eurofins | Genomics

Data Analysis Report: Microbiome Profiling

Project / Study: NG-33317

Date: May 15, 2023

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 my 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	571 411	100.0%
Remaining sequences after preprocessing and quality filtering	571 318	100.0%
Remaining sequences after chimera detection and filtering	569 491	99.7%
Total number of sequences assigned to OTUs	432 957	75.8%
Total number of sequences assigned to taxa	431 196	75.5%
Copy-number corrected total count	242 852	-
Total number of OTUs	519	100.0%
Number of OTUs assigend to taxa	517	99.6%

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)**
R1.Ende.V3V4a	71 568	99.9%	73.0%	73.0%	37 789	417
R1.Mitte.V3V4a	71 786	99.9%	73.4%	73.4%	38 913	417
R1.Start.V3V4a	71 300	99.9%	72.7%	72.7%	36 696	419
R1.Sumpf.V3V4a	71 458	99.9%	75.6%	75.5%	38 869	417
R2.Ende.V3V4a	71 230	99.5%	79.6%	78.6%	22 213	424
R2.Mitte.V3V4a	71 047	99.5%	78.6%	77.8%	22 428	424
R2.Start.V3V4a	71 455	99.5%	77.5%	77.0%	23 648	424
R2.Sumpf.V3V4a	71 567	99.1%	75.7%	75.6%	22 296	424

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 to 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 perfor 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)

Taxonomic Level	Taxonomic Unit	Normalized Fraction	Raw Fraction
R1.Ende.V3V4a (37 789 reads)			
g	**Nitrospira** (14 OTUs with 98-100% identity in 417bp to: 3 unclassified Nitrospira strains, Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira defluvii, Nitrospira marina)	**33.0%**	23.9%
s	**Nitrospira japonica** (9 OTUs with 97-99% identity in 419bp to: Nitrospira japonica)	**10.4%**	7.5%
g	**Nitrosomonas** (8 OTUs with 99-100% identity in 424bp to: 3 unclassified Nitrosomonas strains, Nitrosomonas mobilis, Nitrosomonas ureae)	**9.6%**	7.0%
s	**Nitrosomonas sp. Nm86** (2 OTUs with 98% identity in 424bp to: Nitrosomonas sp. Nm86)	**3.6%**	2.6%
s	**Paludibaculum sp.** (9 OTUs with 86-95% identity in 399bp to: Paludibaculum sp.)	**2.7%**	2.3%
g	**Nitrosospira** (3 OTUs with 92-99% identity in 424bp to: 9 unclassified Nitrosospira strains, Nitrosospira multiformis)	**2.7%**	1.9%
s	**Chryseolinea serpens** (6 OTUs with 90-95% identity in 421-423bp to: Chryseolinea serpens)	**2.2%**	4.2%
s	**Gemmatimonas aurantiaca** (8 OTUs with 89-95% identity in 402-413bp to: Gemmatimonas aurantiaca)	**2.1%**	1.5%
s	**Holophaga sp. WY42** (4 OTUs with 89-97% identity in 423-425bp to: Holophaga sp. WY42)	**1.6%**	1.2%
s	**Halochromatium sp. MTK6IM088** (3 OTUs with 88-89% identity in 424-426bp to: Halochromatium sp. MTK6IM088)	**1.6%**	2.1%
s	**Candidatus Amoebophilus asiaticus** (2 OTUs with 93% identity in 421bp to: Candidatus Amoebophilus asiaticus)	**1.5%**	2.8%
s	**Nitrospira moscoviensis** (3 OTUs with 98-99% identity in 417-419bp to: Nitrospira moscoviensis)	**1.2%**	0.9%
s	**Acidobacterium sp. WY65** (3 OTUs with 90-95% identity in 399bp to: Acidobacterium sp. WY65)	**1.1%**	0.9%
s	**Nitrosomonas ureae** (1 OTU with 100% identity in 424bp to: Nitrosomonas ureae)	**0.9%**	0.6%
s	**Elioraea tepidiphila** (2 OTUs with 100% identity in 399-413bp to: Elioraea tepidiphila)	**0.9%**	1.9%
s	**Gemmatimonas phototrophica** (2 OTUs with 91-92% identity in 416bp to: Gemmatimonas phototrophica)	**0.8%**	0.6%
g	**Pedomicrobium** (1 OTU with 96% identity in 399bp to: 2 unclassified Pedomicrobium strains, Pedomicrobium americanum)	**0.7%**	0.7%
s	**Ignavibacterium album** (4 OTUs with 80-84% identity in 422-423bp to: Ignavibacterium album)	**0.7%**	1.2%
s	**Roseisolibacter agri** (2 OTUs with 91-93% identity in 416bp to: Roseisolibacter agri)	**0.7%**	0.5%
s	**Hypericibacter terrae** (1 OTU with 94% identity in 399bp to: Hypericibacter terrae)	**0.7%**	1.8%
g	**Geobacter** (3 OTUs with 88-90% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.7%**	1.4%
g	**Hyphomicrobium** (2 OTUs with 95-99% identity in 399bp to: 5 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.6%**	0.6%
s	**Dongia sp. URHE0060** (2 OTUs with 94% identity in 399-413bp to: Dongia sp. URHE0060)	**0.6%**	1.7%

