

Data Analysis Report: Microbiome Profiling

Project / Study: NG-A2417

Date: July 17, 2024



1 Microbiome Analysis Pipeline

The microbiome analysis pipeline consists of three major steps and some intermediate filtering steps. Each major pipeline step is described in more detail in its respective report section. The following list provides an overview of the full pipeline, while the **main results** of the microbiome analysis are presented in section *Microbiome Profiling*.

Demultiplexing All reads passing the standard Illumina chastity filter (PF reads) are demultiplexed according to their index sequences.

Primer clipping The target region specific forward and reverse primer sequences are identified and clipped from the starts of the raw forward and reverse reads. If primer sequences could not be perfectly matched (no mismatches allowed), read pairs are removed at this step to retain only high-quality reads. The information on the remaining read pairs are provided in section *FASTQ Read Statistics*. The files with clipped reads are provided in the FASTQ directory and are named **trimmed_1.fastq.gz* and **trimmed_2.fastq.gz*. These files are not directly used as inputs for the final microbiome profiling, but are further processed as described in the following steps.

Merging If the ends of forward and reverse reads overlap, the reads are merged (assembled) to obtain a single, longer read that covers the full target region. If the target region is longer than two times the read length, merging should be impossible. If in such a case a read pair can still be merged, it is considered as an artifact and will be removed in the following quality filtering step. If the target region is only slightly shorter than two times the read length, merging may fail due to an insufficiently long high-quality overlap of the read ends. In such a case, typically only a fraction of the read pairs can be merged. In all abovementioned cases where some read pairs can't be merged, the forward read is retained and processed in the following steps instead.

In short, reads are merged if possible, and as a fallback the high quality forward read is used. No read pair is completely discarded in this step. See section *Read Merging* for additional details.

Quality filtering Merged reads are length filtered according to the expected length and known length variations of the target region (see table 1). Merged reads that are significantly shorter than the expected minimal target region length, or that are significantly longer than the expected maximal target region length, are discarded at this step. Merged and retained reads containing ambiguous bases ("N") are discarded.

The files with filtered reads are provided in the FASTQ directory and are named **_merged_for_profiling_1.fastq.gz*. These files are used as inputs for the following microbiome profiling.

Microbiome profiling The length filtered merged reads and the quality clipped retained forward reads are used as input for the microbiome profiling, where as a first step chimeric reads are identified and removed. All details of the microbiome step can be found in section *Microbiome Profiling*:

- Methods description of chimera removal, OTU picking, taxonomic assignment, etc.
- Tables with statistics describing the results of microbiome profiling
- Overview of the taxonomic composition of samples
- Detailed descriptions of delivered result files

Region code	Expected length	Merging efficiency
MI16Sa	ca. 395 bp	high
COIa	ca. 650 bp	not expected
CYTBa	(highly variable)	(highly variable)
Fu18Sa	ca. 290 bp	high
ITS1b	(highly variable)	high
PITS1a	ca. 445 bp	high
ITS2a	ca. 350 bp	high
TRNLa	(highly variable)	high
V1V3a	ca. 490 bp	moderate
V3V4a	ca. 445 bp	high
V3V5b	ca. 535 bp	high

Table 1: Standard target regions, expected lengths (rough average), and expected merging efficiency.

2 Microbiome Profiling

2.1 Results

This section summarizes the results of read preprocessing, OTU picking, and taxonomic assignment. A description of the applied methodology and according literature references are provided in the section *Methods*. Descriptions of result files and visualizations are provided in the section *Output Files and Descriptions*.

2.1.1 Statistics

Total number of input sequences	246 992	100.0%
Remaining sequences after preprocessing and quality filtering	246 962	100.0%
Remaining sequences after chimera detection and filtering	246 665	99.9%
Total number of sequences assigned to OTUs	159 605	64.6%
Total number of sequences assigned to taxa	159 605	64.6%
Copy-number corrected total count	68 808	-
Total number of OTUs	800	100.0%
Number of OTUs assignend to taxa	800	100.0%

Table 2: Summarized statistics

The number of OTUs correlates with the diversity of the data set. Sequences that were considered as noise by the OTU picking algorithm were not assigned to an OTU. The fraction of OTUs that could be assigned to taxa indicates how well the microbiome is represented in the used reference database. A copy-number correction was performed for bacterial species only, see Angly FE et al. (2014). To do so, the number of reads assigned to a species was divided by the known or assumed copy-number of marker genes/regions. The resulting corrected total count may be significantly lower than the (raw) total number of assigned reads.

Sample	1)	2)	3)	4)	5)	6)
952043.V3V4a	73 362	99.8%	73.6%	73.6%	21 537	422
952044.V3V4a	60 635	99.9%	50.6%	50.6%	14 478	407
952045.V3V4a	53 642	99.9%	73.9%	73.9%	16 862	422
952046.V3V4a	59 353	99.8%	59.4%	59.4%	15 931	420

Table 3: **1)** Input sequences. **2)** Sequences after preprocessing and chimera removal. **3)** Sequences assigned to OTUs. **4)** Sequences assigned to taxa. **5)** Count after lineage-specific copy-number correction. **6)** Median sequence length after preprocessing.

The tables can be found as files in the results directory. Please see the according section for details about result files.

2.1.2 Taxonomic Composition of Samples

The following table provides an overview of the identified taxonomic units in each sample. The most specific taxonomic units are listed with their taxonomy level and fraction (k...kingdom, p...phylum, c...class, o...order, f...family, g...genus, s...species). The most specific taxonomic unit is the lowest common taxonomic unit of the listed species (small font). These species came up as best hits of the OTUs representative sequences during the database comparison.

Next to each sample name, the corrected total number of reads of this sample that were assigned to OTUs is given. All taxonomic units with less than 0.1% of reads are collapsed in the category "Other". If the representative sequence of an OTU had no significant database match, no taxonomic unit could be assigned. The total number of reads of these unclassified OTUs is stated as category "Unclassified".

Depending on the type of analysis, some taxonomic units might be removed as they do not match the expected clade, e.g. eukaryotes in a bacterial microbiome analysis. The number of removed reads is stated as category "Filtered". If this category is not listed, no filtering was performed.

A copy-number correction was performed for bacterial species only, see Angly FE et al. (2014). If the listed normalized fraction and raw fraction are identical, either no copy-number correction factor was available in the database or the factor is exactly one.

Sample Name (copy-number corrected read counts)		Normalized Fraction	Raw Fraction
Taxonomic Level	Taxonomic Unit		
952043.V3V4a (21 537 reads)			
s	Flavobacterium sp. S13 (25 OTUs with 99-100% identity in 419-422bp to: Flavobacterium sp. S13)	12.6%	18.1%
s	Cellvibrio gandavensis (7 OTUs with 99-100% identity in 422-424bp to: Cellvibrio gandavensis)	8.6%	7.2%
s	Fluviicola kyonggii (11 OTUs with 99-100% identity in 420-422bp to: Fluviicola kyonggii)	7.9%	8.4%
g	Acidovorax (5 OTUs with 99-100% identity in 422-424bp to: 25 unclassified Acidovorax strains, Acidovorax delafieldii, Acidovorax facilis, Acidovorax radicus)	6.7%	5.2%
g	Cellvibrio (12 OTUs with 99-100% identity in 424bp to: 13 unclassified Cellvibrio strains, Cellvibrio fibrivorans, Cellvibrio mixtus)	5.3%	4.4%
s	Cellvibrio sp. R-1(2011) (2 OTUs with 99% identity in 424bp to: Cellvibrio sp. R-1(2011))	3.1%	2.6%
g	Flavobacterium (6 OTUs with 99-100% identity in 420-422bp to: 15 unclassified Flavobacterium strains, Flavobacterium johnsoniae)	2.9%	4.1%
s	Rheinheimera sp. NT-1 (2 OTUs with 99-100% identity in 424bp to: Rheinheimera sp. NT-1)	2.6%	1.9%
s	Flavobacterium sp. WB 1.1-23 (1 OTU with 99% identity in 422bp to: Flavobacterium sp. WB 1.1-23)	2.5%	3.6%
g	Asticcacaulis (1 OTU with 100% identity in 399bp to: Asticcacaulis benevestitus, Asticcacaulis excentricus)	2.2%	1.6%
g	Pseudomonas (4 OTUs with 99-100% identity in 424bp to: 42 unclassified Pseudomonas strains, Pseudomonas alcaligenes, Pseudomonas lini, Pseudomonas mediterranea, Pseudomonas nitroreducens, Pseudomonas oleovorans, Pseudomonas taiwanensis)	2.1%	3.7%
s	Cellvibrio zantedeschiae (2 OTUs with 99% identity in 424bp to: Cellvibrio zantedeschiae)	1.8%	1.5%
f	Rhizobiaceae (6 OTUs with 99-100% identity in 399bp to: 10 unclassified Rhizobium strains, 2 unclassified Sinorhizobium strains, Agrobacterium sp. Emb7, Ensifer collicola, Neorhizobium algalisoli, Neorhizobium galegae, Neorhizobium sp., Rhizobium vignae, Shinella yambaruensis)	1.8%	1.5%
g	Methylophilus (1 OTU with 100% identity in 424bp to: 3 unclassified Methylophilus strains, Methylophilus leisingeri)	1.6%	0.6%
s	Vulgatibacter incompus (2 OTUs with 79% identity in 427-428bp to: Vulgatibacter incompus)	1.4%	1.3%
s	Cellvibrio sp. R-4075 (3 OTUs with 99-100% identity in 424bp to: Cellvibrio sp. R-4075)	1.3%	1.1%
s	Acidovorax sp. R40 (1 OTU with 99% identity in 424bp to: Acidovorax sp. R40)	1.2%	0.9%
g	Bdellovibrio (6 OTUs with 95-99% identity in 405bp to: 5 unclassified Bdellovibrio strains)	1.2%	1.3%
g	Thermomonas (1 OTU with 99% identity in 427bp to: 2 unclassified Thermomonas strains, Thermomonas brevis, Thermomonas haemolytica)	1.1%	1.3%
f	Microbacteriaceae (1 OTU with 100% identity in 404bp to: 5 unclassified Leifsonia strains, Lysinimonas soli, Lysinimonas sp. 762H99, Salinibacterium sp.)	1.1%	0.9%
g	Desulfuromonas (2 OTUs with 83-86% identity in 425-427bp to: Desulfuromonas carbonis, Desulfuromonas sp. DDH964)	1.0%	0.4%

s	Sterolibacterium sp. AMi17 (2 OTUs with 97% identity in 424bp to: Sterolibacterium sp. AMi17)	1.0%	0.4%
g	Devosia (2 OTUs with 100% identity in 399bp to: 20 unclassified Devosia strains, Devosia crocina, Devosia humi, Devosia riboflavina, Devosia soli)	1.0%	0.8%
k	Bacteria (1 OTU with 100% identity in 424bp to: Flavobacterium sp. A1-26, Rhodanobacter sp. 7B-393, Simplicispira sp.)	0.9%	0.9%
g	Rhizobium (3 OTUs with 99-100% identity in 399bp to: 20 unclassified Rhizobium strains, Rhizobium gallicum, Rhizobium genosp. TUXTLAS-24, Rhizobium grahamii, Rhizobium oryzae, Rhizobium pseudoryzae, Rhizobium rhizoryzae, Rhizobium tibeticum)	0.9%	1.0%
s	Dechlorobacter hydrogenophilus (1 OTU with 99% identity in 424bp to: Dechlorobacter hydrogenophilus)	0.9%	1.1%
g	Chryseobacterium (1 OTU with 100% identity in 422bp to: 22 unclassified Chryseobacterium strains, Chryseobacterium indoltheticum, Chryseobacterium mulctrae, Chryseobacterium piscium, Chryseobacterium scopthalmum)	0.9%	0.9%
g	Chromobacterium (1 OTU with 100% identity in 424bp to: 6 unclassified Chromobacterium strains, Chromobacterium aquaticum, Chromobacterium rhizoryzae, Chromobacterium violaceum)	0.8%	1.4%
g	Phenylobacterium (1 OTU with 100% identity in 399bp to: Phenylobacterium immobile, Phenylobacterium sp.)	0.7%	0.3%
s	Novispirillum itersonii (2 OTUs with 99-100% identity in 399bp to: Novispirillum itersonii)	0.7%	0.9%
s	Rhizomicrobium sp. (2 OTUs with 98% identity in 402bp to: Rhizomicrobium sp.)	0.7%	0.6%
s	Racemicystis persica (2 OTUs with 88-94% identity in 423-426bp to: Racemicystis persica)	0.7%	1.1%
s	Pseudoduganella sp. (1 OTU with 100% identity in 424bp to: Pseudoduganella sp.)	0.7%	0.7%
f	Oxalobacteraceae (2 OTUs with 99% identity in 424bp to: 2 unclassified Massilia strains, 4 unclassified Duganella strains, Herminiimonas glaciei, Herminiimonas saxosidens, Herminiimonas sp. CUW1, Janthinobacterium sp. Marseille)	0.7%	0.7%
s	Kofleria flava (2 OTUs with 90% identity in 423bp to: Kofleria flava)	0.7%	0.6%
O	Enterobacterales (2 OTUs with 99-100% identity in 424bp to: 2 unclassified Enterobacter strains, Enterobacter asburiae, Enterobacter cloacae, Enterobacter hormaechei, Enterobacter ludwigii, Enterobacter mori, Enterobacter roggenkampii, Enterobacter sichuanensis, Leclercia adecarboxylata, Pantoea agglomerans)	0.6%	0.5%
g	Caulobacter (1 OTU with 100% identity in 399bp to: 37 unclassified Caulobacter strains, Caulobacter flavus, Caulobacter hibisci, Caulobacter radialis, Caulobacter segnis, Caulobacter vibrioides, Caulobacter zeae)	0.6%	0.5%
s	Asticcacaulis sp. VA7 (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. VA7)	0.6%	0.4%
s	Bacteriovorax stolpii (2 OTUs with 97% identity in 425-426bp to: Bacteriovorax stolpii)	0.6%	0.6%
s	Cellvibrio mixtus (2 OTUs with 99-100% identity in 424bp to: Cellvibrio mixtus)	0.5%	0.5%
g	Rheinheimera (1 OTU with 100% identity in 424bp to: 2 unclassified Rheinheimera strains)	0.5%	0.4%
s	Gemmatimonas aurantiaca (1 OTU with 89% identity in 402bp to: Gemmatimonas aurantiaca)	0.5%	0.2%
f	Rhodospirillaceae (3 OTUs with 99% identity in 399bp to: 3 unclassified Azospirillum strains, 3 unclassified Phaeospirillum strains, 8 unclassified Niveispirillum strains, Ferrovibrio soli, Ferrovibrio sp., Ferrovibrio terrae, Niveispirillum cyanobacteriorum, Rhodocista sp. W38, Rhodospirillum sp. SL38)	0.5%	0.7%
s	Acidovorax sp. Sy22HNO (1 OTU with 100% identity in 424bp to: Acidovorax sp. Sy22HNO)	0.5%	0.4%
s	Bdellovibrio sp. MPA (2 OTUs with 97% identity in 405bp to: Bdellovibrio sp. MPA)	0.5%	0.5%
O	Flavobacteriales (1 OTU with 100% identity in 422bp to: 28 unclassified Chryseobacterium strains, Chryseobacterium indologenes, Chryseobacterium rhizoplanae, Chryseobacterium sediminis, Chryseobacterium tructae, Flavobacterium sp. PH1)	0.5%	0.5%
s	Hydrogenophaga flava (1 OTU with 100% identity in 424bp to: Hydrogenophaga flava)	0.5%	0.4%
s	Thermomonas brevis (1 OTU with 97% identity in 427bp to: Thermomonas brevis)	0.5%	0.5%
g	Pseudoxanthomonas (1 OTU with 100% identity in 424bp to: 4 unclassified Pseudoxanthomonas strains, Pseudoxanthomonas japonensis, Pseudoxanthomonas mexicana)	0.5%	0.2%
f	Sphingomonadaceae (3 OTUs with 100% identity in 399bp to: 10 unclassified Sphingopyxis strains, 8 unclassified Sphingomonas strains, Sphingomonas adhaesiva, Sphingopyxis soli, Sphingopyxis taejoniensis, Sphingopyxis terrae)	0.5%	0.4%
s	Flavobacterium sp. A6AT7 (1 OTU with 100% identity in 422bp to: Flavobacterium sp. A6AT7)	0.4%	0.6%

