00000000	AXOG00000000	AXOH00000000
AXOJ00000000	AXOK00000000	AXOL00000000	AXOM00000000	AXON00000000	AXOO00000000	AXPB00000000	AXPO00000000
AXPY00000000	AXQG00000000	AXQO00000000	AXQQ00000000	AXQW00000000	AXSX00000000	AXSY00000000	AXTH00000000
AXTK00000000	AXTL00000000	AXTR00000000	AXTX00000000	AXUA00000000	AXUD00000000	AXUF00000000	AXUG00000000
AXUI00000000	AXUM00000000	AXUN00000000	AXUO00000000	AXUQ00000000	AXUR00000000	AXUS00000000	AXUZ00000000
AXVC00000000	AXVF00000000	AXVG00000000	AXVK00000000	AXVN00000000	AXVO00000000	AXVY00000000	AXWB00000000
AXWE00000000	AXWF00000000	AXWG00000000	AXWS00000000	AXWT00000000	AXWZ00000000	AXXE00000000	AXXF00000000
AXXG00000000	AXXJ00000000	AXXL00000000	AXXW00000000	AXYF00000000	AXYI00000000	AXYM00000000	AXYP00000000
AXYR00000000	AXYV00000000	AXYW00000000	AXYX00000000	AXZB00000000	AXZF00000000	AXZG00000000	AXZJ00000000
AXZK00000000	AXZL00000000	AXZP00000000	AXZX00000000	AXZY00000000	AXZZ00000000	AYEQ00000000	AYES00000000
AYET00000000	AYEU00000000	AYEV00000000	AYFN00000000	AYGU00000000	AYGX00000000	AYGZ00000000	AYHA00000000
AYHL00000000	AYHT00000000	AYIC00000000	AYID00000000	AYIE00000000	AYIF00000000	AYIG00000000	AYIK00000000
AYIM00000000	AYIO00000000	AYIP00000000	AYIQ00000000	AYIR00000000	AYIU00000000	AYIV00000000	AYIW00000000
AYIX00000000	AYIY00000000	AYJA00000000	AYJC00000000	AYJD00000000	AYJE00000000	AYJF00000000	AYJG00000000
AYJH00000000	AYJI00000000	AYJJ00000000	AYJK00000000	AYJM00000000	AYJO00000000	AYJP00000000	AYJU00000000
AYJW00000000	AYKQ00000000	AYKS00000000	AYKV00000000	AYKY00000000	AYLA00000000	AYLG00000000	AYLO00000000
AYLQ00000000	AYLW00000000	AYLZ00000000	AYME00000000	AYMG00000000	AYMH00000000	AYMI00000000	AYMJ00000000
AYMN00000000	AYMP00000000	AYMQ00000000	AYMR00000000	AYMT00000000	AYMV00000000	AYMZ00000000	AYNA00000000
AYNE00000000	AYOP00000000	AYOX00000000	AYPR00000000	AYPV00000000	AYPW00000000	AYPY00000000	AYPZ00000000
AYQA00000000	AYQB00000000	AYQC00000000	AYQE00000000	AYRA00000000	AYRB00000000	AYRC00000000	AYRN00000000
AYRO00000000	AYRP00000000	AYRQ00000000	AYRR00000000	AYRS00000000	AYSC00000000	AYSD00000000	AYSE00000000
AYSF00000000	AYSG00000000	AYSH00000000	AYSN00000000	AYSU00000000	AYSV00000000	AYSW00000000	AYTB00000000
AYTD00000000	AYTE00000000	AYTK00000000	AYTM00000000	AYTO00000000	AYTT00000000	AYTU00000000	AYTY00000000
AYUA00000000	AYUC00000000	AYUD00000000	AYUF00000000	AYXG00000000	AYXM00000000	AYXR00000000	AYXS00000000
AYXZ00000000	AYYA00000000	AYYC00000000	AYYD00000000	AYYE00000000	AYYF00000000	AZAE00000000	AZAP00000000
AZBL00000000	AZBY00000000	AZGQ00000000	AZHI00000000	AZHJ00000000	AZHL00000000	AZHP00000000	AZHR00000000
AZHT00000000	AZIA00000000	AZIO00000000	AZIR00000000	AZIS00000000	AZJC00000000	AZJD00000000	AZJE00000000
AZJF00000000	AZJG00000000	AZJH00000000	AZJI00000000	AZJJ00000000	AZJS00000000	AZJT00000000	AZKM00000000
AZLV00000000	AZLW00000000	AZLX00000000	AZLZ00000000	AZMB00000000	AZMC00000000	AZMD00000000	AZME00000000
AZMF00000000	AZMG00000000	AZMH00000000	AZMI00000000	AZMJ00000000	AZMK00000000	AZML00000000	AZMV00000000
AZND00000000	AZNU00000000	AZOE00000000	AZQN00000000	AZQP00000000	AZQQ00000000	AZQS00000000	AZQY00000000
AZQZ00000000	AZRA00000000	AZRH00000000	AZRT00000000	AZRX00000000	AZRY00000000	AZRZ00000000	AZSA00000000
AZSB00000000	AZSC00000000	AZSI00000000	AZSN00000000	AZSU00000000	AZTE00000000	AZTJ00000000	AZTK00000000
AZTL00000000	AZTM00000000	AZUA00000000	AZUB00000000	AZUH00000000	AZUI00000000	AZUO00000000	AZUP00000000
AZUR00000000	AZUZ00000000	AZVA00000000	AZVB00000000	AZVC00000000	AZVH00000000	AZVI00000000	AZVS00000000
AZVX00000000	AZXO00000000	AZXP00000000	AZXQ00000000	AZXR00000000	AZXS00000000	AZXU00000000	AZXW00000000
AZXX00000000	AZXZ00000000	AZYA00000000	AZYE00000000	AZYF00000000	AZYH00000000	AZYK00000000	AZYN00000000
AZYP00000000	AZYU00000000	AZYV00000000	AZYW00000000	AZYX00000000	AZYY00000000	AZYZ00000000	AZZD00000000
AZZK00000000	AZZO00000000	AZZQ00000000	AZZW00000000	AZZZ00000000	BABR00000000	BABS00000000	BABV00000000
BABW00000000	BACF00000000	BACG00000000	BACM00000000	BACN00000000	BACO00000000	BACP00000000	BACQ00000000
BACR00000000	BACS00000000	BACU00000000	BACV00000000	BACW00000000	BACY00000000	BACZ00000000	BADA00000000
BADB00000000	BADC00000000	BADD00000000	BADE00000000	BADF00000000	BADH00000000	BADI00000000	BADJ00000000
BADK00000000	BADL00000000	BADM00000000	BADP00000000	BADT00000000	BADW00000000	BADX00000000	BADY00000000
BADZ00000000	BAEC00000000	BAEG00000000	BAEH00000000	BAEI00000000	BAEV00000000	BAEW00000000	BAEX00000000
BAEY00000000	BAFC00000000	BAFD00000000	BAFE00000000	BAFF00000000	BAFO00000000	BAGA00000000	BAGB00000000
BAGF00000000	BAGL00000000	BAGT00000000	BAGW00000000	BAGY00000000	BAGZ00000000	BAHD00000000	BAHE00000000
BAHP00000000	BAHQ00000000	BAHR00000000	BAHT00000000	BAHU00000000	BAHV00000000	BAHW00000000	BAHY00000000
BAHZ00000000	BAIA00000000	BAIB00000000	BAIC00000000	BAID00000000	BAIE00000000	BAIF00000000	BAIG00000000
BAIG00000000	BAIH00000000	BAIJ00000000	BAIK00000000	BAIL00000000	BAIO00000000	BAIP00000000	BAIQ00000000
BAIR00000000	BAIS00000000	BAIT00000000	BAIU00000000	BAIV00000000	BAIW00000000	BAIX00000000	BAIY00000000
BAIZ00000000	BAJA00000000	BAJB00000000	BAJC00000000	BAJD00000000	BAJE00000000	BAJF00000000	BAJG00000000
BAJH00000000	BAJI00000000	BAJJ00000000	BAJK00000000	BAJL00000000	BAJM00000000	BAJN00000000	BAJO00000000
BAJP00000000	BAJQ00000000	BAJR00000000	BAJS00000000	BAJT00000000	BAJU00000000	BAJV00000000	BAJW00000000
BAJX00000000	BAJY00000000	BAJZ00000000	BAKD00000000	BAKE00000000	BAKF00000000	BAKG00000000	BAKH00000000
BAKI00000000	BAKJ00000000	BAKK00000000	BAKL00000000	BAKM00000000	BAKN00000000	BAKO00000000	BAKP00000000
BAKQ00000000	BAKR00000000	BAKS00000000	BAKW00000000	BAKX00000000	BAKZ00000000	BALB00000000	BALG00000000
BALJ00000000	BALK00000000	BALP00000000	BALQ00000000	BALR00000000	BALS00000000	BALT00000000	BALU00000000
BALV00000000	BALW00000000	BALX00000000	BAMC00000000	BAMD00000000	BAMF00000000	BAMG00000000	BAMH00000000
BAMI00000000	BAMJ00000000	BAMK00000000	BAML00000000	BAMO00000000	BAMQ00000000	BANK00000000	BANM00000000
BANN00000000	BANR00000000	BANS00000000	BANX00000000	BAOC00000000	BAOK00000000	BAOL00000000	BAON00000000
BAOQ00000000	BAOR00000000	BAOT00000000	BAOU00000000	BAOV00000000	BARD00000000	BARE00000000	BARL00000000
BARM00000000	BARY00000000	BASG00000000	BASM00000000	BASQ00000000	BATY00000000	BAUO00000000	BAUS00000000
BAUT00000000	BAUU00000000	BAUV00000000	BAUW00000000	BAVL00000000	BAVR00000000	BAVZ00000000	BAWE00000000
BAWG00000000	BAWH00000000	BAWI00000000	BAWJ00000000	BAWK00000000	BAWL00000000	BAWO00000000	BAWP00000000
BAWQ00000000	BAWR00000000	BAWT00000000	BAWU00000000	BAXY00000000	BAYA00000000	BAYD00000000	BAYT00000000
BAZB00000000	BAZH00000000	BBIK00000000	BBIS00000000	BBIT00000000	BBIU00000000	BBIV00000000	BBIW00000000
BBIX00000000	BBJJ00000000	BBJK00000000	BBJR00000000	BBJS00000000	BBJT00000000	BBJV00000000	BBJW00000000
BBJZ00000000	CABY00000000	CACD00000000	CACH00000000	CACR00000000	CACS00000000	CACY00000000	CADP00000000
CADR00000000	CADS00000000	CADT00000000	CAEF00000000	CAEG00000000	CAEI00000000	CAEK00000000	CAEL00000000
CAEM00000000	CAEN00000000	CAEO00000000	CAEP00000000	CAES00000000	CAEU00000000	CAEV00000000	CAFE00000000
CAFN00000000	CAFW00000000	CAGD00000000	CAGE00000000	CAGF00000000	CAGG00000000	CAGH00000000	CAGT00000000
CAGU00000000	CAGV00000000	CAGX00000000	CAGY00000000	CAGZ00000000	CAHA00000000	CAHB00000000	CAHC00000000
CAHD00000000	CAHE00000000	CAHF00000000	CAHI00000000	CAHJ00000000	CAHK00000000	CAHL00000000	CAHP00000000
CAHU00000000	CAIA00000000	CAIB00000000	CAIE00000000	CAIG00000000	CAIK00000000	CAIT00000000	CAIU00000000

Genome ID							
CAIZ00000000	CAJB00000000	CAJD00000000	CAJF00000000	CAJH00000000	CAJJ000000000	CAJK00000000	CAJL00000000
CAJN000000000	CAJO00000000	CAJP00000000	CAJQ00000000	CAJU000000000	CAJZ00000000	CAKD00000000	CAKE00000000
CAKO00000000	CAKP00000000	CAKQ00000000	CAKT00000000	CAKU00000000	CAKV00000000	CAKW00000000	CAKX00000000
CAKZ00000000	CALA00000000	CALB00000000	CALC00000000	CALD00000000	CALE00000000	CALF00000000	CALG00000000
CALH00000000	CALI00000000	CALK00000000	CALM00000000	CALN00000000	CALX00000000	CALY00000000	CAMA00000000
CAMB00000000	CAME00000000	CAMH00000000	CAML00000000	CAMQ00000000	CAMU00000000	CAMV00000000	CAMX00000000
CAMZ00000000	CANE00000000	CANG00000000	CANL00000000	CANO00000000	CANP00000000	CANQ00000000	CANU00000000
CANV00000000	CAOB00000000	CAOD00000000	CAOG00000000	CAOR00000000	CAOS00000000	CAPF00000000	CAPH00000000
CAPI00000000	CAPL00000000	CAQI00000000	CAQO00000000	CAQT00000000	CAQU00000000	CAQV00000000	CAQW00000000
CAQX00000000	CAQY00000000	CAQZ00000000	CARA00000000	CAUA00000000	CAUC00000000	CAUJ00000000	CAUK00000000
CAVF000000000	CAVI000000000	CAVN000000000	CAVO000000000	CAVS000000000	CAWE000000000	CAWF000000000	CAWG000000000
CAWJ000000000	CAWK000000000	CAWL000000000	CAWM000000000	CAWN000000000	CAWO000000000	CAWP000000000	CAWQ000000000
CAWR000000000	CAWS000000000	CAWT000000000	CAWU000000000	CAWV000000000	CAWW000000000	CAWX000000000	CAWY000000000
CAXA000000000	CAXB000000000	CAXC000000000	CAXD000000000	CAXE000000000	CAXF000000000	CAXG000000000	CAXH000000000
CAXI000000000	CAXJ000000000	CAXK000000000	CAXL000000000	CAXM000000000	CAXN000000000	CAXO000000000	CAXP000000000
CAXQ000000000	CAXR000000000	CAXS000000000	CAXT000000000	CAXU000000000	CAXV000000000	CAXW000000000	CAXX000000000
CAXY000000000	CAYA000000000	CAYB000000000	CAYC000000000	CAYD000000000	CAYE000000000	CAYF000000000	CAYG000000000
CAYH000000000	CAYI000000000	CAYJ000000000	CAYK000000000	CAYL000000000	CAYM000000000	CAYN000000000	CAYO000000000
CAYP000000000	CAYQ000000000	CAYR000000000	CAYS000000000	CAYT000000000	CAYU000000000	CAYW000000000	CAYX000000000
CAYY000000000	CAZB000000000	CAZC000000000	CAZD000000000	CAZE000000000	CAZF000000000	CAZG000000000	CAZH000000000
CAZI000000000	CAZJ000000000	CAZK000000000	CAZL000000000	CAZM000000000	CAZN000000000	CAZO000000000	CAZP000000000
CAZR000000000	CAZS000000000	CAZT000000000	CAZU000000000	CAZV000000000	CAZW000000000	CAZX000000000	CAZY000000000
CBAA000000000	CBAB000000000	CBAC000000000	CBAD000000000	CBAE000000000	CBAF000000000	CBAG000000000	CBAH000000000
CBAI000000000	CBAJ000000000	CBAK000000000	CBAL000000000	CBAM000000000	CBAN000000000	CBAO000000000	CBAP000000000
CBAQ000000000	CBAR000000000	CBAS000000000	CBAT000000000	CBAU000000000	CBAV000000000	CBAW000000000	CBAX000000000
CBAY000000000	CBAZ000000000	CBBA000000000	CBBB000000000	CBBC000000000	CBBD000000000	CBBE000000000	CBBF000000000
CBBG000000000	CBBH000000000	CBBI000000000	CBBJ000000000	CBBL000000000	CBBM000000000	CBBN000000000	CBBO000000000
CBBP000000000	CBBQ000000000	CBBR000000000	CBBS000000000	CBBT000000000	CBBU000000000	CBBV000000000	CBBW000000000
CBBX000000000	CBBY000000000	CBBZ000000000	CBCA000000000	CBCB000000000	CBCC000000000	CBCD000000000	CBCE000000000
CBCF000000000	CBCG000000000	CBCH000000000	CBCI000000000	CBCJ000000000	CBCK000000000	CBCL000000000	CBCM000000000
CBCN000000000	CBCO000000000	CBCP000000000	CBCQ000000000	CBCR000000000	CBCS000000000	CBCT000000000	CBCU000000000
CBCV000000000	CBCW000000000	CBCX000000000	CBCY000000000	CBCZ000000000	CBDA000000000	CBDB000000000	CBDC000000000
CBDD000000000	CBDE000000000	CBDF000000000	CBDG000000000	CBDH000000000	CBDI000000000	CBDJ000000000	CBDK000000000
CBDL000000000	CBDM000000000	CBDN000000000	CBDO000000000	CBDP000000000	CBDQ000000000	CBDR000000000	CBDS000000000
CBDT000000000	CBDU000000000	CBDV000000000	CBDW000000000	CBDX000000000	CBDY000000000	CBDZ000000000	CBEA000000000
CBEB000000000	CBEC000000000	CBED000000000	CBEE000000000	CBEF000000000	CBEG000000000	CBEH000000000	CBEI000000000
CBEJ000000000	CBEK000000000	CBEL000000000	CBEM000000000	CBEN000000000	CBEO000000000	CBEP000000000	CBEQ000000000
CBER000000000	CBES000000000	CBET000000000	CBEU000000000	CBEV000000000	CBEW000000000	CBEX000000000	CBEY000000000
CBEZ000000000	CBFA000000000	CBFB000000000	CBFC000000000	CBFD000000000	CBFE000000000	CBFF000000000	CBFG000000000
CBFH000000000	CBFI000000000	CBFJ000000000	CBFK000000000	CBFL000000000	CBFM000000000	CBFN000000000	CBFO000000000
CBFP000000000	CBFQ000000000	CBFR000000000	CBFS000000000	CBFT000000000	CBFU000000000	CBFV000000000	CBFW000000000
CBFX000000000	CBFY000000000	CBFZ000000000	CBGA000000000	CBGB000000000	CBGC000000000	CBGD000000000	CBGE000000000
CBGF000000000	CBGG000000000	CBGH000000000	CBGI000000000	CBGJ000000000	CBGK000000000	CBGL000000000	CBGM000000000
CBGN000000000	CBGO000000000	CBGP000000000	CBGQ000000000	CBGR000000000	CBGS000000000	CBGT000000000	CBGU000000000
CBGV000000000	CBGW000000000	CBGX000000000	CBGY000000000	CBGZ000000000	CBHA000000000	CBHB000000000	CBHC000000000
CBHD000000000	CBHE000000000	CBHF000000000	CBHG000000000	CBHH000000000	CBHI000000000	CBHJ000000000	CBHK000000000
CBHL000000000	CBHM000000000	CBHN000000000	CBHO000000000	CBHP000000000	CBHQ000000000	CBHR000000000	CBHS000000000
CBHT000000000	CBHU000000000	CBHV000000000	CBHW000000000	CBHX000000000	CBHY000000000	CBIA000000000	CBIB000000000
CBIC000000000	CBID000000000	CBIE000000000	CBIF000000000	CBIG000000000	CBIH000000000	CBII000000000	CBIJ000000000
CBIK000000000	CBIL000000000	CBIM000000000	CBIN000000000	CBIO000000000	CBIP000000000	CBIQ000000000	CBIR000000000
CBIS000000000	CBIT000000000	CBIU000000000	CBIW000000000	CBIX000000000	CBIY000000000	CBIZ000000000	CBJA000000000
CBJB000000000	CBJC000000000	CBJD000000000	CBJE000000000	CBJF000000000	CBJG000000000	CBJH000000000	CBJI000000000
CBJJ000000000	CBJK000000000	CBJL000000000	CBJM000000000	CBJN000000000	CBJO000000000	CBJP000000000	CBJQ000000000
CBJR000000000	CBJS000000000	CBJT000000000	CBJU000000000	CBJV000000000	CBJW000000000	CBJX000000000	CBJY000000000
CBJZ000000000	CBKA000000000	CBKB000000000	CBKC000000000	CBKD000000000	CBKE000000000	CBKF000000000	CBKG000000000
CBKI000000000	CBKJ000000000	CBKK000000000	CBKL000000000	CBKM000000000	CBKN000000000	CBKO000000000	CBKP000000000
CBKR000000000	CBKS000000000	CBKT000000000	CBKU000000000	CBKV000000000	CBLF000000000	CBLJ000000000	CBLK000000000
CBLL000000000	CBLM000000000	CBLP000000000	CBLQ000000000	CBLR000000000	CBLS000000000	CBLT000000000	CBLU000000000
CBLX000000000	CBLY000000000	CBMK000000000	CBMO000000000	CBQO000000000	CBQQ000000000	CBQR000000000	CBQT000000000
CBQU000000000	CBRO000000000	CBSD000000000	CBST000000000	CBSU000000000	CBSV000000000	CBSW000000000	CBSY000000000
CBSZ000000000	CBTA000000000	CBTB000000000	CBTK000000000	CBTQ000000000	CBTR000000000	CBTU000000000	CBTZ000000000
CBUB000000000	CBUH000000000	CBUJ000000000	CBUK000000000	CBUL000000000	CBUM000000000	CBUN000000000	CBVB000000000
CBVI000000000	CBVJ000000000	CBVR000000000	CBWB000000000	CBWE000000000	CBWF000000000	CBWG000000000	CBWL000000000
CBWO000000000	CBWP000000000	CBWT000000000	CBWU000000000	CBWX000000000	CBXG000000000	CBXH000000000	CBXI000000000
CBXL000000000	CBXW000000000	CBYL000000000	CBYM000000000	CBYN000000000	CBYP000000000	CBZM000000000	CBZR000000000
CCAB000000000	CCAC000000000	CCAE000000000	CCAH000000000	CCAI000000000	CCAJ000000000	CCAL000000000	CCAQ000000000
CCAR000000000	CCAT000000000	CCAU000000000	CCAV000000000	CCAW000000000	CCAY000000000	CCAZ000000000	CCBA000000000
CCBD000000000	CCBE000000000	CCBF000000000	CCBG000000000	CCBI000000000	CCCL000000000	CCCO000000000	CCCQ000000000
CCDF00000000	CCDG000000000	CCDJ000000000	CCDM000000000	CCDN000000000	CCDO000000000	CCDQ000000000	CCDR000000000
CCDS00000000	CCDT00000000	CCDU00000000	CCDV00000000	CCET00000000	CCEU00000000	CCEZ00000000	CCFB00000000
CCFE00000000	CCMM00000000	CCMN00000000	JAAC00000000	JAAD00000000	JAAE00000000	JAAF00000000	JAAG00000000
JAAH00000000	JAAJ00000000	JAAN00000000	JAAO00000000	JACA00000000	JACB00000000	JACC00000000	JACD00000000
JACG00000000	JACR00000000	JACS00000000	JACW00000000	JACX00000000	JACY00000000	JADB00000000	JADC00000000
JADK00000000	JADO00000000	JADU00000000	JADV00000000	JADW00000000	JAEA00000000	JAEB00000000	JAED00000000
JAEE00000000	JAEF00000000	JAEI00000000	JAEM00000000	JAEN00000000	JAEO00000000	JAER00000000	JAES00000000
JAET00000000	JAEV00000000	JAEW00000000	JAEX00000000	JAEZ00000000	JAFD00000000	JAFH00000000	JAFU00000000
JAFW00000000	JAGB00000000	JAGD00000000	JAGE00000000	JAGF00000000	JAGG00000000	JAGH00000000	JAGI00000000
JAGJ00000000	JAGK00000000	JAGQ00000000	JAGR00000000	JAGS00000000	JAGT00000000	JAGU00000000	JAHA00000000
JAII00000000	JAIK00000000	JAIL00000000	JAIM00000000	JAIN00000000	JAJA00000000	JAJB00000000	JAJD00000000
JAKX00000000	JAKY00000000	JAKZ00000000	JALA00000000	JALB00000000	JALC00000000	JALD00000000	JALJ00000000

