Rhinology Online

ORIGINAL CONTRIBUTION

OPEN ACCESS

Assessment of microbial DNA enrichment techniques from sino-nasal swab samples for metagenomics*

Brett Wagner Mackenzie[1], David W. Waite[2,3], Kristi Biswas[1], Richard G. Douglas[1], Michael W. Taylor[3,4]

[1]School of Medicine, Department of Surgery, The University of Auckland, Auckland, New Zealand

Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Australia

[3]School of Biological Sciences, The University of Auckland, Auckland, New Zealand

[4]Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Auckland, New Zealand

Rhinology Online, Vol 1: 160 - 193, 2018

http://doi.org/10.4193/RHINOL/18.052

***Received for publication:**

August 14, 2018

Accepted: October 17, 2018

Abstract

Background: The role of the sinus microbiota, including bacteria, fungi and viruses, in health and disease remains unclear despite the application of molecular microbiological techniques to describe the microbiome. This is due, in part, to the overwhelming proportion of contaminating host DNA compared with recovered microbial DNA.

Methods: In this study, three techniques were assessed for microbial DNA enrichment: 1. A series of centrifugation steps, 2. Enrichment of microbial DNA using the NEBNext® Microbiome DNA Enrichment kit, and 3. Whole-genome amplification following the previous enrichment strategies. A no-treatment control and a whole-genome amplified control were also included. Swab samples from three adult patients undergoing functional endoscopic sinus surgery for chronic rhinosinusitis (CRS) were collected intraoperatively for this study. Paired-end shotgun metagenome sequencing was conducted using Illumina HiSeq and bacterial 16S rRNA gene amplicon sequencing using Illumina MiSeq to assess bacterial community composition.

Results: After quality filtering of metagenomic sequences, the centrifugation method returned the highest proportion of microbial reads (1.1±1.7%) compared to the no-treatment control (0.15±0.07%). However, this result was neither reproducible nor was centrifugation significantly different to the other methods. Despite low recovery of total microbial DNA from metagenomic sequencing, a *Propionibacterium acnes* genome (97% complete) was recovered, suggesting metagenomic sequencing techniques can still be successfully applied to investigate the microbial component of CRS.

Conclusions: Based on these results, we recommend omitting microbial DNA enrichment steps and sequencing fewer samples per metagenomic sequencing run to increase the depth of sequencing without altering *in situ* microbial community structure.

Key words: Human microbiome, chronic rhinosinusitis, bacterial genome, metagenomics, microbiota, *Propionibacterium acnes*

Introduction

Contemporary sequencing technologies, such as targeted amplicon sequencing and shotgun metagenomics approaches, have provided unprecedented insights into the structure and function of the human microbiome [1–3]. One aspect of human microbiome study that has particularly benefitted from the application of high-throughput sequencing techniques is that of chronic sinus disease. In chronic rhinosinusitis (CRS) long-term inflammation of the sino-nasal mucosal lining severely impacts patient quality of life and places a substantial burden on health care systems [4–6]. Molecular technologies have transformed the way we view the role of the microbiota in CRS pathogenesis [7–10]. It now appears that these communities are imbalanced (dysbiotic), and characterized by a significant increase in bacterial load accompanied by a decrease in overall bacterial diversity [9,11,12].

Shotgun metagenomic sequencing can provide comprehensive, strain-level identification and functional information about viral, fungal, bacterial and archaeal diversity within a sample. Metagenomic sequencing thus offers great potential to deepen our

understanding of CRS, yet its application to sino-nasal samples is not without challenges. The size of the human genome (~3 billion bp) compared to that of a typical bacterium (~4 million bp for *Escherichia coli)* means that the presence of relatively few human cells can lead to overwhelming proportions of contaminating human DNA in a sample [13,14]. In one study of the human skin microbiome, nares samples had the highest proportion of reads, on average, that mapped to the human genome (98.2%), when compared with 17 other skin sites [15]. Metagenomic studies of samples originating from the middle meatus, which acts as a reservoir for mucous drainage within the human sinuses, have not yet been conducted, and we therefore carried out a pilot study (unpublished data) to assess levels of recovered microbial versus human DNA. We found that less than 1% of quality-filtered sequencing reads were of microbial (viral, bacterial, archaeal, or fungal) origin. Based on this pilot study, we designed a methods study to test a variety of host DNA depletion and microbial DNA enrichment strategies.

Enrichment techniques can be applied to capture genomes of interest or remove contaminating DNA [16,17]. Probe-based methods that target a single organism have been adopted, but are not suitable for studying entire microbial communities [18]. Other methods such as differential cell lysis, filtration and centrifugation separate host and microbial cells based on physical properties, but may have varying results between samples depending on community composition and sample consistency [19–22]. A range of commercially produced kits are available for enriching microbial DNA from human-derived samples including MolYsis®, Pureprover®, LOOXSTER®, Molyzm Ultra-Deep Microbiome Prep, and the NEBNext® Microbiome DNA Enrichment Kit, but all are associated with increased processing costs when compared with lysis, filtration and centrifugation methods. The efficacy of enrichment associated with these methods is variable and likely sample-type specific [23–28].

Whole genome amplification refers to a process in which segments of entire genomes originating from any type of DNA, microbial or human, are amplified (unlike traditional PCR, in which primers target specific regions of DNA within genomes). Whole genome amplification (WGA) by multiple displacement amplification (MDA) involves binding of random hexamers to denatured DNA for the initial amplification followed by strand displacement with Phi29 polymerase. WGA MDA is very useful for low biomass samples; however, due to its non-targeted nature contaminating host DNA must be significantly reduced prior to WGA [20].

A number of limitations are associated with each of the aforementioned enrichment strategies, such as enzymatic treatments applied to preferentially lyse human cells may also lyse bacterial cells [23–25,29]. Very few methods studies incorporating enrichment techniques and metagenomics sequencing from human mixed microbial communities are available and the majority of method enrichment comparisons focus on the detection of specific pathogens [17,20,30,31]. Additionally, many studies do not include non-spiked samples from patients which would validate the efficacy of these techniques on microbial communities in the clinical setting [24,25]. To date, no studies exist comparing enrichment techniques for metagenomic sequencing from sino-nasal samples. Based on known limitations and bias of a variety of enrichment techniques, we chose two different techniques for removing human DNA: a series of centrifugation steps prior to nucleic acids extraction and the NEBNext® Microbiome DNA Enrichment Kit, each in conjunction with whole genome amplification, in an attempt to enrich the total amount of microbial DNA before metagenomic sequencing. Additionally, we amplified the bacterial 16S rRNA gene to investigate the effects of the chosen methods on the recovered bacterial community profiles and our ability to describe them accurately.

Materials and Methods

Three male, adult patients undergoing functional endoscopic sinus surgery for idiopathic CRS by a single surgeon (RD) were recruited from Auckland City Hospital, Auckland, New Zealand. Exclusion criteria included age <18 years, current smoker, symptoms of asthma, aspirin sensitivity, and antibiotic and prednisone usage within the four weeks prior to surgery (Additional file 1). Written consent from the patients and ethical approval (NTX/08/12/126) from the New Zealand Health and Disability Ethics Committee was obtained for this study. Sterile rayon-tipped swabs (Copan Diagnostics, Inc., Murrieta, CA, USA) were used to collect a total of 12 mucosal samples from the middle meatus of each patient (6 left, 6 right) at the time of induction of anaesthesia. Swabs were immediately placed in 1 mL RNAlater® solution and stored the same day at -20°C until DNA extraction. A diagram outlining this study is found in Figure 1.

Microbial DNA enrichment methods

No-treatment Control (N)

Samples were thawed on ice, and DNA was extracted from pairs of swabs from the same patient (1 left, 1 right) using the Qiagen® AllPrep DNA/RNA Mini Kit (Bio-Strategy LTD., Auckland, New Zealand) as previously described [9]. Elution Buffer EB (55 µL) was added to the spin column filter and incubated for 5 min before DNA was eluted by centrifuging for 1 min at 11,200 x g. The eluate was centrifuged through the spin column filter a second time to increase DNA concentration. Triplicate negative extractions of sterile water were performed to test the DNA extraction kit for contamination.

Yield (ng/µL) and purity (260/280 nm absorbance ratio) of extracted DNA were determined spectrophotometrically using the NanoDrop® ND-1000 (NanoDrop Technologies Inc., Wilmington, USA). DNA yield was also determined fluorometrically using the High Sensitivity (HS) kit on the Qubit® Fluorometer 1.0 (Invi-

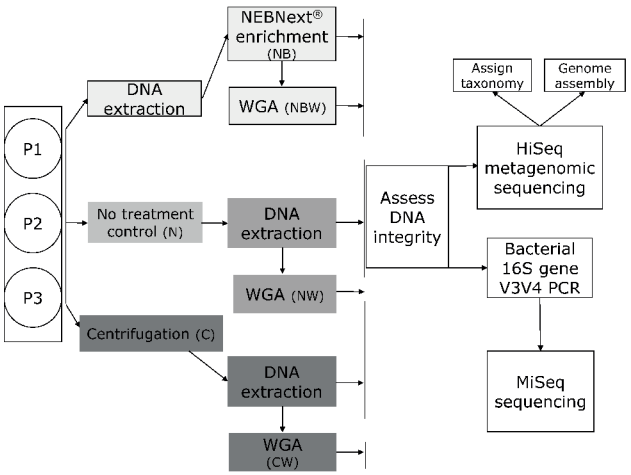

Figure 1. Schematic diagram showing the experimental design applied in this study. P1, P2, P3 refer to the individual CRS patients from whom samples were obtained.

trogen Co., Carlsbad, CA, USA). Integrity of genomic DNA was determined by visualizing 3 µL of extracted DNA on a 0.8% agarose gel (w/v) containing SYBR Safe DNA Gel Stain (Invitrogen Co., Carlsbad, USA) run in 0.5X TBE buffer at 90 V for 45 min.

Centrifugation Enrichment Technique (C)

Samples were thawed on ice, and pairs of swabs from the same patient (1 left, 1 right) were briefly vortexed to release bacterial cells from the swab matrix into the RNAlater® solution. Both swabs were removed from the microcentrifuge tube, and the RNAlater® solution was centrifuged for 3 min at 500 x g in order to pellet the large (heavy) human cells while leaving smaller, microbial cells in the supernatant. The supernatant was removed into a fresh microcentrifuge tube, which was then centrifuged for 7 min at 8000 x g to pellet microbial cells. The supernatant was discarded, and pelleted cells were resuspended in 600 µL of Buffer RLT Plus. DNA was extracted from the resuspended pellet as described above. DNA yield, quality and integrity of extracted DNA were determined as described above.

NEBNext® Microbiome DNA Enrichment Kit (NB)

The NEBNext® Microbiome DNA Enrichment Kit selectively removes human DNA from samples by binding double-stranded DNA containing 5-methyl CpG dinucleotides (which are common in vertebrate DNA) to a magnetic bead [17]. Briefly, samples were thawed on ice, and DNA was extracted from pairs of swabs from the same patient (1 left, 1 right) using the Qiagen® AllPrep DNA/RNA Mini Kit as described above. Total DNA concentration was calculated and 2 µg input DNA was used for enrichment of microbial DNA using the NEBNext® Microbiome DNA Enrichment Kit, following the manufacturer's instructions (New England BioLabs® Inc., Thermo Fisher Scientific, Auckland, New Zealand). All volumes were adjusted to allow for 2 µg of input DNA for each of the three samples.

After microbial and host DNA were selectively captured using the NEBNext® Microbiome DNA Enrichment Kit, Agencourt AMPure XP Bead Clean-up (Beckman Coulter Inc., Brea, CA, USA) was used to purify the enriched samples. Briefly, all sample volumes were split into 160 µL volumes, if necessary, and 1.8X volumes of AMPure beads were added to each sample. After several ethanol wash steps, DNA was eluted from the magnetic beads in 15-25 µL (depending on initial input volume) of TE Buffer, pH 7.5.

Whole genome amplification ('WGA')

Samples were first subjected to one of the two enrichment techniques, or originated from the no-enrichment control DNA extraction. The Qiagen® REPLI-g Mini Kit (Bio-Strategy LTD., Auckland, New Zealand) was used to amplify 5 µL of template DNA from each sample according to the manufacturer's instructions. For each reaction, a positive control of *E. coli* genomic DNA and a negative control of PCR-grade water was used. Quality, integrity and yield of amplified DNA were assessed as previously described.

Sequencing preparation

PCR amplification and Illumina sequencing

In order to compare the recovery of bacterial community composition profiles based on metagenomic sequencing to those based on the usual approach employed by CRS researchers (i.e. 16S rRNA gene-targeted sequencing), we amplified the V3-V4 hypervariable region of the bacterial 16S rRNA gene for each sample in this project. Amplifications were carried out as described previously [9], with minor adjustments (Additional file 2). The triplicate negative extractions from the DNA extraction kits were amplified and verified for lack of contamination.

Replicate PCR products from each sample were pooled and purified using Agencourt AMPure beads according to manufacturer instructions. Bacterial 16S rRNA gene amplicons were submitted to New Zealand Genomics Limited for library preparation using a dual-indexing approach with Nextera technology and sequencing (2 x 300 bp, paired-end) on the Illumina MiSeq. Metagenomic samples were submitted as is to New Zealand Genomics Limited for Thruplex DNA library preparation and sequencing (2 x 125 bp, paired-end) on one lane of the Illumina HiSeq.

Data analyses

Bacterial 16S rRNA gene sequence analysis

Bioinformatic processing of amplicon sequencing data involved a combination of USEARCH (version 7.0.1090, 64-bit built for Linux) and QIIME version 1.8 (Additional file 2) [32,33]. Samples were rarefied to 1,678 sequences, and rarefied tables were used for all downstream analyses. Alpha diversity measures Chao1, Shannon, Simpson and observed species (OTUs), and a Bray-Curtis dissimilarity matrix were assessed and generated using QIIME.

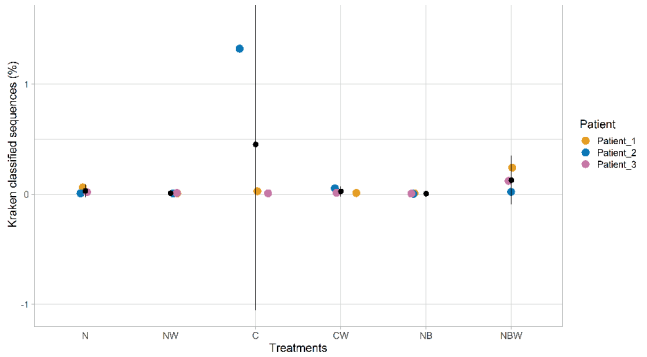

Figure 2. Proportion of metagenomic sequences that were classified using the Kraken database after quality filtering and removal of human-aligned sequences. Mean values are indicated by black dots and standard deviations are defined by the extent of the black lines.

PRIMER6 version 6.1.13 was used to identify microbial community similarities between groups of samples based on enrichment method for each type of sequencing data (SIMPER), to assess variation within the data due to inter-individual variation and enrichment method (PERMANOVA), and to analyse patterns in microbial community composition between methods for both amplicon and metagenomics datasets (ANOSIM) [34]. Multivariate dispersion (MVDISP) was used to quantify relative multivariate variability between methods.

Metagenomic sequencing analysis

Raw reads were quality filtered using trimmomatic v0.33 with default settings [35]. Reads that aligned to the human genome were removed, and the remaining sequences from each sample were assigned taxonomy using Kraken v0.10.5-beta and a custom-built database compiled with all archaeal, fungal, bacterial, viral and protozoan genomes available on NCBI as of September 13, 2016 (Additional file 2) [36]. The Bracken bioinformatics program was used to calculate abundances from Kraken-assigned taxonomy at the phylum, family, and species levels, and these data were used for all downstream analyses (Additional file 2) [37].