s	**Woodsholea maritima** (3 OTUs with 93-98% identity in 399bp to: Woodsholea maritima)	**0.6%**	0.6%
s	**Reyranella soli** (2 OTUs with 99% identity in 399bp to: Reyranella soli)	**0.6%**	1.4%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.6%**	0.4%
s	**Hyphomicrobium sp.** (3 OTUs with 97-99% identity in 399bp to: Hyphomicrobium sp.)	**0.6%**	0.6%
o	**Fragilariales** (1 OTU with 98% identity in 402bp to: Fragilaria pinnata, Nanofrustulum shiloi)	**0.5%**	0.4%
s	**Nitrospirillum amazonense** (4 OTUs with 89-97% identity in 399-400bp to: Nitrospirillum amazonense)	**0.5%**	1.3%
s	**Micavibrio aeruginosavorus** (2 OTUs with 91% identity in 402bp to: Micavibrio aeruginosavorus)	**0.5%**	0.8%
g	**Thauera** (6 OTUs with 99-100% identity in 424bp to: 11 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**0.4%**	0.9%
s	**Pelobacter carbinolicus** (2 OTUs with 87-89% identity in 425-426bp to: Pelobacter carbinolicus)	**0.4%**	0.3%
s	**Vermiphilus pyriformis** (2 OTUs with 90-91% identity in 427bp to: Vermiphilus pyriformis)	**0.4%**	0.7%
s	**Nitrospira calida** (2 OTUs with 96% identity in 415-417bp to: Nitrospira calida)	**0.4%**	0.3%
s	**Pandoraea sp.** (3 OTUs with 92-94% identity in 424bp to: Pandoraea sp.)	**0.4%**	1.5%
g	**Phenylobacterium** (1 OTU with 89% identity in 403bp to: Phenylobacterium koreense, Phenylobacterium sp. S140)	**0.4%**	0.3%
s	**Parvularcula sp. P33** (5 OTUs with 94-95% identity in 399bp to: Parvularcula sp. P33)	**0.4%**	0.5%
s	**Nitrosomonas sp.** (2 OTUs with 94-95% identity in 424-425bp to: Nitrosomonas sp.)	**0.3%**	0.2%
s	**Bauldia consociata** (2 OTUs with 96-99% identity in 401bp to: Bauldia consociata)	**0.3%**	0.6%
s	**Conexibacter sp.** (1 OTU with 92% identity in 424bp to: Conexibacter sp.)	**0.3%**	0.2%
o	**Rhodocyclales** (1 OTU with 91% identity in 424bp to: 3 unclassified Azoarcus strains, Aromatoleum aromaticum, Aromatoleum buckelii, Azoarcus olearius)	**0.3%**	0.7%
s	**Endomicrobium sp.** (1 OTU with 80% identity in 420bp to: Endomicrobium sp.)	**0.3%**	0.2%
s	**Enhydrobacter sp.** (3 OTUs with 91-94% identity in 399-400bp to: Enhydrobacter sp.)	**0.3%**	0.7%