O	Rhizobiales (3 OTUs with 100% identity in 399bp to: 14 unclassified Boreia strains, 2 unclassified Afipia strains, 2 unclassified Chelatococcus strains, 2 unclassified Rhizobium strains, 2 unclassified Starkeya strains, 5 unclassified Agrobacterium strains, Agrobacterium deltaense, Agrobacterium tumefaciens, Beijerinckia fluminensis, Boreia minatitlanensis, Boreia robiniae, Candidatus Rhizobium massiliae, Chelatococcus asaccharovorans, Methylobacterium sp. S120, Rhizobium leguminosarum, Rhizobium nepotum, Rhizobium pusense)	0.4%	0.4%
S	Haematococcus lacustris (1 OTU with 86% identity in 422bp to: Haematococcus lacustris)	0.4%	0.2%
S	Sediminibacterium sp. (1 OTU with 99% identity in 422bp to: Sediminibacterium sp.)	0.4%	0.4%
S	Pseudolabrys sp. W2.09-308 (2 OTUs with 97% identity in 399bp to: Pseudolabrys sp. W2.09-308)	0.4%	0.3%
S	Aquella oligotrophica (1 OTU with 94% identity in 424bp to: Aquella oligotrophica)	0.4%	0.5%
g	Hydrogenophaga (1 OTU with 100% identity in 424bp to: 3 unclassified Hydrogenophaga strains, Hydrogenophaga borbori)	0.3%	0.3%
g	Bradyrhizobium (1 OTU with 100% identity in 399bp to: 4 unclassified Bradyrhizobium strains, Bradyrhizobium elkanii, Bradyrhizobium jicamae, Bradyrhizobium paxllaeri)	0.3%	0.2%
S	Duganella sp. RCP6 (2 OTUs with 99-100% identity in 424bp to: Duganella sp. RCP6)	0.3%	0.3%
S	Methylothermobacter mobilis (1 OTU with 99% identity in 424bp to: Methylothermobacter mobilis)	0.3%	0.1%
C	Actinobacteria (1 OTU with 100% identity in 404bp to: 11 unclassified Mycobacterium strains, Micromonospora sp. NEAU-JG13, Mycolicibacterium rhodesiae)	0.3%	0.3%
S	Spirochaeta sp. MWH-HuW24 (1 OTU with 97% identity in 427bp to: Spirochaeta sp. MWH-HuW24)	0.3%	0.2%
S	Methylophilus sp. ECd5 (1 OTU with 99% identity in 424bp to: Methylophilus sp. ECd5)	0.3%	0.1%
S	Sphingomonas sp. IJ1 (1 OTU with 99% identity in 402bp to: Sphingomonas sp. IJ1)	0.3%	0.2%
S	Hyphomicrobium sp. LAT3 (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. LAT3)	0.3%	0.1%
f	Enterobacteriaceae (1 OTU with 100% identity in 424bp to: Enterobacter asburiae, Klebsiella pneumoniae)	0.3%	0.2%
p	Proteobacteria (2 OTUs with 99-100% identity in 399bp to: 3 unclassified Shinella strains, 8 unclassified Rhizobium strains, Ochrobactrum anthropi, Ochrobactrum sp., Rhizobium gallicum, Rhizobium yanglingense, Shinella curvata, Shinella kummerowiae, Shinella yambaruensis, Shinella zoogloeoides, Variovorax sp. 4LI(+)_OTU30)	0.3%	0.3%
S	Cytophaga hutchinsonii (2 OTUs with 99% identity in 421bp to: Cytophaga hutchinsonii)	0.2%	0.2%
C	Gammaproteobacteria (1 OTU with 100% identity in 424bp to: Alishewanella sp. 114NP19, Pararheinheimera texasensis, Rheinheimera sp. 114NP12)	0.2%	0.2%
S	Cohnella sp. (1 OTU with 98% identity in 426bp to: Cohnella sp.)	0.2%	0.2%
S	Pararheinheimera texasensis (1 OTU with 100% identity in 424bp to: Pararheinheimera texasensis)	0.2%	0.1%
S	Dongia sp. URHE0060 (1 OTU with 99% identity in 399bp to: Dongia sp. URHE0060)	0.2%	0.3%
g	Chitinimonas (1 OTU with 100% identity in 424bp to: 4 unclassified Chitinimonas strains, Chitinimonas viridis)	0.2%	0.4%
S	Oligoflexus tunisiensis (1 OTU with 95% identity in 428bp to: Oligoflexus tunisiensis)	0.2%	0.2%
S	Cellvibrio ostraviensis (1 OTU with 99% identity in 424bp to: Cellvibrio ostraviensis)	0.2%	0.2%
S	Devosia psychrophila (1 OTU with 100% identity in 399bp to: Devosia psychrophila)	0.2%	0.1%
f	Mycobacteriaceae (2 OTUs with 100% identity in 404bp to: 13 unclassified Mycobacterium strains, Mycobacterium neglectum, Mycobacterium petroleophilum, Mycolicibacterium gilvum, Mycolicibacterium helvum, Mycolicibacterium rhodesiae, Mycolicibacterium sp., Mycolicibacterium sphagni, Mycolicibacterium tusciae)	0.2%	0.1%
g	Streptomyces (1 OTU with 99% identity in 404bp to: 28 unclassified Streptomyces strains, Streptomyces canarius, Streptomyces clavuligerus, Streptomyces corchorusii, Streptomyces galbus, Streptomyces kagawaensis, Streptomyces lucensis, Streptomyces mirabilis, Streptomyces monashensis, Streptomyces novaecaesareae, Streptomyces olivochromogenes, Streptomyces rhizosphaerihabitans, Streptomyces tsukubensis)	0.2%	0.1%
g	Azospirillum (2 OTUs with 100% identity in 399bp to: 45 unclassified Azospirillum strains, Azospirillum humicireducens, Azospirillum largimobile, Azospirillum lipoferum, Azospirillum melinis, Azospirillum oryzae, Azospirillum zeae, Candidatus Azospirillum massiliensis)	0.2%	0.4%
S	Pseudomonas sp. EA S 49 (1 OTU with 100% identity in 424bp to: Pseudomonas sp. EA.S.49)	0.2%	0.3%
S	Desulfobulbus mediterraneus (1 OTU with 78% identity in 430bp to: Desulfobulbus mediterraneus)	0.2%	0.1%
g	Lentzea (1 OTU with 100% identity in 404bp to: 11 unclassified Lentzea strains, Lentzea albida, Lentzea albidocapitata, Lentzea californiensis, Lentzea flaviverrucosa, Lentzea violacea, Lentzea waywayandensis)	0.1%	0.1%

s	Flavobacterium sp. (1 OTU with 98% identity in 422bp to: Flavobacterium sp.)	0.1%	0.2%
g	Ferrovibrio (1 OTU with 100% identity in 399bp to: Ferrovibrio soli, Ferrovibrio terrae)	0.1%	0.2%
f	Xanthomonadaceae (1 OTU with 100% identity in 424bp to: 3 unclassified Xanthomonas strains, Pseudoxanthomonas japonensis, Pseudoxanthomonas mexicana, Pseudoxanthomonas sp. Lo13)	0.1%	0.2%
g	Massilia (1 OTU with 99% identity in 424bp to: Massilia sp. Sco-D23, Naxibacter intermedius, Naxibacter sp. KF-26)	0.1%	0.1%
s	Anaeromyxobacter dehalogenans (1 OTU with 88% identity in 425bp to: Anaeromyxobacter dehalogenans)	0.1%	0.1%
s	Lacibacter cauensis (1 OTU with 94% identity in 422bp to: Lacibacter cauensis)	0.1%	0.1%
s	Chryseolinea soli (1 OTU with 92% identity in 421bp to: Chryseolinea soli)	0.1%	0.1%
s	Novosphingobium sp. TrD22 (1 OTU with 99% identity in 399bp to: Novosphingobium sp. TrD22)	0.1%	0.1%
s	Telluria mixta (1 OTU with 100% identity in 424bp to: Telluria mixta)	0.1%	0.1%
s	Haliscomenobacter hydrossis (1 OTU with 99% identity in 422bp to: Haliscomenobacter hydrossis)	0.1%	0.1%
s	Niveispirillum sp. (1 OTU with 97% identity in 399bp to: Niveispirillum sp.)	0.1%	0.1%
s	Azospirillum picis (1 OTU with 100% identity in 399bp to: Azospirillum picis)	0.1%	0.3%
	Other	0.6%	0.7%
	Unclassified (0 reads)		
	Filtered (0 reads)		

952044.V3V4a (14 478 reads)

g	Streptomyces (10 OTUs with 100% identity in 404bp to: 79, Streptomyces)	9.7%	9.7%
s	Devosia sp. (7 OTUs with 99-100% identity in 399-400bp to: Devosia sp.)	2.8%	2.6%
s	Candidatus Solibacter usitatus (4 OTUs with 90-96% identity in 399bp to: Candidatus Solibacter usitatus)	2.5%	1.4%
s	Flavobacterium sp. S13 (4 OTUs with 99-100% identity in 422bp to: Flavobacterium sp. S13)	2.4%	4.1%
s	Conexibacter sp. (3 OTUs with 93-98% identity in 424bp to: Conexibacter sp.)	2.2%	1.1%
s	Asprobacter aquaticus (4 OTUs with 97-99% identity in 399bp to: Asprobacter aquaticus)	2.0%	1.6%
s	Flavobacterium sp. WB 1.1-23 (2 OTUs with 99% identity in 419-422bp to: Flavobacterium sp. WB 1.1-23)	2.0%	3.3%
g	Flavobacterium (6 OTUs with 95-100% identity in 422bp to: 13 unclassified Flavobacterium strains, Flavobacterium fluminis, Flavobacterium frigidimar, Flavobacterium johnsoniae, Flavobacterium resistens)	1.9%	3.3%
s	Solirubrobacter sp. URHE0046 (3 OTUs with 90% identity in 424bp to: Solirubrobacter sp. URHE0046)	1.9%	0.9%
g	Methylosinus (3 OTUs with 96-97% identity in 399bp to: Methylosinus sp. NCIMB 13214, Methylosinus sporium)	1.7%	1.2%
s	Telmatobacter bradus (2 OTUs with 98-99% identity in 399bp to: Telmatobacter bradus)	1.6%	0.9%
s	Nocardioides sp. (2 OTUs with 99-100% identity in 404-420bp to: Nocardioides sp.)	1.6%	1.6%
f	Sphingomonadaceae (5 OTUs with 99-100% identity in 399-402bp to: 10 unclassified Sphingopyxis strains, 13 unclassified Sphingobium strains, 21 unclassified Sphingomonas strains, Sphingobium bisphenolivorans, Sphingobium chlorophenolicum, Sphingobium chungbukense, Sphingobium estrovivorans, Sphingomonas adhaesiva, Sphingomonas agrestis, Sphingopyxis soli, Sphingopyxis taejonensis, Sphingopyxis terrae)	1.5%	1.5%
s	Rhizomicrobium sp. (4 OTUs with 98-99% identity in 402bp to: Rhizomicrobium sp.)	1.5%	1.5%
s	Flavobacterium sp. 71B1 (1 OTU with 100% identity in 422bp to: Flavobacterium sp. 71B1)	1.4%	2.4%
g	Nitrospira (2 OTUs with 90-94% identity in 424bp to: 5 unclassified Nitrospira strains)	1.3%	0.6%
s	Haematococcus lacustris (1 OTU with 88% identity in 422bp to: Haematococcus lacustris)	1.3%	0.6%
f	Rhizobiaceae (4 OTUs with 99-100% identity in 399bp to: Agrobacterium sp. Emb7, Ensifer sp., Rhizobium azooxidifex, Rhizobium sp., Shinella yambaruensis)	1.2%	1.2%
f	Microbacteriaceae (1 OTU with 100% identity in 404bp to: 5 unclassified Leifsonia strains, Lysinimonas soli, Lysinimonas sp. 762H99, Salinibacterium sp.)	1.2%	1.2%