Genome ID							
JALK00000000	JALN00000000	JALO00000000	JALQ00000000	JALR00000000	JALS00000000	JALT00000000	JALU00000000
JALV00000000	JALW00000000	JALX00000000	JANA00000000	JANB00000000	JANC00000000	JAND00000000	JANE00000000
JANL00000000	JANS00000000	JANT00000000	JANV00000000	JANZ00000000	JAOE00000000	JAOI00000000	JAOQ00000000
JAOR00000000	JAOZ00000000	JAPA00000000	JAPB00000000	JAPQ00000000	JAPW00000000	JAPX00000000	JAPY00000000
JAQC00000000	JAQG00000000	JAQK00000000	JAQM00000000	JAQS00000000	JAQT00000000	JAQU00000000	JAQV00000000
JARC00000000	JARP00000000	JARQ00000000	JASW00000000	JASX00000000	JASY00000000	JASZ00000000	JATL00000000
JATP00000000	JATQ00000000	JATR00000000	JATS00000000	JATT00000000	JATV00000000	JAUF00000000	JAUG00000000
JAUH00000000	JAUI00000000	JAUJ00000000	JBEZ00000000	JBOK00000000	JBOY00000000	JBTQ00000000	JCEY00000000
JCGQ00000000	JCKH00000000	JCKI00000000	JCKK00000000	JCKL00000000	JCKM00000000	JCKN00000000	JCKO00000000
JCKP00000000	JCKQ00000000	JCKR00000000	JCKS00000000	JCKT00000000	JCKU00000000	JCKV00000000	JCKW00000000
JCKX00000000	JCKY00000000	JCKZ00000000	JCLA00000000	JCLB00000000	JCLC00000000	JCLD00000000	JCLE00000000
JCLF00000000	JCLG00000000	JCLH00000000	JCLI00000000	JCLJ00000000	JCLK00000000	JCLL00000000	JCLM00000000
JCLN00000000	JCLO00000000	JCLP00000000	JCLQ00000000	JCLR00000000	JCLS00000000	JCLT00000000	JCLU00000000
JCLV00000000	JCLW00000000	JCMD00000000	JCML00000000	JCNU00000000	JCNZ00000000	JDFF00000000	JDFG00000000
JDFH00000000	JDFI00000000	JDFJ00000000	JDFL00000000	JDFP00000000	JDFQ00000000	JDFR00000000	JDFT00000000
JDRS00000000	JDRW00000000	JDSI00000000	JDTI00000000	JDVA00000000	JDVC00000000	JDVE00000000	JDVW00000000
JDWB00000000	JDWC00000000	JDWD00000000	JDWG00000000	JDWH00000000	JDWN00000000	JDWQ00000000	JDWU00000000
JDYJ00000000	JDYK00000000	JELS00000000	JEMF00000000	JEMH00000000	JEMJ00000000	JEMK00000000	JEML00000000
JEMQ00000000	JEMV00000000	JEMY00000000	JENA00000000	JENC00000000	JEND00000000	JENK00000000	JENL00000000
JENM00000000	JENN00000000	JENU00000000	JENV00000000	JENW00000000	JENX00000000	JENY00000000	JEOA00000000
JEOB00000000	JEOD00000000	JEOF00000000	JEWP00000000	JEYL00000000	JEZH00000000	JFAF00000000	JFBQ00000000
JFBR00000000	JFBS00000000	JFBT00000000	JFBU00000000	JFBV00000000	JFBW00000000	JFBZ00000000	JFCA00000000
JFCF00000000	JFFS00000000	JFGP00000000	JFGS00000000	JFHC00000000	JFHD00000000	JFHE00000000	JFHN00000000
JFHO00000000	JFHR00000000	JFHT00000000	JFHU00000000	JFHV00000000	JFHW00000000	JFHX00000000	JFHY00000000
JFHZ00000000	JFJK00000000	JFJO00000000	JFJV00000000	JFJW00000000	JFLT00000000	JFOJ00000000	JFVK00000000
JFVQ00000000	JFVW00000000	JFWF00000000	JFYM00000000	JFYN00000000	JFYP00000000	JFYQ00000000	JFYR00000000
JFYS00000000	JFYT00000000	JFYU00000000	JFYV00000000	JFYW00000000	JFYX00000000	JFZB00000000	JFZC00000000
JFZG00000000	JFZJ00000000	JFZK00000000	JFZR00000000	JFZV00000000	JFZW00000000	JGAG00000000	JGBF00000000
JGBR00000000	JGCJ00000000	JGCP00000000	JGCR00000000	JGCW00000000	JGCZ00000000	JGDA00000000	JGDB00000000
JGDF00000000	JGDI00000000	JGDL00000000	JGDM00000000	JGDN00000000	JGDQ00000000	JGDR00000000	JGDS00000000
JGDT00000000	JGDU00000000	JGED00000000	JGEF00000000	JGEK00000000	JGEN00000000	JGEP00000000	JGUS00000000
JGVB00000000	JGVI00000000	JGVN00000000	JGVP00000000	JGVR00000000	JGVU00000000	JGWC00000000	JGWG00000000
JGWH00000000	JGWJ00000000	JGWL00000000	JGWM00000000	JGWO00000000	JGWP00000000	JGWQ00000000	JGWR00000000
JGWS00000000	JGWT00000000	JGWV00000000	JGWW00000000	JGWX00000000	JGWZ00000000	JGXB00000000	JGXC00000000
JGXD00000000	JGXF00000000	JGXG00000000	JGXJ00000000	JGXM00000000	JGXU00000000	JGXV00000000	JGXW00000000
JGXY00000000	JGXZ00000000	JGYA00000000	JGYB00000000	JGYF00000000	JGYH00000000	JGYJ00000000	JGYK00000000
JGYL00000000	JGYM00000000	JGYN00000000	JGYO00000000	JGYP00000000	JGYQ00000000	JGYR00000000	JGYS00000000
JGYT00000000	JGYU00000000	JGYV00000000	JGYW00000000	JGYX00000000	JGYY00000000	JGYZ00000000	JGZA00000000
JGZB00000000	JGZC00000000	JGZD00000000	JGZE00000000	JGZF00000000	JGZG00000000	JGZH00000000	JGZI00000000
JGZJ00000000	JGZK00000000	JGZL00000000	JGZM00000000	JGZN00000000	JGZO00000000	JGZP00000000	JGZQ00000000
JGZR00000000	JGZS00000000	JGZT00000000	JGZU00000000	JGZV00000000	JGZW00000000	JHAC00000000	JHAL00000000
JHAM00000000	JHBO00000000	JHBT00000000	JHBU00000000	JHBV00000000	JHBZ00000000	JHCX00000000	JHDF00000000
JHDK00000000	JHDN00000000	JHDV00000000	JHDW00000000	JHDX00000000	JHEB00000000	JHEL00000000	JHEM00000000
JHEN00000000	JHEO00000000	JHER00000000	JHES00000000	JHET00000000	JHOF00000000	JHOK00000000	JHPN00000000
JHPQ00000000	JHQB00000000	JHQH00000000	JHQI00000000	JHQQ00000000	JHTS00000000	JHUA00000000	JHUB00000000
JHUR00000000	JHUS00000000	JHUW00000000	JHUX00000000	JHVC00000000	JHVD00000000	JHVE00000000	JHVI00000000
JHVK00000000	JHVL00000000	JHVO00000000	JHVQ00000000	JHVU00000000	JHVX00000000	JHWA00000000	JHWB00000000
JHWC00000000	JHWH00000000	JHWI00000000	JHWJ00000000	JHWK00000000	JHWL00000000	JHWM00000000	JHWN00000000
JHWO00000000	JHWP00000000	JHWQ00000000	JHWR00000000	JHWS00000000	JHWU00000000	JHWV00000000	JHWW00000000
JHWY00000000	JHWZ00000000	JHXA00000000	JHXB00000000	JHXC00000000	JHXD00000000	JHXE00000000	JHXF00000000
JHXG00000000	JHXH00000000	JHXI00000000	JHXJ00000000	JHXK00000000	JHXL00000000	JHXM00000000	JHXO00000000
JHXQ00000000	JHXV00000000	JHXW00000000	JHXX00000000	JHXY00000000	JHXZ00000000	JHYD00000000	JHYF00000000
JHYG00000000	JHYH00000000	JHYI00000000	JHYJ00000000	JHYK00000000	JHYL00000000	JHYR00000000	JHYT00000000
JHYZ00000000	JHZC00000000	JHZG00000000	JHZH00000000	JHZO00000000	JHZS00000000	JHZT00000000	JHZU00000000
JHZV00000000	JHZW00000000	JHZX00000000	JHZY00000000	JHZZ00000000	JIAA00000000	JIAB00000000	JIAC00000000
JIAD00000000	JIAE00000000	JIAF00000000	JIAH00000000	JIAJ00000000	JIAL00000000	JIAM00000000	JIAR00000000
JIAS00000000	JIAT00000000	JIAU00000000	JIAV00000000	JIAW00000000	JIBA00000000	JIBB00000000	JIBD00000000
JIBE00000000	JIBF00000000	JIBG00000000	JIBJ00000000	JIBO00000000	JIBP00000000	JIBQ00000000	JIBR00000000
JIBS00000000	JIBU00000000	JIBX00000000	JICI00000000	JIDX00000000	JIEC00000000	JIEL00000000	JIEY00000000
JILJ00000000	JIUK00000000	JIXM00000000	JIZS00000000	JJAO00000000	JJDX00000000	JJFU00000000	JJJD00000000
JJJL00000000	JJLU00000000	JJLV00000000	JJMB00000000	JJMG00000000	JJMK00000000	JJMM00000000	JJMN00000000
JJMR00000000	JJMS00000000	JJMT00000000	JJMW00000000	JJMZ00000000	JJNA00000000	JJNL00000000	JJNT00000000
JJNW00000000	JJNX00000000	JJNY00000000	JJOA00000000	JJOP00000000	JJRY00000000	JMCD00000000	JMDW00000000
JMEA00000000	JMEC00000000	JMEF00000000	JMEG00000000	JMEL00000000	JMGB00000000	JMGO00000000	JMGQ00000000
JMGR00000000	JMGX00000000	JMHU00000000	JMHW00000000	JMIG00000000	JMIH00000000	JMIL00000000	JMIQ00000000
JMIR00000000	JMIS00000000	JMIW00000000	JMIX00000000	JMJF00000000	JMJL00000000	JMKI00000000	JMKP00000000
JMKQ00000000	JMKS00000000	JMKU00000000	JMKV00000000	JMLB00000000	JMLC00000000	JMLG00000000	JMLH00000000
JMLI00000000	JMLK00000000	JMLQ00000000	JMLU00000000	JMLX00000000	JMLZ00000000	JMMA00000000	JMMB00000000
JMMC00000000	JMMD00000000	JMMU00000000	JMNT00000000	JMNU00000000	JMOU00000000	JMOW00000000	JMPB00000000
JMPD00000000	JMPF00000000	JMPG00000000	JMPJ00000000	JMPK00000000	JMPL00000000	JMPM00000000	JMPN00000000
JMPO00000000	JMPP00000000	JMPS00000000	JMQF00000000	JMQQ00000000	JMQR00000000	JMRL00000000	JMRT00000000
JMRV00000000	JMRW00000000	JMRX00000000	JMSI00000000	JMSQ00000000	JMTA00000000	JMTB00000000	JMTZ00000000
JMUJ00000000	JMUK00000000	JMUL00000000	JMUM00000000	JMUN00000000	JMUO00000000	JMUP00000000	JMUQ00000000
JMUR00000000	JMUS00000000	JMUT00000000	JMUU00000000	JMUV00000000	JMUX00000000	JMVG00000000	JMVT00000000
JMVV00000000	JMWB00000000	JMWE00000000	JMWF00000000	JMYW00000000	JMYX00000000	JMZA00000000	JMZD00000000
JMZJ00000000	JMZK00000000	JMZM00000000	JMZP00000000	JMZV00000000	JMZX00000000	JMZY00000000	JMZZ00000000
JNAA00000000	JNAB00000000	JNAE00000000	JNBN00000000	JNBY00000000	JNCA00000000	JNCB00000000	JNCC00000000
JNCG00000000	JNCH00000000	JNCM00000000	JNCP00000000	JNCU00000000	JNCV00000000	JNFE00000000	JNFF00000000
JNFG00000000	JNFL00000000	JNFS00000000	JNFT00000000	JNGC00000000	JNGD00000000	JNGW00000000	JNHD00000000
JNHE00000000	JNHF00000000	JNHG00000000	JNHH00000000	JNHI00000000	JNHJ00000000	JNHK00000000	JNHL00000000

Genome ID							
JNHM00000000	JNHN00000000	JNHO00000000	JNHP00000000	JNHQ00000000	JNHR00000000	JNHT00000000	JNHU00000000
JNIG00000000	JNII00000000	JNIJ00000000	JNIK00000000	JNIL00000000	JNIN00000000	JNIO00000000	JNIP00000000
JNIQ00000000	JNIR00000000	JNIS00000000	JNIW00000000	JNIX00000000	JNJA00000000	JNJB00000000	JNJC00000000
JNJD00000000	JNJE00000000	JNJF00000000	JNJH00000000	JNJI00000000	JNJK00000000	JNJM00000000	JNJN00000000
JNJO00000000	JNJP00000000	JNJQ00000000	JNJS00000000	JNKB00000000	JNKC00000000	JNKE00000000	JNKF00000000
JNKG00000000	JNKH00000000	JNKI00000000	JNKJ00000000	JNKK00000000	JNKM00000000	JNKN00000000	JNKO00000000
JNKP00000000	JNKQ00000000	JNKR00000000	JNKS00000000	JNKT00000000	JNKU00000000	JNKW00000000	JNKX00000000
JNKY00000000	JNKZ00000000	JNLA00000000	JNLE00000000	JNLF00000000	JNLG00000000	JNLH00000000	JNLK00000000
JNLM00000000	JNLN00000000	JNLP00000000	JNLQ00000000	JNLR00000000	JNLX00000000	JNMD00000000	JNMG00000000
JNMH00000000	JNMJ00000000	JNMK00000000	JNMN00000000	JNMO00000000	JNMR00000000	JNMS00000000	JNMX00000000
JNNC00000000	JNNH00000000	JNNT00000000	JNNV00000000	JNOE00000000	JNOG00000000	JNOL00000000	JNOS00000000
JNOZ00000000	JNPD00000000	JNPE00000000	JNPK00000000	JNPO00000000	JNPT00000000	JNQD00000000	JNQK00000000
JNQP00000000	JNQX00000000	JNRG00000000	JNRH00000000	JNRJ00000000	JNRO00000000	JNRU00000000	JNRV00000000
JNSA00000000	JNVA00000000	JNVC00000000	JNVM00000000	JNVP00000000	JNVX00000000	JNVY00000000	JNVZ00000000
JNWA00000000	JNWB00000000	JNWC00000000	JNWD00000000	JNWH00000000	JNWI00000000	JOKF00000000	JOKJ00000000
JOKP00000000	JOKX00000000	JOMH00000000	JOMI00000000	JOMS00000000	JONC00000000	JONE00000000	JONJ00000000
JONM00000000	JONN00000000	JONS00000000	JONT00000000	JONU00000000	JONV00000000	JONW00000000	JOOA00000000
JOOE00000000	JOOH00000000	JOPN00000000	JOPR00000000	JOPS00000000	JOPU00000000	JOPV00000000	JOQR00000000
JOQS00000000	JOQX00000000	JOQY00000000	JORF00000000	JORS00000000	JORW00000000	JOSB00000000	JOSE00000000
JOSK00000000	JOSQ00000000	JOST00000000	JOSU00000000	JOTM00000000	JOUB00000000	JOUE00000000	JOVI00000000
JOVP00000000	JPDJ00000000	JPDK00000000	JPDL00000000	JPDM00000000	JPDP00000000	JPEF00000000	JPEG00000000
JPEH00000000	JPEJ00000000	JPEK00000000	JPEL00000000	JPEP00000000	JPEQ00000000	JPFK00000000	JPFN00000000
JPFS00000000	JPFT00000000	JPFU00000000	JPFV00000000	JPFW00000000	JPFX00000000	JPFY00000000	JPFZ00000000
JPGA00000000	JPGB00000000	JPGC00000000	JPGL00000000	JPGM00000000	JPGN00000000	JPGQ00000000	JPGW00000000
JPGX00000000	JPGY00000000	JPHC00000000	JPHD00000000	JPHE00000000	JPHG00000000	JPHP00000000	JPHS00000000
JPHU00000000	JPIC00000000	JPJC00000000	JPJI00000000	JPKG00000000	JPKN00000000	JPLY00000000	JPLZ00000000
JPMC00000000	JPMD00000000	JPME00000000	JPMI00000000	JPNB00000000	JPNN00000000	JPNZ00000000	JPOB00000000
JPOD00000000	JPOE00000000	JPOF00000000	JPOI00000000	JPOL00000000	JPOO00000000	JPOQ00000000	JPOU00000000
JPQG00000000	JPQU00000000	JPQV00000000	JPQW00000000	JPQY00000000	JPRG00000000	JPRH00000000	JPRI00000000
JPRJ00000000	JPRL00000000	JPRM00000000	JPRN00000000	JPRO00000000	JPRP00000000	JPRS00000000	JPRT00000000
JPRU00000000	JPSM00000000	JPSN00000000	JPSO00000000	JPUG00000000	JPUI00000000	JPUX00000000	JPVW00000000
JPVX00000000	JPWY00000000	JPXI00000000	NC_000907	NC_000914	NC_000958	NC_000964	NC_001988
NC_002662	NC_002696	NC_002936	NC_002940	NC_002944	NC_002947	NC_002950	NC_002967
NC_002971	NC_003042	NC_003080	NC_003228	NC_003240	NC_003450	NC_003454	NC_003888
NC_003902	NC_004129	NC_004252	NC_004307	NC_004319	NC_004567	NC_004578	NC_004604
NC_004663	NC_004923	NC_005090	NC_005125	NC_005241	NC_005244	NC_005296	NC_005362
NC_005363	NC_005773	NC_005835	NC_005966	NC_006087	NC_006138	NC_006177	NC_006297
NC_006361	NC_006448	NC_006449	NC_006461	NC_006529	NC_006576	NC_006814	NC_006834
NC_006958	NC_007005	NC_007086	NC_007146	NC_007168	NC_007204	NC_007298	NC_007333
NC_007350	NC_007404	NC_007406	NC_007410	NC_007481	NC_007486	NC_007488	NC_007492
NC_007498	NC_007503	NC_007509	NC_007576	NC_007595	NC_007681	NC_007705	NC_007722
NC_007760	NC_007777	NC_007794	NC_007901	NC_007907	NC_007968	NC_007971	NC_008010
NC_008011	NC_008023	NC_008027	NC_008036	NC_008054	NC_008060	NC_008146	NC_008148
NC_008255	NC_008261	NC_008262	NC_008343	NC_008346	NC_008378	NC_008385	NC_008496
NC_008500	NC_008502	NC_008503	NC_008525	NC_008529	NC_008530	NC_008536	NC_008537
NC_008542	NC_008554	NC_008570	NC_008578	NC_008593	NC_008595	NC_008596	NC_008618
NC_008686	NC_008697	NC_008703	NC_008726	NC_008738	NC_008752	NC_008765	NC_008825
NC_009004	NC_009009	NC_009012	NC_009040	NC_009077	NC_009142	NC_009338	NC_009342
NC_009425	NC_009439	NC_009441	NC_009445	NC_009454	NC_009466	NC_009467	NC_009475
NC_009507	NC_009512	NC_009513	NC_009515	NC_009516	NC_009566	NC_009567	NC_009614
NC_009615	NC_009617	NC_009633	NC_009656	NC_009660	NC_009663	NC_009713	NC_009717
NC_009718	NC_009719	NC_009726	NC_009767	NC_009785	NC_009792	NC_009795	NC_009800
NC_009943	NC_009952	NC_010001	NC_010002	NC_010070	NC_010080	NC_010115	NC_010123
NC_010162	NC_010168	NC_010172	NC_010175	NC_010320	NC_010322	NC_010337	NC_010371
NC_010373	NC_010380	NC_010381	NC_010466	NC_010485	NC_010501	NC_010502	NC_010528
NC_010545	NC_010549	NC_010551	NC_010554	NC_010572	NC_010582	NC_010609	NC_010610
NC_010617	NC_010655	NC_010678	NC_010688	NC_010717	NC_010721	NC_010729	NC_010801
NC_010813	NC_010935	NC_010943	NC_010999	NC_011004	NC_011071	NC_011126	NC_011134
NC_011143	NC_011146	NC_011281	NC_011295	NC_011352	NC_011365	NC_011366	NC_011527
NC_011561	NC_011593	NC_011662	NC_011666	NC_011740	NC_011745	NC_011757	NC_011768
NC_011830	NC_011835	NC_011836	NC_011879	NC_011883	NC_011898	NC_011916	NC_011958
NC_011992	NC_011995	NC_012032	NC_012108	NC_012121	NC_012489	NC_012526	NC_012552
NC_012660	NC_012669	NC_012673	NC_012704	NC_012778	NC_012781	NC_012791	NC_012803
NC_012807	NC_012814	NC_012815	NC_012848	NC_012849	NC_012891	NC_012913	NC_012914
NC_012923	NC_012984	NC_012987	NC_013037	NC_013061	NC_013062	NC_013132	NC_013159
NC_013162	NC_013164	NC_013165	NC_013166	NC_013169	NC_013170	NC_013172	NC_013173
NC_013192	NC_013198	NC_013199	NC_013203	NC_013204	NC_013216	NC_013235	NC_013316
NC_013406	NC_013410	NC_013441	NC_013515	NC_013520	NC_013521	NC_013656	NC_013714
NC_013715	NC_013716	NC_013721	NC_013722	NC_013729	NC_013730	NC_013739	NC_013740
NC_013757	NC_013798	NC_013850	NC_013851	NC_013853	NC_013854	NC_013893	NC_013895
NC_013921	NC_013946	NC_014005	NC_014011	NC_014033	NC_014034	NC_014041	NC_014100
NC_014103	NC_014106	NC_014107	NC_014131	NC_014147	NC_014158	NC_014168	NC_014169
NC_014209	NC_014215	NC_014221	NC_014246	NC_014248	NC_014258	NC_014259	NC_014304
NC_014307	NC_014319	NC_014323	NC_014330	NC_014355	NC_014363	NC_014366	NC_014370
NC_014375	NC_014376	NC_014387	NC_014391	NC_014393	NC_014410	NC_014483	NC_014484
NC_014494	NC_014498	NC_014501	NC_014506	NC_014538	NC_014548	NC_014554	NC_014614
NC_014616	NC_014624	NC_014632	NC_014638	NC_014643	NC_014644	NC_014656	NC_014724
NC_014727	NC_014734	NC_014738	NC_014752	NC_014811	NC_014815	NC_014816	NC_014824
NC_014828	NC_014829	NC_014830	NC_014834	NC_014837	NC_014844	NC_014908	NC_014920
NC_014922	NC_014924	NC_014925	NC_014931	NC_014932	NC_014933	NC_014964	NC_014972