Population genome assembly and binning

After quality filtering and removal of human DNA from the metagenomic dataset, high-quality sequences from all samples were pooled, assembled, and assessed using SPAdes v3.7.1 [38], BamM v1.7.3 (https://github.com/Ecogenomics/BamM), GroopM v0.3.4 [39], and CheckM v1.0.7 [40] (Additional file 2). A single near-complete genome (97.35% complete) with no contamination was reconstructed. This population bin, identified as the bacterium *Propionibacterium acnes*, was examined for contigs with abnormal coverage or composition profile using RefineM v0.0.13 (https://github.com/dparks1134/refinem), then gaps filled and the refined genome bin assembled into scaffolds using FinishM (https://github.com/wwood/finishm). Taxonomic identification and phylogenetic inference were performed using Genome Taxonomy Database [41] and FastTree v2.1.9 [42], respectively (Additional file 2). Gene prediction, annotation and metabolic reconstruction for the recovered *P. acnes* genome were carried out using Rapid Annotations using Subsystems Technology (RAST) online server and all default settings [43]. The RAST server was used to compare the recovered *P. acnes* genome to that of its closest phylogenetic neighbour, *P. acnes* strain KPA171202.

Results

Total DNA was extracted from 18 sino-nasal middle meatus swab samples (3 patients × 6 methods) using a variety of bacterial enrichment techniques (N, N+WGA, C, C+WGA, NB, NB+WGA) (Figure 1). Bacterial 16S rRNA gene amplicons were sequenced using Illumina MiSeq, and metagenomic DNA samples were subjected to shotgun metagenomic sequencing using Illumina HiSeq. Quality filtering and rarefaction of MiSeq-derived reads resulted in 355,080 16S rRNA gene sequences from 14 samples. Four samples, including the centrifugation-treated sample from Patient 2, and all samples treated using the NB enrichment method, resulted in insufficient sequences for downstream analyses. Metagenomic sequencing returned a total of 539,433,708 sequences across 18 samples. Quality filtering and removal of host-associated DNA from the HiSeq-derived metagenome samples resulted in a total of 611,151 Kraken-classified sequences (<1 % of total sequences).

Efficacy of enrichment methods: Metagenomic data

Method C recovered the highest mean percentage of microbial sequences (mean 0.45% ± SD 0.75); however, this method did not yield reproducible results (coefficient of variation (CV) = 166%) (Figure 2). All other methods had CV values less than 100%, which indicate reproducibility (method N = 0.03% ± 0.029, CV=95.20%; method N+WGA = 0.0093% ± 0.0016, CV=17.40%; method C+WGA = 0.026% ± 0.024, CV= 90.54; method NB = 0.0064% ± 0.0028, CV=44.36%; method NB+WGA = 0.13 ± 0.11%, CV=85.97%). The Kruskal-Wallis group test revealed no significant differences among any of the enrichment methods regarding the proportion of recovered microbial sequence reads (p = 0.056).

Effect of DNA enrichment method on profiling of microbial community structure

Microbial communities from both amplicon and metagenomic datasets were dominated by the bacterial phyla *Actinobacteria, Firmicutes*, and *Proteobacteria*. Metagenomic sequencing revealed a large relative abundance of the archaeal phylum *Euryarchaeota*, including members of the family *Methanobacteriaceae*. On average, *Propionibacterium acnes*, *Staphylococcus aureus, Methanobacterium formicicum, Salmonella enterica* and *Staphylococcus epidermidis* were the five species that dominated

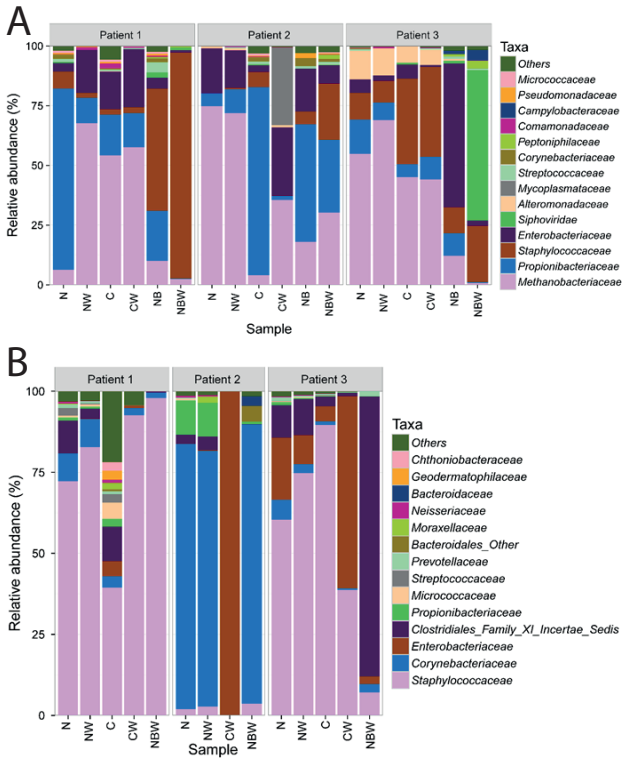

Figure 3. Relative abundance taxa plot of the top 14 family level taxon-classified sequences from (A) Metagenomic sequencing (n = 18), and (B) Bacterial 16S rRNA gene amplicon sequencing (n = 14).

the metagenome datasets. Evaluation of community composition at family level revealed differences in relative abundances between types of sequencing and enrichment methods (Figure 3). Differences in community composition between bacterial amplicon and metagenomic sequencing was expected due to the targeted nature of 16S rRNA gene sequencing to bacteria. The metagenomic sequencing results indicated the presence of over 500 species, including a variety of archaea, viruses and bacteriophages, bacteria, and fungi (Additional file 3). Relatively few types of fungi (all belonging to division *Ascomycota*) were recovered, and these were only in the centrifugation sample from Patient 2. Furthermore, metagenomic sequencing recovered 39 species of *Corynebacterium*, 5 species of *Propionibacterium*, and 12 species of *Staphylococcus* (Additional file 4). Several types of double-stranded DNA viruses were also detected, the majority of which were members of the viral family *Papillomaviridae* and many of which were found only in Patient 2 (Additional file 3). Spearman correlation coefficients were used to compute the correlation between amplicon and metagenomic sequencing samples for only family-assigned bacterial taxa (all viral, fungal, and members of the phylum *Euryarchaeota* were removed from the metagenomic dataset) (rs = 0.38, p = 0.001). These results suggest that the two types of sequencing recover dissimilar microbial community profiles. This may be due to the relatively few sequencing reads from metagenomic sequencing; however, we encourage future studies to incorporate both amplicon and metagenomic sequencing, where possible, for further comparisons.

Effect of DNA enrichment method on comparisons of beta-diversity

PERMANOVA results from amplicon and metagenomic (for species and family-level taxa) Bray-Curtis dissimilarity matrices revealed no significant influence of enrichment methods on variation in either MiSeq or HiSeq datasets ($p > 0.05$). PERMANOVA results from the amplicon data suggested that differences between patients' bacterial diversity contributed 36.5% to the variation in this dataset ($p = 0.001$, $R2=0.365$). This result was not evident in the metagenomic PERMANOVA results, however. These clustering patterns are not surprising considering previous research reporting the heavy influence of inter-subject variability [8].

Both the metagenomics and amplicon approaches revealed similar bacterial community compositions in regards to the effect of enrichment method on beta-diversity (Figure 4). In the metagenomic dataset, analysis of similarity (ANOSIM) pair-wise tests between all methods and the no-enrichment control revealed no significant differences. This finding is visualized in the nMDS, with the no-enrichment control samples positioned in the center of all samples (Figure 4A). Comparisons of variability between methods, using MVDISP, revealed that the NW samples clustered closest to the N samples, and the NBW method showed the largest dispersion from the no-enrichment control samples, suggesting this method recovers slightly (although not significantly) different microbial community compositions. Similarity percentages (SIMPER) results revealed samples treated with the NBW method were the most dissimilar to the no-treatment control (average dissimilarity = 71%), and samples from the NW method were the most similar to the no-treatment control (average dissimilarity = 44.6%).

Samples tended to cluster more by patient in the amplicon sequencing nMDS plot when compared to the metagenomic nMDS. Similar to the metagenomic dataset, ANOSIM results revealed no significant differences between microbial community compositions recovered by the different methods. However, samples treated using the CW method clustered further away from the no-enrichment control samples (Figure 4B). SIMPER results suggested samples treated with this method (CW) were most dissimilar to the no-treatment control (average dissimilarity = 78.5%), and samples from the NW method were the most similar to the no-treatment control (average dissimilarity = 45.6%).

Propionibacterium acnes genome

Phylogenomic inference of the concatenated marker genes

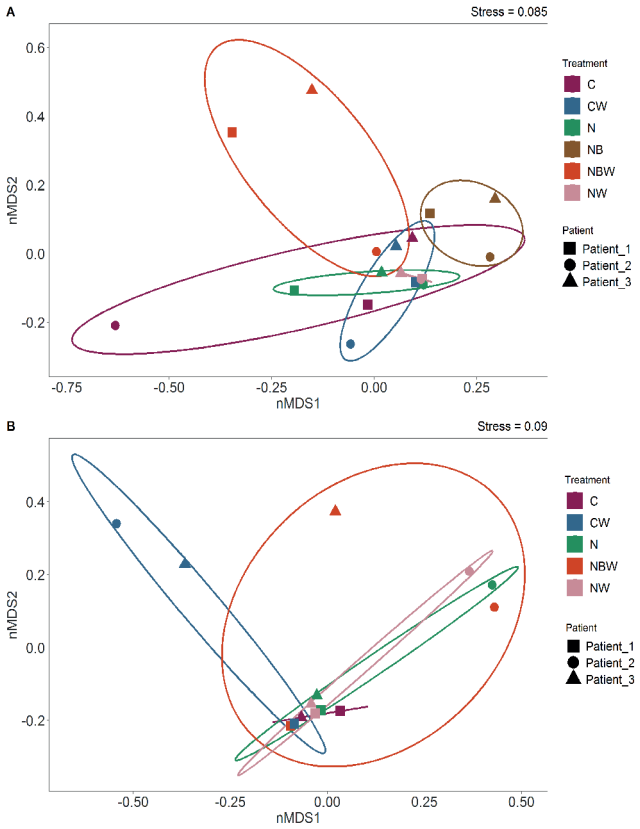

Figure 4. Non-metric multidimensional scaling (nMDS) plot of (A) Metagenomic sequencing data, (B) Bacterial 16S rRNA gene targeted amplicon data. Patients are represented by shapes, and treatment categories are represented by colours. Ellipses represent the mean of the description coordinates at the centre, and the dispersion of the ellipses were calculated using the standard error of the weighted average of covariance matrix group scores.

identified the genome as *Propionibacterium acnes*, a common skin commensal bacterium. The recovered *P. acnes* genome is 2,588,344 bp, with 60.0% GC content, and contains 2,510 coding sequences. A comparative analysis to characterize the pan-genome of 69 *P. acnes* isolates (67 were isolated from human skin) reported an average genome size of 2.5 Mb, 60% GC content, and 2,626 open reading frames [44]. The recovered *P. acnes* genome formed a strong supported monophyletic clade with *P. acnes* strains KPA171202 (GenBank assembly GCF_000008345.1, [45]) and JCM 18 (82.96% complete, 0.78% contamination), (GenBank assembly GB_GCA_000521405.1) (Additional file 5).

Annotation and pathway mapping of protein-coding genes (CDs) classified into subsystem categories revealed that a majority of CDs belonged to amino acid and derivative production, followed by genes coding for carbohydrates (Additional file 6). Due to incompleteness of the recovered *P. acnes* genome, the apparent presence or absence of a single copy or particular coding genes should be regarded with caution. Comparisons of metabolic reconstruction between our recovered *P. acnes* genome and the genome of *P. acnes* strain KPA171202, isolated from human skin, revealed 112 genes that are present in the CRS *P. acnes* genome but not the KPA171202 reference strain (Additional file 7). A majority of the genes unique to the recovered *P. acnes* genome were related to carbohydrate production.

Discussion

The results from this methods study suggest that the enrichment techniques unpredictably alter microbial community profiles when comparing them to non-treated samples. Despite these effects and the ineffective removal of contaminating human DNA, a diverse range of viral, fungal, archaeal and bacterial species were reported, as well as a near-complete genome of the bacterium *Propionibacterium acnes*.

Microbial enrichment techniques

The application of metagenomic sequencing to sino-nasal research is promising, but greater sequencing depths will be required to garner useful information. In this study, the centrifugation method was the most promising enrichment method as it recovered the largest proportion of microbial sequences, however it was inconsistent between samples. Commercial tools, such as the NEBNext® Microbiome DNA Enrichment Kit are specifically designed to remove human DNA, however we observed no significant increase in the proportion of microbial assigned sequences when compared with the no treatment control. Furthermore, the incorporation of WGA MDA prior to sequencing did not improve the recovery of microbial sequences in any of the enrichment methods.

This study has several limitations. First, the results from this study provide limited insights into the function of the microbiome in CRS due to small sample sizes and the lack of healthy controls. Additionally, future metagenomic studies with low biomass samples should include sequencing results from negative controls throughout the experiment to assess contamination in low biomass samples [46]. Other enrichment techniques with sino-nasal samples should be explored, including modifying the sample type and testing a wider range of methods. For example, mucous lavage samples provide more starting volume than swab samples and multiple enrichments from the same starting material may be applied. Finally, where possible, we recommend processing samples fresh from collection to prevent lysing of human cells and the uncontrolled release of human DNA which may affect enrichment outcomes.

The application of metagenomic sequencing to CRS research will first have to overcome the substantial challenge of the overwhelming proportion of host-associated DNA. Careful consideration of enrichment techniques more generally, regarding sample type (tissue versus swab versus lavage), biases and limi-

tations of enrichment methods, costs (sample processing and sequencing), as well as desired outcomes, is necessary. Based on these results, we recommend sequencing fewer sino-nasal derived samples per metagenomic sequencing run to increase the depth of sequencing without altering in situ microbial community structure.

Microbial composition revealed by metagenomic sequencing

The bacterial genera *Corynebacterium*, *Propionibacterium*, and *Staphylococcus* are frequent colonizers of the sino-nasal cavity, however their role in health and contribution to CRS pathogenesis remains unclear [47–49]. High species- and strain-level variability of these genera typically goes uncharacterized due to the limited taxonomic resolution of 16S rRNA gene-targeted sequencing. Future studies should focus on characterizing the presence of these key bacteria at increased resolution in patients with and without CRS.

The results from this study identified 55 *Propionibacterium* and 14 *Staphylococcus* phages. The extensive diversity and presence of bacteriophages, especially related to the genera *Propionibacterium* and *Staphylococcus*, is consistent with results from a previous metagenomic study which reported a significant abundance of viral DNA, including bacteriophages, in the nares of healthy subjects when compared to other skin sites (mean relative abundance 51.0% ± 11.8 S.E.) [15]. The presence and diversity of dsDNA viruses in our results are consistent with findings from others [50,51].

High recovered proportions of the methanogenic archaeal species *Methanobacterium* formicicum, belonging to the family *Methanobacteriaceae*, were unexpected and are not well studied elsewhere in the sino-nasal literature. Whether such high relative abundances are typical for the sino-nasal cavity during health or disease should be addressed in future studies. Additionally, such low levels of fungi were in agreement with the pilot data, yet unexpected, as previous amplicon studies have identified several fungal species in both healthy patients and those with CRS [52,53]. Taken together, these results suggest that fungi are present in the sino-nasal cavity, although at very low relative abundances, and that amplification of fungal DNA may be necessary to capture total diversity.