f	Mycobacteriaceae (5 OTUs with 99-100% identity in 404bp to: 73 unclassified Mycobacterium strains, Arthrobacter sp. SLBN-53, Mycobacterium aquiterrae, Mycobacterium hackensackense, Mycobacterium marinum, Mycobacterium neglectum, Mycobacterium sacrum, Mycolicibacterium aichiense, Mycolicibacterium anyangense, Mycolicibacterium aurum, Mycolicibacterium bacteremicum, Mycolicibacterium cosmeticum, Mycolicibacterium crocinum, Mycolicibacterium diernhoferi, Mycolicibacterium fluoranthenvivorans, Mycolicibacterium fortuitum, Mycolicibacterium frederiksbergense, Mycolicibacterium gilvum, Mycolicibacterium goodii, Mycolicibacterium helvum, Mycolicibacterium hodleri, Mycolicibacterium moriokaense, Mycolicibacterium neoaurum, Mycolicibacterium peregrinum, Mycolicibacterium sarraeeniae, Mycolicibacterium smegmatis, Mycolicibacterium sp., Mycolicibacterium sphagni)	1.1%	1.1%
	Aciditerrimonas ferrireducens (4 OTUs with 91-93% identity in 401bp to: Aciditerrimonas ferrireducens)	1.1%	1.0%
s	Candidatus Nitrosarchaeum limnium (1 OTU with 95% identity in 400bp to: Candidatus Nitrosarchaeum limnium)	1.1%	0.7%
s	Nocardioides sp. URHD0098 (3 OTUs with 98-100% identity in 420bp to: Nocardioides sp. URHD0098)	1.1%	1.1%
f	Bradyrhizobiaceae (2 OTUs with 100% identity in 399bp to: 2 unclassified Bradyrhizobium strains, 2 unclassified Rhodopseudomonas strains, Afipia massiliensis, Bradyrhizobium canariense, Bradyrhizobium lupini, Bradyrhizobium macuxiense, Nitrobacter sp. NS5-5, Rhodopseudomonas faecalis)	1.1%	0.8%
	Actinobacteria (1 OTU with 100% identity in 404bp to: 9 unclassified Streptomyces strains, Actinomyces sp., Streptomyces atriruber, Streptomyces costaricanus, Streptomyces endophyticus, Streptomyces lanatus, Streptomyces prasinopilosus, Streptomyces rochei, Streptomyces venezuelae, Streptomyces xiangtanensis)	1.0%	1.0%
s	Acidobacterium sp. WY65 (2 OTUs with 93-94% identity in 399bp to: Acidobacterium sp. WY65)	1.0%	0.5%
g	Bradyrhizobium (2 OTUs with 100% identity in 399bp to: 15 unclassified Bradyrhizobium strains, Bradyrhizobium arachidis, Bradyrhizobium betae, Bradyrhizobium canariense, Bradyrhizobium centrosematis, Bradyrhizobium daqingense, Bradyrhizobium diazoefficiens, Bradyrhizobium erythrophlei, Bradyrhizobium guangxiense, Bradyrhizobium huanghuaihaiense, Bradyrhizobium japonicum, Bradyrhizobium liaoningense, Bradyrhizobium lupini, Bradyrhizobium ottawaense, Bradyrhizobium yuanmingense)	1.0%	0.6%
	Racemicystis persica (4 OTUs with 88-94% identity in 423-426bp to: Racemicystis persica)	1.0%	1.8%
g	Mesorhizobium (2 OTUs with 100% identity in 399bp to: 34 unclassified Mesorhizobium strains, Mesorhizobium cantuariense, Mesorhizobium ciceri, Mesorhizobium norvegicum, Mesorhizobium qingshengii)	1.0%	0.8%
s	Chryseolinea soli (2 OTUs with 91-92% identity in 421bp to: Chryseolinea soli)	0.9%	1.1%
s	Niastella sp. HME8655 (2 OTUs with 96% identity in 422bp to: Niastella sp. HME8655)	0.9%	1.1%
g	Phenyllobacterium (1 OTU with 100% identity in 399bp to: Phenyllobacterium immobile, Phenyllobacterium sp.)	0.8%	0.4%
s	Pseudolabrys sp. W2.09-308 (2 OTUs with 97-98% identity in 399bp to: Pseudolabrys sp. W2.09-308)	0.8%	0.8%
s	Mesorhizobium sp. (2 OTUs with 99-100% identity in 399bp to: Mesorhizobium sp.)	0.8%	0.7%
s	Ferruginibacter paludis (1 OTU with 97% identity in 422bp to: Ferruginibacter paludis)	0.8%	1.0%
g	Hyphomicrobium (2 OTUs with 98-99% identity in 399bp to: 13 unclassified Hyphomicrobium strains, Hyphomicrobium denitrificans, Hyphomicrobium facile)	0.8%	0.5%
s	Sandaracinus amylolyticus (3 OTUs with 86-89% identity in 425-429bp to: Sandaracinus amylolyticus)	0.8%	0.9%
s	Steroidobacter sp. JC2953 (4 OTUs with 93-95% identity in 424bp to: Steroidobacter sp. JC2953)	0.8%	0.7%
s	Paludibaculum fermentans (2 OTUs with 91-99% identity in 399bp to: Paludibaculum fermentans)	0.7%	0.4%
s	Methyloligella solikamskensis (2 OTUs with 97-98% identity in 399bp to: Methyloligella solikamskensis)	0.7%	0.7%
s	Terrimonas sp. (1 OTU with 97% identity in 422bp to: Terrimonas sp.)	0.7%	0.8%
s	Nitrospira sp. (1 OTU with 98% identity in 417bp to: Nitrospira sp.)	0.7%	0.3%
s	Terrimonas sp. CB 286475 (1 OTU with 97% identity in 422bp to: Terrimonas sp. CB 286475)	0.6%	0.8%
s	Gemmatimonas sp. (1 OTU with 92% identity in 416bp to: Gemmatimonas sp.)	0.6%	0.3%
g	Cellvibrio (3 OTUs with 99-100% identity in 424bp to: 4 unclassified Cellvibrio strains, Cellvibrio fibrivorans, Cellvibrio mixtus)	0.6%	0.6%
o	Rhizobiales (3 OTUs with 98-100% identity in 399bp to: 2 unclassified Aurantimonas strains, 7 unclassified Mesorhizobium strains, 7 unclassified Rhizobium strains, Biserrula isolate, Bradyrhizobium sp. AMF3693, Jiella sp., Mesorhizobium amorphae, Mesorhizobium huakuii, Mesorhizobium japonicum, Mesorhizobium loti, Mesorhizobium opportunistum, Mesorhizobium robiniae, Mesorhizobium silamurunense, Rhizobium genosp. TUXTLAS-27, Rhizobium grahamii, Rhizobium tibeticum)	0.6%	0.6%
	Gemmatimonas aurantiaca (1 OTU with 93% identity in 402bp to: Gemmatimonas aurantiaca)	0.6%	0.3%

g	Rhizobium (3 OTUs with 100% identity in 399bp to: 24 unclassified Rhizobium strains, Rhizobium gallicum, Rhizobium genosp. TUXTLAS-24, Rhizobium grahamii, Rhizobium oryzae, Rhizobium pseudoryzae, Rhizobium rhizoryzae, Rhizobium tibeticum)	0.6%	0.7%
s	Bauldia sp. (3 OTUs with 98% identity in 402bp to: Bauldia sp.)	0.6%	0.6%
s	Sideroxydans lithotrophicus (1 OTU with 92% identity in 424bp to: Sideroxydans lithotrophicus)	0.5%	0.3%
s	Candidatus Koribacter versatilis (1 OTU with 93% identity in 399bp to: Candidatus Koribacter versatilis)	0.5%	0.3%
s	Edaphobacter dinghuensis (1 OTU with 99% identity in 399bp to: Edaphobacter dinghuensis)	0.5%	0.3%
c	Gammaproteobacteria (2 OTUs with 90% identity in 424bp to: Candidatus Thiodiazotropha endolucinida, Ectothiorhodospira mobilis, Thiohalocapsa halophila)	0.5%	0.5%
s	Methylosinus sporium (1 OTU with 98% identity in 399bp to: Methylosinus sporium)	0.5%	0.4%
g	Acidovorax (2 OTUs with 99-100% identity in 424bp to: 11 unclassified Acidovorax strains, Acidovorax delafieldii, Acidovorax facilis)	0.5%	0.4%
f	Erythrobacteraceae (2 OTUs with 98% identity in 400bp to: 2 unclassified Erythrobacter strains, Altererythrobacter salegens)	0.5%	0.3%
s	Sediminibacterium sp. (1 OTU with 99% identity in 422bp to: Sediminibacterium sp.)	0.4%	0.6%
s	Taibaiella yonginensis (2 OTUs with 98-99% identity in 423bp to: Taibaiella yonginensis)	0.4%	0.5%
g	Afipia (1 OTU with 100% identity in 399bp to: 5 unclassified Afipia strains, Afipia felis)	0.4%	0.3%
s	Pseudolabrys sp. (2 OTUs with 98% identity in 399bp to: Pseudolabrys sp.)	0.4%	0.4%
s	Mesorhizobium plurifarum (1 OTU with 100% identity in 399bp to: Mesorhizobium plurifarum)	0.4%	0.3%
f	Oxalobacteraceae (2 OTUs with 95-99% identity in 424bp to: Herbaspirillum sp. S20-88, Herminiimonas glaciei, Herminiimonas saxosidens, Herminiimonas sp. CUW1, Janthinobacterium sp. Marseille, Lacisediminimonas profundus)	0.4%	0.5%
s	Candidatus Ovatusbacter abovo (1 OTU with 93% identity in 424bp to: Candidatus Ovatusbacter abovo)	0.4%	0.4%
f	Micrococcaceae (2 OTUs with 99-100% identity in 408-411bp to: 10 unclassified Arthrobacter strains, Arthrobacter keyseri, Glutamicibacter protophormiae, Micrococcus sp. YGJ1, Paenarthrobacter nicotinovorans, Paenarthrobacter nitroguajacolicus, Paenarthrobacter ureafaciens)	0.4%	0.4%
p	Proteobacteria (2 OTUs with 95-100% identity in 399-402bp to: 2 unclassified Caulobacter strains, 3 unclassified Methylosinus strains, Hydrogenophaga intermedia, Methylomonas sp. GYJ3, Methylosinus trichosporium, Phenylbacterium conjunctum)	0.4%	0.5%
s	Hyphomicrobium sp. (1 OTU with 100% identity in 399bp to: Hyphomicrobium sp.)	0.4%	0.3%
f	Rhodospirillaceae (2 OTUs with 99% identity in 399bp to: 2 unclassified Azospirillum strains, 3 unclassified Phaeosporillum strains, Ferrovibrio soli, Ferrovibrio sp., Ferrovibrio terrae, Oleisolibacter albus)	0.4%	0.7%
g	Nannocystis (1 OTU with 95% identity in 420bp to: 3 unclassified Nannocystis strains, Nannocystis exedens, Nannocystis pusilla)	0.4%	0.4%
s	Cellvibrio sp. R-1(2011) (1 OTU with 99% identity in 424bp to: Cellvibrio sp. R-1(2011))	0.4%	0.4%
s	Ignavibacterium album (2 OTUs with 81-83% identity in 421-423bp to: Ignavibacterium album)	0.4%	0.5%
o	Myxococcales (1 OTU with 87% identity in 425bp to: Chondromyces pediculatus, Sandaracinus amylolyticus)	0.4%	0.5%
f	Polyangiaceae (5 OTUs with 86-94% identity in 423-427bp to: Aetherobacter rufus, Aetherobacter sp. SBSr008, Chondromyces apiculatus, Chondromyces lanuginosus, Chondromyces pediculatus, Chondromyces robustus, Jahnella thaxteri, Racemicystis persica)	0.4%	0.8%
k	Bacteria (2 OTUs with 97-99% identity in 399-422bp to: 33 unclassified Rhizobium strains, Arthrobacter sp., Fluviicola sp. YS 1.2.4, Glaciecola sp. THG-3.7, Rhizobium alamii, Rhizobium etli, Rhizobium leguminosarum, Rhizobium mesosinicum, Rhizobium sullae, Rhizobium viscosum)	0.4%	0.5%
g	Jatrophihabitans (2 OTUs with 98-99% identity in 404bp to: 2 unclassified Jatrophihabitans strains, Jatrophihabitans soli)	0.4%	0.4%
s	Taibaiella sp. (1 OTU with 98% identity in 422bp to: Taibaiella sp.)	0.4%	0.5%
s	Hyphomicrobium sp. LAT3 (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. LAT3)	0.4%	0.3%
g	Nocardioides (2 OTUs with 98-100% identity in 404-413bp to: 5 unclassified Nocardioides strains)	0.4%	0.4%
s	Ferruginibacter sp. (2 OTUs with 95-98% identity in 422-424bp to: Ferruginibacter sp.)	0.4%	0.5%
g	Dehalococcoides (1 OTU with 83% identity in 404bp to: Dehalococcoides mccartyi, Dehalococcoides sp. UCH007)	0.4%	0.4%
s	Hyphomicrobium sp. D3 (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. D3)	0.4%	0.3%
s	Asticcacaulis sp. SA7 (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. SA7)	0.4%	0.3%
s	Methylocystis sp. SH31p (1 OTU with 96% identity in 400bp to: Methylocystis sp. SH31p)	0.4%	0.4%