Genome ID							
NC_014974	NC_014976	NC_015052	NC_015061	NC_015067	NC_015125	NC_015138	NC_015160
NC_015161	NC_015164	NC_015172	NC_015178	NC_015183	NC_015213	NC_015215	NC_015275
NC_015278	NC_015291	NC_015311	NC_015312	NC_015321	NC_015385	NC_015389	NC_015390
NC_015420	NC_015422	NC_015437	NC_015460	NC_015500	NC_015501	NC_015514	NC_015519
NC_015555	NC_015556	NC_015563	NC_015565	NC_015566	NC_015567	NC_015571	NC_015576
NC_015579	NC_015588	NC_015589	NC_015598	NC_015600	NC_015601	NC_015634	NC_015635
NC_015656	NC_015660	NC_015663	NC_015671	NC_015677	NC_015678	NC_015686	NC_015690
NC_015693	NC_015696	NC_015697	NC_015723	NC_015733	NC_015737	NC_015738	NC_015756
NC_015760	NC_015846	NC_015859	NC_015873	NC_015875	NC_015913	NC_015930	NC_015947
NC_015958	NC_015963	NC_015964	NC_015974	NC_015975	NC_015977	NC_015978	NC_016001
NC_016012	NC_016021	NC_016023	NC_016046	NC_016052	NC_016077	NC_016147	NC_016510
NC_016514	NC_016584	NC_016589	NC_016593	NC_016594	NC_016603	NC_016604	NC_016605
NC_016609	NC_016610	NC_016612	NC_016616	NC_016627	NC_016630	NC_016633	NC_016640
NC_016641	NC_016749	NC_016776	NC_016791	NC_016805	NC_016808	NC_016809	NC_016812
NC_016816	NC_016818	NC_016826	NC_016887	NC_016906	NC_016935	NC_016943	NC_017022
NC_017027	NC_017045	NC_017047	NC_017067	NC_017068	NC_017075	NC_017079	NC_017138
NC_017175	NC_017194	NC_017195	NC_017199	NC_017214	NC_017215	NC_017216	NC_017217
NC_017218	NC_017219	NC_017220	NC_017249	NC_017267	NC_017271	NC_017272	NC_017294
NC_017304	NC_017317	NC_017353	NC_017442	NC_017448	NC_017451	NC_017452	NC_017456
NC_017464	NC_017465	NC_017467	NC_017469	NC_017470	NC_017473	NC_017474	NC_017477
NC_017479	NC_017482	NC_017483	NC_017490	NC_017491	NC_017492	NC_017530	NC_017532
NC_017549	NC_017558	NC_017563	NC_017568	NC_017569	NC_017573	NC_017576	NC_017581
NC_017583	NC_017587	NC_017594	NC_017595	NC_017617	NC_017626	NC_017633	NC_017639
NC_017646	NC_017671	NC_017672	NC_017764	NC_017768	NC_017769	NC_017771	NC_017803
NC_017834	NC_017860	NC_017866	NC_017867	NC_017905	NC_017911	NC_017927	NC_017949
NC_017955	NC_017999	NC_018000	NC_018011	NC_018017	NC_018022	NC_018066	NC_018073
NC_018079	NC_018106	NC_018142	NC_018221	NC_018289	NC_018405	NC_018485	NC_018513
NC_018515	NC_018520	NC_018525	NC_018580	NC_018594	NC_018604	NC_018605	NC_018607
NC_018609	NC_018631	NC_018641	NC_018665	NC_018673	NC_018684	NC_018704	NC_018708
NC_018712	NC_018720	NC_018742	NC_018866	NC_018867	NC_019425	NC_019430	NC_019567
NC_019670	NC_019675	NC_019682	NC_019695	NC_019703	NC_019733	NC_019738	NC_019744
NC_019771	NC_019896	NC_019897	NC_019903	NC_019905	NC_019954	NC_019956	NC_019960
NC_020063	NC_020125	NC_020127	NC_020134	NC_020164	NC_020180	NC_020209	NC_020229
NC_020230	NC_020244	NC_020291	NC_020302	NC_020418	NC_020450	NC_020506	NC_020507
NC_020515	NC_020516	NC_020517	NC_020519	NC_020526	NC_020542	NC_020546	NC_020813
NC_020819	NC_020829	NC_020832	NC_020887	NC_020990	NC_020995	NC_021008	NC_021009
NC_021010	NC_021011	NC_021012	NC_021013	NC_021014	NC_021015	NC_021016	NC_021017
NC_021018	NC_021019	NC_021020	NC_021021	NC_021022	NC_021023	NC_021024	NC_021030
NC_021031	NC_021035	NC_021038	NC_021039	NC_021040	NC_021041	NC_021042	NC_021043
NC_021044	NC_021046	NC_021047	NC_021064	NC_021066	NC_021081	NC_021082	NC_021171
NC_021175	NC_021181	NC_021182	NC_021184	NC_021191	NC_021200	NC_021219	NC_021224
NC_021235	NC_021237	NC_021277	NC_021280	NC_021284	NC_021351	NC_021352	NC_021353
NC_021354	NC_021362	NC_021487	NC_021491	NC_021492	NC_021494	NC_021499	NC_021505
NC_021507	NC_021514	NC_021591	NC_021593	NC_021658	NC_021661	NC_021721	NC_021723
NC_021725	NC_021741	NC_021744	NC_021821	NC_021872	NC_021900	NC_021987	NC_022000
NC_022040	NC_022041	NC_022097	NC_022111	NC_022112	NC_022115	NC_022196	NC_022198
NC_022234	NC_022236	NC_022237	NC_022238	NC_022239	NC_022244	NC_022245	NC_022246
NC_022356	NC_022369	NC_022513	NC_022523	NC_022526	NC_022532	NC_022549	NC_022567
NC_022571	NC_022582	NC_022584	NC_022587	NC_022592	NC_022600	NC_022601	NC_022737
NC_022738	NC_022780	NC_022794	NC_022878	NC_022898	NC_022909	NC_022997	NC_023024
NC_023036	NC_023061	NC_023064	NC_023075	NC_023076	NC_023134	NC_023144	

b, Reference genomes collected from ref. 124

Genomes from reference 22
Clostridiales bacterium 2E1
Clostridiales bacterium 1E3
Clostridiales bacterium 2E3
Clostridiales bacterium 1E11
Clostridiales bacterium 1A9
Bifidobacterium pseudolongum 1B11
Clostridiales bacterium 1C12
Bifidobacterium breve 1C2
Clostridiales bacterium 1D10
Clostridiales bacterium 1D1
Clostridiales bacterium 1D2
Clostridiales bacterium 1D4
Clostridiales bacterium 1F7
Clostridiales bacterium 1F8
Clostridiales bacterium 2D9
Clostridiales bacterium 2F7
Bacteroides dorei 2G11
Clostridiales bacterium 2G4
Clostridiales bacterium 2H11
Clostridiales bacterium 2H6

c, reference genomes sequenced in my laboratory

Genomes sequenced in my laboratory
Slackia sp. 2F
Fusobacterium varium 70

Appendix 3. Details of deposited sequence files

File name	Individual	Time point	Sequencer	Accession number
apr01.1.MiSeq.fastq	'apr01	S00	MiSeq	DRR042264
apr01.2.MiSeq.fastq	'apr01	S00	MiSeq	DRR042265
apr01S00.454.fastq	'apr01	S00	454	DRR042266
apr01S00.PGM.fastq	'apr01	S00	Ion PGM	DRR042267
apr01S02.454.fastq	'apr01	S02	454	DRR042268
apr01S04.454.fastq	'apr01	S04	454	DRR042269
apr01S06.454.fastq	'apr01	S06	454	DRR042270
apr01S08.454.fastq	'apr01	S08	454	DRR042271
apr02.1.MiSeq.fastq	'apr02	S00	MiSeq	DRR042272
apr02.2.MiSeq.fastq	'apr02	S00	MiSeq	DRR042273
apr02S00.454.fastq	'apr02	S00	454	DRR042274
apr02S00.PGM.fastq	'apr02	S00	Ion PGM	DRR042275
apr02S02.454.fastq	'apr02	S02	454	DRR042276
apr02S04.454.fastq	'apr02	S04	454	DRR042277
apr02S06.454.fastq	'apr02	S06	454	DRR042278
apr02S08.454.fastq	'apr02	S08	454	DRR042279
apr03.1.MiSeq.fastq	'apr03	S00	MiSeq	DRR042280
apr03.2.MiSeq.fastq	'apr03	S00	MiSeq	DRR042281
apr03S00.454.fastq	'apr03	S00	454	DRR042282
apr03S00.PGM.fastq	'apr03	S00	Ion PGM	DRR042283
apr03S02.454.fastq	'apr03	S02	454	DRR042284
apr03S04.454.fastq	'apr03	S04	454	DRR042285
apr03S06.454.fastq	'apr03	S06	454	DRR042286
apr03S08.454.fastq	'apr03	S08	454	DRR042287
apr05.1.MiSeq.fastq	'apr05	S00	MiSeq	DRR042288
apr05.2.MiSeq.fastq	'apr05	S00	MiSeq	DRR042289
apr05S00.454.fastq	'apr05	S00	454	DRR042290
apr05S08.454.fastq	'apr05	S08	454	DRR042291
apr06S00.454.fastq	'apr06	S00	454	DRR042292
apr06S00.PGM.fastq	'apr06	S00	Ion PGM	DRR042293
apr06S00.Proton.1.fastq	'apr06	S00	Ion Proton	DRR042294
apr06S08.454.fastq	'apr06	S08	454	DRR042295
apr07S00.454.fastq	'apr07	S00	454	DRR042296
apr07S00.PGM.fastq	'apr07	S00	Ion PGM	DRR042297
apr07S00.Proton.1.fastq	'apr07	S00	Ion Proton	DRR042298
apr07S08.454.fastq	'apr07	S08	454	DRR042299
apr08S00.454.fastq	'apr08	S00	454	DRR042300
apr08S00.PGM.fastq	'apr08	S00	Ion PGM	DRR042301
apr08S00.Proton.1.fastq	'apr08	S00	Ion Proton	DRR042302
apr08S08.454.fastq	'apr08	S08	454	DRR042303
apr09.1.MiSeq.fastq	'apr09	S00	MiSeq	DRR042304
apr09.2.MiSeq.fastq	'apr09	S00	MiSeq	DRR042305
apr09S00.454.fastq	'apr09	S00	454	DRR042306
apr09S00.PGM.fastq	'apr09	S00	Ion PGM	DRR042307
apr09S02.454.fastq	'apr09	S02	454	DRR042308
apr09S04.454.fastq	'apr09	S04	454	DRR042309
apr09S06.454.fastq	'apr09	S06	454	DRR042310
apr09S08.454.fastq	'apr09	S08	454	DRR042311
apr10.1.MiSeq.fastq	'apr10	S00	MiSeq	DRR042312
apr10.2.MiSeq.fastq	'apr10	S00	MiSeq	DRR042313
apr10S00.454.fastq	'apr10	S00	454	DRR042314
apr10S08.454.fastq	'apr10	S08	454	DRR042315
apr11.1.MiSeq.fastq	'apr11	S00	MiSeq	DRR042316
apr11.2.MiSeq.fastq	'apr11	S00	MiSeq	DRR042317
apr11S00.454.fastq	'apr11	S00	454	DRR042318
apr11S08.454.fastq	'apr11	S08	454	DRR042319
apr12.1.MiSeq.fastq	'apr12	S00	MiSeq	DRR042320
apr12.2.MiSeq.fastq	'apr12	S00	MiSeq	DRR042321
apr12S00.454.fastq	'apr12	S00	454	DRR042322
apr12S00.PGM.fastq	'apr12	S00	Ion PGM	DRR042323
apr12S02.454.fastq	'apr12	S02	454	DRR042324
apr12S03.454.fastq	'apr12	S03	454	DRR042325
apr12S06.454.fastq	'apr12	S06	454	DRR042326
apr12S08.454.fastq	'apr12	S08	454	DRR042327
apr15.1.MiSeq.fastq	'apr15	S00	MiSeq	DRR042328
apr15.2.MiSeq.fastq	'apr15	S00	MiSeq	DRR042329
apr15S00.454.fastq	'apr15	S00	454	DRR042330
apr15S08.454.fastq	'apr15	S08	454	DRR042331

File name	Individual	Time point	Sequencer	Accession number
apr16.1.MiSeq.fastq	'apr16	S00	MiSeq	DRR042332
apr16.2.MiSeq.fastq	'apr16	S00	MiSeq	DRR042333
apr16S00.454.fastq	'apr16	S00	454	DRR042334
apr16S00.PGM.fastq	'apr16	S00	Ion PGM	DRR042335
apr16S02.454.fastq	'apr16	S02	454	DRR042336
apr16S04.454.fastq	'apr16	S04	454	DRR042337
apr16S06.454.fastq	'apr16	S06	454	DRR042338
apr16S08.454.fastq	'apr16	S08	454	DRR042339
apr17.1.MiSeq.fastq	'apr17	S00	MiSeq	DRR042340
apr17.2.MiSeq.fastq	'apr17	S00	MiSeq	DRR042341
apr17S00.454.fastq	'apr17	S00	454	DRR042342
apr17S00.PGM.fastq	'apr17	S00	Ion PGM	DRR042343
apr17S02.454.fastq	'apr17	S02	454	DRR042344
apr17S04.454.fastq	'apr17	S04	454	DRR042345
apr17S06.454.fastq	'apr17	S06	454	DRR042346
apr17S08.454.fastq	'apr17	S08	454	DRR042347
apr18.1.MiSeq.fastq	'apr18	S00	MiSeq	DRR042348
apr18.2.MiSeq.fastq	'apr18	S00	MiSeq	DRR042349
apr18S00.454.fastq	'apr18	S00	454	DRR042350
apr18S08.454.fastq	'apr18	S08	454	DRR042351
apr19.1.MiSeq.fastq	'apr19	S00	MiSeq	DRR042352
apr19.2.MiSeq.fastq	'apr19	S00	MiSeq	DRR042353
apr19S00.454.fastq	'apr19	S00	454	DRR042354
apr19S08.454.fastq	'apr19	S08	454	DRR042355
apr21.1.MiSeq.fastq	'apr21	S00	MiSeq	DRR042356
apr21.2.MiSeq.fastq	'apr21	S00	MiSeq	DRR042357
apr21S00.454.fastq	'apr21	S00	454	DRR042358
apr21S08.454.fastq	'apr21	S08	454	DRR042359
apr22.1.MiSeq.fastq	'apr22	S00	MiSeq	DRR042360
apr22.2.MiSeq.fastq	'apr22	S00	MiSeq	DRR042361
apr22S00.454.fastq	'apr22	S00	454	DRR042362
apr22S08.454.fastq	'apr22	S08	454	DRR042363
apr23.1.MiSeq.fastq	'apr23	S00	MiSeq	DRR042364
apr23.2.MiSeq.fastq	'apr23	S00	MiSeq	DRR042365
apr23S00.454.fastq	'apr23	S00	454	DRR042366
apr23S08.454.fastq	'apr23	S08	454	DRR042367
apr24S00.454.fastq	'apr24	S00	454	DRR042368
apr24S08.454.fastq	'apr24	S08	454	DRR042369
apr25S00.454.fastq	'apr25	S00	454	DRR042370
apr25S08.454.fastq	'apr25	S08	454	DRR042371
apr26S00.454.fastq	'apr26	S00	454	DRR042372
apr26S08.454.fastq	'apr26	S08	454	DRR042373
apr27S00.454.fastq	'apr27	S00	454	DRR042374
apr27S08.454.fastq	'apr27	S08	454	DRR042375
apr28S00.454.fastq	'apr28	S00	454	DRR042376
apr28S08.454.fastq	'apr28	S08	454	DRR042377
apr30S00.454.fastq	'apr30	S00	454	DRR042378
apr30S08.454.fastq	'apr30	S08	454	DRR042379
apr31S00.454.fastq	'apr31	S00	454	DRR042380
apr31S08.454.fastq	'apr31	S08	454	DRR042381
apr32S00.454.fastq	'apr32	S00	454	DRR042382
apr32S08.454.fastq	'apr32	S08	454	DRR042383
apr33S00.454.fastq	'apr33	S00	454	DRR042384
apr33S08.454.fastq	'apr33	S08	454	DRR042385
apr34S00.454.fastq	'apr34	S00	454	DRR042386
apr34S08.454.fastq	'apr34	S08	454	DRR042387
apr35S00.454.fastq	'apr35	S00	454	DRR042388
apr35S08.454.fastq	'apr35	S08	454	DRR042389
apr36.1.MiSeq.fastq	'apr36	S00	MiSeq	DRR042390
apr36.2.MiSeq.fastq	'apr36	S00	MiSeq	DRR042391
apr36S00.454.fastq	'apr36	S00	454	DRR042392
apr36S08.454.fastq	'apr36	S08	454	DRR042393
apr37.1.MiSeq.fastq	'apr37	S00	MiSeq	DRR042394
apr37.2.MiSeq.fastq	'apr37	S00	MiSeq	DRR042395
apr37S00.454.fastq	'apr37	S00	454	DRR042396
apr37S08.454.fastq	'apr37	S08	454	DRR042397
apr38.1.MiSeq.fastq	'apr38	S00	MiSeq	DRR042398
apr38.2.MiSeq.fastq	'apr38	S00	MiSeq	DRR042399
apr38S00.454.fastq	'apr38	S00	454	DRR042400
apr38S08.454.fastq	'apr38	S08	454	DRR042401

File name	Individual	Time point	Sequencer	Accession number
apr39.1.MiSeq.fastq	'apr39	S00	MiSeq	DRR042402
apr39.2.MiSeq.fastq	'apr39	S00	MiSeq	DRR042403
apr39S00.454.fastq	'apr39	S00	454	DRR042404
apr39S00.PGM.fastq	'apr39	S00	Ion PGM	DRR042405
apr39S02.454.fastq	'apr39	S02	454	DRR042406
apr39S04.454.fastq	'apr39	S04	454	DRR042407
apr39S06.454.fastq	'apr39	S06	454	DRR042408
apr39S08.454.fastq	'apr39	S08	454	DRR042409
apr40.1.MiSeq.fastq	'apr40	S00	MiSeq	DRR042410
apr40.2.MiSeq.fastq	'apr40	S00	MiSeq	DRR042411
apr40S00.454.fastq	'apr40	S00	454	DRR042412
apr40S00.PGM.fastq	'apr40	S00	Ion PGM	DRR042413
apr40S02.454.fastq	'apr40	S02	454	DRR042414
apr40S04.454.fastq	'apr40	S04	454	DRR042415
apr40S06.454.fastq	'apr40	S06	454	DRR042416
apr40S08.454.fastq	'apr40	S08	454	DRR042417
FAKO01.454.fastq	'FAKO01	-	454	DRR042418
FAKO01.PGM.fastq	'FAKO01	-	Ion PGM	DRR042419
FAKO02.1.MiSeq.fastq	'FAKO02	-	MiSeq	DRR042420
FAKO02.2.MiSeq.fastq	'FAKO02	-	MiSeq	DRR042421
FAKO02.454.fastq	'FAKO02	-	454	DRR042422
FAKO02.PGM.fastq	'FAKO02	-	Ion PGM	DRR042423
FAKO03.1.MiSeq.fastq	'FAKO03	-	MiSeq	DRR042424
FAKO03.2.MiSeq.fastq	'FAKO03	-	MiSeq	DRR042425
FAKO03.454.fastq	'FAKO03	-	454	DRR042426
FAKO03.PGM.fastq	'FAKO03	-	Ion PGM	DRR042427
FAKO05.1.MiSeq.fastq	'FAKO05	-	MiSeq	DRR042428
FAKO05.2.MiSeq.fastq	'FAKO05	-	MiSeq	DRR042429
FAKO05.454.fastq	'FAKO05	-	454	DRR042430
FAKO05.PGM.fastq	'FAKO05	-	Ion PGM	DRR042431
FAKO06.454.fastq	'FAKO06	-	454	DRR042432
FAKO06.PGM.fastq	'FAKO06	-	Ion PGM	DRR042433
FAKO07.454.fastq	'FAKO07	-	454	DRR042434
FAKO07.PGM.fastq	'FAKO07	-	Ion PGM	DRR042435
FAKO08.1.MiSeq.fastq	'FAKO08	-	MiSeq	DRR042436
FAKO08.2.MiSeq.fastq	'FAKO08	-	MiSeq	DRR042437
FAKO08.454.fastq	'FAKO08	-	454	DRR042438
FAKO08.PGM.fastq	'FAKO08	-	Ion PGM	DRR042439
FAKO09.454.fastq	'FAKO09	-	454	DRR042440
FAKO09.PGM.fastq	'FAKO09	-	Ion PGM	DRR042441
FAKO10.454.fastq	'FAKO10	-	454	DRR042442
FAKO10.PGM.fastq	'FAKO10	-	Ion PGM	DRR042443
FAKO11.454.fastq	'FAKO11	-	454	DRR042444
FAKO11.PGM.fastq	'FAKO11	-	Ion PGM	DRR042445
FAKO12.454.fastq	'FAKO12	-	454	DRR042446
FAKO12.PGM.fastq	'FAKO12	-	Ion PGM	DRR042447
FAKO13.454.fastq	'FAKO13	-	454	DRR042448
FAKO13.PGM.fastq	'FAKO13	-	Ion PGM	DRR042449
FAKO14.454.fastq	'FAKO14	-	454	DRR042450
FAKO14.PGM.fastq	'FAKO14	-	Ion PGM	DRR042451
FAKO15.1.MiSeq.fastq	'FAKO15	-	MiSeq	DRR042452
FAKO15.2.MiSeq.fastq	'FAKO15	-	MiSeq	DRR042453
FAKO15.454.fastq	'FAKO15	-	454	DRR042454
FAKO15.PGM.fastq	'FAKO15	-	Ion PGM	DRR042455
FAKO16.454.fastq	'FAKO16	-	454	DRR042456
FAKO16.PGM.fastq	'FAKO16	-	Ion PGM	DRR042457
FAKO17.454.fastq	'FAKO17	-	454	DRR042458
FAKO17.PGM.fastq	'FAKO17	-	Ion PGM	DRR042459
FAKO18.454.fastq	'FAKO18	-	454	DRR042460
FAKO18.PGM.fastq	'FAKO18	-	Ion PGM	DRR042461
FAKO19.1.MiSeq.fastq	'FAKO19	-	MiSeq	DRR042462
FAKO19.2.MiSeq.fastq	'FAKO19	-	MiSeq	DRR042463
FAKO19.454.fastq	'FAKO19	-	454	DRR042464
FAKO19.PGM.fastq	'FAKO19	-	Ion PGM	DRR042465
FAKO21.454.fastq	'FAKO21	-	454	DRR042466
FAKO21.PGM.fastq	'FAKO21	-	Ion PGM	DRR042467
FAKO22.1.MiSeq.fastq	'FAKO22	-	MiSeq	DRR042468
FAKO22.2.MiSeq.fastq	'FAKO22	-	MiSeq	DRR042469
FAKO22.454.fastq	'FAKO22	-	454	DRR042470
FAKO22.PGM.fastq	'FAKO22	-	Ion PGM	DRR042471

File name	Individual	Time point	Sequencer	Accession number
FAKO23.1.MiSeq.fastq	'FAKO23	-	MiSeq	DRR042472
FAKO23.2.MiSeq.fastq	'FAKO23	-	MiSeq	DRR042473
FAKO23.454.fastq	'FAKO23	-	454	DRR042474
FAKO23.PGM.fastq	'FAKO23	-	Ion PGM	DRR042475
FAKO24.454.fastq	'FAKO24	-	454	DRR042476
FAKO24.PGM.fastq	'FAKO24	-	Ion PGM	DRR042477
FAKO25.454.fastq	'FAKO25	-	454	DRR042478
FAKO25.PGM.fastq	'FAKO25	-	Ion PGM	DRR042479
FAKO26.454.fastq	'FAKO26	-	454	DRR042480
FAKO26.PGM.fastq	'FAKO26	-	Ion PGM	DRR042481
FAKO27.1.MiSeq.fastq	'FAKO27	-	MiSeq	DRR042482
FAKO27.2.MiSeq.fastq	'FAKO27	-	MiSeq	DRR042483
FAKO27.454.fastq	'FAKO27	-	454	DRR042484
FAKO27.PGM.fastq	'FAKO27	-	Ion PGM	DRR042485
FAKO28.454.fastq	'FAKO28	-	454	DRR042486
FAKO28.PGM.fastq	'FAKO28	-	Ion PGM	DRR042487
FAKO29.1.MiSeq.fastq	'FAKO29	-	MiSeq	DRR042488
FAKO29.2.MiSeq.fastq	'FAKO29	-	MiSeq	DRR042489
FAKO29.454.fastq	'FAKO29	-	454	DRR042490
FAKO29.PGM.fastq	'FAKO29	-	Ion PGM	DRR042491
FAKO30.454.fastq	'FAKO30	-	454	DRR042492
FAKO30.PGM.fastq	'FAKO30	-	Ion PGM	DRR042493
FBAN01.454.fastq	'FBAN01	-	454	DRR042494
FBAN01.PGM.fastq	'FBAN01	-	Ion PGM	DRR042495
FBAN02-m80c-Fast.Proton.fastq	'FBAN02	-	Ion Proton	DRR042497
FBAN02-m80c-mAc.Proton.fastq	'FBAN02	-	Ion Proton	DRR042498
FBAN02-m80c-NF.Proton.fastq	'FBAN02	-	Ion Proton	DRR042499
FBAN02-m80c-Power.Proton.fastq	'FBAN02	-	Ion Proton	DRR042500
FBAN02-m80c-QIA.Proton.fastq	'FBAN02	-	Ion Proton	DRR042502
FBAN02-m80c.Proton.fastq	'FBAN02	-	Ion Proton	DRR042501
FBAN02.454.fastq	'FBAN02	-	454	DRR042496
FBAN02.PGM.fastq	'FBAN02	-	Ion PGM	DRR042503
FBAN04.454.fastq	'FBAN04	-	454	DRR042504
FBAN04.PGM.fastq	'FBAN04	-	Ion PGM	DRR042505
FBAN05.454.fastq	'FBAN05	-	454	DRR042506
FBAN05.PGM.fastq	'FBAN05	-	Ion PGM	DRR042507
FBAN06.454.fastq	'FBAN06	-	454	DRR042508
FBAN06.PGM.fastq	'FBAN06	-	Ion PGM	DRR042509
FBAN07-m80c-Fast.Proton.fastq	'FBAN07	-	Ion Proton	DRR042511
FBAN07-m80c-mAc.Proton.fastq	'FBAN07	-	Ion Proton	DRR042512
FBAN07-m80c-NF.Proton.fastq	'FBAN07	-	Ion Proton	DRR042513
FBAN07-m80c-Power.Proton.fastq	'FBAN07	-	Ion Proton	DRR042514
FBAN07-m80c-QIA.Proton.fastq	'FBAN07	-	Ion Proton	DRR042516
FBAN07-m80c.Proton.fastq	'FBAN07	-	Ion Proton	DRR042515
FBAN07.454.fastq	'FBAN07	-	454	DRR042510
FBAN07.PGM.fastq	'FBAN07	-	Ion PGM	DRR042517
FBAN08.454.fastq	'FBAN08	-	454	DRR042518
FBAN08.PGM.fastq	'FBAN08	-	Ion PGM	DRR042519
FBAN09.454.fastq	'FBAN09	-	454	DRR042520
FBAN09.PGM.fastq	'FBAN09	-	Ion PGM	DRR042521
FBAN10.454.fastq	'FBAN10	-	454	DRR042522
FBAN10.PGM.fastq	'FBAN10	-	Ion PGM	DRR042523
FMOR01.454.fastq	'FMOR01	-	454	DRR042524
FMOR01.PGM.fastq	'FMOR01	-	Ion PGM	DRR042525
FMOR02.454.fastq	'FMOR02	-	454	DRR042526
FMOR02.PGM.fastq	'FMOR02	-	Ion PGM	DRR042527
FMOR03.454.fastq	'FMOR03	-	454	DRR042528
FMOR03.PGM.fastq	'FMOR03	-	Ion PGM	DRR042529
FMOR04.454.fastq	'FMOR04	-	454	DRR042530
FMOR04.PGM.fastq	'FMOR04	-	Ion PGM	DRR042531
FMOR11-m80c-Fast.Proton.fastq	'FMOR11	-	Ion Proton	DRR042533
FMOR11-m80c-mAc.Proton.fastq	'FMOR11	-	Ion Proton	DRR042534
FMOR11-m80c-NF.Proton.fastq	'FMOR11	-	Ion Proton	DRR042535
FMOR11-m80c-Power.Proton.fastq	'FMOR11	-	Ion Proton	DRR042536
FMOR11-m80c-QIA.Proton.fastq	'FMOR11	-	Ion Proton	DRR042538
FMOR11-m80c.Proton.fastq	'FMOR11	-	Ion Proton	DRR042537
FMOR11.454.fastq	'FMOR11	-	454	DRR042532
FMOR11.PGM.fastq	'FMOR11	-	Ion PGM	DRR042539
FMOR14.454.fastq	'FMOR14	-	454	DRR042540
FMOR14.PGM.fastq	'FMOR14	-	Ion PGM	DRR042541