Bacterial taxa dominated the reference database, with lower representation of fungal, archaeal and viral genomes, so it is likely that these latter microbes are underrepresented in our results. Additionally, the general lack of metagenomic and viral data from patients with CRS makes it difficult to contextualize the results from this study, which included data from only three CRS patients. A study examining the eukaryotic double-stranded DNA and single-stranded DNA viruses from the Human Microbiome Project cohort [50] reported unique viral fingerprints among subjects (much like host-associated bacterial communities), a combination of stable and transient viral carriage, and an average diversity of 5.5 viruses per individual. Of clinical importance, carriage of a known disease-causing virus was not associated with symptoms or apparent clinical consequences (50), which may suggest that onset of disease is an interaction of events involving the host immune system, bacterial and active viral infection, and fungal communities. The CRS-associated virome warrants further investigation, with the few published studies being somewhat contradictory [54,55].

Functional insights from metagenomic sequencing

Although only negligible improvements were made to the recovery of microbial DNA, we nevertheless sought to explore the potential of these data to deliver useful genomic insights. We succeeded in reconstructing a near-complete *Propionibacterium acnes* genome (97.35% complete, 0% contamination). Interestingly, our *P. acnes* genome contains the cas1 gene, which is part of the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas locus which helps to protect the bacterium from bacteriophages and other mobile genetic elements [56]. The presence of CRISPR/Cas genes in *P. acnes* is not uncommon, but is typically identified in type II *P. acnes* strains [44,45]. If our genome is in fact a type II strain, it may be functionally more similar to *P. acnes* ATCC_11828 strain. Some evidence suggests people carry different *P. acnes* strains in the same environment, and the role of *P. acnes* in CRS, and the sino-nasal cavity more generally, should be investigated further.

Conclusion

Existing research in CRS microbiology is limited to culture-based and gene-targeted approaches. Expanding our knowledge base from identifying which bacteria are present in the sinuses towards a view which includes archaeal, fungal, and viral species, and describing their functional importance and impact on health status, is the next logical step for studying CRS pathogenesis. Taken together, the results from this study support the application of metagenomic sequencing techniques in the study of microbial communities associated with CRS, however we do not recommend enriching samples for microbial DNA using the techniques described here. We encourage continued research that focuses on limiting the proportion of recovered human DNA in order to increase the resolution of in situ microbial communities.

Acknowledgements

The authors would like to thank the patients who took part in this study. Many thanks to Philip Hugenholtz and ACE for computing support, David Wood for his help operating RemoveM, and Brian Kemish for providing computing support. The research in this study was supported by The Garnett Passe and

Rodney Williams Memorial Foundation Charitable Trust Fund.
Sequencing data are available upon request.

Authorship contribution

BWM analysed and interpreted the data and wrote the manuscript. DWW helped with analysis of the dataset. KB, MWT helped with study design and interpretation of the dataset. RGD collected the samples and contributed to interpretation of the results. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References:

1. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature.;486(7402):207–14.
2. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, et al. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature. 2017;550(7674):61–6.
3. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol. 2014;32(8).
4. Bhattacharyya N. Incremental health care utilization and expenditures for chronic rhinosinusitis in the United States. Ann Otol Rhinol Laryngol. 2011;120(7):423–7.
5. Fokkens WJ, Lund VJ, Mullol J, Bachert C, Alobid I, Baroody F. European Position Paper on Rhinosinusitus and Nasal Polys 2012. Rhinology. 2012;(23):1–20.
6. Orlandi RR, Kingdom TT, Hwang PH, Smith TL, Alt JA, Baroody FM, et al. International Consensus Statement on Allergy and Rhinology: Rhinosinusitis. Int Forum Allergy Rhinol. 2016;6(November 2015):S22–209.
7. Abreu NA, Nagalingam NA, Song Y, Roediger FC, Pletcher SD, Goldberg AN, et al. Sinus microbiome diversity depletion and Corynebacterium tuberculostearicum enrichment mediates rhinosinusitis. Sci Transl Med. 2012;4(151):151ra24.
8. Biswas K, Hoggard M, Jain R, Taylor MW, Douglas RG. The nasal microbiota in health and disease: variation within and between subjects. Front Microbiol. 2015 Jan;9(March):134.
9. Hoggard M, Biswas K, Zoing M, Wagner Mackenzie B, Taylor MW, Douglas RG. Evidence of microbiota dysbiosis in chronic rhinosinusitis. Int Forum Allergy Rhinol. 2016;7(3):7–11.
10. Ramakrishnan VR, Hauser LJ, Feazel LM, Ir D, Robertson CE, Frank DN. Sinus microbiota varies among chronic rhinosinusitis phenotypes and predicts surgical outcome. J Allergy Clin Immunol. 2015 Aug;136(2):334–42.e1.
11. Hoggard M, Mackenzie BW, Jain R, Taylor MW, Biswas K, Douglas RG. Chronic rhinosinusitis and the evolving understanding of microbial ecology in chronic inflammatory mucosal disease. Clin Microbiol Rev. 2017;30(1):321–48.
12. Wagner Mackenzie B, Waite DW, Hoggard M, Douglas RG, Taylor MW, Biswas K. Bacterial community collapse: A meta-analysis of the sinonasal microbiota in chronic rhinosinusitis. Environ Microbiol. 2016;19:381–92.
13. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of Escherichia coli K-12. Sci Mag. 1997;277(September):1453–62.
14. Venter JC, Adams MDM, Myers EEW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001;291(5507):1304-51.
15. Oh J, Byrd AL, Deming C, Conlan S, Barnabas B, Blakesley R, et al. Biogeography and individuality shape function in the human skin metagenome. Nature. 2014;514(7520):59.
16. Wylie KM, Mihindukulasuriya K a, Zhou Y, Sodergren E, Storch G a, Weinstock GM. Metagenomic analysis of double-stranded DNA viruses in healthy adults. BMC Biol. 2014;12:71.
17. Feehery GR, Yigit E, Oyola SO, Langhorst BW, Schmidt VT, Stewart FJ, et al. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. PLoS One. 2013 Jan;8(10):e76096.
18. Bright A, Tewhey R, Abeles S, Chuquiyauri R, Llanos-Cuentas A, Ferreira MU, et al. Whole genome sequencing analysis of Plasmodium vivax using whole genome capture. BMC Genomics. 2012;13(1):262.
19. Burke RM, McKenna JP, Cox C, Coyle P V., Shields MD, Fairley DJ. A comparison of different pre-lysis methods and extraction kits for recovery of Streptococcus agalacticae (Lancefield group B Streptococcus) DNA from whole blood. J Microbiol Methods. 2016;129:103–8.
20. Thurber R V., Haynes M, Breitbart M, Wegley L, Rohwer F. Laboratory procedures to generate viral metagenomes. Nat Protoc. 2009;4(4):470–83.
21. Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome. 2018;6(1):1–9.
22. Clarke EL, Lauder AP, Hofstaedter CE, Hwang Y, Fitzgerald AS, Imai I, et al. Microbial lineages in sarcoidosis a metagenomic analysis tailored for low-microbial content samples. Am J Respir Crit Care Med. 2018;197(2):225–34.
23. Horz HP, Scheer S, Vianna ME, Conrads G. New methods for selective isolation of bacterial DNA from human clinical specimens. Anaerobe. 2010;16(1):47–53.
24. Thoendel M, Jeraldo PR, Greenwood-Quaintance KE, Yao JZ, Chia N, Hanssen AD, et al. Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. J Microbiol Methods. 2016;127:141–5.
25. Hasan MR, Rawat A, Tang P, Jithesh P V, Thomas E, Tan R, et al. Depletion of human DNA in spiked clinical specimens to improve the sensitivity of pathogen detection by next generation sequencing. J Clin Microbiol. 2016;54(January):JCM.03050-15-.
26. Votintseva AA, Pankhurst LJ, Anson LW, Morgan MR, Gascoyne-Binzi D, Walker TM, et al. Mycobacterial DNA extraction for whole-genome sequencing from early positive liquid (MGIT) cultures. J Clin Microbiol. 2015;53(4):1137–43.
27. Glassing A, Dowd SE, Galandiuk S, Davis B, Jorden JR, Chiodini RJ. Changes in 16s RNA gene microbial community profiling by concentration of prokaryotic DNA. J Microbiol Methods. 2015;119:239–42.
28. Leo S, Gaïa N, Ruppé E, Emonet S, Girard M, Lazarevic V, et al. Detection of bacterial pathogens from broncho-alveolar lavage by next-generation sequencing. Int J Mol Sci. 2017;18(9):2011.
29. Oyola SO, Gu Y, Manske M, Otto TD, O'Brien J, Alcock D, et al. Efficient depletion of host DNA contamination in malaria clinical sequencing. J Clin Microbiol. 2013;51(3):745–51.
30. Liu P, Fang X, Feng Z, Guo Y-M, Peng R-J, Liu T, et al. Direct sequencing and characterization of a clinical isolate of Epstein-Barr virus from nasopharyngeal carcinoma tissue by using next-generation sequencing technology. J Virol. 2011;85(21):11291–9.
31. Loonen AJM, Bos MP, van Meerbergen B, Neerken S, Catsburg A, Dobbelaer I, et al. Comparison of pathogen DNA isolation methods from large volumes of whole blood to improve molecular diagnosis of bloodstream infections. PLoS One. 2013;8(8):1–7.
32. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013 Oct;10(10):996–8.
33. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput

community sequencing data. Nat Methods. 2010;7(5):335.
34. Anderson MJ, Gorley RN, Clarke KR. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. Plymouth, U.K.: PRIMER-E Ltd; 2008.
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114.
36. Wood DE, Salzberg SL. Kraken : ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15(3).
37. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: Estimating species abundance in metagenomics data. bioRxiv. 2016;1–14.
38. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
39. Imelfort M, Parks D, Woodcroft BJ, Dennis P, Hugenholtz P, Tyson GW. GroopM: An automated tool for the recovery of population genomes from related metagenomes. PeerJ. 2014;2:e409v1.
40. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043–55.
41. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A proposal for a standardized bacterial taxonomy based on genome phylogeny. bioRxiv. 2018. doi: https://doi.org/10.1101/256800.
42. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. 2009 Jul;26(7):1641–50.
43. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: Rapid annotations using subsystems technology. BMC Genomics. 2008;9:1–15.
44. Tomida S, Nguyen L, Chiu B, Analyses CG, Its R, Diversity G, et al. Pan-genome and comparative genome analyses of Propionibacterium acnes reveal its genomic diversity in the healthy and diseased human skin microbiome. MBio. 2013;4(3):e00003-13.
45. Brüggemann H, Lomholt HB, Tettelin H, Kilian M. CRISPR/cas loci of type II Propionibacterium acnes confer immunity against acquisition of mobile elements present in type I P. acnes. PLoS One. 2012;7(3):e34171.
46. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov;12(1):87.
47. Liu CM, Soldanova K, Nordstrom L, Dwan MG, Moss OL, Contente-Cuomo TL, et al. Medical therapy reduces microbiota diversity and evenness in surgically recalcitrant chronic rhinosinusitis. Int Forum Allergy Rhinol. 2013 Oct;3(10):775–81.
48. Kaspar U, Kriegeskorte A, Schubert T, Peters G, Rudack C, Pieper DH, et al. The culturome of the human nose habitats reveals individual bacterial fingerprint patterns. Environ Microbiol. 2015 Apr 28;18(7):2130–42.
49. Ramakrishnan VR, Frank DN. Impact of cigarette smoking on the middle meatus microbiome in health and chronic rhinosinusitis. Int Forum Allergy Rhinol. 2015 Aug 14;5(11):981–9.
50. Wylie KM, Weinstock GM, Storch GA. Emerging view of the human virome. Transl Res. 2012;160(4):283–90.
51. Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, Silva J, et al. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. PLoS One. 2009;4(10):e7370.
52. Zhao YC, Bassiouni A, Tanjararak K, Vreugde S, Wormald P-J, James Psaltis A. Role of fungi in chronic rhinosinusitis through ITS sequencing. Laryngoscope. 2017;1–7.
53. Cleland EJ, Bassioni A, Boase S, Dowd S, Vreugde S, Wormald PJ. The fungal microbiome in chronic rhinosinusitis: Richness, diversity, postoperative changes and patient outcomes. Int Forum Allergy Rhinol. 2014;4(4):259–65.
54. Cho GS, Moon BJ, Lee BJ, Gong CH, Kim NH, Kim YS, et al. High rates of detection of respiratory viruses in the nasal washes and mucosae of patients with chronic rhinosinusitis. J Clin Microbiol. 2013;51(3):979–84.
55. Wood AJ, Antoszewska H, Fraser J, Douglas RG. Is chronic rhinosinusitis caused by persistent respiratory virus infection? Int Forum Allergy Rhinol. 2011;1(2):95–100.
56. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science. 2010;327(5962):167–70.

Professor Richard G. Douglas
School of Medicine
Department of Surgery
The University of Auckland
Auckland
New Zealand

E-mail:
Richard.douglas@auckland.ac.nz

Additional files

Table S1. Baseline characteristics of patients at the time of surgery and sample collection.

Patient	Sex	Age	Ethnicity	Diagnosis	Procedure	Smoker	Asthma/ Aspirin Sensitivity	Antibiotics	Prednisone
1	M	59	NZE	CRS	FHFESSa, ITb	Ex	No	No	No
2	M	30	NZE	CRS	FHFESS, Sc, IT	Ex	No	No	No
3	M	55	NZE	CRS	FHFESS, FDd	No	No	No	No

aFHFESS: full-house functional endoscopic sinus surgery

bIT: indicates removal of inferior turbinates

cS: septoplasty procedure. Used to surgically modify the nasal septal cartilage.

dFD: frontal drill out, also known as Lothrop procedure. Used in revision cases to create the largest possible frontal ostium.

Supplementary methods

Sequencing preparation

PCR amplification and Illumina sequencing

In order to compare the recovery of bacterial community composition profiles based on metagenomic sequencing to those based on the usual approach employed by CRS researchers (i.e. 16S rRNA gene-targeted sequencing), we amplified the V3-V4 hypervariable region of the bacterial 16S rRNA gene for each sample in this project. Briefly, up to 3 µL of template DNA from each sample was added to the PCR master mix, and as many as three PCR replicates were completed for each sample in order to generate sufficient amplicon product for sequencing. The triplicate negative extractions from the DNA extraction kits were amplified and verified for lack of contamination.

Replicate PCR products from each sample were pooled and purified using Agencourt AMPure beads according to manufacturer instructions. PCR products were quantified fluorometrically using the High Sensitivity (HS) kit on the Qubit® Fluorometer 1.0 (Invitrogen Co., Carlsbad, CA, USA) and qualitatively assessed using the Agilent High Sensitivity DNA chip (Agilent Technologies, Santa Clara, CA, USA).

Bacterial 16S rRNA gene amplicons were submitted to New Zealand Genomics Limited for library preparation using a dual-indexing approach with Nextera technology and sequencing (2 x 300 bp, paired-end) on the Illumina MiSeq. Metagenomic samples were submitted to New Zealand Genomics Limited for Thruplex DNA library preparation and sequencing (2 x 125 bp, paired-end) on one lane of the Illumina HiSeq.

Data analyses

Bacterial 16S rRNA gene sequence analysis

Bioinformatic processing of amplicon sequencing data involved a combination of USEARCH (version 7.0.1090, 64-bit built for Linux) and QIIME (version 1.8) [1,2]. Briefly, reads less than 200 bp after merging were removed from the dataset. USEARCH was used to cluster reads into de novo operational taxonomic units (OTUs) at 97% sequence similarity, singleton OTUs were removed, and taxonomy was assigned in QIIME using RDP v2.2 and SILVA v111 as a reference [3,4]. Sequences that aligned to the human mitochondrial genome using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) were removed from the dataset. Samples were rarefied to 1,678 sequences, and rarefied tables were used for all downstream analyses [5]. Alpha diversity measures Chao1, Shannon, Simpson and observed species (OTUs) diversity indices, and a Bray-Curtis dissimilarity matrix were assessed and generated using QIIME.