S	Pseudomonas sp. EA S 49 (1 OTU with 100% identity in 424bp to: Pseudomonas sp. EA_S.49)	0.4%	0.8%
g	Rhodanobacter (2 OTUs with 99-100% identity in 424bp to: 3 unclassified Rhodanobacter strains, Rhodanobacter fulvus, Rhodanobacter ginsenosidimutans, Rhodanobacter spatiphylli)	0.4%	0.6%
C	Alphaproteobacteria (1 OTU with 100% identity in 399bp to: 34 unclassified Rhizobium strains, Rhizobium endophyticum, Rhizobium etli, Rhizobium grahamii, Rhizobium leguminosarum, Rhizobium metalidurans, Rhodospirillum sp. kmd.475)	0.4%	0.4%
S	Candidatus Thiodiazotropha endolucinida (1 OTU with 90% identity in 424bp to: Candidatus Thiodiazotropha endolucinida)	0.4%	0.4%
S	Aquihabitans daechungensis (2 OTUs with 93-94% identity in 401-402bp to: Aquihabitans daechungensis)	0.3%	0.3%
S	Iamia majanohamensis (2 OTUs with 92% identity in 401bp to: Iamia majanohamensis)	0.3%	0.3%
S	Cytophaga hutchinsonii (2 OTUs with 99% identity in 421bp to: Cytophaga hutchinsonii)	0.3%	0.4%
S	Flavilitoribacter nigricans (1 OTU with 87% identity in 422bp to: Flavilitoribacter nigricans)	0.3%	0.4%
S	Asticcacaulis sp. VA7 (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. VA7)	0.3%	0.3%
S	Sterolibacterium sp. AMi17 (1 OTU with 98% identity in 424bp to: Sterolibacterium sp. AMi17)	0.3%	0.2%
S	Cellvibrio sp. PSBB023 (1 OTU with 98% identity in 424bp to: Cellvibrio sp. PSBB023)	0.3%	0.3%
g	Devosia (2 OTUs with 99-100% identity in 399bp to: 17 unclassified Devosia strains, Devosia humi, Devosia insulae)	0.3%	0.3%
S	Blastochloris sp. AT2101 (1 OTU with 94% identity in 399bp to: Blastochloris sp. AT2101)	0.3%	0.2%
g	Chelativorans (1 OTU with 95% identity in 399bp to: Chelativorans composti, Chelativorans sp. A52C2)	0.3%	0.3%
g	Kofleria (1 OTU with 89% identity in 424bp to: Kofleria flava, Kofleria sp. DSM 53797)	0.3%	0.4%
S	Burkholderia sp. (2 OTUs with 98% identity in 424bp to: Burkholderia sp.)	0.3%	0.7%
g	Actinospica (1 OTU with 99% identity in 404bp to: 3 unclassified Actinospica strains, Actinospica robiniae)	0.3%	0.3%
S	Bradyrhizobium sp. URBM1 (1 OTU with 97% identity in 399bp to: Bradyrhizobium sp. URBM1)	0.3%	0.2%
g	Massilia (1 OTU with 98% identity in 424bp to: Massilia agilis, Massilia solisilvae, Massilia sp.)	0.3%	0.3%
g	Thiopfundum (1 OTU with 92% identity in 424bp to: Thiopfundum hispidum, Thiopfundum lithotrophicum)	0.3%	0.2%
f	Streptomyces (2 OTUs with 99-100% identity in 404bp to: 2 unclassified Streptomyces strains, Kitasatospora psammotica, Kitasatospora sp., Streptacidiphilus luteoalbus, Streptomyces gilvifuscus, Streptomyces lucensis, Streptomyces miharaensis, Streptomyces novae-caesareae, Streptomyces olivicoloratus, Streptomyces panaciradicis, Streptomyces recifensis, Streptomyces rhizophilus, Streptomyces seoulensis, Streptomyces similanensis, Streptomyces triostinicus)	0.3%	0.3%
S	Cellvibrio mixtus (1 OTU with 100% identity in 424bp to: Cellvibrio mixtus)	0.3%	0.3%
S	Pseudolabrys taiwanensis (1 OTU with 98% identity in 399bp to: Pseudolabrys taiwanensis)	0.3%	0.3%
S	Methylobacillus sp. zju323 (1 OTU with 96% identity in 424bp to: Methylobacillus sp. zju323)	0.3%	0.1%
S	Phaselicystis flava (1 OTU with 87% identity in 423bp to: Phaselicystis flava)	0.3%	0.3%
S	Steroidobacter sp. JC2986 (1 OTU with 97% identity in 424bp to: Steroidobacter sp. JC2986)	0.3%	0.3%
g	Arthrobacter (1 OTU with 100% identity in 412bp to: 2 unclassified Arthrobacter strains)	0.3%	0.3%
f	Pseudonocardiaceae (1 OTU with 100% identity in 404bp to: Lentzea aerocolonigenes, Lentzea kentuckyensis)	0.3%	0.3%
S	Parviterribacter sp. (1 OTU with 95% identity in 424bp to: Parviterribacter sp.)	0.3%	0.1%
S	Niabella sp. UI6 (1 OTU with 96% identity in 422bp to: Niabella sp. UI6)	0.3%	0.3%
S	Haliangium tepidum (1 OTU with 91% identity in 424bp to: Haliangium tepidum)	0.3%	0.3%
S	Candidatus Halomonas phosphatis (2 OTUs with 96-97% identity in 424bp to: Candidatus Halomonas phosphatis)	0.3%	0.5%
g	Rhodoplanes (1 OTU with 97% identity in 399bp to: Rhodoplanes piscinae, Rhodoplanes sp.)	0.2%	0.2%
g	Pedobacter (1 OTU with 98% identity in 422bp to: Pedobacter heparinus, Pedobacter sp.)	0.2%	0.3%
S	Mesorhizobium sp. M3(2013) (1 OTU with 100% identity in 399bp to: Mesorhizobium sp. M3(2013))	0.2%	0.2%
O	Burkholderiales (1 OTU with 99% identity in 424bp to: Aquabacterium sp. DNG29, Aquincola sp., Rubrivivax albus)	0.2%	0.4%
S	Achromatium minus (1 OTU with 85% identity in 424bp to: Achromatium minus)	0.2%	0.2%
S	Gemmatimonas sp. URHD0086 (1 OTU with 93% identity in 417bp to: Gemmatimonas sp. URHD0086)	0.2%	0.1%

s	Nordella sp. (1 OTU with 98% identity in 399bp to: Nordella sp.)	0.2%	0.2%
f	Chitinophagaceae (1 OTU with 94% identity in 422bp to: Flav- isolibacter sp. X7X, Lacibacter nakdongensis, Lacibacter sp. JJ009)	0.2%	0.3%
g	Chryseolinea (1 OTU with 92% identity in 422bp to: Chryseolinea serpens, Chryseolinea soli)	0.2%	0.3%
s	Chryseobacterium sp. (1 OTU with 100% identity in 422bp to: Chryseobacterium sp.)	0.2%	0.3%
s	Nocardioides sp. 805 (1 OTU with 98% identity in 418bp to: Nocardioides sp. 805)	0.2%	0.2%
s	Roseisolibacter agri (1 OTU with 90% identity in 416bp to: Roseisolibacter agri)	0.2%	0.1%
s	Terrimonas terrae (1 OTU with 97% identity in 422bp to: Terrimonas terrae)	0.2%	0.2%
s	Actinoplanes globisporus (1 OTU with 98% identity in 404bp to: Actinoplanes globisporus)	0.2%	0.2%
s	Polyangium sp. xj4 (2 OTUs with 87-94% identity in 424- 425bp to: Polyangium sp. xj4)	0.2%	0.3%
f	Rhodanobacteraceae (1 OTU with 98% identity in 424bp to: Dokdonella sp. TSY06, Tahibacter caeni)	0.2%	0.3%
s	Rhodovastum atsumiense (1 OTU with 97% identity in 399bp to: Rhodovastum atsumiense)	0.2%	0.2%
c	Betaproteobacteria (1 OTU with 91% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Glaciimonas alpina, Nitrosomonas mobilis)	0.2%	0.2%
s	Silvanigrella aquatica (1 OTU with 90% identity in 427bp to: Silvanigrella aquatica)	0.2%	0.2%
s	Lentimicrobium saccharophilum (1 OTU with 86% identity in 420bp to: Lentimicrobium saccharophilum)	0.2%	0.2%
g	Pseudolabrys (1 OTU with 100% identity in 399bp to: 3 unclas- sified Pseudolabrys strains)	0.2%	0.2%
s	Steroidobacter sp. (1 OTU with 93% identity in 424bp to: Steroidobacter sp.)	0.2%	0.2%
s	Arthrobacter sp. (1 OTU with 100% identity in 406bp to: Arthrobacter sp.)	0.2%	0.2%
s	Asticcacaulis sp. 148B (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. 148B)	0.2%	0.1%
s	Jahnella thaxteri (2 OTUs with 95-97% identity in 423-424bp to: Jahnella thaxteri)	0.2%	0.3%
s	Lacibacter cauensis (1 OTU with 95% identity in 422bp to: Lacibacter cauensis)	0.2%	0.2%
f	Acetobacteraceae (1 OTU with 95% identity in 400bp to: Acidisphaera sp. NO-15, Rhodopila globiformis)	0.2%	0.2%
s	Dokdonella soli (1 OTU with 99% identity in 424bp to: Dok- donella soli)	0.2%	0.2%
g	Mycobacterium (1 OTU with 100% identity in 404bp to: 2 un- classified Mycobacterium strains, Mycobacterium sacrum)	0.2%	0.1%
g	Mucilaginibacter (1 OTU with 99% identity in 422bp to: Mu- cilaginibacter sp. HME9300, Mucilaginibacter yixingensis)	0.2%	0.3%
s	Candidatus Scalindua marina (1 OTU with 94% identity in 400bp to: Candidatus Scalindua marina)	0.2%	0.2%
s	Reyranella soli (1 OTU with 99% identity in 399bp to: Reyranella soli)	0.2%	0.2%
s	Lysobacter sp. (1 OTU with 100% identity in 424bp to: Lysobac- ter sp.)	0.2%	0.2%
s	Desulforegula conservatrix (1 OTU with 86% identity in 425bp to: Desulforegula conservatrix)	0.1%	0.2%
s	Labilithrix luteola (1 OTU with 92% identity in 424bp to: La- bilithrix luteola)	0.1%	0.2%
s	Nakamurella sp. (1 OTU with 96% identity in 404bp to: Naka- murella sp.)	0.1%	0.1%
s	Flavobacterium sp. WB2.1-83 (1 OTU with 99% iden- tity in 422bp to: Flavobacterium sp. WB2.1-83)	0.1%	0.2%
s	Azospirillum sp. T2-YC6788 (1 OTU with 97% identity in 399bp to: Azospirillum sp. T2-YC6788)	0.1%	0.4%
g	Nocardia (1 OTU with 99% identity in 404bp to: 4 unclassified Nocardia strains, Nocardia veterana)	0.1%	0.1%
s	Dongia sp. URHE0060 (1 OTU with 98% identity in 399bp to: Dongia sp. URHE0060)	0.1%	0.2%
s	Chondromyces robustus (1 OTU with 88% identity in 423bp to: Chondromyces robustus)	0.1%	0.2%
s	Tepidisphaera mucosa (1 OTU with 82% identity in 402bp to: Tepidisphaera mucosa)	0.1%	0.2%
f	Micropepsaceae (1 OTU with 94% identity in 399bp to: Mi- cropepsis pineolensis, Rhizomicrobium electricum)	0.1%	0.1%
s	Edaphobaculum flavum (1 OTU with 96% identity in 422bp to: Edaphobaculum flavum)	0.1%	0.1%
s	Alicyclobacillus fodiniaquatilis (1 OTU with 100% iden- tity in 428bp to: Alicyclobacillus fodiniaquatilis)	0.1%	0.2%
s	Anaeromyxobacter dehalogenans (1 OTU with 88% identity in 426bp to: Anaeromyxobacter dehalogenans)	0.1%	0.1%
f	Xanthomonadaceae (1 OTU with 100% identity in 424bp to: Luteimonas sp., Lysobacter pocheonensis, Lysobacter sp. Is39)	0.1%	0.2%
s	Limisphaera ngatamarikiensis (1 OTU with 85% identity in 433bp to: Limisphaera ngatamarikiensis)	0.1%	0.2%