s	**Blastochloris viridis** (1 OTU with 95% identity in 399bp to: Blastochloris viridis)	**0.3%**	0.3%
s	**Thiobacter sp. SL-1** (1 OTU with 92% identity in 425bp to: Thiobacter sp. SL-1)	**0.3%**	0.7%
c	**Betaproteobacteria** (1 OTU with 92% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Azospira sp. AH, Glaciimonas alpina)	**0.3%**	0.5%
s	**Desulfovibrio cavernae** (2 OTUs with 88-89% identity in 429bp to: Desulfovibrio cavernae)	**0.3%**	0.5%
s	**Desulfuromonas sp. AOP6** (1 OTU with 87% identity in 427bp to: Desulfuromonas sp. AOP6)	**0.3%**	0.2%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.3%**	0.3%
g	**Bradyrhizobium** (1 OTU with 100% identity in 399bp to: 4 unclassified Bradyrhizobium strains, Bradyrhizobium elkanii, Bradyrhizobium jicamae, Bradyrhizobium paxllaeri)	**0.2%**	0.2%
s	**Nitrospira sp. Ecomares 2.1** (1 OTU with 95% identity in 417bp to: Nitrospira sp. Ecomares 2.1)	**0.2%**	0.2%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.2%**	0.6%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.2%**	0.2%
g	**Lewinella** (1 OTU with 87% identity in 424bp to: Lewinella marina, Lewinella sp. SRO_346)	**0.2%**	0.4%
s	**Mesorhizobium sp. PETBA10** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp. PETBA10)	**0.2%**	0.3%
s	**Microcystis sp. SAG 43.90** (1 OTU with 91% identity in 416bp to: Microcystis sp. SAG 43.90)	**0.2%**	0.2%
s	**Fodinicurvata sp.** (1 OTU with 92% identity in 399bp to: Fodinicurvata sp.)	**0.2%**	0.6%
g	**Rhodovibrio** (1 OTU with 93% identity in 400bp to: 3 unclassified Rhodovibrio strains)	**0.2%**	0.1%
s	**Blastochloris sp. AT2101** (2 OTUs with 95% identity in 400bp to: Blastochloris sp. AT2101)	**0.2%**	0.2%
s	**Hyphomicrobium sp. Pd-S-(I)-e-D-8(2)** (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. Pd-S-(I)-e-D-8(2))	**0.2%**	0.2%
o	**Rhizobiales** (2 OTUs with 95-100% identity in 399-400bp to: 10 unclassified Bradyrhizobium strains, 3 unclassified Nordella strains, Bradyrhizobium genosp. Y, Bradyrhizobium vignae, Methylocapsa aurea, Nordella oligomobilis, Rhodoligotrophus sp. RIPI3)	**0.2%**	0.3%
s	**Asprobacter aquaticus** (1 OTU with 99% identity in 399bp to: Asprobacter aquaticus)	**0.2%**	0.2%