Metagenomic sequencing analysis

Raw reads were quality filtered using trimmomatic v0.33 with default settings [6]. To remove contaminating human DNA, the human reference genome GRCh38 was downloaded from the NCBI Genome database and quality-filtered reads aligned to the reference using bwa [7]. Reads that aligned to the human genome were removed, and the remaining sequences from each sample were assigned taxonomy using Kraken v0.10.5-beta and a custom-built database compiled with all archaeal, fungal, bacterial, viral and protozoan genomes available on NCBI as of September 13, 2016 [8]. The Bracken bioinformatics program was used to calculate abundances from Kraken-assigned taxonomy at the phylum, family, and species levels, and these data were used for all downstream analyses [9].

Efficacy of enrichment techniques was assessed as the proportion of microbial classified sequencing reads, once human-assigned reads were removed, to the total number of reads prior to classification. The mean value and standard deviation of classified sequences for each method were calculated and visualized as strip plots using the program ggplot2 in R version 3.2.5 [10,11]. Coefficient of variation, tests for normality, and pairwise tests were calculated using the native 'stats' package in R for each method to give an indication of the reproducibility, distribution,

and quantifiable differences between methods, respectively [11]. Relative abundances of the 22 most abundant microbial families within the amplicon and metagenomics datasets were calculated and visualized in R. Beta diversity metrics for metagenomics data were calculated using the species-level Bracken abundance taxon table using a Bray-Curtis dissimilarity matrix generated with the 'vegdist' function in the vegan package [12]. Non-metric multidimensional scaling (nMDS) plots for both amplicon and metagenomics datasets were generated as previously described [13]. Comparisons of total microbial, and bacterial only, diversity at family level, as assessed by the amplicon and metagenomics approaches, were calculated in QIIME v1.9 using Spearman correlations and all other default settings in the compare_taxa_summaries.py command.

Population genome assembly and binning

After quality filtering and removal of human DNA from the HiSeq dataset, high-quality sequences from all samples were pooled and assembled using SPAdes v3.7.1 [14]. Reads from each sample were separately mapped to the resulting assembly using BamM v1.7.3 (https://github.com/Ecogenomics/BamM) and differential coverage binning performed using GroopM v0.3.4 [15]. Completeness and contamination of each population genome bin were assessed using the presence or absence of 120 single copy marker genes using CheckM v1.0.7 [16]. Several bins were reported, including one bin that identified as a member of the genus *Staphylococcus*. However, this genome reported only 45.88% completeness with 1.30% contamination, and was not pursued for reconstruction. A single near-complete genome (97.35% complete) with no contamination was reconstructed. This population bin, identified as the bacterium *Propionibacterium acnes,* was examined for contigs with abnormal coverage or composition profile using RefineM v0.0.13 (https://github.com/dparks1134/refinem), then gaps filled and the refined genome bin assembled into scaffolds using FinishM (https://github.com/wwood/finishm).

Analysis of the Propionibacterium acnes genome

Taxonomic identification of the refined genome bin was performed against a reference set of 14,256 high-quality bacterial genomes downloaded from the Genome Taxonomy Database (http://gtdb.ecogenomic.org/) using a concatenated protein sequence obtained from 120 marker genes [17]. Phylogenetic inference was performed using FastTree v2.1.9 [18] with the WAG+Γ model of amino acid evolution and 100 bootstrap iterations to assess node support. A high-resolution tree for species-level identification and publication purposes was constructed from a subset of 44 reference genomes, consisting of closely related and outgroup genomes from different phyla, using RAxML v8.1.4 [19] under the same evolution model and bootstrap criteria, for display purposes.

Gene prediction, annotation, and metabolic reconstruction for the recovered *P. acnes* genome were carried out using Rapid Annotations using Subsystems Technology (RAST) online server and all default settings [20]. The RAST server was used to compare the recovered *P. acnes* genome to that of its closest phylogenetic neighbour, *P. acnes* strain KPA171202.

Statistical analyses

The similarity percentage (SIMPER) approach was used to identify microbial community similarities between groups of samples based on enrichment method for each type of sequencing data. The species-level metagenome taxa summary and the amplicon taxon-assigned OTU tables were square root-transformed and SIMPER analyses were conducted in PRIMER6 version 6.1.13 using Bray-Curtis similarities [21]. Permutational analysis of variance (PERMANOVA) was used to partition variation within the data due to inter-individual variation and enrichment method, and analysis of similarity (ANOSIM) was used to assess patterns in microbial community composition between methods for both amplicon and metagenomics datasets. Multivariate dispersion (MVDISP) was used to quantify relative multivariate variability between methods. PERMANOVA, ANOSIM, and MVDISP analyses were performed in PRIMER6 version 6.1.13 using Bray-Curtis dissimilarity matrices generated from each dataset [22].

References

1. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10(10):996–8.
2. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nature. 2010;7(5):335–6.
3. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Glo FO, et al. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. Nucleic Acids Res. 2013;41:590–6.
4. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261–7.
5. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res. 2008;36:5–9.
6. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114.
7. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
8. Wood DE, Salzberg SL. Kraken : ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15(3).
9. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: Estimating species abundance in metagenomics data. bioRxiv. 2016;1–14.
10. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verglad, New York; 2009. Available from: http://ggplot2.org
11. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2012. Available from: http://

www.r-project.org/
12. Oksanen AJ, Blanchet FG, Kindt R, Legen- P, Minchin PR, Hara RBO, et al. Vegan: Community ecology package. R Package Version 2.0.10. 2011.
13. Wagner Mackenzie B, Waite DW, Hoggard M, Douglas RG, Taylor MW, Biswas K. Bacterial community collapse: A meta-analysis of the sinonasal microbiota in chronic rhinosinusitis. Environ Microbiol. 2016;19:381–92.
14. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
15. Imelfort M, Parks D, Woodcroft BJ, Dennis P, Hugenholtz P, Tyson GW. GroopM: An automated tool for the recovery of population genomes from related metagenomes. PeerJ Prepr. 2014;2:e409v1.
16. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from. Genome Res. 2015;1:1–31.
17. Ormerod KL, Wood DLA, Lachner N, Gellatly SL, Daly JN, Parsons JD, et al. Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. Microbiome. 2016;4:1–17.
18. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. 2009;26(7):1641–50.
19. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.
20. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: Rapid annotations using subsystems technology. BMC Genomics. 2008;9:1–15.
21. Clarke KR. Non-parametric multivariate analyses of changes in community structure. Aust J Ecol. 1993;18:117–43.
22. Anderson MJ, Gorley RN, Clarke KR. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. Plymouth, U.K.: PRIMER-E Ltd; 2008.

Table S2. Bracken normalised relative abundances of recovered (A) Bacteria, (B) Fungi and (C) Viruses in samples from metagenomic sequencing. Page 171 - 185.

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Achromobacter_xylosoxidans	0	0	0	0	0	0	0	0	9.87E-05	0	0	0	0	0	0	0	0	0
Acidovorax_avenae	0	0	0	0	0	0	0	0	5.78E-05	0	0	0	0	0	0	0	0	0
Acidovorax_ebreus	0	0	0	0	0	0	0	0	0.000173297	0	0	0	0	0	0	0	0	0
Acidovorax_sp._JS42	0	0	0	0	0	0	0	0	0.000194959	0	0	0	0	0	0	0	0	0
Acidovorax_sp._KKS102	0	0	0	0	0	0	0	0	0.000125159	0	0	0	0	0	0	0	0	0
Acidovorax_sp._RAC01	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Acinetobacter_baumannii	0	0	0	0	0	0	0	0	0.000315304	0	0	0	0	0	0	0	0	0
Acinetobacter_johnsonii	0	0	0	0	0	0	0	0	0.000368256	0	0	0	0	0	0	0	0	0
Acinetobacter_nosocomialis	0	0	0	0	0	0	0	0	0.000368256	0	0	0	0	0	0	0	0	0
Acinetobacter_pittii	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Acinetobacter_sp._NCu2D-2	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Actinomyces_meyeri	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Actinomyces_oris	0.000780553	0	0	0	0	0	0	0	0.001047002	0	0	0	0	0	0	0	0	0
Actinomyces_radicidentis	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0
Actinomyces_sp._oral_taxon_414	0	0	0	0	0	0	0	0	0.000182924	0	0	0	0	0	0	0	0	0
Actinoplanes_friuliensis	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Actinoplanes_missouriensis	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Actinoplanes_sp._N902-109	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Actinoplanes_sp._SE50/110	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Actinosynnema_mirum	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Aerococcus_christensenii	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Aerococcus_urinaeequi	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Aeromicrobium_erythreum	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Agrobacterium_fabrum	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Agrobacterium_tumefaciens	0	0	0	0	0	0	0	0	0.000108311	0	0	0	0	0	0	0	0	0
Agromyces_sp._AR33	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Alicycliphilus_denitrificans	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Alteromonas_macleodii	0	0	0	0	0	0	0	0	0	0.003841794	0	0	0	0	0	0	0	0
Alteromonas_mediterranea	0	0	0.002067065	0	0	0	0.009666506	0.007481297	0.000601725	0.004597556	0	0.002611367	0.123337766	0.116496087	0.072788828	0.067056945	0.00877193	0.000999273
Amycolatopsis_mediterranei	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Amycolatopsis_methanolica	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Amycolatopsis_orientalis	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Anaerococcus_prevotii	0	0	0	0	0	0	0	0	0.000284014	0	0	0	0	0	0	0	0	0.001302083
Arsenicicoccus_sp._oral_taxon_190	0	0	0	0	0	0	0	0	9.15E-05	0	0	0	0	0	0	0	0	0
Arthrobacter_sp._ERGS1:01	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Arthrobacter_sp._PAMC_25486	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Atopobium_parvulum	0	0	0	0	0	0	0	0	0.000127566	0	0	0	0	0	0	0	0	0
Bacteroides_fragilis	0	0	0	0	0	0	0	0	9.87E-05	0	0	0	0	0	0	0	0	0.000545058