	Other	1.4%	2.3%
	Unclassified (0 reads)		
	Filtered (0 reads)		
952045.V3V4a (16 862 reads)			
s	Cellvibrio ostraviensis (9 OTUs with 99% identity in 422-424bp to: <i>Cellvibrio ostraviensis</i>)	11.3%	10.0%
s	Flavobacterium sp. S13 (17 OTUs with 99-100% identity in 419-422bp to: <i>Flavobacterium sp. S13</i>)	9.8%	14.9%
c	Gammaproteobacteria (8 OTUs with 99-100% identity in 422-424bp to: 10 unclassified <i>Pseudomonas</i> strains, <i>Enterobacter sp.</i> , <i>Pseudomonas aeruginosa</i> , <i>Pseudomonas citronellolis</i> , <i>Pseudomonas denitrificans</i> (nom. rej.), <i>Pseudomonas guguanensis</i> , <i>Pseudomonas knackmussii</i> , <i>Pseudomonas nitritireducens</i> , <i>Pseudomonas nitroreducens</i> , <i>Pseudomonas putida</i> , <i>Pseudomonas taiwanensis</i> , <i>Stenotrophomonas rhizophila</i>)	7.3%	6.5%
s	Cellvibrio mixtus (8 OTUs with 99-100% identity in 422-424bp to: <i>Cellvibrio mixtus</i>)	6.1%	5.4%
s	Flavobacterium sp. 71B1 (7 OTUs with 99-100% identity in 420-422bp to: <i>Flavobacterium sp. 71B1</i>)	3.3%	5.1%
g	Acidovorax (5 OTUs with 99-100% identity in 424bp to: 31 unclassified <i>Acidovorax</i> strains, <i>Acidovorax cattleyae</i> , <i>Acidovorax delafieldii</i> , <i>Acidovorax facilis</i> , <i>Acidovorax soli</i>)	3.0%	2.5%
g	Cellvibrio (5 OTUs with 99-100% identity in 424bp to: 2 unclassified <i>Cellvibrio</i> strains, <i>Cellvibrio fibrivorans</i>)	2.9%	2.6%
s	Haematococcus lacustris (1 OTU with 88% identity in 422bp to: <i>Haematococcus lacustris</i>)	2.9%	1.2%
f	Sphingomonadaceae (8 OTUs with 99-100% identity in 399-402bp to: 3 unclassified <i>Sphingopyxis</i> strains, 4 unclassified <i>Sphingobium</i> strains, 7 unclassified <i>Sphingomonas</i> strains, <i>Sphingomonas adhaesiva</i> , <i>Sphingopyxis soli</i> , <i>Sphingopyxis taejonensis</i>)	2.6%	2.3%
g	Flavobacterium (6 OTUs with 97-100% identity in 422bp to: 6 unclassified <i>Flavobacterium</i> strains, <i>Flavobacterium arsenitoxidans</i> , <i>Flavobacterium fluminis</i> , <i>Flavobacterium gilvum</i> , <i>Flavobacterium johnsoniae</i> , <i>Flavobacterium resistens</i>)	2.5%	3.8%
o	Rhizobiales (5 OTUs with 99-100% identity in 399bp to: 14 unclassified <i>Bosea</i> strains, 2 unclassified <i>Afipia</i> strains, 2 unclassified <i>Starkeya</i> strains, 5 unclassified <i>Agrobacterium</i> strains, 8 unclassified <i>Mesorhizobium</i> strains, 8 unclassified <i>Rhizobium</i> strains, <i>Agrobacterium deltaense</i> , <i>Agrobacterium tumefaciens</i> , <i>Beijerinckia fluminensis</i> , <i>Bosea minatitlanensis</i> , <i>Bosea robiniae</i> , <i>Bradyrhizobium sp. AMF3693</i> , <i>Candidatus Rhizobium massiliae</i> , <i>Prosthecomicrobium sp. ATCC 27825</i> , <i>Rhizobium genosp. TUXTLAS-27</i> , <i>Rhizobium grahamii</i> , <i>Rhizobium leguminosarum</i> , <i>Rhizobium nepotum</i> , <i>Rhizobium pusense</i> , <i>Rhizobium tibeticum</i>)	2.3%	2.2%
p	Proteobacteria (6 OTUs with 99-100% identity in 399-424bp to: 14 unclassified <i>Acidovorax</i> strains, 17 unclassified <i>Caulobacter</i> strains, 18 unclassified <i>Shinella</i> strains, 2 unclassified <i>Pseudomonas</i> strains, 4 unclassified <i>Rhizobium</i> strains, <i>Acidovorax anthurii</i> , <i>Acidovorax delafieldii</i> , <i>Agrobacterium sp. HY-35</i> , <i>Candidatus Nardonella dryophthoridicola</i> , <i>Caulobacter endosymbiont of Tamarixia radiata</i> , <i>Caulobacter segnis</i> , <i>Caulobacter vibrioides</i> , <i>Delftia acidovorans</i> , <i>Ensifer adhaerens</i> , <i>Ensifer sp. S75</i> , <i>Hydrogenophaga intermedia</i> , <i>Mycoplana sp. NC100</i> , <i>Phenylobacterium conjunctum</i> , <i>Rhizobium gallicum</i> , <i>Rhizobium yanglingense</i> , <i>Shinella granuli</i> , <i>Shinella zoogloeoides</i> , <i>Zoogloea ramigera</i> , <i>Zoogloea sp. 42HD3</i>)	1.9%	2.2%
s	Acidovorax sp. (1 OTU with 100% identity in 424bp to: <i>Acidovorax sp.</i>)	1.6%	1.4%
s	Cellvibrio sp. R-4001 (4 OTUs with 98-99% identity in 424bp to: <i>Cellvibrio sp. R-4001</i>)	1.5%	1.3%
s	Gemmatimonas aurantiaca (2 OTUs with 89-90% identity in 402bp to: <i>Gemmatimonas aurantiaca</i>)	1.3%	0.6%
f	Rhodospirillaceae (7 OTUs with 99-100% identity in 399bp to: 3 unclassified <i>Phaeospirillum</i> strains, 5 unclassified <i>Azospirillum</i> strains, 8 unclassified <i>Niveispirillum</i> strains, <i>Ferrovibrio soli</i> , <i>Ferrovibrio sp.</i> , <i>Ferrovibrio terrae</i> , <i>Niveispirillum cyanobacteriorum</i> , <i>Oleisolibacter albus</i> , <i>Rhodocista sp. W38</i> , <i>Rhodospirillum sp. SL38</i>)	1.3%	2.0%
s	Cellvibrio zantedeschiae (2 OTUs with 99% identity in 424bp to: <i>Cellvibrio zantedeschiae</i>)	1.3%	1.1%
f	Rhizobiaceae (5 OTUs with 99-100% identity in 399bp to: 17 unclassified <i>Rhizobium</i> strains, 2 unclassified <i>Ensifer</i> strains, 8 unclassified <i>Agrobacterium</i> strains, <i>Agrobacterium fabrum</i> , <i>Agrobacterium rosae</i> , <i>Agrobacterium rubi</i> , <i>Agrobacterium salinitolerans</i> , <i>Agrobacterium tumefaciens</i> , <i>Candidatus Rhizobium massiliae</i> , <i>Rhizobium aquaticum</i> , <i>Rhizobium nepotum</i> , <i>Rhizobium pusense</i> , <i>Rhizobium skierniewicense</i> , <i>Shinella fusca</i> , <i>Shinella zoogloeoides</i>)	1.2%	1.1%
g	Rhizobium (2 OTUs with 99-100% identity in 399bp to: 4 unclassified <i>Rhizobium</i> strains, <i>Rhizobium gallicum</i> , <i>Rhizobium oryzae</i> , <i>Rhizobium pseudoryzae</i> , <i>Rhizobium rhizoryzae</i>)	1.2%	1.3%
o	Burkholderiales (3 OTUs with 97-100% identity in 424bp to: 2 unclassified <i>Ideonella</i> strains, 2 unclassified <i>Variovorax</i> strains, 2 unclassified <i>Xylophilus</i> strains, 4 unclassified <i>Xenophilus</i> strains, <i>Aquabacterium sp. AKB-2008-KU6</i> , <i>Aquicola sp.</i> , <i>Aquicola tertiaricarbonis</i> , <i>Pseudorivibacter rhizosphaerae</i> , <i>Variovorax paradoxus</i> , <i>Variovorax soli</i> , <i>Xenophilus aerolatus</i> , <i>Xylophilus ampelinus</i>)	1.1%	1.6%
g	Bradyrhizobium (1 OTU with 100% identity in 399bp to: 4 unclassified <i>Bradyrhizobium</i> strains, <i>Bradyrhizobium elkanii</i> , <i>Bradyrhizobium jicamae</i> , <i>Bradyrhizobium paxllaeri</i>)	1.1%	0.6%

g	Pseudoxanthomonas (1 OTU with 100% identity in 424bp to: 4 unclassified Pseudoxanthomonas strains, Pseudoxanthomonas japonensis, Pseudoxanthomonas mexicana)	1.0%	0.4%
g	Devosia (3 OTUs with 100% identity in 399bp to: 22 unclassified Devosia strains, Devosia crocina, Devosia humi, Devosia riboflavina, Devosia soli)	1.0%	0.8%
s	Roseisolibacter agri (1 OTU with 90% identity in 416bp to: Roseisolibacter agri)	0.9%	0.4%
s	Sterolibacterium sp. AMi17 (1 OTU with 99% identity in 424bp to: Sterolibacterium sp. AMi17)	0.9%	0.4%
g	Phenylobacterium (1 OTU with 100% identity in 399bp to: Phenylobacterium immobile, Phenylobacterium sp.)	0.9%	0.4%
f	Micrococcaceae (2 OTUs with 100% identity in 411bp to: 47 unclassified Arthrobacter strains, Arthrobacter keyseri, Paenarthrobacter nicotinovorans, Paenarthrobacter ureafaciens)	0.8%	0.8%
s	Cellvibrio sp. R-4075 (2 OTUs with 99-100% identity in 424bp to: Cellvibrio sp. R-4075)	0.8%	0.7%
g	Nitrospira (1 OTU with 94% identity in 424bp to: 3 unclassified Nitrospira strains)	0.8%	0.3%
s	Rhizomicrobium sp. (1 OTU with 98% identity in 402bp to: Rhizomicrobium sp.)	0.7%	0.7%
s	Asticcacaulis sp. VA7 (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. VA7)	0.7%	0.6%
g	Thermomonas (1 OTU with 99% identity in 427bp to: 2 unclassified Thermomonas strains, Thermomonas brevis, Thermomonas haemolytica)	0.7%	0.8%
s	Pseudolabrys sp. W2.09-308 (2 OTUs with 97% identity in 399bp to: Pseudolabrys sp. W2.09-308)	0.6%	0.5%
g	Methylophilus (1 OTU with 100% identity in 424bp to: 3 unclassified Methylophilus strains, Methylophilus leisingeri)	0.6%	0.2%
s	Flavobacterium urocaniciphilum (2 OTUs with 97% identity in 422bp to: Flavobacterium urocaniciphilum)	0.6%	0.8%
f	Microbacteriaceae (1 OTU with 100% identity in 404bp to: 5 unclassified Leifsonia strains, Lysinimonas soli, Lysinimonas sp. 762H99, Salinibacterium sp.)	0.5%	0.5%
g	Hyphomicrobium (3 OTUs with 99-100% identity in 399bp to: 7 unclassified Hyphomicrobium strains, Hyphomicrobium denitrificans, Hyphomicrobium zavarzinii)	0.5%	0.3%
g	Arthrobacter (1 OTU with 100% identity in 412bp to: 2 unclassified Arthrobacter strains)	0.5%	0.4%
g	Pseudomonas (5 OTUs with 100% identity in 424bp to: 45 unclassified Pseudomonas strains, Pseudomonas alcaligenes, Pseudomonas chlororaphis, Pseudomonas entomophila, Pseudomonas fluorescens, Pseudomonas fulva, Pseudomonas granadensis, Pseudomonas humanensis, Pseudomonas jessenii, Pseudomonas koreensis, Pseudomonas kribbensis, Pseudomonas lutea, Pseudomonas mohnii, Pseudomonas moorei, Pseudomonas moraviensis, Pseudomonas plecoglossicida, Pseudomonas putida, Pseudomonas vancouverensis, Pseudomonas viridiflava)	0.5%	0.9%
s	Altererythrobacter sp. CB 286495 (1 OTU with 99% identity in 399bp to: Altererythrobacter sp. CB 286495)	0.5%	0.3%
s	Acidovorax sp. Sy22HNO (1 OTU with 100% identity in 424bp to: Acidovorax sp. Sy22HNO)	0.4%	0.4%
C	Alphaproteobacteria (2 OTUs with 99-100% identity in 399bp to: 34 unclassified Rhizobium strains, Rhizobium endophyticum, Rhizobium etli, Rhizobium grahamii, Rhizobium leguminosarum, Rhizobium metallidurans, Rhodospirillum sp. kmd.475)	0.4%	0.4%
s	Ignavibacterium album (3 OTUs with 81-83% identity in 423-424bp to: Ignavibacterium album)	0.4%	0.4%
C	Actinobacteria (1 OTU with 100% identity in 404bp to: 11 unclassified Mycobacterium strains, Micromonospora sp. NEAU-JG13, Mycolicibacterium rhodesiae)	0.4%	0.3%
g	Desulfuromonas (1 OTU with 86% identity in 425bp to: Desulfuromonas carbonis, Desulfuromonas sp. DDH964)	0.3%	0.1%
g	Nitrobacter (1 OTU with 100% identity in 399bp to: 4 unclassified Nitrobacter strains, Nitrobacter vulgaris)	0.3%	0.2%
s	Lautropia sp. CB 286320 (1 OTU with 99% identity in 424bp to: Lautropia sp. CB 286320)	0.3%	0.1%
g	Streptomyces (1 OTU with 100% identity in 404bp to: 9 unclassified Streptomyces strains, Streptomyces achromogenes, Streptomyces atrocyaneus, Streptomyces bungoensis, Streptomyces dioscori, Streptomyces ederensis, Streptomyces glomeroaurantiacus, Streptomyces humidus, Streptomyces lacrimifluminis, Streptomyces ossamyceticus, Streptomyces rishiriensis, Streptomyces ryensis, Streptomyces spinicoumarensis, Streptomyces tauricus, Streptomyces torulosus, Streptomyces turgidiscabies)	0.3%	0.3%
g	Nemorincola (1 OTU with 97% identity in 422bp to: Nemorincola caseinilytica, Nemorincola sp.)	0.3%	0.4%
g	Massilia (2 OTUs with 99% identity in 424bp to: Massilia sp. ScoD23, Naxibacter intermedius, Naxibacter sp. KF-26)	0.3%	0.3%
s	Acidobacterium sp. WY65 (1 OTU with 94% identity in 399bp to: Acidobacterium sp. WY65)	0.3%	0.2%
g	Sphingomonas (1 OTU with 99% identity in 399bp to: Sphingomonas soli, Sphingomonas sp. Cra4)	0.3%	0.3%
s	Rheinheimera sp. 03SU9 (1 OTU with 98% identity in 424bp to: Rheinheimera sp. 03SU9)	0.3%	0.2%

S	Lentimicrobium saccharophilum (1 OTU with 84% identity in 422bp to: Lentimicrobium saccharophilum)	0.3%	0.4%
S	Duganella sp. RCP6 (1 OTU with 100% identity in 424bp to: Duganella sp. RCP6)	0.3%	0.3%
S	Novosphingobium arabidopsis (1 OTU with 99% identity in 399bp to: Novosphingobium arabidopsis)	0.3%	0.2%
S	Terrimonas sp. CB 286475 (1 OTU with 97% identity in 422bp to: Terrimonas sp. CB 286475)	0.3%	0.3%
S	Devosia psychrophila (1 OTU with 100% identity in 399bp to: Devosia psychrophila)	0.3%	0.2%
S	Niastella sp. HME8655 (1 OTU with 93% identity in 422bp to: Niastella sp. HME8655)	0.3%	0.3%
S	Novosphingobium sp. TrD22 (2 OTUs with 99% identity in 399bp to: Novosphingobium sp. TrD22)	0.3%	0.2%
S	Enhydrobacter sp. (3 OTUs with 94-97% identity in 399bp to: Enhydrobacter sp.)	0.3%	0.3%
S	Curvibacter sp. R-36930 (1 OTU with 99% identity in 424bp to: Curvibacter sp. R-36930)	0.3%	0.2%
S	Haliangium tepidum (1 OTU with 90% identity in 424bp to: Haliangium tepidum)	0.3%	0.3%
S	Hydrogenophaga intermedia (1 OTU with 99% identity in 424bp to: Hydrogenophaga intermedia)	0.2%	0.2%
g	Bdellovibrio (2 OTUs with 93-98% identity in 405bp to: 24 unclassified Bdellovibrio strains, Bdellovibrio bacteriovorus)	0.2%	0.3%
S	Rhizobium sp. BZ3 (1 OTU with 100% identity in 399bp to: Rhizobium sp. BZ3)	0.2%	0.3%
S	Flavobacterium ginsenosidimutans (1 OTU with 100% identity in 422bp to: Flavobacterium ginsenosidimutans)	0.2%	0.4%
S	Reyranella sp. B6.10-109 (1 OTU with 98% identity in 399bp to: Reyranella sp. B6.10-109)	0.2%	0.3%
S	Desulfobulbus mediterraneus (1 OTU with 78% identity in 430bp to: Desulfobulbus mediterraneus)	0.2%	0.2%
S	Aquincola amnicola (2 OTUs with 99-100% identity in 424bp to: Aquincola amnicola)	0.2%	0.3%
S	Chitinophaga pinensis (1 OTU with 95% identity in 422bp to: Chitinophaga pinensis)	0.2%	0.2%
S	Taibaiella yonginensis (1 OTU with 98% identity in 423bp to: Taibaiella yonginensis)	0.2%	0.2%
S	Chryseolinea soli (1 OTU with 92% identity in 421bp to: Chryseolinea soli)	0.2%	0.2%
S	Prochlorococcus sp. SCGC AAA298-C17 (1 OTU with 86% identity in 422bp to: Prochlorococcus sp. SCGC AAA298-C17)	0.2%	0.1%
S	Pedobacter sp. JJ3205 (1 OTU with 85% identity in 422bp to: Pedobacter sp. JJ3205)	0.2%	0.3%
S	Paracoccus sp. 'CJSPY1 (P-I)' (1 OTU with 98% identity in 399bp to: Paracoccus sp. 'CJSPY1 (P-I)')	0.2%	0.2%
S	Paracoccus sp. 13636B (1 OTU with 98% identity in 399bp to: Paracoccus sp. 13636B)	0.2%	0.2%
S	Telluria mixta (1 OTU with 100% identity in 424bp to: Telluria mixta)	0.2%	0.2%
S	Sediminibacterium sp. (1 OTU with 99% identity in 422bp to: Sediminibacterium sp.)	0.2%	0.2%
g	Fluviicola (1 OTU with 99% identity in 422bp to: 2 unclassified Fluviicola strains, Fluviicola hefeinensis)	0.2%	0.2%
S	Mucibacter soli (1 OTU with 96% identity in 422bp to: Mucibacter soli)	0.2%	0.2%
g	Bacillus (1 OTU with 100% identity in 428bp to: 2 unclassified Bacillus strains, Bacillus amyloliquefaciens, Bacillus atrophaeus, Bacillus mojavensis, Bacillus siamensis, Bacillus subtilis, Bacillus velezensis)	0.2%	0.7%
g	Mesorhizobium (2 OTUs with 99-100% identity in 399bp to: 3 unclassified Mesorhizobium strains, alpha proteobacterium WG1)	0.2%	0.1%
f	Comamonadaceae (1 OTU with 99% identity in 424bp to: 2 unclassified Albidiferax strains, Curvibacter sp. R-36930)	0.2%	0.2%
S	Ohtaekwangia kribbensis (1 OTU with 93% identity in 421bp to: Ohtaekwangia kribbensis)	0.2%	0.2%
g	Azospirillum (1 OTU with 100% identity in 399bp to: 44 unclassified Azospirillum strains, Azospirillum humicireducens, Azospirillum largimobile, Azospirillum lipoferum, Azospirillum melinis, Azospirillum oryzae, Azospirillum zeae)	0.2%	0.5%
S	Paucibacter sp. (1 OTU with 98% identity in 424bp to: Paucibacter sp.)	0.2%	0.3%
S	Aquicella siphonis (1 OTU with 90% identity in 424bp to: Aquicella siphonis)	0.2%	0.1%
g	Nitrospira (1 OTU with 99% identity in 417bp to: Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira sp.)	0.2%	0.1%
S	Tepidisphaera mucosa (1 OTU with 82% identity in 402bp to: Tepidisphaera mucosa)	0.2%	0.2%
g	Sphingobium (1 OTU with 98% identity in 402bp to: Sphingobium sp. UR 7-05, Sphingobium vulgare)	0.2%	0.2%
g	Ideonella (2 OTUs with 99% identity in 424bp to: 3 unclassified Ideonella strains)	0.2%	0.2%
S	Frankia sp. (1 OTU with 83% identity in 405bp to: Frankia sp.)	0.2%	0.2%