File name	Individual	Time point	Sequencer	Accession number
FMOR21.454.fastq	'FMOR21	-	454	DRR042542
FMOR21.PGM.fastq	'FMOR21	-	Ion PGM	DRR042543
FPR01-Fast.Proton.fastq	'FPR01	-	Ion Proton	DRR042544
FPR01-m80c-NF.Proton.fastq	'FPR01	-	Ion Proton	DRR042545
FPR01-m80c.Proton.fastq	'FPR01	-	Ion Proton	DRR042546
FPR01-mAc.Proton.fastq	'FPR01	-	Ion Proton	DRR042547
FPR01-Power.Proton.fastq	'FPR01	-	Ion Proton	DRR042548
FPR01-QIA.Proton.fastq	'FPR01	-	Ion Proton	DRR042551
FPR01.Proton.1.fastq	'FPR01	-	Ion Proton	DRR042549
FPR01.Proton.2.fastq	'FPR01	-	Ion Proton	DRR042550
FPR02-0d.PGM.fastq	'FPR02	-	Ion PGM	DRR042552
FPR02-1d-air.PGM.fastq	'FPR02	-	Ion PGM	DRR042553
FPR02-1d-ane.PGM.fastq	'FPR02	-	Ion PGM	DRR042554
FPR02-3d-air.PGM.fastq	'FPR02	-	Ion PGM	DRR042555
FPR02-3d-ane.PGM.fastq	'FPR02	-	Ion PGM	DRR042556
FPR03-Fast.Proton.fastq	'FPR03	-	Ion Proton	DRR042557
FPR03-m80c-NF.Proton.fastq	'FPR03	-	Ion Proton	DRR042558
FPR03-m80c.Proton.fastq	'FPR03	-	Ion Proton	DRR042559
FPR03-mAc.Proton.fastq	'FPR03	-	Ion Proton	DRR042560
FPR03-Power.Proton.fastq	'FPR03	-	Ion Proton	DRR042561
FPR03-QIA.Proton.fastq	'FPR03	-	Ion Proton	DRR042564
FPR03.Proton.1.fastq	'FPR03	-	Ion Proton	DRR042562
FPR03.Proton.2.fastq	'FPR03	-	Ion Proton	DRR042563
FPR04-0d.PGM.fastq	'FPR04	-	Ion PGM	DRR042565
FPR04-1d-air.PGM.fastq	'FPR04	-	Ion PGM	DRR042566
FPR04-1d-ane.PGM.fastq	'FPR04	-	Ion PGM	DRR042567
FPR04-3d-air.PGM.fastq	'FPR04	-	Ion PGM	DRR042568
FPR04-3d-ane.PGM.fastq	'FPR04	-	Ion PGM	DRR042569
FPR04-Fast.Proton.fastq	'FPR04	-	Ion Proton	DRR042570
FPR04-m80c-NF.Proton.fastq	'FPR04	-	Ion Proton	DRR042571
FPR04-m80c.Proton.fastq	'FPR04	-	Ion Proton	DRR042572
FPR04-mAc.Proton.fastq	'FPR04	-	Ion Proton	DRR042573
FPR04-Power.Proton.fastq	'FPR04	-	Ion Proton	DRR042574
FPR04-QIA.Proton.fastq	'FPR04	-	Ion Proton	DRR042577
FPR04.Proton.1.fastq	'FPR04	-	Ion Proton	DRR042575
FPR04.Proton.2.fastq	'FPR04	-	Ion Proton	DRR042576
FPR05-0d.PGM.fastq	'FPR05	-	Ion PGM	DRR042578
FPR05-1d-air.PGM.fastq	'FPR05	-	Ion PGM	DRR042579
FPR05-1d-ane.PGM.fastq	'FPR05	-	Ion PGM	DRR042580
FPR05-3d-air.PGM.fastq	'FPR05	-	Ion PGM	DRR042581
FPR05-3d-ane.PGM.fastq	'FPR05	-	Ion PGM	DRR042582
FPR05-Fast.Proton.fastq	'FPR05	-	Ion Proton	DRR042583
FPR05-m80c-NF.Proton.fastq	'FPR05	-	Ion Proton	DRR042584
FPR05-m80c.Proton.fastq	'FPR05	-	Ion Proton	DRR042585
FPR05-mAc.Proton.fastq	'FPR05	-	Ion Proton	DRR042586
FPR05-Power.Proton.fastq	'FPR05	-	Ion Proton	DRR042587
FPR05-QIA.Proton.fastq	'FPR05	-	Ion Proton	DRR042590
FPR05.Proton.1.fastq	'FPR05	-	Ion Proton	DRR042588
FPR05.Proton.2.fastq	'FPR05	-	Ion Proton	DRR042589
FTAG01.1.MiSeq.fastq	'FTAG01	-	MiSeq	DRR042591
FTAG01.2.MiSeq.fastq	'FTAG01	-	MiSeq	DRR042592
FTAG01.454.fastq	'FTAG01	-	454	DRR042593
FTAG01.PGM.fastq	'FTAG01	-	Ion PGM	DRR042594
FTAG02.454.fastq	'FTAG02	-	454	DRR042595
FTAG02.PGM.fastq	'FTAG02	-	Ion PGM	DRR042596
FTAG03-m80c-Fast.Proton.fastq	'FTAG03	-	Ion Proton	DRR042598
FTAG03-m80c-mAc.Proton.fastq	'FTAG03	-	Ion Proton	DRR042599
FTAG03-m80c-NF.Proton.fastq	'FTAG03	-	Ion Proton	DRR042600
FTAG03-m80c-Power.Proton.fastq	'FTAG03	-	Ion Proton	DRR042601
FTAG03-m80c-QIA.Proton.fastq	'FTAG03	-	Ion Proton	DRR042603
FTAG03-m80c.Proton.fastq	'FTAG03	-	Ion Proton	DRR042602
FTAG03.454.fastq	'FTAG03	-	454	DRR042597
FTAG03.PGM.fastq	'FTAG03	-	Ion PGM	DRR042604
FTAG06.454.fastq	'FTAG06	-	454	DRR042605
FTAG06.PGM.fastq	'FTAG06	-	Ion PGM	DRR042606
FTAG07.454.fastq	'FTAG07	-	454	DRR042607
FTAG07.PGM.fastq	'FTAG07	-	Ion PGM	DRR042608
FTAG08.454.fastq	'FTAG08	-	454	DRR042609
FTAG08.PGM.fastq	'FTAG08	-	Ion PGM	DRR042610
FTAG09-m80c-Fast.Proton.fastq	'FTAG09	-	Ion Proton	DRR042612

File name	Individual	Time point	Sequencer	Accession number
FTAG09-m80c-mAc.Proton.fastq	'FTAG09	-	Ion Proton	DRR042613
FTAG09-m80c-NF.Proton.fastq	'FTAG09	-	Ion Proton	DRR042614
FTAG09-m80c-Power.Proton.fastq	'FTAG09	-	Ion Proton	DRR042615
FTAG09-m80c-QIA.Proton.fastq	'FTAG09	-	Ion Proton	DRR042617
FTAG09-m80c.Proton.fastq	'FTAG09	-	Ion Proton	DRR042616
FTAG09.454.fastq	'FTAG09	-	454	DRR042611
FTAG09.PGM.fastq	'FTAG09	-	Ion PGM	DRR042618
FTAG10.454.fastq	'FTAG10	-	454	DRR042619
FTAG10.PGM.fastq	'FTAG10	-	Ion PGM	DRR042620
FTAG12.454.fastq	'FTAG12	-	454	DRR042621
FTAG12.PGM.fastq	'FTAG12	-	Ion PGM	DRR042622
FTAG13.454.fastq	'FTAG13	-	454	DRR042623
FTAG13.PGM.fastq	'FTAG13	-	Ion PGM	DRR042624
FTAG14.454.fastq	'FTAG14	-	454	DRR042625
FTAG14.PGM.fastq	'FTAG14	-	Ion PGM	DRR042626
FTAG15.454.fastq	'FTAG15	-	454	DRR042627
FTAG15.PGM.fastq	'FTAG15	-	Ion PGM	DRR042628
FTAG16-m80c-Fast.Proton.fastq	'FTAG16	-	Ion Proton	DRR042630
FTAG16-m80c-mAc.Proton.fastq	'FTAG16	-	Ion Proton	DRR042631
FTAG16-m80c-NF.Proton.fastq	'FTAG16	-	Ion Proton	DRR042632
FTAG16-m80c-Power.Proton.fastq	'FTAG16	-	Ion Proton	DRR042633
FTAG16-m80c-QIA.Proton.fastq	'FTAG16	-	Ion Proton	DRR042635
FTAG16-m80c.Proton.fastq	'FTAG16	-	Ion Proton	DRR042634
FTAG16.454.fastq	'FTAG16	-	454	DRR042629
FTAG16.PGM.fastq	'FTAG16	-	Ion PGM	DRR042636
FTAG17.454.fastq	'FTAG17	-	454	DRR042637
FTAG17.PGM.fastq	'FTAG17	-	Ion PGM	DRR042638
FTAG18.454.fastq	'FTAG18	-	454	DRR042639
FTAG18.PGM.fastq	'FTAG18	-	Ion PGM	DRR042640
FTAG19.454.fastq	'FTAG19	-	454	DRR042641
FTAG19.PGM.fastq	'FTAG19	-	Ion PGM	DRR042642
FTAG20-m80c-Fast.Proton.fastq	'FTAG20	-	Ion Proton	DRR042644
FTAG20-m80c-mAc.Proton.fastq	'FTAG20	-	Ion Proton	DRR042645
FTAG20-m80c-NF.Proton.fastq	'FTAG20	-	Ion Proton	DRR042646
FTAG20-m80c-Power.Proton.fastq	'FTAG20	-	Ion Proton	DRR042647
FTAG20-m80c-QIA.Proton.fastq	'FTAG20	-	Ion Proton	DRR042649
FTAG20-m80c.Proton.fastq	'FTAG20	-	Ion Proton	DRR042648
FTAG20.454.fastq	'FTAG20	-	454	DRR042643
FTAG20.PGM.fastq	'FTAG20	-	Ion PGM	DRR042650
FTAG21-m80c-Fast.Proton.fastq	'FTAG21	-	Ion Proton	DRR042652
FTAG21-m80c-mAc.Proton.fastq	'FTAG21	-	Ion Proton	DRR042653
FTAG21-m80c-NF.Proton.fastq	'FTAG21	-	Ion Proton	DRR042654
FTAG21-m80c-Power.Proton.fastq	'FTAG21	-	Ion Proton	DRR042655
FTAG21-m80c-QIA.Proton.fastq	'FTAG21	-	Ion Proton	DRR042657
FTAG21-m80c.Proton.fastq	'FTAG21	-	Ion Proton	DRR042656
FTAG21.454.fastq	'FTAG21	-	454	DRR042651
FTAG21.PGM.fastq	'FTAG21	-	Ion PGM	DRR042658
TS-11.Proton.fastq	'TS-11	-	Ion Proton	DRR042659
TS-21.Proton.fastq	'TS-21	-	Ion Proton	DRR042660
TS-29.Proton.fastq	'TS-29	-	Ion Proton	DRR042661
TS-33.Proton.fastq	'TS-33	-	Ion Proton	DRR042662
TS-41.Proton.fastq	'TS-41	-	Ion Proton	DRR042663

Appendix 4. KOs unique to the JP gene set

KEGG orthology	Function
K00035	D-galactose 1-dehydrogenase [EC:1.1.1.48]
K00193	acetyl-CoA decarbonylase/synthase complex subunit beta [EC:2.3.1.-]
K00201	formylmethanofuran dehydrogenase subunit B [EC:1.2.99.5]
K00271	valine dehydrogenase [EC:1.4.1.8]
K00292	saccharopine dehydrogenase (NAD+, L-glutamate forming) [EC:1.5.1.9]
K00319	methylenetetrahydromethanopterin dehydrogenase [EC:1.5.99.9]
K00360	nitrate reductase (NADH) [EC:1.7.1.1]
K00401	methyl-coenzyme M reductase beta subunit [EC:2.8.4.1]
K00406	cytochrome c oxidase cbb3-type subunit III
K00407	cytochrome c oxidase cbb3-type subunit IV
K00410	ubiquinol-cytochrome c reductase cytochrome b/c1 subunit
K00413	ubiquinol-cytochrome c reductase cytochrome c1 subunit
K00440	coenzyme F420 hydrogenase alpha subunit [EC:1.12.98.1]
K00456	cysteine dioxygenase [EC:1.13.11.20]
K00496	alkane 1-monooxygenase [EC:1.14.15.3]
K00499	choline monooxygenase [EC:1.14.15.7]
K00507	stearoyl-CoA desaturase (delta-9 desaturase) [EC:1.14.19.1]
K00514	zeta-carotene desaturase [EC:1.3.5.6]
K00524	ribonucleotide reductase, class II [EC:1.17.4.1]
K00555	tRNA (guanine26-N2/guanine27-N2)-dimethyltransferase [EC:2.1.1.215 2.1.1.216]
K00577	tetrahydromethanopterin S-methyltransferase subunit A [EC:2.1.1.86]
K00578	tetrahydromethanopterin S-methyltransferase subunit B [EC:2.1.1.86]
K00579	tetrahydromethanopterin S-methyltransferase subunit C [EC:2.1.1.86]
K00580	tetrahydromethanopterin S-methyltransferase subunit D [EC:2.1.1.86]
K00582	tetrahydromethanopterin S-methyltransferase subunit F [EC:2.1.1.86]
K00583	tetrahydromethanopterin S-methyltransferase subunit G [EC:2.1.1.86]
K00586	diphthine synthase [EC:2.1.1.98]
K00624	carnitine O-acetyltransferase [EC:2.3.1.7]
K00635	diacylglycerol O-acyltransferase [EC:2.3.1.20]
K00862	erythritol kinase [EC:2.7.1.27]
K00887	undecaprenol kinase [EC:2.7.1.66]
K00988	ATP adenylyltransferase [EC:2.7.7.53]
K01001	UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase [EC:2.7.8.15]
K01084	glucose-6-phosphatase [EC:3.1.3.9]
K01117	sphingomyelin phosphodiesterase [EC:3.1.4.12]
K01158	deoxyribonuclease II [EC:3.1.22.1]
K01170	tRNA-intron endonuclease, archaea type [EC:3.1.27.9]
K01172	NA
K01230	mannosyl-oligosaccharide alpha-1,2-mannosidase [EC:3.2.1.113]
K01280	tripeptidyl-peptidase II [EC:3.4.14.10]
K01385	thermopsin [EC:3.4.23.42]
K01504	glucosamine-6-phosphate isomerase [EC:3.5.99.6]
K01510	apyrase [EC:3.6.1.5]
K01594	sulfinoalanine decarboxylase [EC:4.1.1.29]
K01622	fructose 1,6-bisphosphate aldolase/phosphatase [EC:4.1.2.13 3.1.3.11]
K01663	glutamine amidotransferase / cyclase [EC:2.4.2.- 4.1.3.-]
K01769	guanylate cyclase, other [EC:4.6.1.2]
K01969	3-methylcrotonyl-CoA carboxylase beta subunit [EC:6.4.1.4]
K02201	pantetheine-phosphate adenylyltransferase [EC:2.7.7.3]
K02288	phycocyanobilin lyase alpha subunit [EC:4.-.-.-]
K02289	phycocyanobilin lyase beta subunit
K02322	DNA polymerase II large subunit [EC:2.7.7.7]
K02497	HemX protein
K02571	periplasmic nitrate reductase NapE
K02595	nitrogenase-stabilizing/protective protein
K02659	twitching motility protein PilI
K02665	type IV pilus assembly protein PilP
K02672	type IV pilus assembly protein PilW
K02676	type IV pilus assembly protein PilZ
K02683	DNA primase [EC:2.7.7.-]
K02685	DNA primase large subunit [EC:2.7.7.-]
K02691	photosystem I subunit VII
K02717	photosystem II oxygen-evolving enhancer protein 2
K02866	large subunit ribosomal protein L10e
K02869	large subunit ribosomal protein L12
K02877	large subunit ribosomal protein L15e
K02883	large subunit ribosomal protein L18e
K02885	large subunit ribosomal protein L19e

KEGG orthology	Function
K02889	large subunit ribosomal protein L21e
K02896	large subunit ribosomal protein L24e
K02910	large subunit ribosomal protein L31e
K02912	large subunit ribosomal protein L32e
K02915	large subunit ribosomal protein L34e
K02921	large subunit ribosomal protein L37Ae
K02922	large subunit ribosomal protein L37e
K02924	large subunit ribosomal protein L39e
K02927	large subunit ribosomal protein L40e
K02929	large subunit ribosomal protein L44e
K02930	large subunit ribosomal protein L4e
K02944	large subunit ribosomal protein LX
K02966	small subunit ribosomal protein S19e
K02974	small subunit ribosomal protein S24e
K02978	small subunit ribosomal protein S27e
K02979	small subunit ribosomal protein S28e
K02984	small subunit ribosomal protein S3Ae
K02987	small subunit ribosomal protein S4e
K02991	small subunit ribosomal protein S6e
K02995	small subunit ribosomal protein S8e
K03044	DNA-directed RNA polymerase subunit B' [EC:2.7.7.6]
K03045	DNA-directed RNA polymerase subunit B" [EC:2.7.7.6]
K03049	DNA-directed RNA polymerase subunit E' [EC:2.7.7.6]
K03050	DNA-directed RNA polymerase subunit E" [EC:2.7.7.6]
K03051	DNA-directed RNA polymerase subunit F [EC:2.7.7.6]
K03053	DNA-directed RNA polymerase subunit H [EC:2.7.7.6]
K03055	DNA-directed RNA polymerase subunit K [EC:2.7.7.6]
K03057	transcription elongation factor
K03058	DNA-directed RNA polymerase subunit N [EC:2.7.7.6]
K03059	DNA-directed RNA polymerase subunit P [EC:2.7.7.6]
K03105	signal recognition particle subunit SRP19
K03120	transcription initiation factor TFIID TATA-box-binding protein
K03136	transcription initiation factor TFIIE subunit alpha
K03166	DNA topoisomerase VI subunit A [EC:5.99.1.3]
K03167	DNA topoisomerase VI subunit B [EC:5.99.1.3]
K03232	elongation factor 1-beta
K03236	translation initiation factor 1A
K03237	translation initiation factor 2 subunit 1
K03239	translation initiation factor eIF-2B subunit alpha
K03242	translation initiation factor 2 subunit 3
K03243	translation initiation factor 5B
K03264	translation initiation factor 6
K03268	benzene 1,2-dioxygenase [EC:1.14.12.3]
K03395	gentamicin 3'-N-acetyltransferase [EC:2.3.1.60]
K03538	ribonuclease P protein subunit POP4 [EC:3.1.26.5]
K03539	ribonuclease P/MRP protein subunit RPP1 [EC:3.1.26.5]
K03540	ribonuclease P protein subunit RPR2 [EC:3.1.26.5]
K03622	archaea-specific DNA-binding protein
K03626	nascent polypeptide-associated complex subunit alpha
K04071	6-pyruvoyltetrahydropterin 2'-reductase [EC:1.1.1.220]
K04090	indolepyruvate ferredoxin oxidoreductase [EC:1.2.7.8]
K04097	glutathione S-transferase [EC:2.5.1.18]
K04340	scyllo-inosamine-4-phosphate amidinotransferase 1 [EC:2.1.4.2]
K04341	dTDP-dihydrostreptose-streptidine-6-phosphate dihydrostreptosyltransferase [EC:2.4.2.27]
K04484	DNA repair protein RadB
K04496	C-terminal binding protein
K04791	mycobactin polyketide synthetase MbtD
K04795	fibrillarin-like pre-rRNA processing protein
K04796	small nuclear ribonucleoprotein
K04797	prefoldin alpha subunit
K04798	prefoldin beta subunit
K04801	replication factor C small subunit
K04802	proliferating cell nuclear antigen
K05301	sulfite dehydrogenase [EC:1.8.2.1]
K05383	CpeT protein
K05551	act minimal PKS ketosynthase (KS/KS alpha) [EC:2.3.1.-]
K05552	act minimal PKS chain-length factor (CLF/KS beta) [EC:2.3.1.-]
K05553	act minimal PKS acyl carrier protein
K05559	multicomponent K+:H+ antiporter subunit A
K05560	multicomponent K+:H+ antiporter subunit C

KEGG orthology	Function
K05561	multicomponent K+:H+ antiporter subunit D
K05562	multicomponent K+:H+ antiporter subunit E
K05563	multicomponent K+:H+ antiporter subunit F
K05564	multicomponent K+:H+ antiporter subunit G
K05573	NAD(P)H-quinone oxidoreductase subunit 2 [EC:1.6.5.3]
K05574	NAD(P)H-quinone oxidoreductase subunit 3 [EC:1.6.5.3]
K05575	NAD(P)H-quinone oxidoreductase subunit 4 [EC:1.6.5.3]
K05576	NAD(P)H-quinone oxidoreductase subunit 4L [EC:1.6.5.3]
K05597	glutamin-(asparagin-)ase [EC:3.5.1.38]
K05716	cyclic 2,3-diphosphoglycerate synthetase [EC:4.6.1.-]
K05797	4-cresol dehydrogenase (hydroxylating) [EC:1.17.99.1]
K05889	polyvinyl alcohol dehydrogenase (cytochrome) [EC:1.1.2.6]
K05908	NA
K05927	quinone-reactive Ni/Fe-hydrogenase small subunit [EC:1.12.5.1]
K05973	poly(3-hydroxybutyrate) depolymerase [EC:3.1.1.75]
K05986	NA
K05994	bacterial leucyl aminopeptidase [EC:3.4.11.10]
K05998	pseudomonalisin [EC:3.4.21.100]
K06034	sulfopyruvate decarboxylase subunit alpha [EC:4.1.1.79]
K06154	Lrp/AsnC family transcriptional regulator, involved in the regulation of lysine biosynthesis
K06237	collagen, type IV, alpha
K06363	response regulator aspartate phosphatase E [EC:3.1.-.-]
K06364	response regulator aspartate phosphatase F [EC:3.1.-.-]
K06377	sporulation-control protein
K06434	small acid-soluble spore protein (thioredoxin-like protein)
K06596	chemosensory pili system protein ChpA (sensor histidine kinase/response regulator)
K06601	flagellar protein FlbT
K06602	flagellar protein FlaF
K06718	L-2,4-diaminobutyric acid acetyltransferase [EC:2.3.1.178]
K06862	energy-converting hydrogenase B subunit Q
K06863	5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase [EC:6.3.4.-]
K06868	Sep-tRNA:Cys-tRNA synthetase [EC:2.5.1.73]
K06869	uncharacterized protein
K06873	NA
K06874	zinc finger protein
K06875	programmed cell death protein 5
K06914	NA
K06930	NA
K06931	NA
K06943	nucleolar GTP-binding protein
K06953	NA
K06961	ribosomal RNA assembly protein
K06963	tRNA acetyltransferase TAN1
K06964	NA
K06982	pantoate kinase [EC:2.7.1.169]
K06984	NA
K07049	TatD-related deoxyribonuclease
K07055	tRNA wybutosine-synthesizing protein 2 [EC:2.1.1.-]
K07060	UPF0271 protein
K07061	NA
K07073	NA
K07086	NA
K07092	NA
K07103	NA
K07108	NA
K07123	NA
K07131	NA
K07135	NA
K07144	NA
K07155	quercetin 2,3-dioxygenase [EC:1.13.11.24]
K07157	NA
K07158	NA
K07159	NA
K07178	RIO kinase 1 [EC:2.7.11.1]
K07179	RIO kinase 2 [EC:2.7.11.1]
K07244	mgtE-like transporter
K07254	tRNA (cytidine56-2'-O)-methyltransferase [EC:2.1.1.206]
K07288	uncharacterized membrane protein
K07333	archaeal flagellar protein FlaJ
K07338	hypothetical protein