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Bacteroides_vulgatus	0	0	0	0	0	0	0	0	7.22E-05	0	0	0	0	0	0	0	0	0
Betapapillomavirus_1	0	0	0	0	0	0	0	0	0.000161262	0	0	0	0	0	0	0	0	0
Betapapillomavirus_2	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Beutenbergia_cavernae	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Bifidobacterium_adolescentis	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Bifidobacterium_animalis	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Bifidobacterium_longum	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Blastococcus_saxobsidens	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Blastomonas_sp._RAC04	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Bosea_sp._RAC05	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Brachybacterium_faecium	0	0	0	0	0	0	0	0	0.00010109	0	0	0	0	0	0	0	0	0
Bradyrhizobium_diazoefficiens	0	0	0	0	0	0	0	0	0.000115531	0	0	0	0	0	0	0	0	0
Bradyrhizobium_icense	0	0	0	0	0	0	0	0	9.39E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_japonicum	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_sp.	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_sp._BTAi1	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_sp._CCGE-LA001	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_sp._ORS_278	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Bradyrhizobium_sp._S23321	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Brevibacterium_linens	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Brevundimonas_naejangsanensis	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Brevundimonas_sp._DS20	0	0	0	0	0	0	0	0	5.54E-05	0	0	0	0	0	0	0	0	0
Brevundimonas_sp._GW460-12-10-14-LB2	0	0	0	0	0	0	0	0	0.000105904	0	0	0	0	0	0	0	0	0
Brevundimonas_subvibrioides	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Buchnera_aphidicola	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Campylobacter_concisus	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Campylobacter_gracilis	0	0	0	0	0	0	0	0	5.78E-05	0	0	0	0	0	0	0	0	0
Campylobacter_hominis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000878149
Campylobacter_ureolyticus	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0.015037594	0.043665213
Candida_dubliniensis	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Capnocytophaga_sp._oral_taxon_323	0	0	0	0	0	0	0	0	8.66E-05	0	0	0	0	0	0	0	0	0
Carnobacterium_maltaromaticum	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Castellaniella_defragrans	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Catenulispora_acidiphila	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Caulobacter_segnis	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Cellulomonas_fimi	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Cellulomonas_flavigena	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Cellulomonas_gilvus	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Chelatococcus_sp._CO-6	0	0	0	0	0	0	0	0	6.98E-05	0	0	0	0	0	0	0	0	0
Chryseobacterium_sp._IHB_B_10212	0	0	0	0	0	0	0	0	6.26E-05	0	0	0	0	0	0	0	0	0
Citrobacter_freundii	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Clavibacter_michiganensis	0	0	0	0	0	0	0	0	0.000185331	0	0	0	0	0	0	0	0	0
Clostridioides_difficile	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0.000635901
Clostridium_baratii	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Clostridium_perfringens	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Comamonas_testosteroni	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Conexibacter_woesei	0	0	0	0	0	0	0	0	5.78E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_argentoratense	0	0	0	0	0	0	0	0	0.000103497	0	0	0	0	0	0	0	0	0
Corynebacterium_atypicum	0	0	0	0	0	0	0	0	0.000146821	0	0	0	0	0	0	0	0	0
Corynebacterium_aurimucosum	0.000892061	0	0	0	0	0	0	0	0.001191416	0	0	0	0	0	0	0	0	0
Corynebacterium_callunae	0	0	0	0	0	0	0	0	6.98E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_camporealensis	0.001393845	0	0	0	0	0	0	0	0.000789464	0	0	0	0	0	0	0	0	0
Corynebacterium_casei	0	0	0	0	0	0	0	0	0.000269573	0	0	0	0	0	0	0	0	0
Corynebacterium_deserti	0	0	0	0	0	0	0	0	5.78E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_diphtheriae	0	0	0	0	0	0	0	0	0.000628201	0	0	0	0	0	0	0	0	0
Corynebacterium_doosanense	0	0	0	0	0	0	0	0	0.000161262	0	0	0	0	0	0	0	0	0
Corynebacterium_efficiens	0	0	0	0	0	0	0	0	0.000120345	0	0	0	0	0	0	0	0	0
Corynebacterium_epidermidicanis	0	0	0	0	0	0	0	0	0.000113124	0	0	0	0	0	0	0	0	0
Corynebacterium_falsenii	0	0	0	0	0	0	0	0	0.000125159	0	0	0	0	0	0	0	0	0
Corynebacterium_glutamicum	0	0	0	0	0	0	0	0	0.000888146	0	0	0	0	0	0	0	0	0
Corynebacterium_glyciniphilum	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_halotolerans	0	0	0	0	0	0	0	0	0.000149228	0	0	0	0	0	0	0	0	0
Corynebacterium_humireducens	0	0	0	0	0	0	0	0	0.000156449	0	0	0	0	0	0	0	0	0
Corynebacterium_imitans	0	0	0	0	0	0	0	0	0.000250318	0	0	0	0	0	0	0	0	0
Corynebacterium_jeikeium	0	0	0	0	0	0	0	0	0.000264759	0	0	0	0	0	0	0	0	0
Corynebacterium_kroppenstedtii	0	0	0	0	0	0	0	0	0.000298456	0	0	0	0	0	0	0	0	0
Corynebacterium_kutscheri	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_lactis	0	0	0	0	0	0	0	0	0.000182924	0	0	0	0	0	0	0	0	0
Corynebacterium_marinum	0	0	0	0	0	0	0	0	0.000125159	0	0	0	0	0	0	0	0	0
Corynebacterium_maris	0	0	0	0	0	0	0	0	0.000129973	0	0	0	0	0	0	0	0	0
Corynebacterium_mustelae	0	0	0	0	0	0	0	0	6.26E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_pseudotuberculosis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_resistens	0	0	0	0	0	0	0	0	0.000235876	0	0	0	0	0	0	0	0	0
Corynebacterium_simulans	0.002230152	0	0	0	0	0	0	0	0.001412851	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Corynebacterium_singulare	0.001115076	0	0	0	0	0	0	0	0.000830381	0	0	0	0	0	0	0	0	0
Corynebacterium_sp._ATCC_6931	0.000669045	0	0	0	0	0	0	0	0.000707629	0	0	0	0	0	0	0	0	0
Corynebacterium_sp._JZ16	0	0	0	0	0	0	0	0	6.98E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_stationis	0	0	0	0	0	0	0	0	0.00024069	0	0	0	0	0	0	0	0	0
Corynebacterium_terpenotabidum	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_testudinoris	0	0	0	0	0	0	0	0	0.000182924	0	0	0	0	0	0	0	0	0
Corynebacterium_ulcerans	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_urealyticum	0	0	0	0	0	0	0	0	0.000158855	0	0	0	0	0	0	0	0	0
Corynebacterium_ureicelerivorans	0.002676182	0	0	0	0	0	0	0	0.002556129	0	0	0.002457757	0	0	0	0	0	0
Corynebacterium_uterequi	0	0	0	0	0	0	0	0	0.00010109	0	0	0	0	0	0	0	0	0
Corynebacterium_variabile	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Corynebacterium_vitaeruminis	0	0	0	0	0	0	0	0	0.000305676	0	0	0	0	0	0	0	0	0
Cryobacterium_arcticum	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Cupriavidus_basilensis	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Cupriavidus_gilardii	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Curtobacterium_sp._MR_MD2014	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Deinococcus_actinosclerus	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Deinococcus_gobiensis	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Deinococcus_proteolyticus	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Deinococcus_radiodurans	0	0	0.01423978	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Deinococcus_swuensis	0	0	0.001148369	0	0	0	0	0	7.22E-05	0	0	0	0	0	0	0	0	0
Delftia_acidovorans	0	0	0	0	0	0	0	0	7.46E-05	0	0	0	0	0	0	0	0	0
Delftia_sp._Cs1-4	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Dermabacter_vaginalis	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0
Dermacoccus_nishinomiyaensis	0	0	0.002870923	0	0	0	0	0	9.39E-05	0	0	0	0	0	0	0	0	0
Draconibacterium_orientale	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Enterobacter_cloacae	0	0	0	0	0	0	0	0	0.000194959	0	0	0	0	0	0	0	0.00877193	0
Enterococcus_faecalis	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Erwinia_billingiae	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0
Erwinia_gerundensis	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Escherichia_coli	0	0	0	0	0	0	0	0	0.000255132	0.105113994	0	0	0	0	0	0	0	0
Filifactor_alocis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Finegoldia_magna	0.001393845	0	0	0	0.007151371	0.005611808	0	0	0.000996457	0	0	0.016897081	0	0	0	0	0.00877193	0.028100775
Flavobacterium_johnsoniae	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Flavobacterium_psychrophilum	0	0	0	0	0	0	0	0	0	0	0	0	0.005817819	0.003612282	0	0.003193188	0	0
Frankia_sp._EAN1pec	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Frankia_sp._Eul1c	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Frondihabitans_sp._PAMC_28766	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Fusobacterium_hwasookii	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Fusobacterium_nucleatum	0	0	0	0	0	0	0	0	0.000262352	0	0	0	0	0	0	0	0	0
Gammapapillomavirus_15	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Gardnerella_vaginalis	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Gemella_sp._oral_taxon_928	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Geodermatophilus_obscurus	0	0	0	0	0	0	0	0	7.22E-05	0	0	0	0	0	0	0	0	0
Gluconacetobacter_diazotrophicus	0	0	0	0	0	0	0	0	0	0.004723517	0	0	0	0	0	0	0	0
Glutamicibacter_arilaitensis	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Gordonia_bronchialis	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Gordonia_polyisoprenivorans	0	0	0	0	0	0	0	0	0.000108311	0	0	0	0	0	0	0	0	0
Gordonia_sp._KTR9	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Gordonia_sp._QH-11	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Gordonia_terrae	0	0	0	0	0	0	0	0	0.000105904	0	0	0	0	0	0	0	0	0
Haemophilus_influenzae	0	0	0	0	0	0	0	0	0.000166076	0	0	0	0	0	0	0	0	0
Haemophilus_parainfluenzae	0	0	0	0	0	0	0	0	0.00045009	0	0	0	0	0	0	0	0	0
Hafnia_alvei	0.001672614	0	0	0	0	0	0	0	0.001388782	0	0	0	0	0	0	0	0	0
Herbaspirillum_seropedicae	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Histophilus_somni	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.001604893
Human_herpesvirus_7	0	0	0	0	0	0	0	0	0	0	0	0.025192012	0	0	0	0	0	0
Human_papillomavirus_type_134	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0
Human_papillomavirus_type_201	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Hydrogenophaga_sp._RAC07	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Hymenobacter_sp._APR13	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Hymenobacter_sp._DG5B	0	0	0	0	0	0	0	0	9.63E-05	0	0	0	0	0	0	0	0	0
Hymenobacter_sp._PAMC_26554	0	0	0	0	0	0	0	0	0.000322525	0	0	0	0	0	0	0	0	0
Hymenobacter_sp._PAMC_26628	0	0	0	0	0	0	0	0	0.000166076	0	0	0	0	0	0	0	0	0
Hymenobacter_swuensis	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Intrasporangium_calvum	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Isoptericola_dokdonensis	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Isoptericola_variabilis	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Jonesia_denitrificans	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Kibdelosporangium_phytohabitans	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Kineococcus_radiotolerans	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Klebsiella_oxytoca	0	0	0	0	0	0	0	0	6.98E-05	0	0	0	0.056017287	0.018061409	0.055014812	0	0.605263158	0.021226986
Kocuria_flava	0	0	0	0	0	0	0	0	0.000125159	0	0	0	0	0	0	0	0	0
Kocuria_palustris	0.001393845	0	0	0	0	0	0	0	0.001201044	0	0	0	0	0	0	0	0	0
Kocuria_rhizophila	0.003568243	0	0	0	0	0	0	0	0.003198772	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Kribbella_flavida	0	0	0	0	0	0	0	0	8.66E-05	0	0	0	0	0	0	0	0	0
Kurthia_sp._11kri321	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Kutzneria_albida	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Kytococcus_sedentarius	0	0	0	0	0	0	0	0	0.000115531	0	0	0	0	0	0	0	0	0
Lactobacillus_acidophilus	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0
Lactobacillus_brevis	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Lactobacillus_curvatus	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Lactobacillus_delbrueckii	0.000613292	0	0.002641249	0	0	0	0	0	0.000385104	0	0	0	0	0	0	0	0	0
Lactobacillus_johnsonii	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Lactobacillus_plantarum	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Lactobacillus_sakei	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Lactococcus_lactis	0	0	0	0	0	0	0	0	0.000356221	0	0	0	0	0	0	0	0	0
Lactococcus_piscium	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Lawsonella_clevelandensis	0.002899197	0	0	0	0.00476758	0	0	0	0.002698136	0	0.013	0.004915515	0	0	0	0	0	0.000787306
Leifsonia_xyli	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Lentzea_sp._DHS_C013	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Leptothrix_cholodnii	0	0	0	0	0	0	0	0	7.22E-05	0	0	0	0	0	0	0	0	0
Leptotrichia_buccalis	0	0	0	0	0	0	0	0	9.39E-05	0	0	0	0	0	0	0	0	0
Leptotrichia_sp._oral_taxon_847	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Leuconostoc_citreum	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Leuconostoc_gelidum	0.000557538	0	0	0	0	0	0	0	0.000421208	0	0	0	0	0	0	0	0	0
Leuconostoc_mesenteroides	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Limnohabitans_sp._63ED37-2	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Listeria_monocytogenes	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Luteipulveratus_mongoliensis	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Macrococcus_caseolyticus	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Magnetospirillum_magneticum	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Massilia_sp._NR_4-1	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Massilia_sp._WG5	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Mesorhizobium_loti	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Methanobacterium_formicicum	0.064618644	0.68678283	0.578433624	0.588033655	0.110846246	0.023920497	0.750120831	0.728595179	0.040746428	0.357223832	0.204	0.311674347	0.558676862	0.69897652	0.45789251	0.451303885	0.127819549	0.007237161
Methylibium_petroleiphilum	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Methylobacterium_aquaticum	0	0	0	0	0	0	0	0	0.000137193	0	0	0	0	0	0	0	0	0
Methylobacterium_extorquens	0	0	0	0	0	0	0	0	0.000486194	0	0	0	0	0	0	0	0	0
Methylobacterium_nodulans	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Methylobacterium_oryzae	0	0	0	0	0	0	0	0	0.000122752	0	0	0	0	0	0	0	0	0
Methylobacterium_populi	0	0	0	0	0	0	0	0	0.000103497	0	0	0	0	0	0	0	0	0
Methylobacterium_radiotolerans	0	0	0	0	0	0	0	0	0.00017089	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Methylobacterium_sp._4-46	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Methylobacterium_sp._AMS5	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Methylocella_silvestris	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Microbacterium_chocolatum	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Microbacterium_sp._CGR1	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Microbacterium_sp._No._7	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Microbacterium_sp._PAMC_28756	0	0	0	0	0	0	0	0	8.66E-05	0	0	0	0	0	0	0	0	0
Microbacterium_sp._XT11	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Microbacterium_testaceum	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Micrococcus_luteus	0.002285905	0	0.006086357	0	0	0	0	0	0.003035102	0	0	0.001536098	0	0	0	0	0	0
Microlunatus_phosphovorus	0	0	0	0	0	0	0	0	9.63E-05	0	0	0	0	0	0	0	0	0
Microterricola_viridarii	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Mitsuaria_sp._7	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Modestobacter_marinus	0	0	0	0	0	0	0	0	0.000103497	0	0	0	0	0	0	0	0	0
Moraxella_osloensis	0.000613292	0	0.003789619	0	0	0	0	0	0.000397139	0	0	0	0	0	0	0	0	0
Mucilaginibacter_sp._PAMC_26640	0	0	0.001607717	0	0	0	0	0	0.000173297	0	0	0	0	0	0	0	0	0
Myceliophthora_thermophila	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_abscessus	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_avium	0	0	0	0	0	0	0	0	0.000257538	0	0	0	0	0	0	0	0	0
Mycobacterium_chubuense	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_gilvum	0	0	0	0	0	0	0	0	9.63E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_goodii	0	0	0	0	0	0	0	0	0.000182924	0	0	0	0	0	0	0	0	0
Mycobacterium_kansasii	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_phlei	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_sinense	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_smegmatis	0	0	0	0	0	0	0	0	0.000144414	0	0	0	0	0	0	0	0	0
Mycobacterium_sp._EPa45	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_sp._JS623	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_sp._YC-RL4	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_sp._djl-10	0	0	0	0	0	0	0	0	5.54E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_vaccae	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Mycobacterium_vanbaalenii	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Mycoplasma_mycoides	0	0	0	0	0	0	0	0	0	0.327812067	0	0	0	0	0	0	0	0
Nakamurella_multipartita	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Neisseria_elongata	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Neisseria_meningitidis	0	0	0	0	0	0	0	0	0.000113124	0	0	0	0	0	0	0	0	0
Neorhizobium_galegae	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Nocardia_brasiliensis	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Nocardia_cyriacigeorgica	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Nocardia_farcinica	0	0	0	0	0	0	0	0	0.000137193	0	0	0	0	0	0	0	0	0
Nocardioides_dokdonensis	0	0	0	0	0	0	0	0	0.000149228	0	0	0	0	0	0	0	0	0
Nocardioides_sp._JS614	0	0	0	0	0	0	0	0	0.000187738	0	0	0	0	0	0	0	0	0
Nocardiopsis_alba	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Nocardiopsis_dassonvillei	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Nostoc_punctiforme	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Novosphingobium_aromatici-vorans	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Olsenella_sp._oral_taxon_807	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Oscillatoria_nigro-viridis	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Pantoea_agglomerans	0	0	0	0	0	0	0	0	7.22E-05	0	0	0	0	0	0	0	0	0
Paraburkholderia_fungorum	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Paracoccus_denitrificans	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Parvimonas_micra	0	0	0	0	0	0	0	0	5.78E-05	0	0	0	0	0	0	0	0	0
Paucibacter_sp._KCTC_42545	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Peptoniphilus_sp._1-1	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0.001877422
Phenylobacterium_zucineum	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Photorhabdus_asymbiotica	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000333091
Pimelobacter_simplex	0	0	0	0	0	0	0	0	0.000144414	0	0	0	0	0	0	0	0	0
Polaromonas_sp._JS666	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Prevotella_dentalis	0	0	0	0	0	0	0	0	0.000103497	0	0	0	0	0	0	0	0	0.003997093
Prevotella_denticola	0	0	0	0	0	0	0	0	0.000127566	0	0	0	0	0	0	0	0	0.000635901
Prevotella_enoeca	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0.001574612
Prevotella_fusca	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0.000817587
Prevotella_intermedia	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Prevotella_melaninogenica	0.000780553	0	0	0	0.00476758	0	0	0	0.000459718	0	0	0	0	0	0	0	0	0.001332364
Prevotella_sp._oral_taxon_299	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0.001090116
Propionibacterium_acidipropionici	0	0	0	0	0	0	0	0	0.000317711	0	0	0	0	0	0	0	0	0
Propionibacterium_acnes	0.778880464	0.107097323	0.180064309	0.145839825	0.222089789	0.002745634	0.05219913	0.099334996	0.786611372	0.016500819	0.55	0.311059908	0.144946809	0.074954846	0.052898857	0.09473124	0.095864662	0.003875969
Propionibacterium_avidum	0.002230152	0	0	0	0.007548669	0	0	0	0.002645184	0	0	0	0	0	0	0	0	0
Propionibacterium_freudenreichii	0	0	0	0	0	0	0	0	0.000108311	0	0	0	0	0	0	0	0	0
Propionibacterium_phage_ ATCC29399B_C	0	0	0	0	0	0	0	0	9.87E-05	0	0	0	0	0	0	0	0	0.002513324
Propionibacterium_phage_ ATCC29399B_T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.003330911
Propionibacterium_phage_Attacne	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0.011809593
Propionibacterium_phage_Kubed	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0.025042393
Propionibacterium_phage_Lau-chelly	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.002149952