f	Xanthomonadaceae (1 OTU with 100% identity in 424bp to: 3 unclassified Xanthomonas strains, Pseudoxanthomonas japonensis, Pseudoxanthomonas mexicana, Pseudoxanthomonas sp. Lo13)	0.2%	0.2%
O	Flavobacteriales (1 OTU with 100% identity in 422bp to: Chryseobacterium taeanense, Flavobacterium sp. AMS2)	0.2%	0.2%
S	Methylocapsa aurea (1 OTU with 96% identity in 400bp to: Methylocapsa aurea)	0.2%	0.1%
S	Novispirillum itersonii (1 OTU with 100% identity in 399bp to: Novispirillum itersonii)	0.2%	0.2%
S	Anaeromyxobacter dehalogenans (1 OTU with 88% identity in 420bp to: Anaeromyxobacter dehalogenans)	0.2%	0.2%
S	Taibaiella koreensis (1 OTU with 100% identity in 422bp to: Taibaiella koreensis)	0.1%	0.2%
S	Fluviicola kyonggii (1 OTU with 99% identity in 422bp to: Fluviicola kyonggii)	0.1%	0.2%
g	Blastochloris (1 OTU with 95% identity in 400bp to: Blastochloris sulfoviridis, Blastochloris viridis)	0.1%	0.1%
S	Mesorhizobium sp. (1 OTU with 100% identity in 399bp to: Mesorhizobium sp.)	0.1%	0.1%
S	Sphingomonas sp. IJ1 (1 OTU with 99% identity in 402bp to: Sphingomonas sp. IJ1)	0.1%	0.1%
k	Bacteria (1 OTU with 100% identity in 422bp to: Brevibacillus sp. KNF6, Dyadobacter soli)	0.1%	0.1%
S	Nocardioides sp. URHD0098 (1 OTU with 99% identity in 420bp to: Nocardioides sp. URHD0098)	0.1%	0.1%
S	Conexibacter sp. (1 OTU with 98% identity in 424bp to: Conexibacter sp.)	0.1%	0.1%
g	Chitinimonas (1 OTU with 100% identity in 424bp to: 4 unclassified Chitinimonas strains, Chitinimonas viridis)	0.1%	0.3%
S	Candidatus Ovatusbacter abovo (1 OTU with 92% identity in 426bp to: Candidatus Ovatusbacter abovo)	0.1%	0.1%
S	Bdellovibrio exovorus (1 OTU with 98% identity in 405bp to: Bdellovibrio exovorus)	0.1%	0.2%
g	Sediminibacterium (1 OTU with 98% identity in 422bp to: 2 unclassified Sediminibacterium strains, Sediminibacterium goheungense)	0.1%	0.2%
S	Flavobacterium hydatidis (1 OTU with 98% identity in 426bp to: Flavobacterium hydatidis)	0.1%	0.2%
g	Pseudolabrys (1 OTU with 100% identity in 399bp to: 3 unclassified Pseudolabrys strains)	0.1%	0.1%
S	Cytophaga sp. SSL03 (1 OTU with 99% identity in 422bp to: Cytophaga sp. SSL03)	0.1%	0.1%
S	Phaeodactylibacter luteus (1 OTU with 88% identity in 422bp to: Phaeodactylibacter luteus)	0.1%	0.1%
S	Dongia sp. URHE0060 (1 OTU with 99% identity in 399bp to: Dongia sp. URHE0060)	0.1%	0.2%
C	Betaproteobacteria (1 OTU with 91% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Glaciimonas alpina, Nitrosomonas mobilis)	0.1%	0.1%
S	Ferruginibacter paludis (1 OTU with 97% identity in 422bp to: Ferruginibacter paludis)	0.1%	0.1%
S	Flavobacterium sp. (1 OTU with 97% identity in 422bp to: Flavobacterium sp.)	0.1%	0.2%
g	Aquabacterium (1 OTU with 98% identity in 424bp to: 5 unclassified Aquabacterium strains)	0.1%	0.2%
S	Legionella sp. KR-14 (1 OTU with 95% identity in 424bp to: Legionella sp. KR-14)	0.1%	0.2%
S	Fluviicola taffensis (1 OTU with 96% identity in 422bp to: Fluviicola taffensis)	0.1%	0.1%
g	Runella (1 OTU with 100% identity in 422bp to: Runella palustris, Runella sp.)	0.1%	0.1%
g	Ferrovibrio (1 OTU with 100% identity in 399bp to: Ferrovibrio soli, Ferrovibrio terrae)	0.1%	0.1%
S	Duganella radidis (1 OTU with 99% identity in 424bp to: Duganella radidis)	0.1%	0.1%
	Other	1.0%	1.7%
	Unclassified (0 reads)		
	Filtered (0 reads)		

952046.V3V4a (15 931 reads)

S	Cellvibrio mixtus (6 OTUs with 99-100% identity in 422-424bp to: Cellvibrio mixtus)	5.9%	5.5%
S	Acidobacterium sp. WY65 (3 OTUs with 93-94% identity in 399bp to: Acidobacterium sp. WY65)	4.9%	2.6%
g	Streptomyces (6 OTUs with 99-100% identity in 404bp to: 13 unclassified Streptomyces strains, Streptomyces coelicoflavus, Streptomyces coeruleus, Streptomyces collinus, Streptomyces diastaticus, Streptomyces griseoaurantiacus, Streptomyces leeuwenhoekii, Streptomyces missionensis, Streptomyces monticola, Streptomyces niveoruber, Streptomyces spinoverrucosus, Streptomyces tendae, Streptomyces thermocarboxydovorans, Streptomyces thermodiastaticus, Streptomyces thermoviolaceus, Streptomyces violarus)	4.6%	4.4%
S	Flavobacterium sp. S13 (9 OTUs with 99-100% identity in 419-422bp to: Flavobacterium sp. S13)	4.3%	7.0%

s	Ignavibacterium album (5 OTUs with 80-83% identity in 420-423bp to: Ignavibacterium album)	2.5%	2.7%
s	Flavobacterium sp. 71B1 (7 OTUs with 99-100% identity in 420-422bp to: Flavobacterium sp. 71B1)	2.4%	3.9%
s	Chryseolinea soli (6 OTUs with 90-92% identity in 421-422bp to: Chryseolinea soli)	2.0%	2.4%
s	Devosia sp. (8 OTUs with 98-100% identity in 399-400bp to: Devosia sp.)	2.0%	1.8%
s	Steroidobacter sp. JC2953 (3 OTUs with 92-95% identity in 424bp to: Steroidobacter sp. JC2953)	1.8%	1.7%
f	Sphingomonadaceae (5 OTUs with 99-100% identity in 399-402bp to: 3 unclassified Sphingopyxis strains, 4 unclassified Sphingobium strains, 7 unclassified Sphingomonas strains, Sphingomonas adhaesiva, Sphingopyxis soli, Sphingopyxis taejonensis)	1.8%	1.8%
s	Asprobacter aquaticus (6 OTUs with 97-99% identity in 399bp to: Asprobacter aquaticus)	1.8%	1.4%
g	Nitrosospira (2 OTUs with 94% identity in 424bp to: 8 unclassified Nitrosospira strains, Nitrosospira lacus)	1.7%	0.8%
s	Cellvibrio ostraviensis (4 OTUs with 99% identity in 424bp to: Cellvibrio ostraviensis)	1.6%	1.5%
s	Haematococcus lacustris (2 OTUs with 88-89% identity in 422bp to: Haematococcus lacustris)	1.5%	0.7%
s	Conexibacter sp. (3 OTUs with 96-98% identity in 424bp to: Conexibacter sp.)	1.5%	0.7%
s	Nocardioides sp. (3 OTUs with 97-100% identity in 404-420bp to: Nocardioides sp.)	1.5%	1.4%
g	Phenylobacterium (1 OTU with 100% identity in 399bp to: Phenylobacterium immobile, Phenylobacterium sp.)	1.4%	0.6%
s	Taibaiella yonginensis (2 OTUs with 98% identity in 423bp to: Taibaiella yonginensis)	1.4%	1.6%
c	Alphaproteobacteria (4 OTUs with 99-100% identity in 399bp to: 34 unclassified Rhizobium strains, Rhizobium endophyticum, Rhizobium etli, Rhizobium grahamii, Rhizobium leguminosarum, Rhizobium metallidurans, Rhodobium sp. 6, Rhodospirillum sp. kmd.475, Skermanella sp. R224-3)	1.4%	1.4%
s	Cellvibrio gandavensis (5 OTUs with 96-99% identity in 424bp to: Cellvibrio gandavensis)	1.4%	1.3%
s	Pseudolabrys sp. W2.09-308 (5 OTUs with 97-98% identity in 399bp to: Pseudolabrys sp. W2.09-308)	1.3%	1.3%
s	Terrimonas sp. CB 286475 (1 OTU with 97% identity in 422bp to: Terrimonas sp. CB 286475)	1.1%	1.3%
s	Rhizomicrobium sp. (2 OTUs with 98-99% identity in 402bp to: Rhizomicrobium sp.)	1.1%	1.0%
f	Microbacteriaceae (2 OTUs with 100% identity in 404bp to: 19 unclassified Agreia strains, 4 unclassified Clavibacter strains, 9 unclassified Leifsonia strains, 9 unclassified Salinibacterium strains, Galbitalea soli, Lysinimonas soli, Lysinimonas sp. 762H99, Microbacterium sp. ANT-2400 S2, Salinibacterium amurskyense)	1.0%	0.9%
g	Flavobacterium (3 OTUs with 98-99% identity in 422bp to: 14 unclassified Flavobacterium strains, Flavobacterium arsenitoxidans, Flavobacterium gilvum, Flavobacterium limi, Flavobacterium quisquiliarum)	0.9%	1.5%
f	Rhodospirillaceae (6 OTUs with 92-100% identity in 399bp to: 2 unclassified Novispirillum strains, 3 unclassified Azospirillum strains, 3 unclassified Phaeospirillum strains, 8 unclassified Niveispirillum strains, Ferroviobrio soli, Ferroviobrio sp., Ferroviobrio terrae, Haematospirillum jordaniae, Niveispirillum cyanobacteriorum, Rhodocista sp. W38, Rhodospirillum sp. SL38)	0.9%	1.5%
g	Lentzea (1 OTU with 100% identity in 404bp to: 11 unclassified Lentzea strains, Lentzea albida, Lentzea albidocapillata, Lentzea californiensis, Lentzea flaviverrucosa, Lentzea violacea, Lentzea waywayandensis)	0.9%	0.8%
c	Actinobacteria (1 OTU with 100% identity in 404bp to: 9 unclassified Streptomyces strains, Actinomyces sp., Streptomyces atriruber, Streptomyces costaricanus, Streptomyces endophyticus, Streptomyces lanatus, Streptomyces prasinopilosus, Streptomyces rochei, Streptomyces venezuelae, Streptomyces xiangtanensis)	0.9%	0.8%
s	Arthrobacter sp. (2 OTUs with 100% identity in 404-406bp to: Arthrobacter sp.)	0.8%	0.8%
g	Caulobacter (2 OTUs with 98-100% identity in 399bp to: 38 unclassified Caulobacter strains, Caulobacter flavus, Caulobacter hibisci, Caulobacter profundus, Caulobacter radialis, Caulobacter segnis, Caulobacter vibrioides, Caulobacter zeae)	0.8%	0.7%
p	Bacteroidetes (1 OTU with 87% identity in 424bp to: Owenweekia hongkongensis, Solitalea canadensis)	0.8%	0.9%
g	Rhizobium (2 OTUs with 99-100% identity in 399bp to: 4 unclassified Rhizobium strains, Rhizobium gallicum, Rhizobium oryzae, Rhizobium pseudoryzae, Rhizobium rhizoryzae)	0.8%	0.9%
s	Aciditerrimonas ferrireducens (3 OTUs with 91-92% identity in 401bp to: Aciditerrimonas ferrireducens)	0.8%	0.7%
s	Gemmatimonas sp. URHD0086 (2 OTUs with 92% identity in 417bp to: Gemmatimonas sp. URHD0086)	0.8%	0.3%
s	Racemicystis persica (3 OTUs with 92-94% identity in 423-425bp to: Racemicystis persica)	0.8%	1.3%
s	Nocardioides sp. URHD0098 (2 OTUs with 99-100% identity in 420bp to: Nocardioides sp. URHD0098)	0.7%	0.7%