s	**Steroidobacter sp.** (1 OTU with 94% identity in 424bp to: Steroidobacter sp.)	**0.2%**	0.3%
s	**Bdellovibrio sp. S1** (1 OTU with 89% identity in 405bp to: Bdellovibrio sp. S1)	**0.2%**	0.4%
o	**Cytophagales** (2 OTUs with 89% identity in 421bp to: Chryseolinea soli, Chryseotalea sanaruensis, Flexibacter flexilis)	**0.2%**	0.3%
s	**Aciditerrimonas ferrireducens** (2 OTUs with 90-92% identity in 401bp to: Aciditerrimonas ferrireducens)	**0.2%**	0.2%
g	**Chondromyces** (2 OTUs with 86% identity in 425bp to: Chondromyces apiculatus, Chondromyces pediculatus, Chondromyces robustus)	**0.2%**	0.5%
s	**Aquicella siphonis** (1 OTU with 93% identity in 425bp to: Aquicella siphonis)	**0.2%**	0.1%
g	**Mesorhizobium** (1 OTU with 100% identity in 399bp to: 2 unclassified Mesorhizobium strains, alpha proteobacterium WG1)	**0.2%**	0.2%
s	**Ilumatobacter fluminis** (2 OTUs with 97% identity in 400bp to: Ilumatobacter fluminis)	**0.2%**	0.2%
s	**Sphingomonas sp.** (1 OTU with 97% identity in 399bp to: Sphingomonas sp.)	**0.2%**	0.2%
s	**Nitrosomonas sp. Is343** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp. Is343)	**0.2%**	0.1%
f	**Sphingomonadaceae** (1 OTU with 98% identity in 399bp to: 3 unclassified Sphingomonas strains, Sphingosinicella cucumeris)	**0.1%**	0.2%
f	**Myxococcaceae** (1 OTU with 87% identity in 424bp to: Corallococcus macrosporus, Corallococcus sp. xj73, Myxococcus sp. 004-95, Myxococcus virescens)	**0.1%**	0.3%
s	**Haematospirillum sp.** (2 OTUs with 96-97% identity in 371bp to: Haematospirillum sp.)	**0.1%**	0.3%
g	**Sphingomonas** (1 OTU with 91% identity in 399bp to: 4 unclassified Sphingomonas strains, Sphingomonas hankyongensis, Sphingomonas sediminicola)	**0.1%**	0.2%
s	**Curvibacter sp. R-36930** (1 OTU with 99% identity in 424bp to: Curvibacter sp. R-36930)	**0.1%**	0.2%
g	**Pseudomonas** (2 OTUs with 100% identity in 424bp to: 10 unclassified Pseudomonas strains, Pseudomonas fluorescens, Pseudomonas hunanensis, Pseudomonas juntendi, Pseudomonas monteilii, Pseudomonas mosselii, Pseudomonas plecoglossicida, Pseudomonas putida, Pseudomonas taiwanensis)	**0.1%**	0.4%
s	**Roseomonas sp.** (2 OTUs with 95% identity in 399bp to: Roseomonas sp.)	**0.1%**	0.2%
s	**Georgfuchsia toluolica** (1 OTU with 93% identity in 424bp to: Georgfuchsia toluolica)	**0.1%**	0.1%
s	**Pseudorhodoplanes sinuspersici** (1 OTU with 99% identity in 399bp to: Pseudorhodoplanes sinuspersici)	**0.1%**	0.2%
s	**Nordella sp.** (1 OTU with 97% identity in 399bp to: Nordella sp.)	**0.1%**	0.2%
s	**Methyloceanibacter caenitepidi** (1 OTU with 96% identity in 399bp to: Methyloceanibacter caenitepidi)	**0.1%**	0.2%
s	**Candidatus Navis piranensis** (2 OTUs with 92-93% identity in 399-400bp to: Candidatus Navis piranensis)	**0.1%**	0.2%
	Other	**3.1%**	6.0%
	Unclassified (0 reads)		
	Filtered (0 reads)		