KEGG orthology	Function
K07342	protein transport protein SEC61 subunit gamma and related proteins
K07394	SM-20-related protein
K07477	translin
K07544	benzylsuccinate CoA-transferase BbsF subunit [EC:2.8.3.15]
K07558	tRNA nucleotidyltransferase (CCA-adding enzyme) [EC:2.7.7.72]
K07562	nonsense-mediated mRNA decay protein 3
K07569	RNA-binding protein
K07572	putative nucleotide binding protein
K07575	PUA domain protein
K07580	hypothetical protein
K07581	hypothetical protein
K07585	hypothetical protein
K07732	riboflavin kinase, archaea type [EC:2.7.1.161]
K07823	3-oxoadipyl-CoA thiolase [EC:2.3.1.174]
K08076	astacin [EC:3.4.24.21]
K08085	type IV fimbrial biogenesis protein FimU
K08096	GTP cyclohydrolase IIa [EC:3.5.4.29]
K08176	MFS transporter, PHS family, inorganic phosphate transporter
K08477	outer membrane protease E [EC:3.4.21.-]
K08598	YopJ protease family
K08604	vibriolysin [EC:3.4.24.25]
K08645	anthrax lethal toxin endopeptidase [EC:3.4.24.83]
K08698	carbon dioxide concentrating mechanism protein CcmM
K08713	potassium channel LctB
K08953	chlorosome envelope protein J
K08971	putative membrane protein
K08973	putative membrane protein
K08975	putative membrane protein
K08979	putative membrane protein
K08983	putative membrane protein
K09119	hypothetical protein
K09139	hypothetical protein
K09140	pre-rRNA-processing protein TSR3
K09148	hypothetical protein
K09152	hypothetical protein
K09154	hypothetical protein
K09713	hypothetical protein
K09721	hypothetical protein
K09722	4-phosphopantoate---beta-alanine ligase [EC:6.3.2.36]
K09723	hypothetical protein
K09724	hypothetical protein
K09725	hypothetical protein
K09727	hypothetical protein
K09732	hypothetical protein
K09735	hypothetical protein
K09736	hypothetical protein
K09737	hypothetical protein
K09738	hypothetical protein
K09739	hypothetical protein
K09741	hypothetical protein
K09744	hypothetical protein
K09746	hypothetical protein
K09796	hypothetical protein
K09845	1-hydroxycarotenoid 3,4-desaturase [EC:1.3.99.27]
K09846	demethylspheroidene O-methyltransferase [EC:2.1.1.210]
K09879	isorenieratene synthase
K09919	hypothetical protein
K09941	hypothetical protein
K09943	hypothetical protein
K09947	hypothetical protein
K09950	hypothetical protein
K09959	hypothetical protein
K09965	hypothetical protein
K09966	hypothetical protein
K09983	hypothetical protein
K10023	arginine/ornithine transport system permease protein
K10024	arginine/ornithine transport system permease protein
K10025	arginine/ornithine transport system ATP-binding protein [EC:3.6.3.-]
K10123	putative ferrous iron transport protein C
K10216	2-hydroxymuconate-semialdehyde hydrolase [EC:3.7.1.9]

KEGG orthology	Function
K10221	2-pyrone-4,6-dicarboxylate lactonase [EC:3.1.1.57]
K10222	2,6-dioxo-6-phenylhexa-3-enoate hydrolase [EC:3.7.1.8]
K10233	alpha-glucoside transport system permease protein
K10623	HOMODA hydrolase [EC:3.7.1.-]
K10725	archaeal cell division control protein 6
K10764	O-succinylhomoserine sulfhydrylase [EC:2.5.1.-]
K10829	iron-chelate-transporting ATPase [EC:3.6.3.34]
K10873	DNA repair and recombination protein RAD52
K10911	two-component system, phosphorelay protein LuxU
K10923	AraC family transcriptional regulator, TCP pilus virulence regulatory protein
K10925	MSHA pilin protein MshB
K10926	MSHA pilin protein MshC
K10929	cholera enterotoxin subunit B
K10930	toxin co-regulated pilin
K10932	toxin co-regulated pilus biosynthesis outer membrane protein C
K10938	accessory colonization factor AcfC
K10966	toxin co-regulated pilus biosynthesis protein J [EC:3.4.23.43 2.1.1.-]
K11006	shiga toxin subunit A
K11007	shiga toxin subunit B
K11014	cytolethal distending toxin subunit B
K11015	cytolethal distending toxin subunit C
K11016	hemolysin
K11023	pertussis toxin subunit 1 [EC:2.4.2.-]
K11028	vacuolating cytotoxin
K11057	beta2-toxin
K11058	enterotoxin Cpe
K11212	LPPG:FO 2-phospho-L-lactate transferase [EC:2.7.8.28]
K11260	formylmethanofuran dehydrogenase subunit G [EC:1.2.99.5]
K11395	2-keto-3-deoxy-gluconate aldolase [EC:4.1.2.-]
K11434	protein arginine N-methyltransferase 1 [EC:2.1.1.-]
K11526	two-component system, chemotaxis family, sensor histidine kinase and response regulator PixL
K11600	exosome complex component RRP41
K11638	two-component system, CitB family, response regulator CitT
K11780	FO synthase subunit 1 [EC:2.5.1.77]
K11814	multidrug resistance protein EbrA
K11913	type VI secretion system protein
K12048	ComB10 competence protein
K12049	ComB9 competence protein
K12054	lytic transglycosylase AtlA
K12080	type IV secretion system protein PtlC
K12247	alpha-2,3 sialyltransferase [EC:2.4.99.-]
K12430	polyketide synthase 1/15
K12513	tight adherence protein E
K12514	tight adherence protein F
K12515	tight adherence protein G
K12534	membrane fusion protein RsaE
K12589	exosome complex component RRP42
K12673	N2-(2-carboxyethyl)arginine synthase [EC:2.5.1.66]
K12786	LEE-encoded effector EspF
K12789	Tir-cytoskeleton coupling protein
K12809	T3SS secreted effector EspG-like protein
K12953	cation-transporting ATPase F [EC:3.6.3.-]
K12978	lipid A 4'-phosphatase [EC:3.1.3.-]
K13008	O-antigen polymerase
K13039	sulfopyruvate decarboxylase subunit beta [EC:4.1.1.79]
K13060	acyl homoserine lactone synthase [EC:2.3.1.184]
K13284	invasin A
K13286	invasin C
K13317	NDP-4-keto-2,6-dideoxyhexose 3-C-methyltransferase
K13450	phosphothreonine lyase [EC:4.2.3.-]
K13455	avirulence protein
K13488	chemotaxis-related protein WspB
K13489	chemotaxis-related protein WspD
K13503	anthranilate synthase [EC:4.1.3.27]
K13520	outer membrane protease [EC:3.4.23.-]
K13583	GcrA cell cycle regulator
K13659	2-beta-glucuronyltransferase [EC:2.4.1.264]
K13661	GumC protein
K13740	secreted effector protein SptP
K13742	protein IpgB1

KEGG orthology	Function
K13790	virulence protein IcsB
K13793	secreted effector OspE
K13797	DNA-directed RNA polymerase subunit beta-beta' [EC:2.7.7.6]
K13798	DNA-directed RNA polymerase subunit B [EC:2.7.7.6]
K13875	L-arabonate dehydrase [EC:4.2.1.25]
K13941	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase / dihydropteroate synthase [EC:2.7.6.3 2.5.1.15]
K14092	energy-converting hydrogenase A subunit A
K14093	energy-converting hydrogenase A subunit B
K14094	energy-converting hydrogenase A subunit C
K14095	energy-converting hydrogenase A subunit D
K14096	energy-converting hydrogenase A subunit E
K14097	energy-converting hydrogenase A subunit F
K14098	energy-converting hydrogenase A subunit G
K14100	energy-converting hydrogenase A subunit I
K14102	energy-converting hydrogenase A subunit K
K14103	energy-converting hydrogenase A subunit L
K14104	energy-converting hydrogenase A subunit M
K14105	energy-converting hydrogenase A subunit N
K14109	energy-converting hydrogenase A subunit R
K14110	energy-converting hydrogenase B subunit A
K14111	energy-converting hydrogenase B subunit B
K14113	energy-converting hydrogenase B subunit D
K14114	energy-converting hydrogenase B subunit E
K14115	energy-converting hydrogenase B subunit F
K14116	energy-converting hydrogenase B subunit G
K14117	energy-converting hydrogenase B subunit H
K14118	energy-converting hydrogenase B subunit I
K14119	energy-converting hydrogenase B subunit J
K14122	energy-converting hydrogenase B subunit M
K14123	energy-converting hydrogenase B subunit N
K14124	energy-converting hydrogenase B subunit O
K14125	energy-converting hydrogenase B subunit P
K14165	dual specificity phosphatase [EC:3.1.3.16 3.1.3.48]
K14166	copper transport protein
K14201	clumping factor A
K14269	glutarate semialdehyde dehydrogenase [EC:1.2.1.20]
K14451	(3S)-malyl-CoA thioesterase
K14465	succinate semialdehyde reductase (NADPH) [EC:1.1.1.-]
K14468	malonyl-CoA reductase / 3-hydroxypropionate dehydrogenase (NADP+) [EC:1.2.1.75 1.1.1.298]
K14561	U3 small nucleolar ribonucleoprotein protein IMP4
K14564	nucleolar protein 56
K14568	essential for mitotic growth 1
K14598	chlorobactene lauroyltransferase
K14628	enoyl reductase
K14653	2-amino-5-formylamino-6-ribosylaminopyrimidin-4(3H)-one 5'-monophosphate deformylase [EC:3.5.1.102]
K14661	nodulation protein F [EC:2.3.1.-]
K14683	solute carrier family 34 (sodium-dependent phosphate cotransporter)
K14974	6-hydroxynicotinate 3-monooxygenase [EC:1.14.13.114]
K14998	surfeit locus 1 family protein
K15226	arogenate dehydrogenase (NADP+) [EC:1.3.1.78]
K15327	polyketide biosynthesis malonyl-CoA-[acyl-carrier-protein] transacylase
K15355	NA
K15366	salmonella plasmid virulence protein B
K15468	cytochrome P450 PksS
K15645	coronafacic acid polyketide synthase Cfa7
K15650	non-haem Fe2+, alpha-ketoglutarate-dependent halogenase
K15676	rhizoxin biosynthesis, polyketide synthase RhiC
K15681	aminotransferase MxcL
K15784	N-alpha-acetyl-L-2,4-diaminobutyrate deacetylase [EC:3.5.1.-]
K15845	outer membrane protein HopZ
K15853	acyl transferase [EC:2.3.1.-]
K15900	tRNA threonylcarbamoyladenosine biosynthesis protein
K15904	bifunctional tRNA threonylcarbamoyladenosine biosynthesis protein [EC:2.7.11.1]
K15918	D-glycerate 3-kinase [EC:2.7.1.31]
K16081	alginate production protein
K16149	1,4-alpha-glucan branching enzyme [EC:2.4.1.18]
K16152	heme acquisition protein HasR
K16190	glucuronokinase [EC:2.7.1.43]

NA indicates not assigned

Appendix 5. KOs unique to the IGC gene set

KEGG orthology	Function
K00035	D-galactose 1-dehydrogenase [EC:1.1.1.48]
K00193	acetyl-CoA decarbonylase/synthase complex subunit beta [EC:2.3.1.-]
K00201	formylmethanofuran dehydrogenase subunit B [EC:1.2.99.5]
K00271	valine dehydrogenase [EC:1.4.1.8]
K00292	saccharopine dehydrogenase (NAD+, L-glutamate forming) [EC:1.5.1.9]
K00319	methylenetetrahydromethanopterin dehydrogenase [EC:1.5.99.9]
K00360	nitrate reductase (NADH) [EC:1.7.1.1]
K00401	methyl-coenzyme M reductase beta subunit [EC:2.8.4.1]
K00406	cytochrome c oxidase cbb3-type subunit III
K00407	cytochrome c oxidase cbb3-type subunit IV
K00410	ubiquinol-cytochrome c reductase cytochrome b/c1 subunit
K00413	ubiquinol-cytochrome c reductase cytochrome c1 subunit
K00440	coenzyme F420 hydrogenase alpha subunit [EC:1.12.98.1]
K00456	cysteine dioxygenase [EC:1.13.11.20]
K00496	alkane 1-monooxygenase [EC:1.14.15.3]
K00499	choline monooxygenase [EC:1.14.15.7]
K00507	stearoyl-CoA desaturase (delta-9 desaturase) [EC:1.14.19.1]
K00514	zeta-carotene desaturase [EC:1.3.5.6]
K00524	ribonucleotide reductase, class II [EC:1.17.4.1]
K00555	tRNA (guanine26-N2/guanine27-N2)-dimethyltransferase [EC:2.1.1.215 2.1.1.216]
K00577	tetrahydromethanopterin S-methyltransferase subunit A [EC:2.1.1.86]
K00578	tetrahydromethanopterin S-methyltransferase subunit B [EC:2.1.1.86]
K00579	tetrahydromethanopterin S-methyltransferase subunit C [EC:2.1.1.86]
K00580	tetrahydromethanopterin S-methyltransferase subunit D [EC:2.1.1.86]
K00582	tetrahydromethanopterin S-methyltransferase subunit F [EC:2.1.1.86]
K00583	tetrahydromethanopterin S-methyltransferase subunit G [EC:2.1.1.86]
K00586	diphthine synthase [EC:2.1.1.98]
K00624	carnitine O-acetyltransferase [EC:2.3.1.7]
K00635	diacylglycerol O-acyltransferase [EC:2.3.1.20]
K00862	erythritol kinase [EC:2.7.1.27]
K00887	undecaprenol kinase [EC:2.7.1.66]
K00988	ATP adenylyltransferase [EC:2.7.7.53]
K01001	UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase [EC:2.7.8.15]
K01084	glucose-6-phosphatase [EC:3.1.3.9]
K01117	sphingomyelin phosphodiesterase [EC:3.1.4.12]
K01158	deoxyribonuclease II [EC:3.1.22.1]
K01170	tRNA-intron endonuclease, archaea type [EC:3.1.27.9]
K01172	NA
K01230	mannosyl-oligosaccharide alpha-1,2-mannosidase [EC:3.2.1.113]
K01280	tripeptidyl-peptidase II [EC:3.4.14.10]
K01385	thermopsin [EC:3.4.23.42]
K01504	glucosamine-6-phosphate isomerase [EC:3.5.99.6]
K01510	apyrase [EC:3.6.1.5]
K01594	sulfinoalanine decarboxylase [EC:4.1.1.29]
K01622	fructose 1,6-bisphosphate aldolase/phosphatase [EC:4.1.2.13 3.1.3.11]
K01663	glutamine amidotransferase / cyclase [EC:2.4.2.- 4.1.3.-]
K01769	guanylate cyclase, other [EC:4.6.1.2]
K01969	3-methylcrotonyl-CoA carboxylase beta subunit [EC:6.4.1.4]
K02201	pantetheine-phosphate adenylyltransferase [EC:2.7.7.3]
K02288	phycocyanobilin lyase alpha subunit [EC:4.-.-.-]
K02289	phycocyanobilin lyase beta subunit
K02322	DNA polymerase II large subunit [EC:2.7.7.7]
K02497	HemX protein
K02571	periplasmic nitrate reductase NapE
K02595	nitrogenase-stabilizing/protective protein
K02659	twitching motility protein PilI
K02665	type IV pilus assembly protein PilP
K02672	type IV pilus assembly protein PilW
K02676	type IV pilus assembly protein PilZ
K02683	DNA primase [EC:2.7.7.-]
K02685	DNA primase large subunit [EC:2.7.7.-]
K02691	photosystem I subunit VII
K02717	photosystem II oxygen-evolving enhancer protein 2
K02866	large subunit ribosomal protein L10e
K02869	large subunit ribosomal protein L12
K02877	large subunit ribosomal protein L15e
K02883	large subunit ribosomal protein L18e
K02885	large subunit ribosomal protein L19e

KEGG orthology	Function
K02889	large subunit ribosomal protein L21e
K02896	large subunit ribosomal protein L24e
K02910	large subunit ribosomal protein L31e
K02912	large subunit ribosomal protein L32e
K02915	large subunit ribosomal protein L34e
K02921	large subunit ribosomal protein L37Ae
K02922	large subunit ribosomal protein L37e
K02924	large subunit ribosomal protein L39e
K02927	large subunit ribosomal protein L40e
K02929	large subunit ribosomal protein L44e
K02930	large subunit ribosomal protein L4e
K02944	large subunit ribosomal protein LX
K02966	small subunit ribosomal protein S19e
K02974	small subunit ribosomal protein S24e
K02978	small subunit ribosomal protein S27e
K02979	small subunit ribosomal protein S28e
K02984	small subunit ribosomal protein S3Ae
K02987	small subunit ribosomal protein S4e
K02991	small subunit ribosomal protein S6e
K02995	small subunit ribosomal protein S8e
K03044	DNA-directed RNA polymerase subunit B' [EC:2.7.7.6]
K03045	DNA-directed RNA polymerase subunit B" [EC:2.7.7.6]
K03049	DNA-directed RNA polymerase subunit E' [EC:2.7.7.6]
K03050	DNA-directed RNA polymerase subunit E" [EC:2.7.7.6]
K03051	DNA-directed RNA polymerase subunit F [EC:2.7.7.6]
K03053	DNA-directed RNA polymerase subunit H [EC:2.7.7.6]
K03055	DNA-directed RNA polymerase subunit K [EC:2.7.7.6]
K03057	transcription elongation factor
K03058	DNA-directed RNA polymerase subunit N [EC:2.7.7.6]
K03059	DNA-directed RNA polymerase subunit P [EC:2.7.7.6]
K03105	signal recognition particle subunit SRP19
K03120	transcription initiation factor TFIID TATA-box-binding protein
K03136	transcription initiation factor TFIIE subunit alpha
K03166	DNA topoisomerase VI subunit A [EC:5.99.1.3]
K03167	DNA topoisomerase VI subunit B [EC:5.99.1.3]
K03232	elongation factor 1-beta
K03236	translation initiation factor 1A
K03237	translation initiation factor 2 subunit 1
K03239	translation initiation factor eIF-2B subunit alpha
K03242	translation initiation factor 2 subunit 3
K03243	translation initiation factor 5B
K03264	translation initiation factor 6
K03268	benzene 1,2-dioxygenase [EC:1.14.12.3]
K03395	gentamicin 3'-N-acetyltransferase [EC:2.3.1.60]
K03538	ribonuclease P protein subunit POP4 [EC:3.1.26.5]
K03539	ribonuclease P/MRP protein subunit RPP1 [EC:3.1.26.5]
K03540	ribonuclease P protein subunit RPR2 [EC:3.1.26.5]
K03622	archaea-specific DNA-binding protein
K03626	nascent polypeptide-associated complex subunit alpha
K04071	6-pyruvoyltetrahydropterin 2'-reductase [EC:1.1.1.220]
K04090	indolepyruvate ferredoxin oxidoreductase [EC:1.2.7.8]
K04097	glutathione S-transferase [EC:2.5.1.18]
K04340	scyllo-inosamine-4-phosphate amidinotransferase 1 [EC:2.1.4.2]
K04341	dTDP-dihydrostreptose-streptidine-6-phosphate dihydrostreptosyltransferase [EC:2.4.2.27]
K04484	DNA repair protein RadB
K04496	C-terminal binding protein
K04791	mycobactin polyketide synthetase MbtD
K04795	fibrillarin-like pre-rRNA processing protein
K04796	small nuclear ribonucleoprotein
K04797	prefoldin alpha subunit
K04798	prefoldin beta subunit
K04801	replication factor C small subunit
K04802	proliferating cell nuclear antigen
K05301	sulfite dehydrogenase [EC:1.8.2.1]
K05383	CpeT protein
K05551	act minimal PKS ketosynthase (KS/KS alpha) [EC:2.3.1.-]
K05552	act minimal PKS chain-length factor (CLF/KS beta) [EC:2.3.1.-]
K05553	act minimal PKS acyl carrier protein
K05559	multicomponent K+:H+ antiporter subunit A
K05560	multicomponent K+:H+ antiporter subunit C

KEGG orthology	Function
K05561	multicomponent K+:H+ antiporter subunit D
K05562	multicomponent K+:H+ antiporter subunit E
K05563	multicomponent K+:H+ antiporter subunit F
K05564	multicomponent K+:H+ antiporter subunit G
K05573	NAD(P)H-quinone oxidoreductase subunit 2 [EC:1.6.5.3]
K05574	NAD(P)H-quinone oxidoreductase subunit 3 [EC:1.6.5.3]
K05575	NAD(P)H-quinone oxidoreductase subunit 4 [EC:1.6.5.3]
K05576	NAD(P)H-quinone oxidoreductase subunit 4L [EC:1.6.5.3]
K05597	glutamin-(asparagin-)ase [EC:3.5.1.38]
K05716	cyclic 2,3-diphosphoglycerate synthetase [EC:4.6.1.-]
K05797	4-cresol dehydrogenase (hydroxylating) [EC:1.17.99.1]
K05889	polyvinyl alcohol dehydrogenase (cytochrome) [EC:1.1.2.6]
K05908	NA
K05927	quinone-reactive Ni/Fe-hydrogenase small subunit [EC:1.12.5.1]
K05973	poly(3-hydroxybutyrate) depolymerase [EC:3.1.1.75]
K05986	NA
K05994	bacterial leucyl aminopeptidase [EC:3.4.11.10]
K05998	pseudomonalisin [EC:3.4.21.100]
K06034	sulfopyruvate decarboxylase subunit alpha [EC:4.1.1.79]
K06154	Lrp/AsnC family transcriptional regulator, involved in the regulation of lysine biosynthesis
K06237	collagen, type IV, alpha
K06363	response regulator aspartate phosphatase E [EC:3.1.-.-]
K06364	response regulator aspartate phosphatase F [EC:3.1.-.-]
K06377	sporulation-control protein
K06434	small acid-soluble spore protein (thioredoxin-like protein)
K06596	chemosensory pili system protein ChpA (sensor histidine kinase/response regulator)
K06601	flagellar protein FlbT
K06602	flagellar protein FlaF
K06718	L-2,4-diaminobutyric acid acetyltransferase [EC:2.3.1.178]
K06862	energy-converting hydrogenase B subunit Q
K06863	5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase [EC:6.3.4.-]
K06868	Sep-tRNA:Cys-tRNA synthetase [EC:2.5.1.73]
K06869	uncharacterized protein
K06873	NA
K06874	zinc finger protein
K06875	programmed cell death protein 5
K06914	NA
K06930	NA
K06931	NA
K06943	nucleolar GTP-binding protein
K06953	NA
K06961	ribosomal RNA assembly protein
K06963	tRNA acetyltransferase TAN1
K06964	NA
K06982	pantoate kinase [EC:2.7.1.169]
K06984	NA
K07049	TatD-related deoxyribonuclease
K07055	tRNA wybutosine-synthesizing protein 2 [EC:2.1.1.-]
K07060	UPF0271 protein
K07061	NA
K07073	NA
K07086	NA
K07092	NA
K07103	NA
K07108	NA
K07123	NA
K07131	NA
K07135	NA
K07144	NA
K07155	quercetin 2,3-dioxygenase [EC:1.13.11.24]
K07157	NA
K07158	NA
K07159	NA
K07178	RIO kinase 1 [EC:2.7.11.1]
K07179	RIO kinase 2 [EC:2.7.11.1]
K07244	mgtE-like transporter
K07254	tRNA (cytidine56-2'-O)-methyltransferase [EC:2.1.1.206]
K07288	uncharacterized membrane protein
K07333	archaeal flagellar protein FlaJ
K07338	hypothetical protein

KEGG orthology	Function
K07342	protein transport protein SEC61 subunit gamma and related proteins
K07394	SM-20-related protein
K07477	translin
K07544	benzylsuccinate CoA-transferase BbsF subunit [EC:2.8.3.15]
K07558	tRNA nucleotidyltransferase (CCA-adding enzyme) [EC:2.7.7.72]
K07562	nonsense-mediated mRNA decay protein 3
K07569	RNA-binding protein
K07572	putative nucleotide binding protein
K07575	PUA domain protein
K07580	hypothetical protein
K07581	hypothetical protein
K07585	hypothetical protein
K07732	riboflavin kinase, archaea type [EC:2.7.1.161]
K07823	3-oxoadipyl-CoA thiolase [EC:2.3.1.174]
K08076	astacin [EC:3.4.24.21]
K08085	type IV fimbrial biogenesis protein FimU
K08096	GTP cyclohydrolase IIa [EC:3.5.4.29]
K08176	MFS transporter, PHS family, inorganic phosphate transporter
K08477	outer membrane protease E [EC:3.4.21.-]
K08598	YopJ protease family
K08604	vibriolysin [EC:3.4.24.25]
K08645	anthrax lethal toxin endopeptidase [EC:3.4.24.83]
K08698	carbon dioxide concentrating mechanism protein CcmM
K08713	potassium channel LctB
K08953	chlorosome envelope protein J
K08971	putative membrane protein
K08973	putative membrane protein
K08975	putative membrane protein
K08979	putative membrane protein
K08983	putative membrane protein
K09119	hypothetical protein
K09139	hypothetical protein
K09140	pre-rRNA-processing protein TSR3
K09148	hypothetical protein
K09152	hypothetical protein
K09154	hypothetical protein
K09713	hypothetical protein
K09721	hypothetical protein
K09722	4-phosphopantoate---beta-alanine ligase [EC:6.3.2.36]
K09723	hypothetical protein
K09724	hypothetical protein
K09725	hypothetical protein
K09727	hypothetical protein
K09732	hypothetical protein
K09735	hypothetical protein
K09736	hypothetical protein
K09737	hypothetical protein
K09738	hypothetical protein
K09739	hypothetical protein
K09741	hypothetical protein
K09744	hypothetical protein
K09746	hypothetical protein
K09796	hypothetical protein
K09845	1-hydroxycarotenoid 3,4-desaturase [EC:1.3.99.27]
K09846	demethylspheroidene O-methyltransferase [EC:2.1.1.210]
K09879	isorenieratene synthase
K09919	hypothetical protein
K09941	hypothetical protein
K09943	hypothetical protein
K09947	hypothetical protein
K09950	hypothetical protein
K09959	hypothetical protein
K09965	hypothetical protein
K09966	hypothetical protein
K09983	hypothetical protein
K10023	arginine/ornithine transport system permease protein
K10024	arginine/ornithine transport system permease protein
K10025	arginine/ornithine transport system ATP-binding protein [EC:3.6.3.-]
K10123	putative ferrous iron transport protein C
K10216	2-hydroxymuconate-semialdehyde hydrolase [EC:3.7.1.9]