s	Roseisolibacter agri (1 OTU with 90% identity in 416bp to: Roseisolibacter agri)	0.7%	0.3%
f	Micrococcaceae (2 OTUs with 100% identity in 411bp to: 47 unclassified Arthrobacter strains, Arthrobacter keyseri, Paenarthrobacter nicotinovorans, Paenarthrobacter ureafaciens)	0.7%	0.7%
s	Pseudolabrys sp. (3 OTUs with 97-100% identity in 399bp to: Pseudolabrys sp.)	0.7%	0.7%
s	Candidatus Solibacter usitatus (2 OTUs with 91-96% identity in 399bp to: Candidatus Solibacter usitatus)	0.7%	0.4%
s	Streptococcus anginosus (1 OTU with 84% identity in 401bp to: Streptococcus anginosus)	0.7%	0.3%
s	Acidobacterium ailaui (2 OTUs with 99-100% identity in 399bp to: Acidobacterium ailaui)	0.7%	0.4%
g	Cellvibrio (3 OTUs with 99-100% identity in 424bp to: 9 unclassified Cellvibrio strains, Cellvibrio fibrivorans)	0.6%	0.6%
g	Bradyrhizobium (1 OTU with 100% identity in 399bp to: 3 unclassified Bradyrhizobium strains, Bradyrhizobium canariense, Bradyrhizobium diazoefficiens, Bradyrhizobium erythrophlei, Bradyrhizobium japonicum, Bradyrhizobium lupini)	0.6%	0.3%
f	Rhizobiaceae (2 OTUs with 100% identity in 399bp to: 2 unclassified Agrobacterium strains, 3 unclassified Rhizobium strains, Agrobacterium fabrum, Agrobacterium salinitolerans, Agrobacterium tumefaciens, Candidatus Rhizobium massiliae, Rhizobium pusense, Shinella fusca, Shinella zoogloeoides)	0.6%	0.6%
s	Nitrosospira sp. REGAU (1 OTU with 94% identity in 424bp to: Nitrosospira sp. REGAU)	0.6%	0.3%
s	Aliidongia dinghuensis (2 OTUs with 95% identity in 400-401bp to: Aliidongia dinghuensis)	0.6%	1.0%
s	Asticcacaulis sp. VA7 (1 OTU with 100% identity in 399bp to: Asticcacaulis sp. VA7)	0.6%	0.5%
g	Nocardia (1 OTU with 100% identity in 404bp to: Nocardia cerra-doensis, Nocardia elegans, Nocardia kruczakiae, Nocardia mikamii, Nocardia sp., Nocardia veterana)	0.6%	0.5%
g	Chryseolinea (1 OTU with 91% identity in 421bp to: Chryseolinea serpens, Chryseolinea soli)	0.5%	0.6%
g	Methylosinus (1 OTU with 97% identity in 399bp to: Methylosinus sp. NCIMB 13214, Methylosinus sporium)	0.5%	0.4%
s	Rhizomicrobium sp. URHD0033 (1 OTU with 97% identity in 402bp to: Rhizomicrobium sp. URHD0033)	0.5%	0.5%
s	Blastochloris sp. AT2101 (1 OTU with 94% identity in 399bp to: Blastochloris sp. AT2101)	0.5%	0.3%
s	Chryseolinea flava (1 OTU with 90% identity in 421bp to: Chryseolinea flava)	0.5%	0.6%
s	Hyphomicrobium sp. LAT3 (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. LAT3)	0.5%	0.3%
s	Paludibaculum fermentans (1 OTU with 99% identity in 399bp to: Paludibaculum fermentans)	0.5%	0.2%
s	Telluria mixta (1 OTU with 100% identity in 424bp to: Telluria mixta)	0.5%	0.5%
s	Sideroxydans lithotrophicus (1 OTU with 92% identity in 424bp to: Sideroxydans lithotrophicus)	0.5%	0.2%
s	Candidatus Thiodiazotropha endolucinida (2 OTUs with 90% identity in 424bp to: Candidatus Thiodiazotropha endolucinida)	0.4%	0.4%
s	Mesorhizobium plurifarum (2 OTUs with 99-100% identity in 399bp to: Mesorhizobium plurifarum)	0.4%	0.3%
s	Pseudolabrys taiwanensis (1 OTU with 98% identity in 399bp to: Pseudolabrys taiwanensis)	0.4%	0.4%
s	Sphingobacterium sp. P-38 (2 OTUs with 99-100% identity in 422bp to: Sphingobacterium sp. P-38)	0.4%	0.2%
s	Cavicella sp. (2 OTUs with 91% identity in 424bp to: Cavicella sp.)	0.4%	0.7%
s	Geothrix sp. (1 OTU with 97% identity in 424bp to: Geothrix sp.)	0.4%	0.2%
s	Candidatus Microthrix calida (1 OTU with 91% identity in 401bp to: Candidatus Microthrix calida)	0.4%	0.3%
f	Chitinophagaceae (2 OTUs with 94% identity in 422bp to: Edaphobaculum flavum, Mucibacter soli, Niabella pedocola, Niabella sp. R2A43-1, Niabella sp. KP03)	0.4%	0.4%
s	Alicyclobacillus kakegawensis (4 OTUs with 99% identity in 427bp to: Alicyclobacillus kakegawensis)	0.4%	0.7%
g	Devosia (2 OTUs with 99-100% identity in 399bp to: 14 unclassified Devosia strains, Devosia humi)	0.4%	0.3%
s	Haliangium tepidum (3 OTUs with 90-91% identity in 424-425bp to: Haliangium tepidum)	0.4%	0.4%
s	Hyphomicrobium nitrativorans (1 OTU with 99% identity in 399bp to: Hyphomicrobium nitrativorans)	0.4%	0.2%
s	Nitrosomonas sp. (1 OTU with 93% identity in 424bp to: Nitrosomonas sp.)	0.4%	0.2%
s	Sterolibacterium sp. AMi17 (1 OTU with 97% identity in 424bp to: Sterolibacterium sp. AMi17)	0.4%	0.2%
s	Candidatus Nitrosarchaeum limnium (1 OTU with 95% identity in 400bp to: Candidatus Nitrosarchaeum limnium)	0.4%	0.2%
s	Lentimicrobium saccharophilum (2 OTUs with 88-90% identity in 415-423bp to: Lentimicrobium saccharophilum)	0.4%	0.5%

S	Reyranella sp. B6.10-109 (1 OTU with 98% identity in 399bp to: Reyranella sp. B6.10-109)	0.4%	0.5%
S	Cytophaga aurantiaca (2 OTUs with 98-99% identity in 421bp to: Cytophaga aurantiaca)	0.4%	0.4%
S	Cellvibrio sp. R-4001 (2 OTUs with 99% identity in 424bp to: Cellvibrio sp. R-4001)	0.4%	0.3%
S	Steroidobacter sp. JC2986 (2 OTUs with 97% identity in 424bp to: Steroidobacter sp. JC2986)	0.4%	0.3%
S	Gemmatimonas aurantiaca (1 OTU with 93% identity in 402bp to: Gemmatimonas aurantiaca)	0.3%	0.2%
S	Bauldia sp. (1 OTU with 98% identity in 402bp to: Bauldia sp.)	0.3%	0.3%
g	Fluviicola (2 OTUs with 99% identity in 422bp to: Fluviicola hefeinensis, Fluviicola sp.)	0.3%	0.4%
S	Candidatus Koribacter versatilis (1 OTU with 93% identity in 399bp to: Candidatus Koribacter versatilis)	0.3%	0.2%
P	Proteobacteria (2 OTUs with 95-100% identity in 399bp to: 18 unclassified Shinella strains, 3 unclassified Methylosinus strains, 4 unclassified Rhizobium strains, Agrobacterium sp. HY-35, Ensifer adhaerens, Ensifer sp. S75, Methylomonas sp. GYJ3, Methylosinus trichosporium, Mycoplasma sp. NC100, Rhizobium gallicum, Rhizobium yanglingense, Shinella granulii, Shinella zoogloeoides, Zoogloea ramigera, Zoogloea sp. 42HD3)	0.3%	0.4%
S	Rhodopila globiformis (2 OTUs with 96% identity in 399-400bp to: Rhodopila globiformis)	0.3%	0.4%
S	Ohtaekwangia kribbensis (1 OTU with 93% identity in 421bp to: Ohtaekwangia kribbensis)	0.3%	0.4%
g	Alicyclobacillus (3 OTUs with 100% identity in 427-428bp to: 18 unclassified Alicyclobacillus strains, Alicyclobacillus acidocaldarius, Alicyclobacillus contaminans, Alicyclobacillus genomosp. 1, Alicyclobacillus mali)	0.3%	0.6%
S	Hyphomicrobium sp. AT3 (1 OTU with 100% identity in 399bp to: Hyphomicrobium sp. AT3)	0.3%	0.2%
C	Betaproteobacteria (1 OTU with 91% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Glaciimonas alpina, Nitrosomonas mobilis)	0.3%	0.3%
S	Phaselicystis flava (2 OTUs with 87-91% identity in 423-424bp to: Phaselicystis flava)	0.3%	0.3%
S	Emticicia sp. (1 OTU with 100% identity in 423bp to: Emticicia sp.)	0.3%	0.3%
O	Rhizobiales (1 OTU with 100% identity in 399bp to: 2 unclassified Rhizobium strains, 5 unclassified Agrobacterium strains, Agrobacterium deltaense, Agrobacterium tumefaciens, Beijerinckia fluminensis, Candidatus Rhizobium massiliae, Rhizobium leguminosarum, Rhizobium nepotum, Rhizobium pusense)	0.3%	0.3%
S	Terrimonas sp. (1 OTU with 97% identity in 422bp to: Terrimonas sp.)	0.3%	0.3%
S	Alicyclobacillus shizuokensis (2 OTUs with 99-100% identity in 427bp to: Alicyclobacillus shizuokensis)	0.3%	0.6%
S	Azospirillum sp. T2-YC6788 (2 OTUs with 97% identity in 399bp to: Azospirillum sp. T2-YC6788)	0.3%	0.9%
S	Cytophaga hutchinsonii (1 OTU with 98% identity in 421bp to: Cytophaga hutchinsonii)	0.3%	0.3%
S	Methylocystis sp. SH31p (1 OTU with 96% identity in 400bp to: Methylocystis sp. SH31p)	0.3%	0.3%
S	Mesorhizobium sp. K6-22 (1 OTU with 99% identity in 399bp to: Mesorhizobium sp. K6-22)	0.3%	0.2%
g	Thiopfundum (1 OTU with 92% identity in 424bp to: Thiopfundum hispidum, Thiopfundum lithotrophicum)	0.3%	0.2%
g	Acidisphaera (2 OTUs with 97% identity in 399bp to: 3 unclassified Acidisphaera strains, Acidisphaera rubrifaciens)	0.3%	0.3%
g	Thermomonas (2 OTUs with 98-99% identity in 427bp to: 4 unclassified Thermomonas strains, Thermomonas brevis, Thermomonas haemolytica)	0.3%	0.3%
g	Nocardioides (1 OTU with 100% identity in 413bp to: 7 unclassified Nocardioides strains, Nocardioides aquiterrae, Nocardioides kribbensis, Nocardioides pyridinolyticus)	0.3%	0.2%
g	Sphingobacterium (1 OTU with 89% identity in 422bp to: Sphingobacterium sp. GR11, Sphingobacterium tabacisoli)	0.2%	0.1%
S	Ferruginibacter paludis (1 OTU with 97% identity in 422bp to: Ferruginibacter paludis)	0.2%	0.3%
S	Dongia sp. URHE0060 (1 OTU with 99% identity in 399bp to: Dongia sp. URHE0060)	0.2%	0.4%
g	Ideonella (2 OTUs with 99-100% identity in 424bp to: 6 unclassified Ideonella strains)	0.2%	0.4%
g	Rhodovibrio (1 OTU with 92% identity in 400bp to: 14 unclassified Rhodovibrio strains, Rhodovibrio salinarum, Rhodovibrio sodomensis)	0.2%	0.1%
S	Reyranella sp. (1 OTU with 99% identity in 400bp to: Reyranella sp.)	0.2%	0.3%
S	Acidisphaera sp. (2 OTUs with 97% identity in 399bp to: Acidisphaera sp.)	0.2%	0.3%
g	Acidovorax (1 OTU with 100% identity in 424bp to: 10 unclassified Acidovorax strains, Acidovorax delafieldii, Acidovorax facilis)	0.2%	0.2%
O	Myxococcales (1 OTU with 87% identity in 425bp to: Chondromyces pediculatus, Sandaracinus amylolyticus)	0.2%	0.2%
S	Frankia sp. (1 OTU with 83% identity in 405bp to: Frankia sp.)	0.2%	0.2%