R1.Mitte.V3V4a (38 913 reads)

g	**Nitrospira** (20 OTUs with 98-100% identity in 417bp to: 3 unclassified Nitrospira strains, Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira defluvii, Nitrospira marina)	**30.6%**	22.6%
s	**Nitrospira japonica** (9 OTUs with 97-99% identity in 419bp to: Nitrospira japonica)	**17.4%**	12.8%
g	**Nitrosomonas** (6 OTUs with 99-100% identity in 424bp to: 3 unclassified Nitrosomonas strains, Nitrosomonas mobilis, Nitrosomonas ureae)	**7.8%**	5.7%
s	**Paludibaculum sp.** (9 OTUs with 86-94% identity in 399bp to: Paludibaculum sp.)	**3.2%**	2.7%
s	**Nitrosomonas sp. Nm86** (1 OTU with 98% identity in 424bp to: Nitrosomonas sp. Nm86)	**2.9%**	2.2%
s	**Chryseolinea serpens** (7 OTUs with 90-95% identity in 421-423bp to: Chryseolinea serpens)	**2.2%**	4.3%
g	**Nitrosospira** (2 OTUs with 92% identity in 424bp to: 6 unclassified Nitrosospira strains)	**2.0%**	1.5%
s	**Halochromatium sp. MTK6IM088** (5 OTUs with 88-89% identity in 424-426bp to: Halochromatium sp. MTK6IM088)	**1.9%**	2.5%
s	**Gemmatimonas aurantiaca** (8 OTUs with 89-95% identity in 402bp to: Gemmatimonas aurantiaca)	**1.7%**	1.3%
s	**Holophaga sp. WY42** (4 OTUs with 89-97% identity in 423-425bp to: Holophaga sp. WY42)	**1.4%**	1.1%
s	**Nitrospira moscoviensis** (3 OTUs with 98-99% identity in 417-419bp to: Nitrospira moscoviensis)	**1.4%**	1.0%
g	**Pedomicrobium** (2 OTUs with 96-97% identity in 399-400bp to: 3 unclassified Pedomicrobium strains, Pedomicrobium americanum, Pedomicrobium australicum)	**1.0%**	1.1%
s	**Candidatus Amoebophilus asiaticus** (1 OTU with 93% identity in 421bp to: Candidatus Amoebophilus asiaticus)	**0.9%**	1.7%