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Propionibacterium_phage_MrAK	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0.006570979
Propionibacterium_phage_ Ouroboros	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0.001665455
Propionibacterium_phage_P1.1	0	0	0	0	0	0.000495553	0	0	8.42E-05	0	0	0	0	0	0	0	0	0.009356831
Propionibacterium_phage_P100D	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0.006328731
Propionibacterium_phage_P100_1	0	0	0	0	0	0	0	0	0.000358628	0	0	0	0	0	0	0	0	0.031159157
Propionibacterium_phage_P100_A	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0.006934351
Propionibacterium_phage_P101A	0	0	0	0	0	0	0	0	0.000117938	0	0	0	0	0	0	0	0	0.026011386
Propionibacterium_phage_P104A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.003966812
Propionibacterium_phage_P105	0	0	0	0	0	0	0	0	0.000149228	0	0	0	0	0	0	0	0	0.004360465
Propionibacterium_phage_P14.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.034156977
Propionibacterium_phage_P9.1	0	0	0	0	0	0	0	0	0.000685967	0	0	0	0	0	0	0	0	0.012718023
Propionibacterium_phage_PA1-14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.013989826
Propionibacterium_phage_PA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.017411579
Propionibacterium_phage_PAC1	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0.005511143
Propionibacterium_phage_PAD20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.008327277
Propionibacterium_phage_PAS50	0	0	0	0	0	0	0	0	0.00010109	0	0	0	0	0	0	0	0	0.009780766
Propionibacterium_phage_PHL009	0	0	0	0	0	0	0	0	0.000163669	0	0	0	0	0	0	0	0	0.006994913
Propionibacterium_phage_ PHL010M04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.006237888
Propionibacterium_phage_PHL025	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0.028918362
Propionibacterium_phage_PHL030	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.004118217
Propionibacterium_phage_PHL041	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0.003694283
Propionibacterium_phage_PHL055	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0.019864341
Propionibacterium_phage_ PHL060L00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.006510417
Propionibacterium_phage_ PHL067M10	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0.010446948
Propionibacterium_phage_PHL070	0	0	0	0	0	0	0	0	5.54E-05	0	0	0	0	0	0	0	0	0.007539971
Propionibacterium_phage_ PHL071N05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.006964632
Propionibacterium_phage_PHL082	0	0	0	0	0	0	0	0	0.00010109	0	0	0	0	0	0	0	0	0.018471415
Propionibacterium_phage_PHL085	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.011022287
Propionibacterium_phage_PHL092	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.018865068
Propionibacterium_phage_PHL095	0	0	0	0	0	0	0	0	0.000368256	0	0	0	0	0	0	0	0	0.020106589
Propionibacterium_phage_ PHL111M01	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0.011537064
Propionibacterium_phage_ PHL112N00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.009084302
Propionibacterium_phage_ PHL113M01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.011264535

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Propionibacterium_phage_PHL114L00	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0.007297723
Propionibacterium_phage_PHL116	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0.010992006
Propionibacterium_phage_PHL132	0	0	0	0	0	0.000589307	0	0	6.02E-05	0	0	0	0	0	0	0	0	0.016382025
Propionibacterium_phage_PHL141	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.002967539
Propionibacterium_phage_PHL150	0	0	0	0	0	0	0	0	7.46E-05	0	0	0	0	0	0	0	0	0.008054748
Propionibacterium_phage_PHL152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.003452035
Propionibacterium_phage_PHL171	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.005965359
Propionibacterium_phage_PHL179	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.008751211
Propionibacterium_phage_PHL199	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.006359012
Propionibacterium_phage_PHL301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.004057655
Propionibacterium_phage_Pacnes_2012-15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.022135417
Propionibacterium_phage_Pirate	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0.004784399
Propionibacterium_phage_Procrass1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.006782946
Propionibacterium_phage_SKKY	0	0	0	0	0	0	0	0	0.000120345	0	0	0	0	0	0	0	0	0.020984738
Propionibacterium_phage_Solid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.010961725
Propionibacterium_phage_Stormborn	0	0	0	0	0	0	0	0	5.54E-05	0	0	0	0	0	0	0	0	0.004148498
Propionibacterium_phage_Wizzo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.039032219
Propionibacterium_propionicum	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Propionibacterium_sp._oral_taxon_193	0.003456735	0	0	0	0	0	0	0	0.002979744	0	0	0.001843318	0	0	0	0	0	0
Proteus_mirabilis	0	0	0	0	0	0	0	0	0.000317711	0	0	0	0	0	0	0	0	0
Pseudarthrobacter_chlorophenolicus	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Pseudarthrobacter_sulfonivorans	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_aeruginosa	0	0	0	0	0	0	0	0	0.000115531	0	0	0	0	0	0	0	0	0
Pseudomonas_antarctica	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_azotoformans	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_chlororaphis	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_fluorescens	0.002007136	0	0	0	0	0	0	0	0.001280471	0	0	0	0	0	0	0	0	0
Pseudomonas_fragi	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_koreensis	0	0	0	0	0	0	0	0	0.000137193	0	0	0	0	0	0	0	0	0
Pseudomonas_mendocina	0	0	0	0	0	0	0	0	0.000178111	0	0	0	0	0	0	0	0	0
Pseudomonas_oryzihabitans	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_poae	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_pseudoalcaligenes	0	0	0.002756086	0	0.004370282	0	0	0	0.000298456	0	0	0	0	0	0	0	0	0
Pseudomonas_putida	0	0	0	0	0	0	0	0	0.000375477	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Pseudomonas_sp._TKP	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_sp._URMO17WK12:I11	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_stutzeri	0	0	0	0	0	0	0	0	0.000125159	0	0	0	0	0	0	0	0	0
Pseudomonas_syringae	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Pseudomonas_trivialis	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Pseudonocardia_dioxanivorans	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Pseudonocardia_sp._AL041005-10	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Pseudonocardia_sp._HH130629-09	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Pseudonocardia_sp._HH130630-07	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Pseudoxanthomonas_suwonensis	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Psychrobacter_alimentarius	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Ralstonia_insidiosa	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Ralstonia_mannitolilytica	0	0	0	0	0	0	0	0	5.05E-05	0	0	0	0	0	0	0	0	0
Ralstonia_pickettii	0	0	0.003330271	0	0	0	0	0	0.000284014	0	0	0	0	0	0	0	0	0
Ralstonia_solanacearum	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Ramlibacter_tataouinensis	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Raoultella_ornithinolytica	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Rathayibacter_tritici	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Rhizobium_leguminosarum	0	0	0	0	0	0	0	0	7.70E-05	0	0	0	0	0	0	0	0	0
Rhodobacter_sphaeroides	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_equi	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_erythropolis	0	0	0	0	0	0	0	0	0.000185331	0	0	0	0	0	0	0	0	0
Rhodococcus_fascians	0	0	0	0	0	0	0	0	0.000149228	0	0	0	0	0	0	0	0	0
Rhodococcus_opacus	0	0	0	0	0	0	0	0	9.63E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_sp._008	0	0	0	0	0	0	0	0	7.46E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_sp._B7740	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_sp._PBTS1	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Rhodococcus_sp._PBTS2	0	0	0	0	0	0	0	0	0.000175704	0	0	0	0	0	0	0	0	0
Rhodopseudomonas_palustris	0	0	0	0	0	0	0	0	0.0001396	0	0	0	0	0	0	0	0	0
Rothia_dentocariosa	0.002230152	0	0	0	0	0	0	0	0.001595775	0	0	0	0	0	0	0	0	0
Rothia_mucilaginosa	0	0	0	0	0	0	0	0	0.000394732	0	0	0	0	0	0	0	0	0
Rubrivivax_gelatinosus	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Saccharomonospora_viridis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Saccharomyces_cerevisiae	0	0	0	0	0	0	0	0	0.00020218	0	0	0	0	0	0	0	0	0
Saccharopolyspora_erythraea	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Saccharothrix_espanaensis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Salmonella_enterica	0.033340767	0.180195495	0.164331649	0.247740729	0.050854191	0.009174435	0.188013533	0.159600998	0.02333972	0.179178738	0.195	0.077880184	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Sanguibacter_keddieii	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Selenomonas_sp._oral_taxon_136	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Selenomonas_sp._oral_taxon_478	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Selenomonas_sputigena	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Serinicoccus_sp._JLT9	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Serratia_liquefaciens	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
Serratia_marcescens	0	0	0	0	0	0	0	0	0.000368256	0	0	0	0	0	0	0	0	0
Shinella_sp._HZN7	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0
Sinomonas_atrocyanea	0	0	0	0	0	0	0	0	5.30E-05	0	0	0	0	0	0	0	0	0
Sinorhizobium_sp._RAC02	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Sphingobacterium_sp._ML3W	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Sphingobium_baderi	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Sphingobium_sp._TKS	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Sphingomonas_hengshuiensis	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Sphingomonas_panacis	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Sphingomonas_sanxanigenens	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Sphingomonas_sp._MM-1	0	0	0	0	0	0	0	0	5.54E-05	0	0	0	0	0	0	0	0	0
Sphingomonas_sp._NIC1	0	0	0	0	0	0	0	0	0.000127566	0	0	0	0	0	0	0	0	0
Sphingomonas_taxi	0	0	0	0	0	0	0	0	0.000142007	0	0	0	0	0	0	0	0	0
Sphingomonas_wittichii	0	0	0	0	0	0	0	0	6.26E-05	0	0	0	0	0	0	0	0	0
Sphingopyxis_alaskensis	0	0	0	0	0	0	0	0	9.63E-05	0	0	0	0	0	0	0	0	0
Sphingopyxis_fribergensis	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Sphingopyxis_granuli	0	0	0	0	0	0	0	0	6.50E-05	0	0	0	0	0	0	0	0	0
Sphingopyxis_macrogoltabida	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Sphingopyxis_terrae	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Spirosoma_radiotolerans	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Stackebrandtia_nassauensis	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_agnetis	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_argenteus	0	0	0	0	0	0	0	0	0.000127566	0	0	0	0	0	0	0	0	0
Staphylococcus_aureus	0.007638269	0.012749681	0.008268259	0	0.54390147	0.951435766	0	0	0.001814803	0	0	0.19078341	0.094913564	0.083684527	0.361404994	0.373336881	0.102756892	0.230044816
Staphylococcus_capitis	0.015722569	0	0.003330271	0.003739483	0.004370282	0	0	0	0.015153849	0	0	0.006758833	0.002659574	0	0	0	0	0
Staphylococcus_carnosus	0	0	0	0	0	0	0	0	7.94E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_condimenti	0	0	0	0	0	0	0	0	3.61E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_epidermidis	0.041369313	0.004249894	0.009186955	0.014646307	0.017481128	0.000669667	0	0.004987531	0.038808873	0	0.038	0.037634409	0.013630319	0.004214329	0	0.007450772	0.006892231	0.007085756
Staphylococcus_equorum	0	0	0	0	0	0	0	0	0.000315304	0	0	0	0	0	0	0	0	0
Staphylococcus_haemolyticus	0.001393845	0	0	0	0	0	0	0	0.000832788	0	0	0	0	0	0	0	0	0.000666182
Staphylococcus_hyicus	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_lugdunensis	0	0	0	0	0	0	0	0	0.000226249	0	0	0	0	0	0	0	0	0

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Staphylococcus_pasteuri	0.001003568	0	0	0	0	0	0	0	0.000965167	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_69	0	0	0	0	0	0.00064288	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_6ec	0	0	0	0	0	0	0	0	0.000375477	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_77	0	0	0	0	0	0.001540234	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_85	0	0	0	0	0	0.000991107	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_Ipla7	0	0	0	0	0	0	0	0	0.000257538	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_Ipla88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.005208333
Staphylococcus_phage_StB20	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_StB20-like	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_StB27	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_ StauST398-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000847868
Staphylococcus_phage_X2	0	0	0	0	0	0.000575913	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_phiETA	0	0	0	0	0	0.000616093	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_phiETA2	0	0	0	0	0	0.000991107	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcus_phage_vB_SepS_ SEP9	0	0	0	0	0	0	0	0	0.000226249	0	0	0	0	0	0	0	0	0
Staphylococcus_pseudintermedius	0	0	0	0	0	0	0	0	9.15E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_saprophyticus	0.000557538	0	0	0	0	0	0	0	0.000464532	0	0	0	0	0	0	0	0	0
Staphylococcus_schleiferi	0	0	0	0	0	0	0	0	9.39E-05	0	0	0	0	0	0	0	0	0
Staphylococcus_simulans	0	0	0	0	0	0	0	0	0.00013238	0	0	0	0	0	0	0	0	0
Staphylococcus_warneri	0.005129349	0	0	0	0	0	0	0	0.005177244	0	0	0.004147465	0	0	0	0	0	0
Staphylococcus_xylosus	0	0	0	0	0	0	0	0	0.000166076	0	0	0	0	0	0	0	0	0
Stenotrophomonas_acidaminiphila	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Stenotrophomonas_maltophilia	0	0	0	0	0	0	0	0	0.00013238	0	0	0	0	0	0	0	0	0
Stenotrophomonas_rhizophila	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Streptococcus_agalactiae	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0.000545058
Streptococcus_anginosus	0	0	0	0	0	0	0	0	0.000142007	0	0	0	0	0	0	0	0	0
Streptococcus_cristatus	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Streptococcus_equi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000454215
Streptococcus_gordonii	0.000613292	0	0	0	0	0	0	0	0.000633015	0	0	0	0	0	0	0	0	0
Streptococcus_infantarius	0	0	0	0	0	0	0	0	6.26E-05	0	0	0	0	0	0	0	0	0
Streptococcus_intermedius	0	0	0	0	0	0	0	0	0.000259945	0	0	0	0	0	0	0	0	0
Streptococcus_mitis	0.00278769	0	0	0	0.009137863	0	0	0	0.002202314	0	0	0	0	0	0	0	0	0
Streptococcus_mutans	0.000669045	0	0	0	0	0	0	0	0.000517484	0	0	0	0	0	0	0	0	0
Streptococcus_oralis	0.001003568	0	0	0	0	0	0	0	0.000645049	0	0	0	0	0	0	0	0	0
Streptococcus_parasanguinis	0.001003568	0	0	0	0	0	0	0	0.000777429	0	0	0	0	0	0	0	0	0
Streptococcus_pneumoniae	0.000780553	0	0	0	0	0	0	0	0.001034967	0	0	0.002918587	0	0	0	0	0	0.000514777

Bacteria	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Streptococcus_pseudopneumoniae	0	0	0	0	0	0	0	0	0.000339373	0	0	0	0	0	0	0	0	0.001968266
Streptococcus_pyogenes	0	0	0	0	0	0	0	0	6.74E-05	0	0	0	0	0	0	0	0	0.001029554
Streptococcus_salivarius	0.002174398	0	0	0	0.007945967	0	0	0	0.001971252	0	0	0	0	0	0	0	0	0
Streptococcus_sanguinis	0	0	0	0	0	0	0	0	0.000356221	0	0	0	0	0	0	0	0	0
Streptococcus_sp._A12	0	0	0	0	0	0	0	0	0.000103497	0	0	0	0	0	0	0	0	0
Streptococcus_sp._I-G2	0	0	0	0	0	0	0	0	4.81E-05	0	0	0	0	0	0	0	0	0
Streptococcus_sp._I-P16	0	0	0	0	0	0	0	0	6.02E-05	0	0	0	0	0	0	0	0	0
Streptococcus_sp._VT_162	0.000669045	0	0	0	0.00476758	0	0	0	0.00075336	0	0	0	0	0	0	0	0	0
Streptococcus_sp._oral_taxon_431	0	0	0	0	0	0	0	0	0.000279201	0	0	0	0	0	0	0	0	0
Streptococcus_thermophilus	0.001003568	0	0	0	0	0	0	0	0.000741326	0	0	0	0	0	0	0	0	0
Streptomyces_albus	0	0	0	0	0	0	0	0	0.000264759	0	0	0	0	0	0	0	0	0
Streptomyces_bingchenggensis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Streptomyces_cattleya	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Streptomyces_fulvissimus	0	0	0	0	0	0	0	0	3.85E-05	0	0	0	0	0	0	0	0	0
Streptomyces_sp._Mg1	0	0	0	0	0	0	0	0	2.65E-05	0	0	0	0	0	0	0	0	0
Streptomyces_sp._SirexAA-E	0	0	0	0	0	0	0	0	4.33E-05	0	0	0	0	0	0	0	0	0
Streptosporangium_roseum	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Tannerella_forsythia	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0
Thermomonospora_curvata	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Thielavia_terrestris	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Torque_teno_virus_15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000423934
Treponema_denticola	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Treponema_sp._OMZ_838	0	0	0	0	0	0	0	0	2.41E-05	0	0	0	0	0	0	0	0	0
Tsukamurella_paurometabola	0	0	0	0	0	0	0	0	3.13E-05	0	0	0	0	0	0	0	0	0
Variovorax_paradoxus	0.001003568	0.008924777	0.014354616	0	0	0	0	0	0.001612624	0.001007684	0	0.001689708	0	0	0	0.002927089	0	0.000484496
Variovorax_sp._PAMC_28711	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Veillonella_parvula	0.002397413	0	0	0	0	0	0	0	0.002539281	0	0	0	0	0	0	0	0	0.000363372
Verminephrobacter_eiseniae	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Xanthobacter_autotrophicus	0	0	0	0	0	0	0	0	3.37E-05	0	0	0	0	0	0	0	0	0
Xanthomonas_campestris	0.000780553	0	0.00149288	0	0	0	0	0	0.001148092	0	0	0	0	0	0	0	0.020050125	0
Xylanimonas_cellulosilytica	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0
[Enterobacter]_aerogenes	0	0	0	0	0	0	0	0	0.000671525	0	0	0	0	0	0	0	0	0
[Eubacterium]_rectale	0	0	0	0	0	0	0	0	4.57E-05	0	0	0	0	0	0	0	0	0