KEGG orthology	Function
K10221	2-pyrone-4,6-dicarboxylate lactonase [EC:3.1.1.57]
K10222	2,6-dioxo-6-phenylhexa-3-enoate hydrolase [EC:3.7.1.8]
K10233	alpha-glucoside transport system permease protein
K10623	HOMODA hydrolase [EC:3.7.1.-]
K10725	archaeal cell division control protein 6
K10764	O-succinylhomoserine sulfhydrylase [EC:2.5.1.-]
K10829	iron-chelate-transporting ATPase [EC:3.6.3.34]
K10873	DNA repair and recombination protein RAD52
K10911	two-component system, phosphorelay protein LuxU
K10923	AraC family transcriptional regulator, TCP pilus virulence regulatory protein
K10925	MSHA pilin protein MshB
K10926	MSHA pilin protein MshC
K10929	cholera enterotoxin subunit B
K10930	toxin co-regulated pilin
K10932	toxin co-regulated pilus biosynthesis outer membrane protein C
K10938	accessory colonization factor AcfC
K10966	toxin co-regulated pilus biosynthesis protein J [EC:3.4.23.43 2.1.1.-]
K11006	shiga toxin subunit A
K11007	shiga toxin subunit B
K11014	cytolethal distending toxin subunit B
K11015	cytolethal distending toxin subunit C
K11016	hemolysin
K11023	pertussis toxin subunit 1 [EC:2.4.2.-]
K11028	vacuolating cytotoxin
K11057	beta2-toxin
K11058	enterotoxin Cpe
K11212	LPPG:FO 2-phospho-L-lactate transferase [EC:2.7.8.28]
K11260	formylmethanofuran dehydrogenase subunit G [EC:1.2.99.5]
K11395	2-keto-3-deoxy-gluconate aldolase [EC:4.1.2.-]
K11434	protein arginine N-methyltransferase 1 [EC:2.1.1.-]
K11526	two-component system, chemotaxis family, sensor histidine kinase and response regulator PixL
K11600	exosome complex component RRP41
K11638	two-component system, CitB family, response regulator CitT
K11780	FO synthase subunit 1 [EC:2.5.1.77]
K11814	multidrug resistance protein EbrA
K11913	type VI secretion system protein
K12048	ComB10 competence protein
K12049	ComB9 competence protein
K12054	lytic transglycosylase AtlA
K12080	type IV secretion system protein PtlC
K12247	alpha-2,3 sialyltransferase [EC:2.4.99.-]
K12430	polyketide synthase 1/15
K12513	tight adherence protein E
K12514	tight adherence protein F
K12515	tight adherence protein G
K12534	membrane fusion protein RsaE
K12589	exosome complex component RRP42
K12673	N2-(2-carboxyethyl)arginine synthase [EC:2.5.1.66]
K12786	LEE-encoded effector EspF
K12789	Tir-cytoskeleton coupling protein
K12809	T3SS secreted effector EspG-like protein
K12953	cation-transporting ATPase F [EC:3.6.3.-]
K12978	lipid A 4'-phosphatase [EC:3.1.3.-]
K13008	O-antigen polymerase
K13039	sulfopyruvate decarboxylase subunit beta [EC:4.1.1.79]
K13060	acyl homoserine lactone synthase [EC:2.3.1.184]
K13284	invasin A
K13286	invasin C
K13317	NDP-4-keto-2,6-dideoxyhexose 3-C-methyltransferase
K13450	phosphothreonine lyase [EC:4.2.3.-]
K13455	avirulence protein
K13488	chemotaxis-related protein WspB
K13489	chemotaxis-related protein WspD
K13503	anthranilate synthase [EC:4.1.3.27]
K13520	outer membrane protease [EC:3.4.23.-]
K13583	GcrA cell cycle regulator
K13659	2-beta-glucuronyltransferase [EC:2.4.1.264]
K13661	GumC protein
K13740	secreted effector protein SptP
K13742	protein IpgB1

KEGG orthology	Function
K13790	virulence protein IcsB
K13793	secreted effector OspE
K13797	DNA-directed RNA polymerase subunit beta-beta' [EC:2.7.7.6]
K13798	DNA-directed RNA polymerase subunit B [EC:2.7.7.6]
K13875	L-arabonate dehydrase [EC:4.2.1.25]
K13941	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase / dihydropteroate synthase [EC:2.7.6.3 2.5.1.15]
K14092	energy-converting hydrogenase A subunit A
K14093	energy-converting hydrogenase A subunit B
K14094	energy-converting hydrogenase A subunit C
K14095	energy-converting hydrogenase A subunit D
K14096	energy-converting hydrogenase A subunit E
K14097	energy-converting hydrogenase A subunit F
K14098	energy-converting hydrogenase A subunit G
K14100	energy-converting hydrogenase A subunit I
K14102	energy-converting hydrogenase A subunit K
K14103	energy-converting hydrogenase A subunit L
K14104	energy-converting hydrogenase A subunit M
K14105	energy-converting hydrogenase A subunit N
K14109	energy-converting hydrogenase A subunit R
K14110	energy-converting hydrogenase B subunit A
K14111	energy-converting hydrogenase B subunit B
K14113	energy-converting hydrogenase B subunit D
K14114	energy-converting hydrogenase B subunit E
K14115	energy-converting hydrogenase B subunit F
K14116	energy-converting hydrogenase B subunit G
K14117	energy-converting hydrogenase B subunit H
K14118	energy-converting hydrogenase B subunit I
K14119	energy-converting hydrogenase B subunit J
K14122	energy-converting hydrogenase B subunit M
K14123	energy-converting hydrogenase B subunit N
K14124	energy-converting hydrogenase B subunit O
K14125	energy-converting hydrogenase B subunit P
K14165	dual specificity phosphatase [EC:3.1.3.16 3.1.3.48]
K14166	copper transport protein
K14201	clumping factor A
K14269	glutarate semialdehyde dehydrogenase [EC:1.2.1.20]
K14451	(3S)-malyl-CoA thioesterase
K14465	succinate semialdehyde reductase (NADPH) [EC:1.1.1.-]
K14468	malonyl-CoA reductase / 3-hydroxypropionate dehydrogenase (NADP+) [EC:1.2.1.75 1.1.1.298]
K14561	U3 small nucleolar ribonucleoprotein protein IMP4
K14564	nucleolar protein 56
K14568	essential for mitotic growth 1
K14598	chlorobactene lauroyltransferase
K14628	enoyl reductase
K14653	2-amino-5-formylamino-6-ribosylaminopyrimidin-4(3H)-one 5'-monophosphate deformylase [EC:3.5.1.102]
K14661	nodulation protein F [EC:2.3.1.-]
K14683	solute carrier family 34 (sodium-dependent phosphate cotransporter)
K14974	6-hydroxynicotinate 3-monooxygenase [EC:1.14.13.114]
K14998	surfeit locus 1 family protein
K15226	arogenate dehydrogenase (NADP+) [EC:1.3.1.78]
K15327	polyketide biosynthesis malonyl-CoA-[acyl-carrier-protein] transacylase
K15355	NA
K15366	salmonella plasmid virulence protein B
K15468	cytochrome P450 PksS
K15645	coronafacic acid polyketide synthase Cfa7
K15650	non-haem Fe2+, alpha-ketoglutarate-dependent halogenase
K15676	rhizoxin biosynthesis, polyketide synthase RhiC
K15681	aminotransferase MxcL
K15784	N-alpha-acetyl-L-2,4-diaminobutyrate deacetylase [EC:3.5.1.-]
K15845	outer membrane protein HopZ
K15853	acyl transferase [EC:2.3.1.-]
K15900	tRNA threonylcarbamoyladenosine biosynthesis protein
K15904	bifunctional tRNA threonylcarbamoyladenosine biosynthesis protein [EC:2.7.11.1]
K15918	D-glycerate 3-kinase [EC:2.7.1.31]
K16081	alginate production protein
K16149	1,4-alpha-glucan branching enzyme [EC:2.4.1.18]
K16152	heme acquisition protein HasR
K16190	glucuronokinase [EC:2.7.1.43]

NA indicates not assigned

Appendix 6. KOs having the highest abundance in the JP cohort among the 12 countries

KEGG orthology	Function
K00005	glycerol dehydrogenase [EC:1.1.1.6]
K00016	L-lactate dehydrogenase [EC:1.1.1.27]
K00020	3-hydroxyisobutyrate dehydrogenase [EC:1.1.1.31]
K00033	6-phosphogluconate dehydrogenase [EC:1.1.1.44]
K00042	2-hydroxy-3-oxopropionate reductase [EC:1.1.1.60]
K00044	estradiol 17beta-dehydrogenase [EC:1.1.1.62]
K00048	lactaldehyde reductase [EC:1.1.1.77]
K00052	3-isopropylmalate dehydrogenase [EC:1.1.1.85]
K00053	ketol-acid reductoisomerase [EC:1.1.1.86]
K00058	D-3-phosphoglycerate dehydrogenase [EC:1.1.1.95]
K00065	2-deoxy-D-gluconate 3-dehydrogenase [EC:1.1.1.125]
K00076	7-alpha-hydroxysteroid dehydrogenase [EC:1.1.1.159]
K00100	NA
K00128	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]
K00129	aldehyde dehydrogenase (NAD(P)+) [EC:1.2.1.5]
K00131	glyceraldehyde-3-phosphate dehydrogenase (NADP) [EC:1.2.1.9]
K00171	pyruvate ferredoxin oxidoreductase, delta subunit [EC:1.2.7.1]
K00196	carbon-monoxide dehydrogenase iron sulfur subunit
K00198	carbon-monoxide dehydrogenase catalytic subunit [EC:1.2.99.2]
K00215	dihydrodipicolinate reductase [EC:1.3.1.26]
K00324	NAD(P) transhydrogenase subunit alpha [EC:1.6.1.2]
K00325	NAD(P) transhydrogenase subunit beta [EC:1.6.1.2]
K00354	NADPH2 dehydrogenase [EC:1.6.99.1]
K00359	NADH oxidase [EC:1.6.-.-]
K00384	thioredoxin reductase (NADPH) [EC:1.8.1.9]
K00386	NA
K00459	nitronate monooxygenase [EC:1.13.12.16]
K00492	NA
K00564	16S rRNA (guanine1207-N2)-methyltransferase [EC:2.1.1.172]
K00588	caffeoyl-CoA O-methyltransferase [EC:2.1.1.104]
K00595	precorrin-6Y C5,15-methyltransferase / precorrin-8W decarboxylase [EC:2.1.1.132 1.-.-.-]
K00614	NA
K00616	transaldolase [EC:2.2.1.2]
K00620	glutamate N-acetyltransferase / amino-acid N-acetyltransferase [EC:2.3.1.35 2.3.1.1]
K00651	homoserine O-succinyltransferase [EC:2.3.1.46]
K00674	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase [EC:2.3.1.117]
K00687	penicillin-binding protein 2B [EC:2.3.2.-]
K00689	dextransucrase [EC:2.4.1.5]
K00690	sucrose phosphorylase [EC:2.4.1.7]
K00762	orotate phosphoribosyltransferase [EC:2.4.2.10]
K00777	NA
K00797	spermidine synthase [EC:2.5.1.16]
K00798	cob(I)alamin adenosyltransferase [EC:2.5.1.17]
K00842	aminotransferase [EC:2.6.1.-]
K00851	gluconokinase [EC:2.7.1.12]
K00852	ribokinase [EC:2.7.1.15]
K00854	xylulokinase [EC:2.7.1.17]
K00864	glycerol kinase [EC:2.7.1.30]
K00868	pyridoxine kinase [EC:2.7.1.35]
K00872	homoserine kinase [EC:2.7.1.39]
K00878	hydroxyethylthiazole kinase [EC:2.7.1.50]
K00886	polyphosphate glucokinase [EC:2.7.1.63]
K00917	tagatose 6-phosphate kinase [EC:2.7.1.144]
K00926	carbamate kinase [EC:2.7.2.2]
K00928	aspartate kinase [EC:2.7.2.4]
K00938	phosphomevalonate kinase [EC:2.7.4.2]
K00942	guanylate kinase [EC:2.7.4.8]
K00949	thiamine pyrophosphokinase [EC:2.7.6.2]
K00960	DNA-directed RNA polymerase [EC:2.7.7.6]
K00965	UDPglucose--hexose-1-phosphate uridylyltransferase [EC:2.7.7.12]
K00982	glutamate-ammonia-ligase adenylyltransferase [EC:2.7.7.42]
K00989	ribonuclease PH [EC:2.7.7.56]
K00997	holo-[acyl-carrier protein] synthase [EC:2.7.8.7]
K01012	biotin synthetase [EC:2.8.1.6]
K01026	propionate CoA-transferase [EC:2.8.3.1]

KEGG orthology	Function
K01055	3-oxoadipate enol-lactonase [EC:3.1.1.24]
K01104	protein-tyrosine phosphatase [EC:3.1.3.48]
K01121	2',3'-cyclic-nucleotide 3'-phosphodiesterase [EC:3.1.4.37]
K01182	oligo-1,6-glucosidase [EC:3.2.1.10]
K01191	alpha-mannosidase [EC:3.2.1.24]
K01193	beta-fructofuranosidase [EC:3.2.1.26]
K01210	glucan 1,3-beta-glucosidase [EC:3.2.1.58]
K01215	glucan 1,6-alpha-glucosidase [EC:3.2.1.70]
K01220	6-phospho-beta-galactosidase [EC:3.2.1.85]
K01223	6-phospho-beta-glucosidase [EC:3.2.1.86]
K01232	maltose-6'-phosphate glucosidase [EC:3.2.1.122]
K01239	purine nucleosidase [EC:3.2.2.1]
K01256	aminopeptidase N [EC:3.4.11.2]
K01261	glutamyl aminopeptidase [EC:3.4.11.7]
K01266	D-aminopeptidase [EC:3.4.11.19]
K01274	D-alanyl-D-alanine dipeptidase [EC:3.4.13.-]
K01305	beta-aspartyl-dipeptidase (metallo-type) [EC:3.4.19.-]
K01322	prolyl oligopeptidase [EC:3.4.21.26]
K01354	oligopeptidase B [EC:3.4.21.83]
K01400	bacillolysin [EC:3.4.24.28]
K01420	CRP/FNR family transcriptional regulator, anaerobic regulatory protein
K01421	putative membrane protein
K01444	N4-(beta-N-acetylglucosaminyl)-L-asparaginase [EC:3.5.1.26]
K01446	N-acetylmuramoyl-L-alanine amidase [EC:3.5.1.28]
K01459	NA
K01462	peptide deformylase [EC:3.5.1.88]
K01466	allantoinase [EC:3.5.2.5]
K01470	creatinine amidohydrolase [EC:3.5.2.10]
K01473	N-methylhydantoinase A [EC:3.5.2.14]
K01478	arginine deiminase [EC:3.5.3.6]
K01488	adenosine deaminase [EC:3.5.4.4]
K01494	dCTP deaminase [EC:3.5.4.13]
K01496	phosphoribosyl-AMP cyclohydrolase [EC:3.5.4.19]
K01515	ADP-ribose pyrophosphatase [EC:3.6.1.13]
K01523	phosphoribosyl-ATP pyrophosphohydrolase [EC:3.6.1.31]
K01567	NA
K01577	oxalyl-CoA decarboxylase [EC:4.1.1.8]
K01589	5-(carboxyamino)imidazole ribonucleotide synthase [EC:6.3.4.18]
K01593	aromatic-L-amino-acid decarboxylase [EC:4.1.1.28]
K01595	phosphoenolpyruvate carboxylase [EC:4.1.1.31]
K01597	diphosphomevalonate decarboxylase [EC:4.1.1.33]
K01599	uroporphyrinogen decarboxylase [EC:4.1.1.37]
K01620	threonine aldolase [EC:4.1.2.5]
K01626	3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]
K01632	fructose-6-phosphate phosphoketolase [EC:4.1.2.22]
K01653	acetolactate synthase I/III small subunit [EC:2.2.1.6]
K01658	anthranilate synthase component II [EC:4.1.3.27]
K01664	para-aminobenzoate synthetase component II [EC:2.6.1.85]
K01697	cystathionine beta-synthase [EC:4.2.1.22]
K01699	propanediol dehydratase large subunit [EC:4.2.1.28]
K01704	3-isopropylmalate/(R)-2-methylmalate dehydratase small subunit [EC:4.2.1.33 4.2.1.35]
K01706	glucarate dehydratase [EC:4.2.1.40]
K01708	galactarate dehydratase [EC:4.2.1.42]
K01739	cystathionine gamma-synthase [EC:2.5.1.48]
K01749	hydroxymethylbilane synthase [EC:2.5.1.61]
K01751	diaminopropionate ammonia-lyase [EC:4.3.1.15]
K01754	threonine dehydratase [EC:4.3.1.19]
K01759	lactoylglutathione lyase [EC:4.4.1.5]
K01760	cystathionine beta-lyase [EC:4.4.1.8]
K01777	proline racemase [EC:5.1.1.4]
K01781	mandelate racemase [EC:5.1.2.2]
K01788	N-acylglucosamine-6-phosphate 2-epimerase [EC:5.1.3.9]
K01792	glucose-6-phosphate 1-epimerase [EC:5.1.3.15]
K01807	ribose 5-phosphate isomerase A [EC:5.3.1.6]
K01814	phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase [EC:5.3.1.16]
K01817	phosphoribosylanthranilate isomerase [EC:5.3.1.24]
K01819	galactose-6-phosphate isomerase [EC:5.3.1.26]
K01820	NA
K01823	isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]
K01826	5-carboxymethyl-2-hydroxymuconate isomerase [EC:5.3.3.10]

KEGG orthology	Function
K01839	phosphopentomutase [EC:5.4.2.7]
K01854	UDP-galactopyranose mutase [EC:5.4.99.9]
K01903	succinyl-CoA synthetase beta subunit [EC:6.2.1.5]
K01906	6-carboxyhexanoate--CoA ligase [EC:6.2.1.14]
K01951	GMP synthase (glutamine-hydrolysing) [EC:6.3.5.2]
K01961	acetyl-CoA carboxylase, biotin carboxylase subunit [EC:6.4.1.2 6.3.4.14]
K02004	putative ABC transport system permease protein
K02007	cobalt/nickel transport system permease protein
K02018	molybdate transport system permease protein
K02020	molybdate transport system substrate-binding protein
K02025	multiple sugar transport system permease protein
K02026	multiple sugar transport system permease protein
K02027	multiple sugar transport system substrate-binding protein
K02030	polar amino acid transport system substrate-binding protein
K02032	peptide/nickel transport system ATP-binding protein
K02033	peptide/nickel transport system permease protein
K02035	peptide/nickel transport system substrate-binding protein
K02036	phosphate transport system ATP-binding protein [EC:3.6.3.27]
K02039	phosphate transport system protein
K02054	putative spermidine/putrescine transport system permease protein
K02055	putative spermidine/putrescine transport system substrate-binding protein
K02068	putative ABC transport system ATP-binding protein
K02074	zinc/manganese transport system ATP-binding protein
K02077	zinc/manganese transport system substrate-binding protein
K02081	DeoR family transcriptional regulator, aga operon transcriptional repressor
K02082	tagatose-6-phosphate ketose/aldose isomerase [EC:5.-.-.-]
K02083	allantoate deiminase [EC:3.5.3.9]
K02100	MFS transporter, SP family, arabinose:H+ symporter
K02103	GntR family transcriptional regulator, arabinose operon transcriptional repressor
K02113	F-type H+-transporting ATPase subunit delta [EC:3.6.3.14]
K02122	V-type H+-transporting ATPase subunit F [EC:3.6.3.14]
K02188	cobalt-precorrin-5B (C1)-methyltransferase [EC:2.1.1.195]
K02189	cobalt-precorrin 5A hydrolase [EC:3.7.1.12]
K02190	sirohydrochlorin cobaltochelatase [EC:4.99.1.3]
K02191	cobalt-precorrin-7 (C15)-methyltransferase [EC:2.1.1.196]
K02224	cobyrinic acid a,c-diamide synthase [EC:6.3.5.9 6.3.5.11]
K02231	adenosylcobinamide kinase / adenosylcobinamide-phosphate guanylyltransferase [EC:2.7.1.156 2.7.7.62]
K02232	adenosylcobyric acid synthase [EC:6.3.5.10]
K02237	competence protein ComEA
K02243	competence protein ComGA
K02246	competence protein ComGD
K02304	precorrin-2 dehydrogenase / sirohydrochlorin ferrochelatase [EC:1.3.1.76 4.99.1.4]
K02424	cystine transport system substrate-binding protein
K02438	glycogen operon protein [EC:3.2.1.-]
K02440	glycerol uptake facilitator protein
K02443	glycerol uptake operon antiterminator
K02444	DeoR family transcriptional regulator, glycerol-3-phosphate regulon repressor
K02501	glutamine amidotransferase [EC:2.4.2.-]
K02525	LacI family transcriptional regulator, kdg operon repressor
K02526	2-keto-3-deoxygluconate permease
K02529	LacI family transcriptional regulator
K02530	DeoR family transcriptional regulator, lactose phosphotransferase system repressor
K02531	transcriptional antiterminator
K02532	MFS transporter, OHS family, lactose permease
K02565	N-acetylglucosamine repressor
K02575	MFS transporter, NNP family, nitrate/nitrite transporter
K02598	nitrite transporter NirC
K02624	IclR family transcriptional regulator, pca regulon regulatory protein
K02647	carbohydrate diacid regulator
K02671	type IV pilus assembly protein PilV
K02688	transcriptional regulator, propionate catabolism operon regulatory protein
K02744	PTS system, N-acetylgalactosamine-specific IIA component [EC:2.7.1.69]
K02745	PTS system, N-acetylgalactosamine-specific IIB component [EC:2.7.1.69]
K02746	PTS system, N-acetylgalactosamine-specific IIC component
K02747	PTS system, N-acetylgalactosamine-specific IID component
K02749	PTS system, arbutin-like IIB component [EC:2.7.1.69]
K02755	PTS system, beta-glucosides-specific IIA component [EC:2.7.1.69]
K02756	PTS system, beta-glucosides-specific IIB component [EC:2.7.1.69]
K02757	PTS system, beta-glucosides-specific IIC component
K02759	PTS system, cellobiose-specific IIA component [EC:2.7.1.69]

KEGG orthology	Function
K02760	PTS system, cellobiose-specific IIB component [EC:2.7.1.69]
K02761	PTS system, cellobiose-specific IIC component
K02763	PTS system, D-glucosamine-specific IIA component [EC:2.7.1.69]
K02764	PTS system, D-glucosamine-specific IIB component [EC:2.7.1.69]
K02765	PTS system, D-glucosamine-specific IIC component
K02771	PTS system, fructose-specific IID component
K02773	PTS system, galactitol-specific IIA component [EC:2.7.1.69]
K02774	PTS system, galactitol-specific IIB component [EC:2.7.1.69]
K02777	PTS system, glucose-specific IIA component [EC:2.7.1.69]
K02786	PTS system, lactose-specific IIA component [EC:2.7.1.69]
K02787	PTS system, lactose-specific IIB component [EC:2.7.1.69]
K02793	PTS system, mannose-specific IIA component [EC:2.7.1.69]
K02794	PTS system, mannose-specific IIB component [EC:2.7.1.69]
K02795	PTS system, mannose-specific IIC component
K02796	PTS system, mannose-specific IID component
K02803	PTS system, N-acetylglucosamine-specific IIB component [EC:2.7.1.69]
K02808	PTS system, sucrose-specific IIA component [EC:2.7.1.69]
K02809	PTS system, sucrose-specific IIB component [EC:2.7.1.69]
K02810	PTS system, sucrose-specific IIC component
K02817	PTS system, trehalose-specific IIA component [EC:2.7.1.69]
K02819	PTS system, trehalose-specific IIC component
K02855	AraC family transcriptional regulator, L-rhamnose operon regulatory protein RhaS
K03147	thiamine biosynthesis protein ThiC
K03148	sulfur carrier protein ThiS adenylyltransferase [EC:2.7.7.73]
K03149	thiamine biosynthesis ThiG
K03182	3-octaprenyl-4-hydroxybenzoate carboxy-lyase UbiD [EC:4.1.1.-]
K03186	3-octaprenyl-4-hydroxybenzoate carboxy-lyase UbiX [EC:4.1.1.-]
K03216	tRNA (cytidine/uridine-2'-O-)-methyltransferase [EC:2.1.1.207]
K03293	amino acid transporter, AAT family
K03294	basic amino acid/polyamine antiporter, APA family
K03295	cation efflux system protein, CDF family
K03299	gluconate:H+ symporter, GntP family
K03322	manganese transport protein
K03332	fructan beta-fructosidase [EC:3.2.1.80]
K03342	para-aminobenzoate synthetase / 4-amino-4-deoxychorismate lyase [EC:2.6.1.85 4.1.3.38]
K03394	precorrin-2/cobalt-factor-2 C20-methyltransferase [EC:2.1.1.130 2.1.1.151]
K03400	long-chain-fatty-acyl-CoA reductase [EC:1.2.1.50]
K03426	NAD+ diphosphatase [EC:3.6.1.22]
K03439	tRNA (guanine-N7-)-methyltransferase [EC:2.1.1.33]
K03446	MFS transporter, DHA2 family, multidrug resistance protein B
K03458	nucleobase:cation symporter-2, NCS2 family
K03475	PTS system, ascorbate-specific IIC component
K03480	transcriptional antiterminator
K03481	RpiR family transcriptional regulator, glv operon transcriptional regulator
K03484	LacI family transcriptional regulator, sucrose operon repressor
K03486	GntR family transcriptional regulator, trehalose operon transcriptional repressor
K03488	beta-glucoside operon transcriptional antiterminator
K03491	lichenan operon transcriptional antiterminator
K03492	GntR family transcriptional regulator
K03518	carbon-monoxide dehydrogenase small subunit [EC:1.2.99.2]
K03519	carbon-monoxide dehydrogenase medium subunit [EC:1.2.99.2]
K03547	exonuclease SbcD
K03604	LacI family transcriptional regulator, purine nucleotide synthesis repressor
K03635	molybdopterin synthase catalytic subunit [EC:2.-.-.-]
K03637	molybdenum cofactor biosynthesis protein C
K03646	colicin import membrane protein
K03647	protein involved in ribonucleotide reduction
K03676	glutaredoxin 3
K03680	translation initiation factor eIF-2B subunit delta
K03684	ribonuclease D [EC:3.1.13.5]
K03688	ubiquinone biosynthesis protein
K03693	penicillin-binding protein
K03707	thiaminase (transcriptional activator TenA) [EC:3.5.99.2]
K03710	GntR family transcriptional regulator
K03721	transcriptional regulator of aroF, aroG, tyrA and aromatic amino acid transport
K03724	ATP-dependent helicase Lhr and Lhr-like helicase [EC:3.6.4.-]
K03727	ATP-dependent RNA helicase HelY [EC:3.6.4.-]
K03743	NA
K03750	molybdopterin molybdotransferase [EC:2.10.1.1]
K03753	molybdopterin-guanine dinucleotide biosynthesis protein B