g	Inquilinus (1 OTU with 95% identity in 399bp to: <i>Inquilinus limosus</i> , <i>Inquilinus</i> sp. RB35)	0.2%	0.3%
s	Anaeromyxobacter dehalogenans (1 OTU with 88% identity in 420bp to: <i>Anaeromyxobacter dehalogenans</i>)	0.2%	0.2%
s	Fluviicola taffensis (1 OTU with 96% identity in 422bp to: <i>Fluviicola taffensis</i>)	0.2%	0.2%
s	Mesorhizobium sp. PETBA10 (1 OTU with 99% identity in 399bp to: <i>Mesorhizobium</i> sp. PETBA10)	0.2%	0.2%
s	Flavobacterium lindanitolerans (1 OTU with 93% identity in 422bp to: <i>Flavobacterium lindanitolerans</i>)	0.2%	0.3%
s	Chondromyces lanuginosus (1 OTU with 91% identity in 423bp to: <i>Chondromyces lanuginosus</i>)	0.2%	0.3%
s	Sediminibacterium sp. (1 OTU with 99% identity in 422bp to: <i>Sediminibacterium</i> sp.)	0.2%	0.2%
s	Inquilinus limosus (1 OTU with 96% identity in 399bp to: <i>Inquilinus limosus</i>)	0.2%	0.2%
s	Bradymonas sp. YN101 (1 OTU with 83% identity in 416bp to: <i>Bradymonas</i> sp. YN101)	0.2%	0.2%
s	Lysobacter sp. (1 OTU with 99% identity in 424bp to: <i>Lysobacter</i> sp.)	0.2%	0.2%
s	Lysobacter sp. C26(2011) (1 OTU with 99% identity in 424bp to: <i>Lysobacter</i> sp. C26(2011))	0.2%	0.2%
s	Sandaracinus amylolyticus (1 OTU with 89% identity in 428bp to: <i>Sandaracinus amylolyticus</i>)	0.2%	0.2%
s	Flavobacterium urocaniciphilum (1 OTU with 97% identity in 422bp to: <i>Flavobacterium urocaniciphilum</i>)	0.2%	0.3%
f	Comamonadaceae (1 OTU with 99% identity in 424bp to: 2 unclassified <i>Albidiferax</i> strains, <i>Curvibacter</i> sp. R-36930)	0.2%	0.1%
g	Amycolatopsis (2 OTUs with 96-100% identity in 404-406bp to: 16 unclassified <i>Amycolatopsis</i> strains, <i>Amycolatopsis balhimycina</i> , <i>Amycolatopsis vastitatis</i>)	0.2%	0.1%
g	Actinospica (1 OTU with 99% identity in 404bp to: 3 unclassified <i>Actinospica</i> strains, <i>Actinospica robiniae</i>)	0.2%	0.1%
g	Sediminibacterium (1 OTU with 98% identity in 422bp to: 2 unclassified <i>Sediminibacterium</i> strains, <i>Sediminibacterium goeungense</i>)	0.2%	0.2%
g	Nemorincola (1 OTU with 97% identity in 422bp to: <i>Nemorincola caseinilytica</i> , <i>Nemorincola</i> sp.)	0.1%	0.2%
s	Phaeodactylibacter luteus (1 OTU with 88% identity in 422bp to: <i>Phaeodactylibacter luteus</i>)	0.1%	0.2%
g	Pseudolabrys (1 OTU with 100% identity in 399bp to: 3 unclassified <i>Pseudolabrys</i> strains)	0.1%	0.1%
s	Duganella sp. RCP6 (1 OTU with 100% identity in 424bp to: <i>Duganella</i> sp. RCP6)	0.1%	0.1%
s	Chondromyces apiculatus (1 OTU with 92% identity in 424bp to: <i>Chondromyces apiculatus</i>)	0.1%	0.2%
o	Burkholderiales (1 OTU with 99% identity in 424bp to: <i>Aquabacterium</i> sp. DNG29, <i>Aquincola</i> sp., <i>Rubrivivax albus</i>)	0.1%	0.2%
g	Mycobacterium (1 OTU with 100% identity in 416bp to: <i>Mycobacterium attenuatum</i> , <i>Mycobacterium avium</i> , <i>Mycobacterium chimera</i> , <i>Mycobacterium colombiense</i> , <i>Mycobacterium intracellulare</i> , <i>Mycobacterium paraintracellulare</i> , <i>Mycobacterium pseudokansasii</i> , <i>Mycobacterium</i> sp.)	0.1%	0.1%
g	Pseudonocardia (1 OTU with 100% identity in 405bp to: 4 unclassified <i>Pseudonocardia</i> strains, <i>Pseudonocardia dioxanivorans</i> , <i>Pseudonocardia sulfidoxydans</i>)	0.1%	0.1%
s	Asticcacaulis sp. 148B (1 OTU with 100% identity in 399bp to: <i>Asticcacaulis</i> sp. 148B)	0.1%	0.1%
s	Methylocystis sp. WI14 (1 OTU with 94% identity in 399bp to: <i>Methylocystis</i> sp. WI14)	0.1%	0.1%
s	Reyranella terrae (1 OTU with 99% identity in 399bp to: <i>Reyranella terrae</i>)	0.1%	0.2%
s	Cellvibrio zantedeschiae (1 OTU with 99% identity in 424bp to: <i>Cellvibrio zantedeschiae</i>)	0.1%	0.1%
s	Bacillus sp. (in (1 OTU with 100% identity in 428bp to: <i>Bacillus</i> sp. (in	0.1%	0.4%
s	Chryseolinea serpens (1 OTU with 92% identity in 421bp to: <i>Chryseolinea serpens</i>)	0.1%	0.1%
s	Vicingus serpentipes (1 OTU with 88% identity in 422bp to: <i>Vicingus serpentipes</i>)	0.1%	0.1%
g	Lewinella (1 OTU with 85% identity in 426bp to: <i>Lewinella litorea</i> , <i>Lewinella marina</i> , <i>Lewinella</i> sp. SRO_346)	0.1%	0.1%
s	Mesorhizobium sp. (1 OTU with 99% identity in 399bp to: <i>Mesorhizobium</i> sp.)	0.1%	0.1%
s	Desulforegula conservatrix (1 OTU with 86% identity in 425bp to: <i>Desulforegula conservatrix</i>)	0.1%	0.2%
g	Kofleria (1 OTU with 89% identity in 424bp to: <i>Kofleria flava</i> , <i>Kofleria</i> sp. DSM 53797)	0.1%	0.1%
s	Lysobacter sp. WF-2 (1 OTU with 100% identity in 424bp to: <i>Lysobacter</i> sp. WF-2)	0.1%	0.1%
s	Iamia sp. T2-YC6790 (1 OTU with 98% identity in 400bp to: <i>Iamia</i> sp. T2-YC6790)	0.1%	0.1%
s	Lewinella cohaerens (1 OTU with 86% identity in 424bp to: <i>Lewinella cohaerens</i>)	0.1%	0.1%

O	Bacillales (1 OTU with 100% identity in 428bp to: 7 unclassified Alicyclobacillus strains, 9 unclassified Bacillus strains, Alicyclobacillus acidoterrestris)	0.1%	0.3%
S	Alicyclobacillus macrosporangiidus (2 OTUs with 99-100% identity in 428bp to: Alicyclobacillus macrosporangiidus)	0.1%	0.2%
S	Caldicoprobacter algeriensis (1 OTU with 84% identity in 401bp to: Caldicoprobacter algeriensis)	0.1%	0.2%
S	Geobacter hydrogenophilus (1 OTU with 86% identity in 426bp to: Geobacter hydrogenophilus)	0.1%	0.1%
S	Candidatus Allobeggiatoa halophila (1 OTU with 88% identity in 424bp to: Candidatus Allobeggiatoa halophila)	0.1%	0.1%
g	Dyella (1 OTU with 99% identity in 424bp to: 3 unclassified Dyella strains, Dyella ginsengisoli)	0.1%	0.2%
S	Limisphaera ngatamarikiensis (1 OTU with 83% identity in 427bp to: Limisphaera ngatamarikiensis)	0.1%	0.2%
S	Ohtaekwangia koreensis (1 OTU with 95% identity in 422bp to: Ohtaekwangia koreensis)	0.1%	0.1%
g	Ferrovibrio (1 OTU with 100% identity in 399bp to: Ferrovibrio soli, Ferrovibrio terrae)	0.1%	0.2%
S	Rhodospirillum sp. (1 OTU with 85% identity in 404bp to: Rhodospirillum sp.)	0.1%	0.1%
g	Massilia (1 OTU with 98% identity in 424bp to: Massilia agilis, Massilia solisilvae, Massilia sp.)	0.1%	0.1%
g	Chondromyces (1 OTU with 88% identity in 425bp to: Chondromyces apiculatus, Chondromyces robustus)	0.1%	0.2%
	Other	1.2%	2.1%
	Unclassified (0 reads)		
	Filtered (0 reads)		

Table 4: Condensed overview of the taxonomic composition of samples.

This table can be found as a file in the results directory. Please see the according section for details about result files.

2.2 Methods

As a first step of the microbiome analysis, all reads with ambiguous bases ("N") were removed. Chimeric reads were identified and removed based on the de-novo algorithm of UCHIME (Edgar RC et al., 2011) as implemented in the VSEARCH package (Rognes T et al., 2016).

The remaining set of high-quality reads was processed using minimum entropy decomposition (Eren AM, 2013 and 2015). Minimum Entropy Decomposition (MED) provides a computationally efficient means to partition marker gene datasets into OTUs (Operational Taxonomic Units). Each OTU represents a distinct cluster with significant sequence divergence to any other cluster. By employing Shannon entropy, MED uses only the information-rich nucleotide positions across reads and iteratively partitions large datasets while omitting stochastic variation. The MED procedure outperforms classical, identity based clustering algorithms. Sequences can be partitioned based on relevant single nucleotide differences without being susceptible to random sequencing errors. **This allows a decomposition of sequence data sets with a single nucleotide resolution.** Furthermore, the MED procedure identifies and filters random "noise" in the dataset, i.e. sequences with a very low abundance (less than $\approx 0.02\%$ of the average sample size).

To assign taxonomic information to each OTU, DC-MEGABLAST alignments of cluster representative sequences to the sequence database were performed. A most specific taxonomic assignment for each OTU was then transferred from the set of best-matching reference sequences (lowest common taxonomic unit of all best hits). Hereby, a sequence identity of 70% across at least 80% of the representative sequence was a minimal requirement for considering reference sequences.

Further processing of OTUs and taxonomic assignments was performed using the QIIME software package (version 1.9.1, <http://qiime.org/>). Abundances of bacterial taxonomic units were normalized using lineage-specific copy numbers of the relevant marker genes to improve estimates (Angly FE, 2014).

OTU-picking strategy: Minimum entropy decomposition

Reference database: /mnt/nsa3/projects/active/bioit_development/ebe_transfer/mdxMicrobiomeProfiling/ncbi_nt/n02-03_well_classified_only/nt.filtered.fa (Release 2020-02-03)

References:

- **OTU picking:** Eren AM et al. (2013). Oligotyping: differentiating between closely related microbial taxa using 16s rRNA gene data. *Methods Ecol Evol* (4), 1111-1119.
Eren AM et al. (2015) Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME Journal* advance online publication, doi: 10.1038/ismej.2014.195.
- **Taxonomic assignment:** Altschul SF et al. (1990) Basic local alignment search tool. *J Mol Biol* 215(3), 403-410.
- **QIIME:** Caporaso JG et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7(5), 335-336.
- **Chimera detection:**
Rognes T et al. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584 <https://doi.org/10.7717/peerj.2584>.
Edgar RC et al. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16), 2194-2200.
- **Copy number correction:** Angly FE et al. (2014) CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* 2:11.

2.3 Output Files and Descriptions

The *MicrobiomeProfiling* directory contains the result files. All relevant files are described below. Some of these descriptions were excerpted from the official QIIME tutorials (<http://qiime.org/tutorials/index.html>).

01_Taxonomy_shortlist.txt: One of the **main results** of the microbiome analysis. This file can be used to get a quick overview of the microbiome. It contains a summarized list of identified taxonomic units for each sample. The first two columns are the sample name and the total number of reads that were assigned to OTUs in this sample. The following columns list all taxonomic units with at least 0.1% of reads assigned to them. The individual columns state:

- The number of reads assigned to the taxonomic unit.
- The number of different OTUs that were classified as this taxonomic unit.
- The taxonomic level of the taxonomic unit. One of k...kingdom, p...phylum, c...class, o...order, f...family, g...genus, s...species.
- The abundance-corrected fraction of reads assigned to the taxonomic unit.
- The fraction of reads assigned to the taxonomic unit.
- The identity and length of the best BLAST hit(s) to the database and a list of species that match with these alignment scores (not for all analysis types).

All taxonomic units with less than 0.1% of reads are collapsed in the category "Other". If the representative sequence of an OTU had no significant database match, no taxonomic unit could be assigned. The total number of reads of these unclassified OTUs is stated as category "Unclassified".

Depending on the type of analysis, some taxonomic units might be removed as they do not match the expected clade, e.g. eukaryotes in a bacterial microbiome analysis. The number of removed reads is stated as category "Filtered". If this category is not listed, no filtering was performed.

Please consider the provided identity and length of the best BLAST hits. The stated taxonomic unit was derived as lowest common ancestor of the best hits, but in case of a low sequence identity, it might be more appropriate to assign a higher taxonomic level than that of the lowest common ancestor.

02_Taxonomy_table.txt: One of the **main results** of the microbiome analysis. There is one line for each taxonomic unit and one column for each sample. The entries of the matrix are the estimated abundances of the respective taxonomic unit/sample combination. The file can be imported into Excel for further processing (sorting, calculations, diagrams).

03_OTU_representative_sequences.fasta: One of the **main results** of the microbiome analysis. Contains all read sequences of OTU representatives in FASTA format. The FASTA header contains the OTU identifier, the read identifier of the representative, the number of reads in the corresponding OTU, and the taxonomic classification. Representatives without taxonomic assignment are marked as "Unassigned", "Unclassified" or as "NOHIT", depending on the OTU picking method. Please note that representative sequences are not sample specific, i.e. a representative read subsumes similar reads of all samples. Thus, the given number of reads is the total number of reads of all samples that were assigned to the corresponding OTU.

Please note that OTUs only subsume sequences with identical lengths. Thus, OTU representatives may be prefixes of other OTU representatives. This occurs if assembled read pairs and (unassembled) single reads are processed together.

04_OTU_table.biom: One of the **main results** of the microbiome analysis. A file in BIOM format (<http://biom-format.org/>). This file is used as input by many QIIME scripts and is useful for downstream processing. OTUs of all samples are contained in this file.

05_OTU_table.txt: There is one line for each OTU and one column for each sample. The entries of the matrix are the estimated abundances of the respective OTU/sample combinations. The last column

contains the taxonomic assignment of the OTU. OTUs without taxonomic assignment are marked as "Unassigned", "Unclassified", or "NOHIT", depending on the OTU picking method. Please see file 02_Taxonomy_table.txt for the abundances per taxonomic unit and sample. The file can be imported into Excel for further processing (sorting, calculations, diagrams).

06_OTU_table_summary.txt: Contains a summary describing 05_OTU_table.txt.

07_OTU_table_per_sample_statistics.txt: Contains statistics for each sample in 05_OTU_table.txt.

08_Processed_reads.fasta.gz: Contains all read sequences in FASTA format that went into the OTU-picking process. Reads that were identified as chimeric are not contained in this file. Processed-read identifiers consist of the sample name and a sequential number, followed by the raw-read identifier and the length of the read. Reads of all samples are contained in this file.

09_OTU_read_assignment.txt: A mapping of OTU identifier to read identifier, i.e. each line represents one OTU, the first column contains the OTU identifier, all other columns contain the identifier of reads that are part of the OTU. OTUs/Reads of all samples are contained in this file.

10_Taxonomy_plots: This directory contains files area_charts.html and bar_charts.html. These files can be opened with any web browser. The data of 02_Taxonomy_table.txt (as relative abundances) will be displayed as either area or bar chart plots. There are several plots, each for a different level of taxonomy: from phylum to species. Hereby, higher level plots give a more coarse-grained view on the data than lower level plots. Mouseover the plots to see which taxa are contributing to the percentage shown, and a click on the hyperlinks in the legend starts a web-search using the most specific taxonomic unit. Charts, legends, and tables can be exported by clicking on the respective hyperlinks.