Fungi	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Thielavia_terrestris	0	0	0	0	0	0	0	0	8.18E-05	0	0	0	0	0	0	0	0	0
Saccharomyces_cerevisiae	0	0	0	0	0	0	0	0	0.00020218	0	0	0	0	0	0	0	0	0
Saccharomonospora_viridis	0	0	0	0	0	0	0	0	2.89E-05	0	0	0	0	0	0	0	0	0
Myceliophthora_thermophila	0	0	0	0	0	0	0	0	8.91E-05	0	0	0	0	0	0	0	0	0
Candida_dubliniensis	0	0	0	0	0	0	0	0	4.09E-05	0	0	0	0	0	0	0	0	0

Viruses	S1N	S1NW	S1C	S1CW	S1NB	S1NBW	S2N	S2NW	S2C	S2CW	S2NB	S2NBW	S3N	S3NW	S3C	S3CW	S3NB	S3NBW
Betapapillomavirus_1	0	0	0	0	0	0	0	0	0.000161262	0	0	0	0	0	0	0	0	0
Betapapillomavirus_2	0	0	0	0	0	0	0	0	3.61035E-05	0	0	0	0	0	0	0	0	0
Gammapapillomavirus_15	0	0	0	0	0	0	0	0	4.33242E-05	0	0	0	0	0	0	0	0	0
Human_herpesvirus_7	0	0	0	0	0	0	0	0	0	0	0	0.025192012	0	0	0	0	0	0
Human_papillomavirus_type_134	0	0	0	0	0	0	0	0	8.90553E-05	0	0	0	0	0	0	0	0	0
Human_papillomavirus_type_201	0	0	0	0	0	0	0	0	4.33242E-05	0	0	0	0	0	0	0	0	0
Torque_teno_virus_15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000423934

Figure S1A. Relative sequence abundance of Corynebacterium species and Corynebacterium-associated phage diversity recovered from metagenomic sequencing.

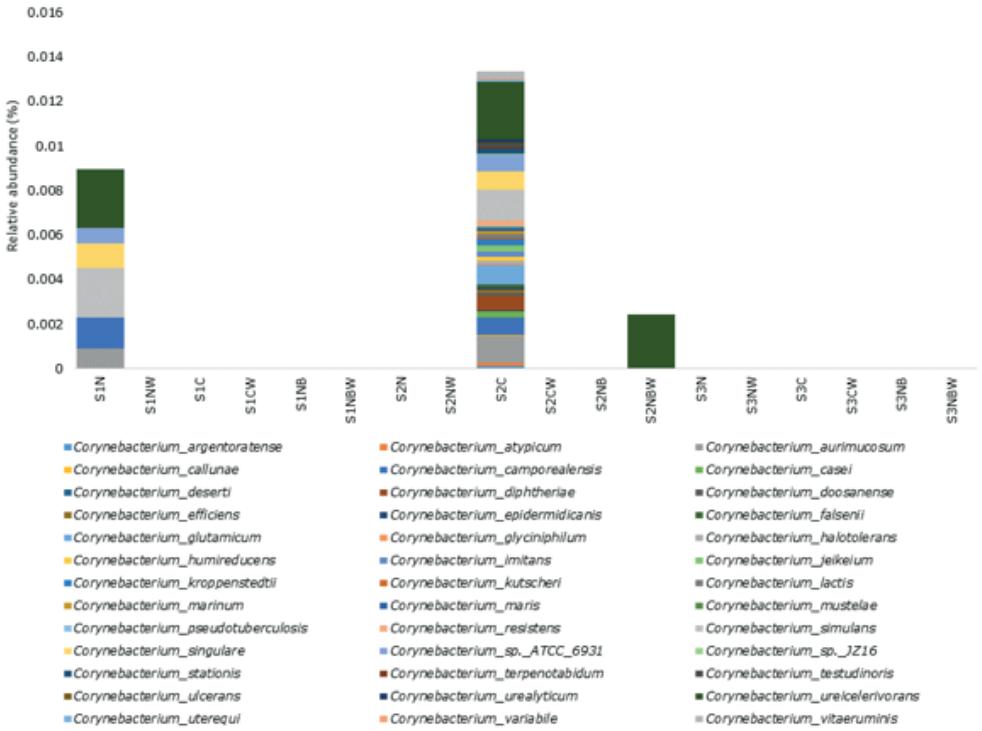

Figure S1B. Relative sequence abundance of Propionibacterium species and Propionibacterium -associated phage diversity recovered from metagenomic sequencing.

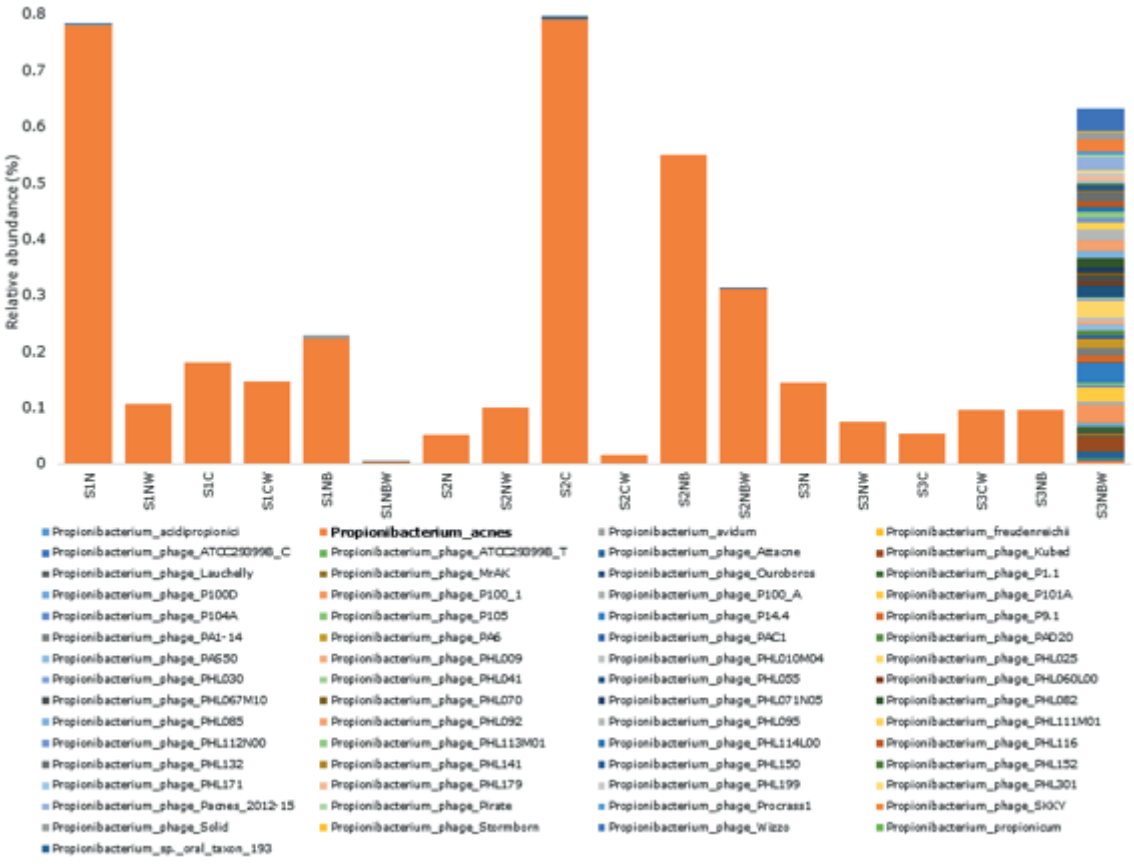

Figure S1C. Relative sequence abundance of Staphylococcus species and Staphylococcus-associated phage diversity recovered from metagenomic sequencing.

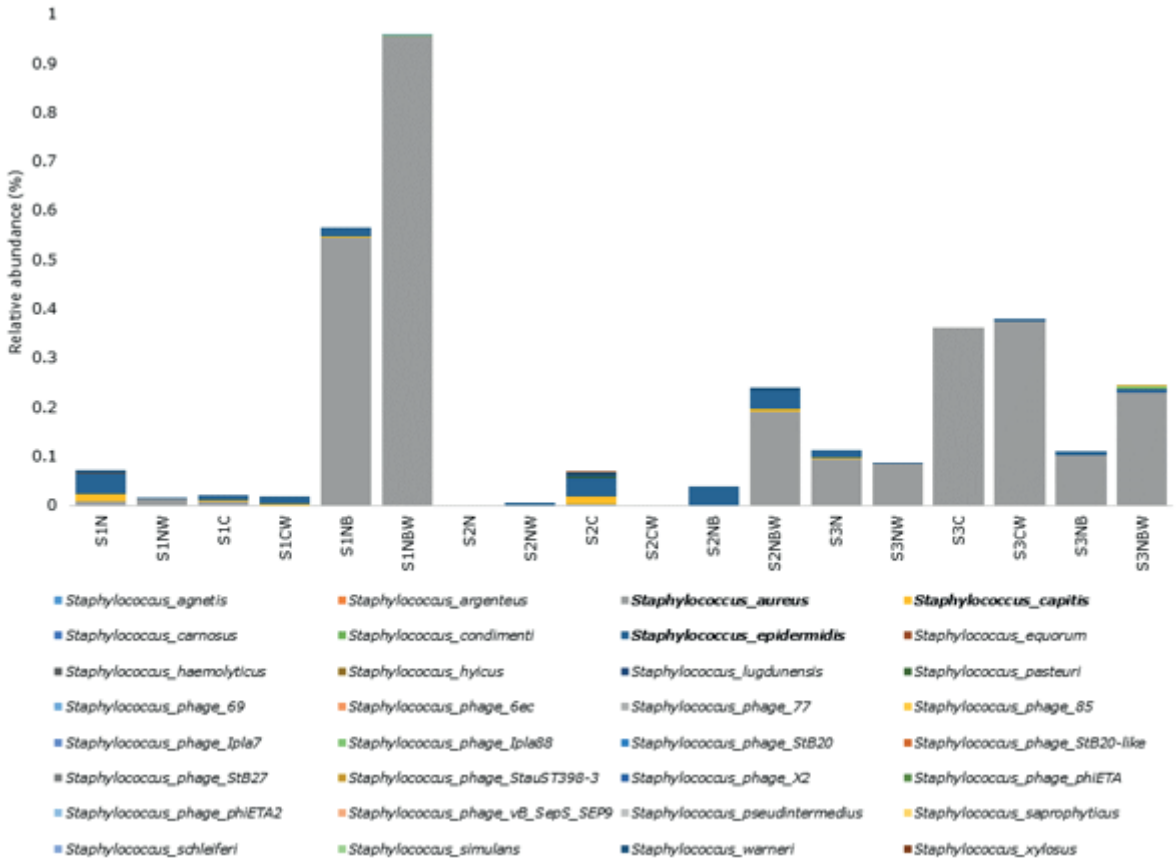

Figure S2. Phylogenomic tree for species-level identification constructed from a subset of 44 reference genomes. Phylogenetic inference was performed using FastTree v2.1.9 with the WAG+Γ model of amino acid evolution and 100 bootstrap iterations to assess node support. The refined genome bin, identified as Propionibacterium acnes is highlighted. Scale bar represents 10% sequence divergence.

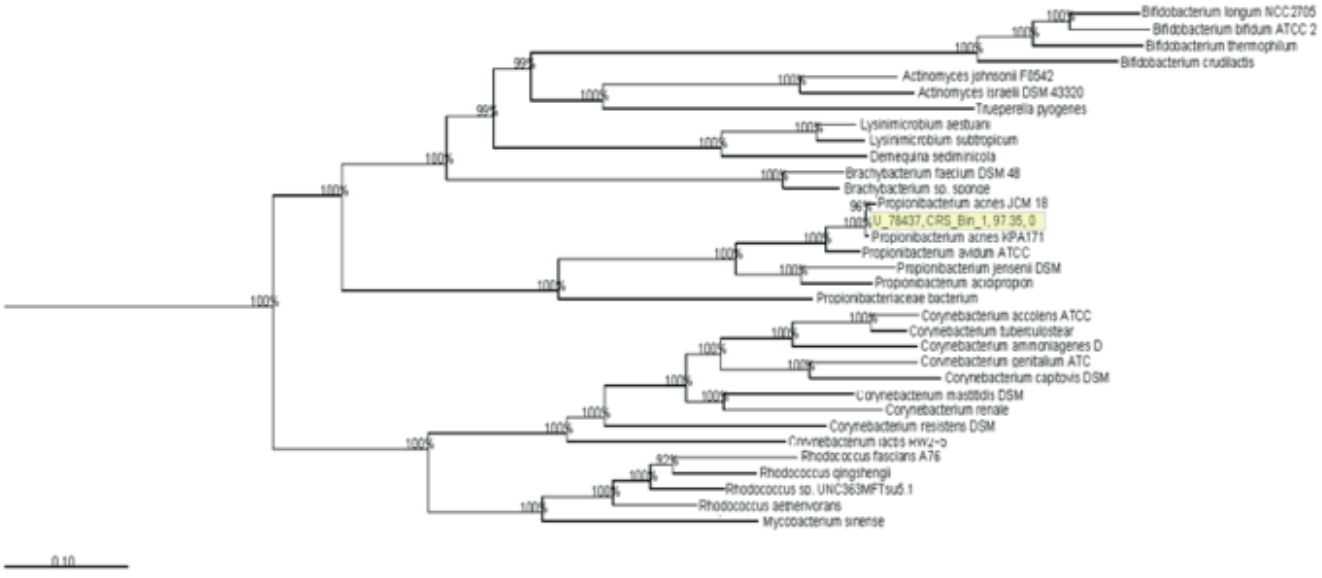

Figure S3. Assignments of gene function to overall metabolic reconstruction categories for the recovered Propionibacterium acnes genome.