KEGG orthology	Function
K03767	peptidyl-prolyl cis-trans isomerase A (cyclophilin A) [EC:5.2.1.8]
K03785	3-dehydroquinate dehydratase I [EC:4.2.1.10]
K03790	ribosomal-protein-alanine N-acetyltransferase [EC:2.3.1.128]
K03799	heat shock protein HtpX [EC:3.4.24.-]
K03823	phosphinothricin acetyltransferase [EC:2.3.1.183]
K03831	molybdopterin adenylyltransferase [EC:2.7.7.75]
K03833	selenocysteine-specific elongation factor
K03851	taurine-pyruvate aminotransferase [EC:2.6.1.77]
K03975	membrane-associated protein
K03980	virulence factor
K04023	ethanolamine transporter
K04034	anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase [EC:4.-.-.-]
K04069	pyruvate formate lyase activating enzyme [EC:1.97.1.4]
K04085	tRNA 2-thiouridine synthesizing protein A [EC:2.8.1.-]
K04086	ATP-dependent Clp protease ATP-binding subunit ClpL
K04092	chorismate mutase [EC:5.4.99.5]
K04477	putative hydrolase
K04565	Cu/Zn superoxide dismutase [EC:1.15.1.1]
K04653	hydrogenase expression/formation protein HypC
K04654	hydrogenase expression/formation protein HypD
K04655	hydrogenase expression/formation protein HypE
K04720	threonine-phosphate decarboxylase [EC:4.1.1.81]
K04748	nitric oxide reductase NorQ protein
K04757	anti-sigma B factor [EC:2.7.11.1]
K04761	LysR family transcriptional regulator, hydrogen peroxide-inducible genes activator
K04767	acetoin utilization protein AcuB
K04783	yersiniabactin salicyl-AMP ligase [EC:6.3.2.-]
K04940	opine dehydrogenase [EC:1.5.1.28]
K05020	glycine betaine transporter
K05275	pyridoxine 4-dehydrogenase [EC:1.1.1.65]
K05303	O-methyltransferase [EC:2.1.1.-]
K05339	holin-like protein LrgB
K05341	amylosucrase [EC:2.4.1.4]
K05350	beta-glucosidase [EC:3.2.1.21]
K05351	D-xylulose reductase [EC:1.1.1.9]
K05362	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-L-lysine ligase [EC:6.3.2.7]
K05499	LacI family transcriptional regulator, repressor for deo operon, udp, cdd, tsx, nupC, and nupG
K05541	tRNA-dihydrouridine synthase C [EC:1.-.-.-]
K05783	1,6-dihydroxycyclohexa-2,4-diene-1-carboxylate dehydrogenase [EC:1.3.1.25]
K05786	chloramphenicol-sensitive protein RarD
K05823	N-acetyldiaminopimelate deacetylase [EC:3.5.1.47]
K05836	GntR family transcriptional regulator, histidine utilization repressor
K05838	putative thioredoxin
K05845	osmoprotectant transport system substrate-binding protein
K05903	NADH dehydrogenase (quinone) [EC:1.6.99.5]
K05936	precorrin-4 C11-methyltransferase [EC:2.1.1.133]
K05937	hypothetical protein
K06016	N-carbamoyl-L-amino-acid hydrolase [EC:3.5.1.87]
K06019	pyrophosphatase PpaX [EC:3.6.1.1]
K06042	precorrin-8X methylmutase [EC:5.4.1.2]
K06046	long-chain-fatty-acid---luciferin-component ligase [EC:6.2.1.19]
K06155	Gnt-I system high-affinity gluconate transporter
K06187	recombination protein RecR
K06189	magnesium and cobalt transporter
K06191	glutaredoxin-like protein NrdH
K06199	CrcB protein
K06201	copper homeostasis protein
K06221	2,5-diketo-D-gluconate reductase A [EC:1.1.1.274]
K06317	inhibitor of the pro-sigma K processing machinery
K06330	spore coat protein H
K06331	spore coat protein I
K06403	stage V sporulation protein AA
K06606	inosose isomerase [EC:5.3.99.-]
K06610	MFS transporter, SP family, inositol transporter
K06864	uncharacterized protein
K06898	NA
K06902	MFS transporter, UMF1 family
K06910	NA
K06924	NA
K06940	NA

KEGG orthology	Function
K06951	NA
K06962	NA
K06971	NA
K06972	NA
K06999	phospholipase/carboxylesterase
K07006	NA
K07008	glutamine amidotransferase
K07013	NA
K07023	putative hydrolases of HD superfamily
K07024	NA
K07038	inner membrane protein
K07046	NA
K07047	NA
K07067	DNA integrity scanning protein
K07104	NA
K07105	NA
K07138	NA
K07173	S-ribosylhomocysteine lyase [EC:4.4.1.21]
K07177	PDZ domain-containing protein
K07230	NA
K07243	high-affinity iron transporter
K07393	putative glutathione S-transferase
K07396	putative protein-disulfide isomerase
K07402	xanthine dehydrogenase accessory factor
K07442	tRNA (adenine57-N1/adenine58-N1)-methyltransferase [EC:2.1.1.219 2.1.1.220]
K07443	methylated-DNA-protein-cysteine methyltransferase related protein
K07457	endonuclease III related protein
K07461	putative endonuclease
K07467	phage replication initiation protein
K07503	hypothetical protein
K07646	two-component system, OmpR family, sensor histidine kinase KdpD [EC:2.7.13.3]
K07649	two-component system, OmpR family, sensor histidine kinase TctE [EC:2.7.13.3]
K07654	two-component system, OmpR family, sensor histidine kinase MtrB [EC:2.7.13.3]
K07664	two-component system, OmpR family, response regulator BaeR
K07680	two-component system, NarL family, sensor histidine kinase ComP [EC:2.7.13.3]
K07685	two-component system, NarL family, nitrate/nitrite response regulator NarP
K07688	two-component system, NarL family, response regulator, fimbrial Z protein, FimZ
K07690	two-component system, NarL family, response regulator EvgA
K07692	two-component system, NarL family, response regulator DegU
K07693	two-component system, NarL family, response regulator DesR
K07710	two-component system, NtrC family, sensor histidine kinase AtoS [EC:2.7.13.3]
K07718	two-component system, sensor histidine kinase YesM [EC:2.7.13.3]
K07719	two-component system, response regulator YcbB
K07720	two-component system, response regulator YesN
K07722	CopG family transcriptional regulator, nickel-responsive regulator
K07739	elongator complex protein 3 [EC:2.3.1.48]
K07749	formyl-CoA transferase [EC:2.8.3.16]
K07768	two-component system, OmpR family, sensor histidine kinase SenX3 [EC:2.7.13.3]
K07770	two-component system, OmpR family, response regulator CssR
K07794	putative tricarboxylic transport membrane protein
K08092	3-dehydro-L-gulonate 2-dehydrogenase [EC:1.1.1.130]
K08094	6-phospho-3-hexuloisomerase [EC:5.3.1.27]
K08139	MFS transporter, SP family, sugar:H+ symporter
K08152	MFS transporter, DHA1 family, multidrug resistance protein B
K08156	MFS transporter, DHA1 family, arabinose polymer transporter
K08164	MFS transporter, DHA1 family, chloramphenicol resistance protein
K08168	MFS transporter, DHA2 family, metal-tetracycline-proton antiporter
K08174	MFS transporter, FHS family, glucose/mannose:H+ symporter
K08177	MFS transporter, OFA family, oxalate/formate antiporter
K08282	non-specific serine/threonine protein kinase [EC:2.7.11.1]
K08296	phosphohistidine phosphatase [EC:3.1.3.-]
K08300	ribonuclease E [EC:3.1.26.12]
K08302	tagatose 1,6-diphosphate aldolase [EC:4.1.2.40]
K08313	fructose-6-phosphate aldolase 1 [EC:4.1.2.-]
K08369	MFS transporter, putative metabolite:H+ symporter
K08372	putative serine protease PepD [EC:3.4.21.-]
K08483	phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]
K08643	zinc metalloprotease ZmpB [EC:3.4.24.-]
K08700	carbon dioxide concentrating mechanism protein CcmO
K08969	aminotransferase [EC:2.6.1.-]

KEGG orthology	Function
K08972	putative membrane protein
K08987	putative membrane protein
K09004	hypothetical protein
K09009	hypothetical protein
K09017	TetR/AcrR family transcriptional regulator
K09117	hypothetical protein
K09118	hypothetical protein
K09121	hypothetical protein
K09133	hypothetical protein
K09155	hypothetical protein
K09681	LysR family transcriptional regulator, transcription activator of glutamate synthase operon
K09684	purine catabolism regulatory protein
K09694	lipooligosaccharide transport system permease protein
K09696	sodium transport system permease protein
K09698	nondiscriminating glutamyl-tRNA synthetase [EC:6.1.1.24]
K09729	hypothetical protein
K09759	nondiscriminating aspartyl-tRNA synthetase [EC:6.1.1.23]
K09767	hypothetical protein
K09773	hypothetical protein
K09925	hypothetical protein
K09931	hypothetical protein
K09936	hypothetical protein
K09962	hypothetical protein
K09963	hypothetical protein
K10005	glutamate transport system substrate-binding protein
K10006	glutamate transport system permease protein
K10007	glutamate transport system permease protein
K10008	glutamate transport system ATP-binding protein [EC:3.6.3.-]
K10041	putative glutamine transport system ATP-binding protein [EC:3.6.3.-]
K10117	multiple sugar transport system substrate-binding protein
K10118	multiple sugar transport system permease protein
K10119	multiple sugar transport system permease protein
K10121	putative sugar transport system permease protein
K10122	putative sugar transport system permease protein
K10189	lactose/L-arabinose transport system permease protein
K10190	lactose/L-arabinose transport system permease protein
K10200	N-acetylglucosamine transport system substrate-binding protein
K10243	cellobiose transport system ATP-binding protein
K10439	ribose transport system substrate-binding protein
K10441	ribose transport system ATP-binding protein [EC:3.6.3.17]
K10546	putative multiple sugar transport system substrate-binding protein
K10561	rhamnose transport system permease protein
K10670	glycine reductase [EC:1.21.4.2]
K10671	sarcosine reductase [EC:1.21.4.3]
K10672	betaine reductase [EC:1.21.4.4]
K10682	two-component system, OmpR family, response regulator SaeR
K10708	fructoselysine 6-phosphate deglycase [EC:3.5.-.-]
K10709	protein FrlC
K10710	fructoselysine 6-kinase [EC:2.7.1.-]
K10711	GntR family transcriptional regulator, frlABCD operon transcriptional regulator
K10793	D-proline reductase (dithiol) PrdA [EC:1.21.4.1]
K10795	D-proline reductase (dithiol) PrdD [EC:1.21.4.1]
K10796	D-proline reductase (dithiol) PrdE [EC:1.21.4.1]
K10805	acyl-CoA thioesterase II [EC:3.1.2.-]
K10823	oligopeptide transport system ATP-binding protein
K10844	DNA excision repair protein ERCC-2 [EC:3.6.4.12]
K10917	PadR family transcriptional regulator, regulatory protein AphA
K11003	hemolysin D
K11041	exfoliative toxin A/B
K11063	toxin A/B
K11068	hemolysin III
K11261	formylmethanofuran dehydrogenase subunit E [EC:1.2.99.5]
K11263	acetyl-/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein [EC:6.3.4.14]
K11358	aspartate aminotransferase [EC:2.6.1.1]
K11384	two-component system, NtrC family, response regulator AlgB
K11521	two-component system, OmpR family, manganese sensing response regulator
K11533	fatty acid synthase, bacteria type [EC:2.3.1.-]
K11616	malate:Na+ symporter
K11617	two-component system, NarL family, sensor histidine kinase LiaS [EC:2.7.13.3]
K11618	two-component system, NarL family, response regulator LiaR

KEGG orthology	Function
K11622	lia operon protein LiaF
K11631	bacitracin transport system ATP-binding protein
K11636	putative ABC transport system permease protein
K11686	chromosome-anchoring protein RacA
K11688	C4-dicarboxylate-binding protein DctP
K11689	C4-dicarboxylate transporter, DctQ subunit
K11690	C4-dicarboxylate transporter, DctM subunit
K11692	two-component system, CitB family, response regulator DctR
K11923	MerR family transcriptional regulator, copper efflux regulator
K12112	evolved beta-galactosidase subunit beta
K12143	hydrogenase-4 component H
K12283	MSHA biogenesis protein MshM
K12296	competence protein ComX
K12297	23S rRNA (guanine2445-N2)-methyltransferase [EC:2.1.1.173]
K12510	tight adherence protein B
K12554	alanine adding enzyme [EC:2.3.2.-]
K12555	penicillin-binding protein 2A [EC:2.4.1.129 2.3.2.-]
K12556	penicillin-binding protein 2X [EC:2.3.2.-]
K12992	rhamnosyltransferase [EC:2.4.1.-]
K13051	beta-aspartyl-peptidase (threonine type) [EC:3.4.19.5]
K13059	N-acetylhexosamine 1-kinase [EC:2.7.1.162]
K13252	putrescine carbamoyltransferase [EC:2.1.3.6]
K13275	major intracellular serine protease [EC:3.4.21.-]
K13288	oligoribonuclease [EC:3.1.-.-]
K13419	serine/threonine-protein kinase PknK [EC:2.7.11.1]
K13527	proteasome-associated ATPase
K13530	AraC family transcriptional regulator, regulatory protein of adaptative response / methylphosphotriester-DNA alkyltransferase methyltransferase [EC:2.1.1.-]
K13541	cobalt-precorrin 5A hydrolase / precorrin-3B C17-methyltransferase [EC:3.7.1.12 2.1.1.131]
K13570	prokaryotic ubiquitin-like protein Pup
K13571	proteasome accessory factor A [EC:6.3.2.-]
K13631	AraC family transcriptional regulator, transcriptional activator of the superoxide response regulon
K13639	MerR family transcriptional regulator, redox-sensitive transcriptional activator SoxR
K13640	MerR family transcriptional regulator, heat shock protein HspR
K13641	IclR family transcriptional regulator, acetate operon repressor
K13771	Rrf2 family transcriptional regulator, nitric oxide-sensitive transcriptional repressor
K13786	cob(II)yrinic acid a,c-diamide reductase [EC:1.16.8.1]
K13787	geranylgeranyl diphosphate synthase, type I [EC:2.5.1.1 2.5.1.10 2.5.1.29]
K13788	phosphate acetyltransferase [EC:2.3.1.8]
K13818	molybdopterin-guanine dinucleotide biosynthesis protein
K13829	shikimate kinase / 3-dehydroquinate synthase [EC:2.7.1.71 4.2.3.4]
K13889	glutathione transport system substrate-binding protein
K13891	glutathione transport system permease protein
K13920	propanediol dehydratase small subunit [EC:4.2.1.28]
K13921	1-propanol dehydrogenase
K13923	phosphotransacylase
K13927	holo-ACP synthase / triphosphoribosyl-dephospho-CoA synthase [EC:2.7.7.61 2.7.8.25]
K13940	dihydroneopterin aldolase / 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase [EC:4.1.2.25 2.7.6.3]
K13955	zinc-binding alcohol dehydrogenase/oxidoreductase
K14082	[methyl-Co(III) methylamine-specific corrinoid protein]:coenzyme M methyltransferase [EC:2.1.1.247]
K14083	trimethylamine---corrinoid protein Co-methyltransferase [EC:2.1.1.250]
K14088	ech hydrogenase subunit C
K14089	ech hydrogenase subunit D
K14090	ech hydrogenase subunit E
K14153	hydroxymethylpyrimidine kinase / phosphomethylpyrimidine kinase / thiamine-phosphate diphosphorylase [EC:2.7.1.49 2.7.4.7 2.5.1.3]
K14205	phosphatidylglycerol lysyltransferase [EC:2.3.2.3]
K14475	inhibitor of cysteine peptidase
K14956	6 kDa early secretory antigenic target
K15066	vanillate/3-O-methylgallate O-demethylase
K15330	phosphoglycerate kinase / triosephosphate isomerase [EC:2.7.2.3 5.3.1.1]
K15545	transcriptional regulator of PTS gene
K15598	putative hydroxymethylpyrimidine transport system substrate-binding protein
K15599	putative hydroxymethylpyrimidine transport system permease protein
K15634	probable phosphoglycerate mutase [EC:5.4.2.1]
K15653	nonribosomal peptide synthetase MxcG
K15835	RpiR family transcriptional regulator, murPQ operon repressor
K15866	2-(1,2-epoxy-1,2-dihydrophenyl)acetyl-CoA isomerase [EC:5.3.3.18]
K15922	alpha-glucosidase [EC:3.2.1.20]

KEGG orthology	Function
K15973	MarR family transcriptional regulator, 2-MHQ and catechol-resistance regulon repressor
K15984	16S rRNA (guanine1516-N2)-methyltransferase [EC:2.1.1.242]
K16012	ATP-binding cassette, subfamily C, bacterial CydC
K16013	ATP-binding cassette, subfamily C, bacterial CydD
K16048	3-hydroxy-9,10-secoandrosta-1,3,5(10)-triene-9,17-dione monooxygenase subunit HsaB (Flavin:NADH reductase)
K16137	TetR/AcrR family transcriptional regulator, transcriptional repressor for nem operon
K16147	starch synthase (maltosyl-transferring) [EC:2.4.99.16]
K16148	starch synthase [EC:2.4.1.21]
K16169	xanthine permease
K16179	dimethylamine corrinoid protein
K16202	dipeptide transport system ATP-binding protein

NA indicates not assigned

Appendix 7. KOs having the lowest abundance in the JP cohort among the 12 countries

KEGG orthology	Function
K00021	hydroxymethylglutaryl-CoA reductase (NADPH) [EC:1.1.1.34]
K00060	threonine 3-dehydrogenase [EC:1.1.1.103]
K00091	dihydroflavonol-4-reductase [EC:1.1.1.219]
K00113	glycerol-3-phosphate dehydrogenase subunit C [EC:1.1.5.3]
K00134	glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]
K00179	indolepyruvate ferredoxin oxidoreductase, alpha subunit [EC:1.2.7.8]
K00186	2-oxoisovalerate ferredoxin oxidoreductase, alpha subunit [EC:1.2.7.7]
K00187	2-oxoisovalerate ferredoxin oxidoreductase, beta subunit [EC:1.2.7.7]
K00201	formylmethanofuran dehydrogenase subunit B [EC:1.2.99.5]
K00203	formylmethanofuran dehydrogenase subunit D [EC:1.2.99.5]
K00273	D-amino-acid oxidase [EC:1.4.3.3]
K00311	electron-transferring-flavoprotein dehydrogenase [EC:1.5.5.1]
K00319	methylenetetrahydromethanopterin dehydrogenase [EC:1.5.99.9]
K00320	coenzyme F420-dependent N5,N10-methenyltetrahydromethanopterin reductase [EC:1.5.99.11]
K00350	Na+-transporting NADH:ubiquinone oxidoreductase subunit E [EC:1.6.5.-]
K00390	phosphoadenosine phosphosulfate reductase [EC:1.8.4.8]
K00399	methyl-coenzyme M reductase alpha subunit [EC:2.8.4.1]
K00400	methyl coenzyme M reductase system, component A2
K00401	methyl-coenzyme M reductase beta subunit [EC:2.8.4.1]
K00440	coenzyme F420 hydrogenase alpha subunit [EC:1.12.98.1]
K00442	coenzyme F420 hydrogenase delta subunit
K00443	coenzyme F420 hydrogenase gamma subunit [EC:1.12.98.1]
K00555	tRNA (guanine26-N2/guanine27-N2)-dimethyltransferase [EC:2.1.1.215 2.1.1.216]
K00558	DNA (cytosine-5-)-methyltransferase [EC:2.1.1.37]
K00571	site-specific DNA-methyltransferase (adenine-specific) [EC:2.1.1.72]
K00575	chemotaxis protein methyltransferase CheR [EC:2.1.1.80]
K00577	tetrahydromethanopterin S-methyltransferase subunit A [EC:2.1.1.86]
K00578	tetrahydromethanopterin S-methyltransferase subunit B [EC:2.1.1.86]
K00579	tetrahydromethanopterin S-methyltransferase subunit C [EC:2.1.1.86]
K00580	tetrahydromethanopterin S-methyltransferase subunit D [EC:2.1.1.86]
K00581	tetrahydromethanopterin S-methyltransferase subunit E [EC:2.1.1.86]
K00583	tetrahydromethanopterin S-methyltransferase subunit G [EC:2.1.1.86]
K00586	diphthine synthase [EC:2.1.1.98]
K00590	site-specific DNA-methyltransferase (cytosine-N4-specific) [EC:2.1.1.113]
K00640	serine O-acetyltransferase [EC:2.3.1.30]
K00641	homoserine O-acetyltransferase [EC:2.3.1.31]
K00672	formylmethanofuran--tetrahydromethanopterin N-formyltransferase [EC:2.3.1.101]
K00683	glutaminyl-peptide cyclotransferase [EC:2.3.2.5]
K00703	starch synthase [EC:2.4.1.21]
K00721	dolichol-phosphate mannosyltransferase [EC:2.4.1.83]
K00737	beta-1,4-mannosyl-glycoprotein beta-1,4-N-acetylglucosaminyltransferase [EC:2.4.1.144]
K00769	xanthine phosphoribosyltransferase [EC:2.4.2.22]
K00801	farnesyl-diphosphate farnesyltransferase [EC:2.5.1.21]
K00809	deoxyhypusine synthase [EC:2.5.1.46]
K00876	uridine kinase [EC:2.7.1.48]
K00908	Ca2+/calmodulin-dependent protein kinase [EC:2.7.11.17]
K00925	acetate kinase [EC:2.7.2.1]
K00929	butyrate kinase [EC:2.7.2.7]
K00962	polyribonucleotide nucleotidyltransferase [EC:2.7.7.8]
K00972	UDP-N-acetylglucosamine pyrophosphorylase [EC:2.7.7.23]
K00981	phosphatidate cytidylyltransferase [EC:2.7.7.41]
K01001	UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase [EC:2.7.8.15]
K01006	pyruvate,orthophosphate dikinase [EC:2.7.9.1]
K01062	1-alkyl-2-acetylglycerophosphocholine esterase [EC:3.1.1.47]
K01067	acetyl-CoA hydrolase [EC:3.1.2.1]
K01068	palmitoyl-CoA hydrolase [EC:3.1.2.2]
K01112	NA
K01126	glycerophosphoryl diester phosphodiesterase [EC:3.1.4.46]
K01139	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase [EC:3.1.7.2]
K01147	exoribonuclease II [EC:3.1.13.1]
K01150	deoxyribonuclease I [EC:3.1.21.1]
K01153	type I restriction enzyme, R subunit [EC:3.1.21.3]
K01156	type III restriction enzyme [EC:3.1.21.5]
K01157	NA
K01167	ribonuclease T1 [EC:3.1.27.3]