s	**Dongia sp. URHE0060** (2 OTUs with 94% identity in 399-413bp to: Dongia sp. URHE0060)	**0.9%**	2.4%
s	**Acidobacterium sp. WY65** (2 OTUs with 90-92% identity in 399bp to: Acidobacterium sp. WY65)	**0.9%**	0.8%
s	**Roseisolibacter agri** (1 OTU with 93% identity in 416bp to: Roseisolibacter agri)	**0.8%**	0.6%
s	**Hypericibacter terrae** (4 OTUs with 94% identity in 399bp to: Hypericibacter terrae)	**0.8%**	2.2%
s	**Gemmatimonas phototrophica** (1 OTU with 92% identity in 416bp to: Gemmatimonas phototrophica)	**0.8%**	0.6%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (2 OTUs with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.7%**	0.5%
s	**Nitrosomonas ureae** (1 OTU with 100% identity in 424bp to: Nitrosomonas ureae)	**0.7%**	0.5%
s	**Elioraea tepidiphila** (1 OTU with 100% identity in 399bp to: Elioraea tepidiphila)	**0.6%**	1.4%
s	**Woodsholea maritima** (4 OTUs with 93-100% identity in 399bp to: Woodsholea maritima)	**0.6%**	0.6%
s	**Reyranella soli** (2 OTUs with 99% identity in 399bp to: Reyranella soli)	**0.5%**	1.2%
g	**Geobacter** (2 OTUs with 88-90% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.5%**	1.1%
s	**Micavibrio aeruginosavorus** (1 OTU with 91% identity in 402bp to: Micavibrio aeruginosavorus)	**0.5%**	0.8%
g	**Thauera** (5 OTUs with 99-100% identity in 424bp to: 11 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**0.5%**	1.1%
g	**Hyphomicrobium** (1 OTU with 99% identity in 399bp to: 4 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.5%**	0.5%
s	**Hyphomicrobium sp.** (3 OTUs with 97-99% identity in 399bp to: Hyphomicrobium sp.)	**0.5%**	0.5%
s	**Desulfuromonas sp. AOP6** (2 OTUs with 87% identity in 427-429bp to: Desulfuromonas sp. AOP6)	**0.5%**	0.3%
s	**Ignavibacterium album** (2 OTUs with 82-84% identity in 422-423bp to: Ignavibacterium album)	**0.4%**	0.8%
s	**Nitrosomonas sp.** (2 OTUs with 95-99% identity in 424-425bp to: Nitrosomonas sp.)	**0.4%**	0.3%
s	**Nitrospirillum amazonense** (3 OTUs with 89-97% identity in 399-400bp to: Nitrospirillum amazonense)	**0.4%**	1.1%
s	**Desulfovibrio cavernae** (1 OTU with 89% identity in 429bp to: Desulfovibrio cavernae)	**0.4%**	0.7%
s	**Pandoraea sp.** (3 OTUs with 92-94% identity in 424bp to: Pandoraea sp.)	**0.4%**	1.4%
s	**Pelobacter carbinolicus** (2 OTUs with 87-89% identity in 425-426bp to: Pelobacter carbinolicus)	**0.4%**	0.3%
o	**Fragilariales** (1 OTU with 98% identity in 402bp to: Fragilaria pinnata, Nanofrustulum shiloi)	**0.3%**	0.3%
s	**Nitrospira calida** (1 OTU with 96% identity in 415bp to: Nitrospira calida)	**0.3%**	0.2%
s	**Vermiphilus pyriformis** (1 OTU with 90% identity in 427bp to: Vermiphilus pyriformis)	**0.3%**	0.6%
s	**Thiobacter sp. SL-1** (1 OTU with 92% identity in 425bp to: Thiobacter sp. SL-1)	**0.3%**	0.8%
g	**Phenylobacterium** (1 OTU with 89% identity in 403bp to: Phenylobacterium koreense, Phenylobacterium sp. S140)	**0.3%**	0.2%
s	**Steroidobacter sp.** (1 OTU with 94% identity in 424bp to: Steroidobacter sp.)	**0.3%**	0.5%
c	**Betaproteobacteria** (2 OTUs with 92-98% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, 3 unclassified Simplicispira strains, Aquaspirillum sp. R-22832, Azospira sp. AH, Glaciimonas alpina)	**0.3%**	0.5%
s	**Endomicrobium sp.** (1 OTU with 80% identity in 420bp to: Endomicrobium sp.)	**0.3%**	0.2%
o	**Rhodocyclales** (1 OTU with 91% identity in 424bp to: 3 unclassified Azoarcus strains, Aromatoleum aromaticum, Aromatoleum buckelii, Azoarcus olearius)	**0.3%**	0.6%
s	**Parvularcula sp. P33** (3 OTUs with 94-95% identity in 399bp to: Parvularcula sp. P33)	**0.3%**	0.5%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.3%**	0.2%
s	**Blastochloris viridis** (1 OTU with 95% identity in 399bp to: Blastochloris viridis)	**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	**Enhydrobacter sp.** (3 OTUs with 91-94% identity in 399-400bp to: Enhydrobacter sp.)	**0.3%**	0.6%
o	**Cytophagales** (3 OTUs with 89-94% identity in 421bp to: Chryseolinea soli, Chryseotalea sanaruensis, Flexibacter flexilis, Ohtaekwangia koreensis)	**0.3%**	0.5%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.2%**	0.7%
s	**Blastochloris sp. AT2101** (2 OTUs with 95% identity in 400bp to: Blastochloris sp. AT2101)	**0.2%**	0.3%