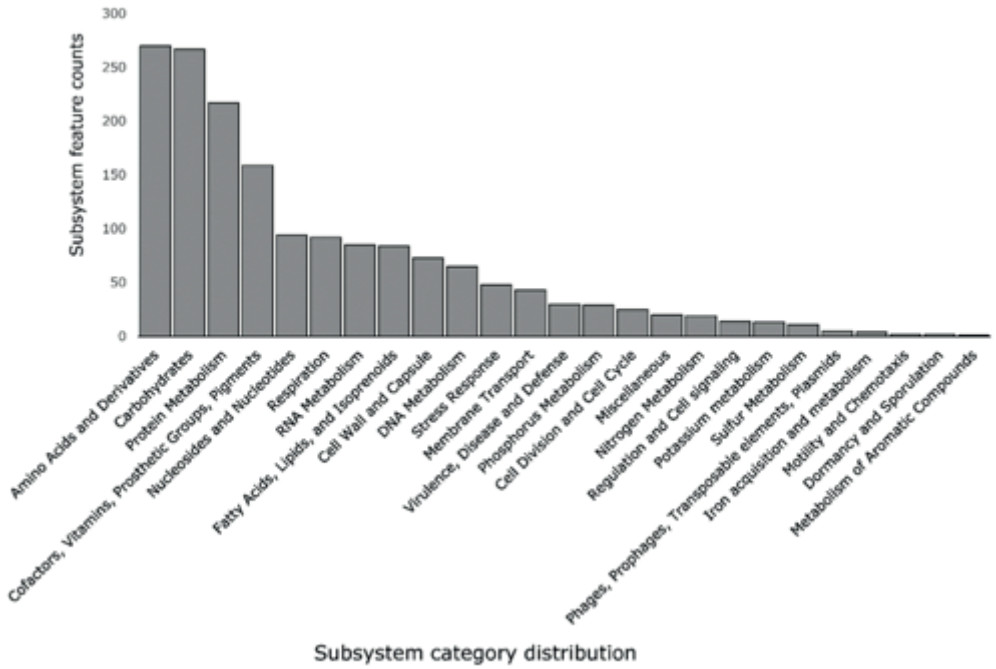

Table S3. List of differential gene pathways associated with the recovered Propionibacterium acnes genome compared with Propionibacterium acnes isolated strain KPA171202.

Category	Subcategory	Subsystem	Role	Presence in recovered *P. acnes* genome	Presence *P. acnes* KPA171202
Amino Acids and Derivatives	Alanine, serine, and glycine	Glycine and Serine Utilization	Serine transporter	yes	no
Amino Acids and Derivatives	Aromatic amino acids and derivatives	Chorismate: Intermediate for synthesis of Tryptophan, PAPA antibiotics, PABA, 3-hydroxyanthranilate and more.	Isochorismate synthase (EC 5.4.4.2) of siderophore biosynthesis	yes	no
Amino Acids and Derivatives	Branched-chain amino acids	Branched-Chain Amino Acid Biosynthesis	2-isopropylmalate synthase (EC 2.3.3.13)	yes	no
Amino Acids and Derivatives	Branched-chain amino acids	Branched-Chain Amino Acid Biosynthesis	3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)	yes	no
Amino Acids and Derivatives	Branched-chain amino acids	Branched-Chain Amino Acid Biosynthesis	3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)	yes	no
Amino Acids and Derivatives	Branched-chain amino acids	Branched-Chain Amino Acid Biosynthesis	3-isopropylmalate dehydrogenase (EC 1.1.1.85)	yes	no
Amino Acids and Derivatives	Branched-chain amino acids	Branched-Chain Amino Acid Biosynthesis	Leucine-responsive regulatory protein, regulator for leucine (or lrp) regulon and high-affinity branched-chain amino acid transport system	yes	no

Category	Subcategory	Subsystem	Role	Presence in recovered *P. acnes* genome	Presence *P. acnes* KPA171202
Amino Acids and Derivatives	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	L-asparaginase I, cytoplasmic (EC 3.5.1.1)	yes	no
Amino Acids and Derivatives	Lysine, threonine, methionine, and cysteine	Lysine Biosynthesis DAP Pathway	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)	yes	no
Carbohydrates	Aminosugars	Chitin and N-acetylglucosamine utilization	Beta-hexosaminidase (EC 3.2.1.52)	yes	no
Carbohydrates	Aminosugars	Chitin and N-acetylglucosamine utilization	N-Acetyl-D-glucosamine ABC transport system, sugar-binding protein	yes	no
Carbohydrates	Central carbohydrate metabolism	Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	Acetyl-coenzyme A synthetase (EC 6.2.1.1)	yes	no
Carbohydrates	Central carbohydrate metabolism	TCA Cycle	Malate dehydrogenase (EC 1.1.1.37)	yes	no
Carbohydrates	Di- and oligosaccharides	Beta-Glucoside Metabolism	Beta-glucosidase (EC 3.2.1.21)	yes	no
Carbohydrates	Di- and oligosaccharides	Beta-Glucoside Metabolism	Beta-glucoside bgl operon antiterminator, BglG family	yes	no
Carbohydrates	Di- and oligosaccharides	Beta-Glucoside Metabolism	PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69)	yes	no
Carbohydrates	Di- and oligosaccharides	Beta-Glucoside Metabolism	PTS system, diacetylchitobiose-specific IIB component (EC 2.7.1.69)	yes	no
Carbohydrates	Di- and oligosaccharides	Lactose and Galactose Uptake and Utilization	Galactokinase (EC 2.7.1.6)	yes	no
Carbohydrates	Di- and oligosaccharides	Maltose and Maltodextrin Utilization	Neopullulanase (EC 3.2.1.135)	yes	no
Carbohydrates	Di- and oligosaccharides	Trehalose Biosynthesis	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	yes	no
Carbohydrates	Fermentation	Acetolactate synthase subunits	Acetolactate synthase large subunit (EC 2.2.1.6)	yes	no
Carbohydrates	Fermentation	Acetolactate synthase subunits	Acetolactate synthase small subunit (EC 2.2.1.6)	yes	no
Carbohydrates	Monosaccharides	D-Tagatose and Galactitol Utilization	PTS system, galactitol-specific IIB component (EC 2.7.1.69)	yes	no
Carbohydrates	Monosaccharides	D-Tagatose and Galactitol Utilization	PTS system, galactitol-specific IIC component (EC 2.7.1.69)	yes	no
Carbohydrates	Monosaccharides	Deoxyribose and Deoxynucleoside Catabolism	Deoxyribonucleoside regulator DeoR (transcriptional repressor)	yes	no
Carbohydrates	Monosaccharides	Deoxyribose and Deoxynucleoside Catabolism	Deoxyribose-phosphate aldolase (EC 4.1.2.4)	yes	no
Carbohydrates	Monosaccharides	Deoxyribose and Deoxynucleoside Catabolism	Thymidine phosphorylase (EC 2.4.2.4)	yes	no
Carbohydrates	Monosaccharides	Mannose Metabolism	Alpha-mannosidase (EC 3.2.1.24)	yes	no
Carbohydrates	Monosaccharides	Mannose Metabolism	Beta-mannosidase (EC 3.2.1.25)	yes	no
Carbohydrates	Monosaccharides	Mannose Metabolism	Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)	yes	no
Carbohydrates	Monosaccharides	Mannose Metabolism	Mannose-6-phosphate isomerase (EC 5.3.1.8)	yes	no

Category	Subcategory	Subsystem	Role	Presence in recovered *P. acnes* genome	Presence *P. acnes* KPA171202
Carbohydrates	Monosaccharides	Mannose Metabolism	Phosphomannomutase (EC 5.4.2.8)	yes	no
Carbohydrates	Polysaccharides	Alpha-Amylase locus in Streptococcus	Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE	yes	no
Carbohydrates	Polysaccharides	Alpha-Amylase locus in Streptococcus	putative esterase	yes	no
Carbohydrates	Sugar alcohols	Glycerol and Glycerol-3-phosphate Uptake and Utilization	Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	yes	no
Carbohydrates	no subcategory	Lacto-N-Biose I and Galacto-N-Biose Metabolic Pathway	UDP-glucose 4-epimerase (EC 5.1.3.2)	yes	no
Cell Wall and Capsule	Capsular and extracellular polysacchrides	Lipid-linked oligosaccharide synthesis related cluster	Exoenzymes regulatory protein AepA in lipid-linked oligosaccharide synthesis cluster	yes	no
Cell Wall and Capsule	Capsular and extracellular polysacchrides	Sialic Acid Metabolism	Sialic acid utilization regulator, RpiR family	yes	no
Cell Wall and Capsule	no subcategory	Peptidoglycan Biosynthesis	Rare lipoprotein A precursor	yes	no
Clustering-based subsystems	Cytochrome biogenesis	CBSS-196164.1.peg.1690	cytochrome oxidase assembly protein	yes	no
Clustering-based subsystems	Isoprenoid/cell wall biosynthesis: PREDICTED UNDECAPRENYL DIPHOSPHATE PHOSPHATASE	CBSS-83331.1.peg.3039	Undecaprenyl diphosphate synthase (EC 2.5.1.31)	yes	no
Clustering-based subsystems	no subcategory	Bacterial Cell Division	Cell division protein FtsQ	yes	no
Clustering-based subsystems	no subcategory	Bacterial Cell Division	Septum formation protein Maf	yes	no
Clustering-based subsystems	no subcategory	Bacterial Cell Division	Septum site-determining protein MinD	yes	no
Clustering-based subsystems	no subcategory	CBSS-228410.1.peg.134	DNA polymerase III epsilon subunit (EC 2.7.7.7)	yes	no
Clustering-based subsystems	no subcategory	CBSS-228410.1.peg.134	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	yes	no
Clustering-based subsystems	no subcategory	CBSS-257314.1.peg.752	Adenine-specific methyltransferase (EC 2.1.1.72)	yes	no
Clustering-based subsystems	no subcategory	CBSS-469378.4.peg.430	FIG002344: Hydrolase (HAD superfamily)	yes	no
Clustering-based subsystems	no subcategory	Conserved gene cluster associated with Met-tRNA formyltransferase	Serine/threonine protein kinase PrkC, regulator of stationary phase	yes	no
Clustering-based subsystems	no subcategory	EC699-706	FIG137478: Hypothetical protein YbgI	yes	no
Cofactors, Vitamins, Prosthetic Groups, Pigments	Biotin	Biotin biosynthesis	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	yes	no
Cofactors, Vitamins, Prosthetic Groups, Pigments	Coenzyme A	Coenzyme A Biosynthesis	2-dehydropantoate 2-reductase (EC 1.1.1.169)	yes	no
Cofactors, Vitamins, Prosthetic Groups, Pigments	Quinone cofactors	Menaquinone and Phylloquinone Biosynthesis -- gjo	1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28) in phylloquinone biosynthesis	yes	no
DNA Metabolism	CRISPs	CRISPRs	CRISPR-associated protein Cas1	yes	no

Category	Subcategory	Subsystem	Role	Presence in recovered *P. acnes* genome	Presence *P. acnes* KPA171202
DNA Metabolism	DNA repair	DNA Repair Base Excision	DNA ligase (EC 6.5.1.2)	yes	no
DNA Metabolism	DNA repair	DNA Repair Base Excision	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	yes	no
DNA Metabolism	DNA repair	DNA repair, bacterial	SOS-response repressor and protease LexA (EC 3.4.21.88)	yes	no
DNA Metabolism	DNA repair	DNA repair, bacterial RecFOR pathway	ATP-dependent DNA helicase RecQ	yes	no
DNA Metabolism	no subcategory	Restriction-Modification System	Putative predicted metal-dependent hydrolase	yes	no
DNA Metabolism	no subcategory	Restriction-Modification System	Type III restriction-modification system methylation subunit (EC 2.1.1.72)	yes	no
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	Fatty Acid Biosynthesis FASII	4'-phosphopantetheinyl transferase (EC 2.7.8.-)	yes	no
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	Fatty Acid Biosynthesis FASII	Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)	yes	no
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	Fatty Acid Biosynthesis FASII	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)	yes	no
Fatty Acids, Lipids, and Isoprenoids	Phospholipids	Cardiolipin synthesis	Cardiolipin synthetase (EC 2.7.8.-)	yes	no
Membrane Transport	ABC transporters	ABC transporter dipeptide (TC 3.A.1.5.2)	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	yes	no
Membrane Transport	ABC transporters	ABC transporter dipeptide (TC 3.A.1.5.2)	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	yes	no
Membrane Transport	ABC transporters	ABC transporter dipeptide (TC 3.A.1.5.2)	Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)	yes	no
Membrane Transport	ABC transporters	ABC transporter oligopeptide (TC 3.A.1.5.1)	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	yes	no
Nitrogen Metabolism	Denitrification	Denitrifying reductase gene clusters	Copper-containing nitrite reductase (EC 1.7.2.1)	yes	no
Nitrogen Metabolism	no subcategory	Ammonia assimilation	Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	yes	no
Nucleosides and Nucleotides	Purines	A hypothetical coupled to de Novo Purine Biosynthesis	FIG021574: Possible membrane protein related to de Novo purine biosynthesis	yes	no
Nucleosides and Nucleotides	Purines	Purine conversions	Nucleotide pyrophosphatase (EC 3.6.1.9)	yes	no
Nucleosides and Nucleotides	Pyrimidines	De Novo Pyrimidine Synthesis	Uracil permease	yes	no
Phages, Prophages, Transposable elements, Plasmids	Phages, Prophages	Phage packaging machinery	Phage terminase, large subunit	yes	no
Phosphorus Metabolism	no subcategory	Phosphate metabolism	Pyrophosphate-energized proton pump (EC 3.6.1.1)	yes	no
Potassium metabolism	no subcategory	Potassium homeostasis	Potassium channel protein	yes	no
Protein Metabolism	Protein biosynthesis	Ribosome SSU bacterial	SSU ribosomal protein S18p, zinc-dependent	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Ala-CGC	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Ala-GGC	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Arg-ACG	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Arg-CCG	yes	no

Category	Subcategory	Subsystem	Role	Presence in recovered *P. acnes* genome	Presence *P. acnes* KPA171202
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Cys-GCA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Gly-CCC	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Gly-GCC	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Leu-CAA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Leu-CAG	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Leu-GAG	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Phe-GAA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Pro-CGG	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Pro-GGG	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Ser-CGA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Ser-GGA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Trp-CCA	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Val-CAC	yes	no
Protein Metabolism	Protein biosynthesis	tRNAs	tRNA-Val-GAC	yes	no
Protein Metabolism	Protein degradation	Aminopeptidases (EC 3.4.11.-)	Membrane alanine aminopeptidase N (EC 3.4.11.2)	yes	no
Protein Metabolism	Protein degradation	Omega peptidases (EC 3.4.19.-)	Isoaspartyl aminopeptidase (EC 3.4.19.5)	yes	no
Protein Metabolism	Protein degradation	Protein degradation	Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)	yes	no
Protein Metabolism	Protein degradation	Protein degradation	Asp-X dipeptidase	yes	no
RNA Metabolism	RNA processing and modification	RNA pseudouridine syntheses	Ribosomal large subunit pseudouridine synthase A (EC 4.2.1.70)	yes	no
Respiration	Electron accepting reactions	Anaerobic respiratory reductases	Electron transfer flavoprotein-ubiquinone oxidoreductase (EC 1.5.5.1)	yes	no
Respiration	Electron accepting reactions	Anaerobic respiratory reductases	Ferredoxin reductase	yes	no
Respiration	no subcategory	Biogenesis of c-type cytochromes	Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)	yes	no
Respiration	no subcategory	Quinone oxidoreductase family	Quinone oxidoreductase (EC 1.6.5.5)	yes	no
Respiration	no subcategory	Soluble cytochromes and functionally related electron carriers	Ferredoxin	yes	no
Stress Response	Heat shock	Heat shock dnaK gene cluster extended	Signal peptidase-like protein	yes	no
Stress Response	Osmotic stress	Choline and Betaine Uptake and Betaine Biosynthesis	L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)	yes	no
Stress Response	Oxidative stress	Oxidative stress	Ferroxidase (EC 1.16.3.1)	yes	no
Stress Response	Oxidative stress	Oxidative stress	Iron-binding ferritin-like antioxidant protein	yes	no
Stress Response	Oxidative stress	Oxidative stress	Non-specific DNA-binding protein Dps	yes	no
Stress Response	no subcategory	SigmaB stress responce regulation	Serine phosphatase RsbU, regulator of sigma subunit	yes	no
Sulfur Metabolism	no subcategory	Galactosylceramide and Sulfatide metabolism	Beta-galactosidase (EC 3.2.1.23)	yes	no