KEGG orthology	Function
K01170	tRNA-intron endonuclease, archaea type [EC:3.1.27.9]
K01174	micrococcal nuclease [EC:3.1.31.1]
K01179	endoglucanase [EC:3.2.1.4]
K01225	cellulose 1,4-beta-cellobiosidase [EC:3.2.1.91]
K01387	microbial collagenase [EC:3.4.24.3]
K01417	NA
K01553	myosin ATPase [EC:3.6.4.1]
K01572	oxaloacetate decarboxylase, beta subunit [EC:4.1.1.3]
K01610	phosphoenolpyruvate carboxykinase (ATP) [EC:4.1.1.49]
K01622	fructose 1,6-bisphosphate aldolase/phosphatase [EC:4.1.2.13 3.1.3.11]
K01667	tryptophanase [EC:4.1.99.1]
K01738	cysteine synthase A [EC:2.5.1.47]
K01791	UDP-N-acetylglucosamine 2-epimerase [EC:5.1.3.14]
K01841	phosphoenolpyruvate phosphomutase [EC:5.4.2.9]
K01844	beta-lysine 5,6-aminomutase [EC:5.4.3.3]
K01866	tyrosyl-tRNA synthetase [EC:6.1.1.1]
K01869	leucyl-tRNA synthetase [EC:6.1.1.4]
K01880	glycyl-tRNA synthetase [EC:6.1.1.14]
K01886	glutaminyl-tRNA synthetase [EC:6.1.1.18]
K01894	glutamyl-Q tRNA(Asp) synthetase [EC:6.1.1.-]
K01895	acetyl-CoA synthetase [EC:6.2.1.1]
K01916	NAD+ synthase [EC:6.3.1.5]
K02042	phosphonate transport system permease protein
K02044	phosphonate transport system substrate-binding protein
K02102	arabinose operon protein AraM
K02201	pantetheine-phosphate adenylyltransferase [EC:2.7.7.3]
K02288	phycocyanobilin lyase alpha subunit [EC:4.-.-.-]
K02314	replicative DNA helicase [EC:3.6.4.12]
K02319	DNA polymerase I [EC:2.7.7.7]
K02322	DNA polymerase II large subunit [EC:2.7.7.7]
K02323	DNA polymerase II small subunit [EC:2.7.7.7]
K02335	DNA polymerase I [EC:2.7.7.7]
K02356	elongation factor P
K02358	elongation factor Tu
K02377	GDP-L-fucose synthase [EC:1.1.1.271]
K02386	flagella basal body P-ring formation protein FlgA
K02387	flagellar basal-body rod protein FlgB
K02388	flagellar basal-body rod protein FlgC
K02389	flagellar basal-body rod modification protein FlgD
K02390	flagellar hook protein FlgE
K02391	flagellar basal-body rod protein FlgF
K02400	flagellar biosynthesis protein FlhA
K02401	flagellar biosynthetic protein FlhB
K02404	flagellar biosynthesis protein FlhF
K02409	flagellar M-ring protein FliF
K02410	flagellar motor switch protein FliG
K02411	flagellar assembly protein FliH
K02412	flagellum-specific ATP synthase [EC:3.6.3.14]
K02414	flagellar hook-length control protein FliK
K02418	flagellar protein FliO/FliZ
K02421	flagellar biosynthetic protein FliR
K02427	23S rRNA (uridine2552-2'-O)-methyltransferase [EC:2.1.1.166]
K02441	GlpG protein
K02445	MFS transporter, OPA family, glycerol-3-phosphate transporter
K02453	general secretion pathway protein D
K02455	general secretion pathway protein F
K02482	two-component system, NtrC family, sensor kinase [EC:2.7.13.3]
K02496	uroporphyrin-III C-methyltransferase [EC:2.1.1.107]
K02549	O-succinylbenzoate synthase [EC:4.2.1.113]
K02556	chemotaxis protein MotA
K02557	chemotaxis protein MotB
K02650	type IV pilus assembly protein PilA
K02655	type IV pilus assembly protein PilE
K02664	type IV pilus assembly protein PilO
K02674	type IV pilus assembly protein PilY1
K02683	DNA primase [EC:2.7.7.-]
K02685	DNA primase large subunit [EC:2.7.7.-]
K02844	UDP-glucose:(heptosyl)LPS alpha-1,3-glucosyltransferase [EC:2.4.1.-]
K02852	UDP-N-acetyl-D-mannosaminuronic acid transferase [EC:2.4.1.-]
K02866	large subunit ribosomal protein L10e

KEGG orthology	Function
K02869	large subunit ribosomal protein L12
K02875	large subunit ribosomal protein L14e
K02877	large subunit ribosomal protein L15e
K02883	large subunit ribosomal protein L18e
K02885	large subunit ribosomal protein L19e
K02889	large subunit ribosomal protein L21e
K02896	large subunit ribosomal protein L24e
K02910	large subunit ribosomal protein L31e
K02912	large subunit ribosomal protein L32e
K02921	large subunit ribosomal protein L37Ae
K02924	large subunit ribosomal protein L39e
K02929	large subunit ribosomal protein L44e
K02930	large subunit ribosomal protein L4e
K02944	large subunit ribosomal protein LX
K02966	small subunit ribosomal protein S19e
K02974	small subunit ribosomal protein S24e
K02978	small subunit ribosomal protein S27e
K02979	small subunit ribosomal protein S28e
K02984	small subunit ribosomal protein S3Ae
K02986	small subunit ribosomal protein S4
K02987	small subunit ribosomal protein S4e
K02991	small subunit ribosomal protein S6e
K02995	small subunit ribosomal protein S8e
K03041	DNA-directed RNA polymerase subunit A' [EC:2.7.7.6]
K03042	DNA-directed RNA polymerase subunit A" [EC:2.7.7.6]
K03044	DNA-directed RNA polymerase subunit B' [EC:2.7.7.6]
K03045	DNA-directed RNA polymerase subunit B" [EC:2.7.7.6]
K03049	DNA-directed RNA polymerase subunit E' [EC:2.7.7.6]
K03050	DNA-directed RNA polymerase subunit E" [EC:2.7.7.6]
K03051	DNA-directed RNA polymerase subunit F [EC:2.7.7.6]
K03053	DNA-directed RNA polymerase subunit H [EC:2.7.7.6]
K03056	DNA-directed RNA polymerase subunit L [EC:2.7.7.6]
K03057	transcription elongation factor
K03087	RNA polymerase nonessential primary-like sigma factor
K03089	RNA polymerase sigma-32 factor
K03105	signal recognition particle subunit SRP19
K03120	transcription initiation factor TFIID TATA-box-binding protein
K03124	transcription initiation factor TFIIB
K03136	transcription initiation factor TFIIE subunit alpha
K03166	DNA topoisomerase VI subunit A [EC:5.99.1.3]
K03167	DNA topoisomerase VI subunit B [EC:5.99.1.3]
K03183	ubiquinone/menaquinone biosynthesis methyltransferase [EC:2.1.1.163 2.1.1.201]
K03203	type IV secretion system protein VirB8
K03231	elongation factor 1-alpha
K03232	elongation factor 1-beta
K03236	translation initiation factor 1A
K03237	translation initiation factor 2 subunit 1
K03238	translation initiation factor 2 subunit 2
K03242	translation initiation factor 2 subunit 3
K03243	translation initiation factor 5B
K03263	translation initiation factor 5A
K03264	translation initiation factor 6
K03271	D-sedoheptulose 7-phosphate isomerase [EC:5.3.1.28]
K03273	D-glycero-D-manno-heptose 1,7-bisphosphate phosphatase [EC:3.1.3.82 3.1.3.83]
K03275	UDP-glucose:(glucosyl)LPS alpha-1,3-glucosyltransferase [EC:2.4.1.-]
K03324	phosphate:Na+ symporter
K03329	hypothetical protein
K03403	magnesium chelatase subunit H [EC:6.6.1.1]
K03405	magnesium chelatase subunit I [EC:6.6.1.1]
K03406	methyl-accepting chemotaxis protein
K03408	purine-binding chemotaxis protein CheW
K03409	chemotaxis protein CheX
K03412	two-component system, chemotaxis family, response regulator CheB [EC:3.1.1.61]
K03420	proteasome regulatory subunit
K03421	methyl-coenzyme M reductase subunit C
K03422	methyl-coenzyme M reductase subunit D
K03427	type I restriction enzyme M protein [EC:2.1.1.72]
K03432	proteasome alpha subunit [EC:3.4.25.1]
K03465	thymidylate synthase (FAD) [EC:2.1.1.148]
K03495	tRNA uridine 5-carboxymethylaminomethyl modification enzyme

KEGG orthology	Function
K03503	DNA polymerase V [EC:3.4.21.-]
K03521	electron transfer flavoprotein beta subunit
K03537	ribonuclease P/MRP protein subunit POP5 [EC:3.1.26.5]
K03538	ribonuclease P protein subunit POP4 [EC:3.1.26.5]
K03539	ribonuclease P/MRP protein subunit RPP1 [EC:3.1.26.5]
K03540	ribonuclease P protein subunit RPR2 [EC:3.1.26.5]
K03546	exonuclease SbcC
K03548	putative permease
K03553	recombination protein RecA
K03555	DNA mismatch repair protein MutS
K03560	biopolymer transport protein TolR
K03562	biopolymer transport protein TolQ
K03567	glycine cleavage system transcriptional repressor
K03570	rod shape-determining protein MreC
K03572	DNA mismatch repair protein MutL
K03583	exodeoxyribonuclease V gamma subunit [EC:3.1.11.5]
K03584	DNA repair protein RecO (recombination protein O)
K03595	GTP-binding protein Era
K03596	GTP-binding protein LepA
K03612	electron transport complex protein RnfG
K03613	electron transport complex protein RnfE
K03614	electron transport complex protein RnfD
K03622	archaea-specific DNA-binding protein
K03625	N utilization substance protein B
K03626	nascent polypeptide-associated complex subunit alpha
K03643	LPS-assembly lipoprotein
K03650	tRNA modification GTPase [EC:3.6.-.-]
K03673	thiol:disulfide interchange protein DsbA
K03679	exosome complex component RRP4
K03726	helicase [EC:3.6.4.-]
K03737	putative pyruvate-flavodoxin oxidoreductase [EC:1.2.7.-]
K03748	SanA protein
K03756	putrescine:ornithine antiporter
K03772	FKBP-type peptidyl-prolyl cis-trans isomerase FkpA [EC:5.2.1.8]
K03796	Bax protein
K03810	virulence factor
K03820	apolipoprotein N-acyltransferase [EC:2.3.1.-]
K03893	arsenical pump membrane protein
K03924	MoxR-like ATPase [EC:3.6.3.-]
K03932	polyhydroxybutyrate depolymerase
K03969	phage shock protein A
K03977	GTP-binding protein
K04067	primosomal replication protein N''
K04071	6-pyruvoyltetrahydropterin 2'-reductase [EC:1.1.1.220]
K04076	Lon-like ATP-dependent protease [EC:3.4.21.-]
K04079	molecular chaperone HtpG
K04084	thiol:disulfide interchange protein DsbD [EC:1.8.1.8]
K04095	cell filamentation protein
K04109	4-hydroxybenzoyl-CoA reductase subunit beta [EC:1.3.7.9]
K04112	benzoyl-CoA reductase subunit [EC:1.3.7.8]
K04484	DNA repair protein RadB
K04754	lipoprotein
K04760	transcription elongation factor GreB
K04795	fibrillarin-like pre-rRNA processing protein
K04797	prefoldin alpha subunit
K04798	prefoldin beta subunit
K04801	replication factor C small subunit
K04802	proliferating cell nuclear antigen
K05365	penicillin-binding protein 1B [EC:2.4.1.129 3.4.-.-]
K05367	penicillin-binding protein 1C [EC:2.4.1.-]
K05384	bilin biosynthesis protein
K05515	penicillin-binding protein 2
K05566	multicomponent Na+:H+ antiporter subunit B
K05569	multicomponent Na+:H+ antiporter subunit E
K05716	cyclic 2,3-diphosphoglycerate synthetase [EC:4.6.1.-]
K05802	potassium efflux system protein KefA
K05837	rod shape determining protein RodA
K05844	ribosomal protein S6 modification protein
K05851	adenylate cyclase, class 1 [EC:4.6.1.1]
K05929	phosphoethanolamine N-methyltransferase [EC:2.1.1.103]

KEGG orthology	Function
K05939	acyl-[acyl-carrier-protein]-phospholipid O-acyltransferase / long-chain-fatty-acid--[acyl-carrier-protein] ligase [EC:2.3.1.40 6.2.1.20]
K06001	tryptophan synthase beta chain [EC:4.2.1.20]
K06027	vesicle-fusing ATPase [EC:3.6.4.6]
K06034	sulfopyruvate decarboxylase subunit alpha [EC:4.1.1.79]
K06190	intracellular septation protein
K06192	paraquat-inducible protein B
K06203	CysZ protein
K06223	DNA adenine methylase [EC:2.1.1.72]
K06296	spore germination protein KB
K06313	spore germination protein
K06343	spore coat protein Y
K06370	morphogenetic protein associated with SpoVID
K06384	stage II sporulation protein M
K06401	stage IV sporulation protein FA
K06402	stage IV sporulation protein FB [EC:3.4.24.-]
K06601	flagellar protein FlbT
K06862	energy-converting hydrogenase B subunit Q
K06863	5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase [EC:6.3.4.-]
K06872	uncharacterized protein
K06874	zinc finger protein
K06875	programmed cell death protein 5
K06877	DEAD/DEAH box helicase domain-containing protein
K06881	phosphoesterase RecJ domain-containing protein
K06883	NA
K06909	phage terminase large subunit
K06914	NA
K06915	NA
K06927	NA
K06932	NA
K06943	nucleolar GTP-binding protein
K06961	ribosomal RNA assembly protein
K06965	protein pelota
K06982	pantoate kinase [EC:2.7.1.169]
K06984	NA
K06985	aspartyl protease family protein
K07022	NA
K07041	NA
K07055	tRNA wybutosine-synthesizing protein 2 [EC:2.1.1.-]
K07058	membrane protein
K07060	UPF0271 protein
K07072	NA
K07082	UPF0755 protein
K07102	NA
K07103	NA
K07108	NA
K07121	NA
K07123	NA
K07135	NA
K07144	NA
K07158	NA
K07159	NA
K07164	NA
K07174	Mn2+-dependent serine/threonine protein kinase [EC:2.7.1.-]
K07178	RIO kinase 1 [EC:2.7.11.1]
K07244	mgtE-like transporter
K07273	lysozyme
K07316	adenine-specific DNA-methyltransferase [EC:2.1.1.72]
K07318	adenine-specific DNA-methyltransferase [EC:2.1.1.72]
K07332	archaeal flagellar protein FlaI
K07333	archaeal flagellar protein FlaJ
K07388	hydrogenase expression/formation protein
K07444	putative N6-adenine-specific DNA methylase [EC:2.1.1.-]
K07459	putative ATP-dependent endonuclease of the OLD family
K07462	single-stranded-DNA-specific exonuclease [EC:3.1.-.-]
K07463	archaea-specific RecJ-like exonuclease
K07466	replication factor A1
K07487	transposase
K07507	putative Mg2+ transporter-C (MgtC) family protein
K07557	archaeosine tRNA-ribosyltransferase [EC:2.4.2.-]

KEGG orthology	Function
K07558	tRNA nucleotidyltransferase (CCA-adding enzyme) [EC:2.7.7.72]
K07562	nonsense-mediated mRNA decay protein 3
K07569	RNA-binding protein
K07572	putative nucleotide binding protein
K07573	exosome complex component CSL4
K07575	PUA domain protein
K07581	hypothetical protein
K07583	tRNA pseudouridine synthase 10 [EC:5.4.99.-]
K07585	hypothetical protein
K07645	two-component system, OmpR family, sensor histidine kinase QseC [EC:2.7.13.3]
K07666	two-component system, OmpR family, response regulator QseB
K07687	two-component system, NarL family, captular synthesis response regulator RcsB
K07732	riboflavin kinase, archaea type [EC:2.7.1.161]
K07769	two-component system, OmpR family, sensor histidine kinase NblS [EC:2.7.13.3]
K07783	MFS transporter, OPA family, sugar phosphate sensor protein UhpC
K07790	putative membrane protein PagO
K08096	GTP cyclohydrolase IIa [EC:3.5.4.29]
K08097	phosphosulfolactate synthase [EC:4.4.1.19]
K08137	MFS transporter, SP family, galactose:H+ symporter
K08259	lysostaphin [EC:3.4.24.75]
K08264	heterodisulfide reductase subunit D [EC:1.8.98.1]
K08309	soluble lytic murein transglycosylase [EC:3.2.1.-]
K08310	dATP pyrophosphohydrolase [EC:3.6.1.-]
K08311	putative (di)nucleoside polyphosphate hydrolase [EC:3.6.1.-]
K08484	phosphotransferase system, enzyme I, PtsP [EC:2.7.3.9]
K08587	clostripain [EC:3.4.22.8]
K08589	gingipain R [EC:3.4.22.37]
K08590	carbon-nitrogen hydrolase family protein
K08641	D-alanyl-D-alanine dipeptidase [EC:3.4.13.22]
K08722	5'-nucleotidase [EC:3.1.3.5]
K08971	putative membrane protein
K08974	putative membrane protein
K08978	putative membrane protein
K08979	putative membrane protein
K09003	hypothetical protein
K09119	hypothetical protein
K09139	hypothetical protein
K09140	pre-rRNA-processing protein TSR3
K09144	hypothetical protein
K09152	hypothetical protein
K09154	hypothetical protein
K09482	glutamyl-tRNA(Gln) amidotransferase subunit D [EC:6.3.5.7]
K09713	hypothetical protein
K09720	hypothetical protein
K09721	hypothetical protein
K09722	4-phosphopantoate---beta-alanine ligase [EC:6.3.2.36]
K09723	hypothetical protein
K09724	hypothetical protein
K09727	hypothetical protein
K09728	hypothetical protein
K09730	hypothetical protein
K09733	hypothetical protein
K09735	hypothetical protein
K09738	hypothetical protein
K09739	hypothetical protein
K09766	hypothetical protein
K09807	hypothetical protein
K09826	Fur family transcriptional regulator, iron response regulator
K09859	hypothetical protein
K09882	cobaltochelatase CobS [EC:6.6.1.2]
K09914	putative lipoprotein
K09942	hypothetical protein
K09968	hypothetical protein
K09973	hypothetical protein
K09987	hypothetical protein
K09989	hypothetical protein
K10212	glycosyl-4,4'-diaponeurosporenoate acyltransferase [EC:2.3.1.-]
K10219	2-hydroxy-4-carboxymuconate semialdehyde hemiacetal dehydrogenase [EC:1.1.1.312]
K10679	nitroreductase / dihydropteridine reductase [EC:1.-.-.- 1.5.1.34]
K10697	two-component system, OmpR family, response regulator RpaA

KEGG orthology	Function
K10702	2-hydroxy-6-oxohepta-2,4-dienoate hydroxylase [EC:3.7.1.-]
K10725	archaeal cell division control protein 6
K10747	DNA ligase 1 [EC:6.5.1.1]
K10806	acyl-CoA thioesterase YciA [EC:3.1.2.-]
K10857	exodeoxyribonuclease X [EC:3.1.11.-]
K10960	geranylgeranyl reductase [EC:1.3.1.83]
K11004	ATP-binding cassette, subfamily B, bacterial HlyB/CyaB
K11005	hemolysin A
K11021	insecticidal toxin complex protein TccC
K11070	spermidine/putrescine transport system permease protein
K11130	H/ACA ribonucleoprotein complex subunit 3
K11131	H/ACA ribonucleoprotein complex subunit 4 [EC:5.4.99.-]
K11212	LPPG:FO 2-phospho-L-lactate transferase [EC:2.7.8.28]
K11260	formylmethanofuran dehydrogenase subunit G [EC:1.2.99.5]
K11434	protein arginine N-methyltransferase 1 [EC:2.1.1.-]
K11600	exosome complex component RRP41
K11693	peptidoglycan pentaglycine glycine transferase (the first glycine) [EC:2.3.2.16]
K11749	regulator of sigma E protease [EC:3.4.24.-]
K11780	FO synthase subunit 1 [EC:2.5.1.77]
K11915	serine/threonine protein phosphatase Stp1 [EC:3.1.3.16]
K11941	glucans biosynthesis protein C [EC:2.1.-.-]
K12071	conjugal transfer pilus assembly protein TraD
K12152	phosphatase NudJ [EC:3.6.1.-]
K12164	ubiquitin-like modifier-activating enzyme 5
K12234	coenzyme F420-0:L-glutamate ligase / coenzyme F420-1:gamma-L-glutamate ligase [EC:6.3.2.31 6.3.2.34]
K12278	MSHA biogenesis protein MshG
K12287	MSHA biogenesis protein MshQ
K12294	two-component system, AgrA family, sensor histidine kinase ComD [EC:2.7.13.-]
K12516	putative surface-exposed virulence protein
K12543	outer membrane protein LapE
K12573	ribonuclease R [EC:3.1.-.-]
K12589	exosome complex component RRP42
K12682	tracheal colonization factor
K12686	outer membrane lipase/esterase
K12975	phosphoethanolamine transferase
K12988	alpha-1,3-rhamnosyltransferase [EC:2.4.1.-]
K13010	perosamine synthetase
K13039	sulfopyruvate decarboxylase subunit beta [EC:4.1.1.79]
K13243	c-di-GMP-specific phosphodiesterase [EC:3.1.4.52]
K13282	cyanophycinase [EC:3.4.15.6]
K13500	chondroitin synthase [EC:2.4.1.175 2.4.1.226]
K13522	bifunctional NMN adenylyltransferase/nudix hydrolase [EC:2.7.7.1 3.6.1.-]
K13583	GcrA cell cycle regulator
K13588	histidine phosphotransferase ChpT
K13730	internalin A
K13735	adhesin/invasin
K13789	geranylgeranyl diphosphate synthase, type II [EC:2.5.1.1 2.5.1.10 2.5.1.29]
K13812	bifunctional enzyme Fae/Hps [EC:4.3.-.- 4.1.2.43]
K13896	microcin C transport system ATP-binding protein
K13929	malonate decarboxylase alpha subunit [EC:2.3.1.187]
K13942	5,10-methenyltetrahydromethanopterin hydrogenase [EC:1.12.98.2]
K14058	tRNA 2-thiocytidine biosynthesis protein TtcA
K14092	energy-converting hydrogenase A subunit A
K14093	energy-converting hydrogenase A subunit B
K14094	energy-converting hydrogenase A subunit C
K14095	energy-converting hydrogenase A subunit D
K14096	energy-converting hydrogenase A subunit E
K14097	energy-converting hydrogenase A subunit F
K14098	energy-converting hydrogenase A subunit G
K14101	energy-converting hydrogenase A subunit J
K14102	energy-converting hydrogenase A subunit K
K14103	energy-converting hydrogenase A subunit L
K14104	energy-converting hydrogenase A subunit M
K14105	energy-converting hydrogenase A subunit N
K14109	energy-converting hydrogenase A subunit R
K14110	energy-converting hydrogenase B subunit A
K14111	energy-converting hydrogenase B subunit B
K14112	energy-converting hydrogenase B subunit C
K14113	energy-converting hydrogenase B subunit D

KEGG orthology	Function
K14115	energy-converting hydrogenase B subunit F
K14116	energy-converting hydrogenase B subunit G
K14117	energy-converting hydrogenase B subunit H
K14118	energy-converting hydrogenase B subunit I
K14119	energy-converting hydrogenase B subunit J
K14121	energy-converting hydrogenase B subunit L
K14122	energy-converting hydrogenase B subunit M
K14123	energy-converting hydrogenase B subunit N
K14124	energy-converting hydrogenase B subunit O
K14125	energy-converting hydrogenase B subunit P
K14126	F420-non-reducing hydrogenase subunit A [EC:1.12.99.-]
K14196	immunoglobulin G-binding protein A
K14415	tRNA-splicing ligase RtcB [EC:6.5.1.3]
K14441	ribosomal protein S12 methylthiotransferase [EC:2.-.-.-]
K14561	U3 small nucleolar ribonucleoprotein protein IMP4
K14564	nucleolar protein 56
K14574	ribosome maturation protein SDO1
K14598	chlorobactene lauroyltransferase
K14623	DNA-damage-inducible protein D
K14653	2-amino-5-formylamino-6-ribosylaminopyrimidin-4(3H)-one 5'-monophosphate deformylase [EC:3.5.1.102]
K14680	RNA ligase [EC:6.5.1.3]
K14682	amino-acid N-acetyltransferase [EC:2.3.1.1]
K15125	filamentous hemagglutinin
K15353	E3 ubiquitin-protein ligase SspH2 [EC:6.3.2.19]
K15359	6-hydroxy-3-succinoylpyridine hydroxylase [EC:3.7.1.-]
K15429	tRNA (guanine37-N1)-methyltransferase [EC:2.1.1.228]
K15525	N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase [EC:3.5.1.103]
K15527	cysteate synthase [EC:2.5.1.76]
K15633	2,3-bisphosphoglycerate-independent phosphoglycerate mutase [EC:5.4.2.1]
K15640	uncharacterized phosphatase
K15665	fengycin family lipopeptide synthetase B
K15770	putative arabinogalactan oligomer transport system substrate-binding protein
K15778	phosphomannomutase / phosphoglucomutase [EC:5.4.2.8 5.4.2.2]
K15888	tritrans,polycis-undecaprenyl-diphosphate synthase [geranylgeranyl-diphosphate specific] [EC:2.5.1.89]
K15904	bifunctional tRNA threonylcarbamoyladenosine biosynthesis protein [EC:2.7.11.1]
K16091	Fe(3+) dicitrate transport protein
K16150	glycogen(starch) synthase [EC:2.4.1.11]
K16183	methylamine methyltransferase corrinoid activation protein

NA indicates not assigned