s	**Bauldia consociata** (1 OTU with 96% identity in 401bp to: Bauldia consociata)	**0.2%**	0.4%
s	**Aquicella siphonis** (2 OTUs with 91-93% identity in 424-425bp to: Aquicella siphonis)	**0.2%**	0.2%
s	**Mesorhizobium sp. PETBA10** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp. PETBA10)	**0.2%**	0.3%
g	**Lewinella** (1 OTU with 87% identity in 424bp to: Lewinella marina, Lewinella sp. SRO_346)	**0.2%**	0.4%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.2%**	0.3%
s	**Hyphomicrobium sp. Pd-S-(I)-e-D-8(2)** (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. Pd-S-(I)-e-D-8(2))	**0.2%**	0.2%
g	**Gaiella** (1 OTU with 96% identity in 423bp to: 2 unclassified Gaiella strains)	**0.2%**	0.1%
s	**Microcystis sp. SAG 43.90** (1 OTU with 91% identity in 416bp to: Microcystis sp. SAG 43.90)	**0.2%**	0.2%
s	**Pseudorhodoplanes sinuspersici** (1 OTU with 99% identity in 399bp to: Pseudorhodoplanes sinuspersici)	**0.2%**	0.3%
g	**Nitrobacter** (1 OTU with 100% identity in 399bp to: 9 unclassified Nitrobacter strains, Nitrobacter alkalicus, Nitrobacter vulgaris, Nitrobacter winogradskyi)	**0.2%**	0.2%
g	**Mesorhizobium** (1 OTU with 100% identity in 399bp to: 2 unclassified Mesorhizobium strains, alpha proteobacterium WG1)	**0.2%**	0.2%
s	**Nitrosomonas sp. Is343** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp. Is343)	**0.2%**	0.1%
s	**Granulicella paludicola** (1 OTU with 85% identity in 407bp to: Granulicella paludicola)	**0.2%**	0.1%
s	**Nitrospira sp. Ecomares 2.1** (1 OTU with 95% identity in 417bp to: Nitrospira sp. Ecomares 2.1)	**0.2%**	0.1%
s	**Bdellovibrio sp. S1** (1 OTU with 89% identity in 405bp to: Bdellovibrio sp. S1)	**0.2%**	0.3%
o	**Bryobacterales** (1 OTU with 90% identity in 399bp to: Candidatus Solibacter usitatus, Paludibaculum sp.)	**0.2%**	0.1%
s	**Fodinicurvata sp.** (1 OTU with 92% identity in 399bp to: Fodinicurvata sp.)	**0.2%**	0.4%
s	**Asprobacter aquaticus** (1 OTU with 99% identity in 399bp to: Asprobacter aquaticus)	**0.2%**	0.2%
o	**Nitrosomonadales** (1 OTU with 91% identity in 424bp to: Methyloversatilis universalis, Nitrosomonas mobilis)	**0.2%**	0.1%
f	**Sphingomonadaceae** (1 OTU with 98% identity in 399bp to: 3 unclassified Sphingomonas strains, Sphingosinicella cucumeris)	**0.2%**	0.2%
o	**Rhizobiales** (2 OTUs with 95-100% identity in 399-400bp to: 10 unclassified Bradyrhizobium strains, 13 unclassified Shinella strains, 2 unclassified Mycoplana strains, 2 unclassified Rhizobium strains, Bradyrhizobium genosp. Y, Bradyrhizobium vignae, Methylocapsa aurea, Rhizobium herbae, Shinella kummerowiae, Shinella zoogloeoides)	**0.1%**	0.2%
f	**Myxococcaceae** (1 OTU with 87% identity in 424bp to: Corallococcus macrosporus, Corallococcus sp. xj73, Myxococcus sp. 004-95, Myxococcus virescens)	**0.1%**	0.3%
g	**Chondromyces** (2 OTUs with 86% identity in 425bp to: Chondromyces apiculatus, Chondromyces pediculatus, Chondromyces robustus)	**0.1%**	0.4%
s	**Ilumatobacter fluminis** (1 OTU with 97% identity in 400bp to: Ilumatobacter fluminis)	**0.1%**	0.2%
g	**Pseudomonas** (1 OTU with 100% identity in 424bp to: 4 unclassified Pseudomonas strains, Pseudomonas hunanensis, Pseudomonas juntendi, Pseudomonas monteilii, Pseudomonas mosselii, Pseudomonas plecoglossicida, Pseudomonas putida, Pseudomonas taiwanensis)	**0.1%**	0.4%
s	**Hyphomicrobium sp. WG6** (1 OTU with 98% identity in 400bp to: Hyphomicrobium sp. WG6)	**0.1%**	0.1%
g	**Rhodovibrio** (1 OTU with 93% identity in 400bp to: 3 unclassified Rhodovibrio strains)	**0.1%**	0.1%
s	**Candidatus Solibacter usitatus** (1 OTU with 95% identity in 399bp to: Candidatus Solibacter usitatus)	**0.1%**	0.1%
g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.1%**	0.3%
s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**0.1%**	0.1%
s	**Enhygromyxa sp. SHK-11** (2 OTUs with 93-94% identity in 420bp to: Enhygromyxa sp. SHK-11)	**0.1%**	0.2%
s	**Sphingomonas sp.** (1 OTU with 97% identity in 399bp to: Sphingomonas sp.)	**0.1%**	0.1%
	Other	**2.8%**	5.7%
	Unclassified (0 reads)		
	Filtered (0 reads)		

R1.Start.V3V4a (36 696 reads)

g	**Nitrospira** (16 OTUs with 98-100% identity in 417bp to: 3 unclassified Nitrospira strains, Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira defluvii, Nitrospira marina)	**29.6%**	20.9%
s	**Nitrospira japonica** (8 OTUs with 97-99% identity in 419bp to: Nitrospira japonica)	**11.7%**	8.3%
g	**Nitrosomonas** (7 OTUs with 99-100% identity in 424bp to: 3 unclassified Nitrosomonas strains, Nitrosomonas mobilis, Nitrosomonas ureae)	**9.4%**	6.7%

s	**Nitrosomonas sp. Nm86** (1 OTU with 98% identity in 424bp to: Nitrosomonas sp. Nm86)	**3.9%**	2.8%
s	**Chryseolinea serpens** (7 OTUs with 90-95% identity in 421-423bp to: Chryseolinea serpens)	**2.6%**	4.9%
g	**Nitrosospira** (2 OTUs with 92% identity in 424bp to: 6 unclassified Nitrosospira strains)	**2.2%**	1.5%
s	**Paludibaculum sp.** (7 OTUs with 86-95% identity in 399bp to: Paludibaculum sp.)	**2.2%**	1.8%
s	**Gemmatimonas aurantiaca** (8 OTUs with 89-95% identity in 402bp to: Gemmatimonas aurantiaca)	**2.1%**	1.5%
s	**Holophaga sp. WY42** (4 OTUs with 89-97% identity in 423-425bp to: Holophaga sp. WY42)	**1.9%**	1.4%
s	**Halochromatium sp. MTK6IM088** (4 OTUs with 88-89% identity in 424-426bp to: Halochromatium sp. MTK6IM088)	**1.8%**	2.3%
s	**Nitrospira moscoviensis** (3 OTUs with 98-99% identity in 417-419bp to: Nitrospira moscoviensis)	**1.3%**	1.0%
s	**Candidatus Amoebophilus asiaticus** (2 OTUs with 93% identity in 421bp to: Candidatus Amoebophilus asiaticus)	**1.1%**	2.0%
s	**Gemmatimonas phototrophica** (2 OTUs with 91-92% identity in 416bp to: Gemmatimonas phototrophica)	**0.9%**	0.6%
s	**Elioraea tepidiphila** (2 OTUs with 100% identity in 399-413bp to: Elioraea tepidiphila)	**0.9%**	1.8%
g	**Pedomicrobium** (1 OTU with 96% identity in 399bp to: 2 unclassified Pedomicrobium strains, Pedomicrobium americanum)	**0.8%**	0.8%
s	**Hypericibacter terrae** (1 OTU with 94% identity in 399bp to: Hypericibacter terrae)	**0.8%**	2.1%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (2 OTUs with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.8%**	0.6%
s	**Nitrosomonas ureae** (1 OTU with 100% identity in 424bp to: Nitrosomonas ureae)	**0.8%**	0.5%
g	**Hyphomicrobium** (2 OTUs with 95-99% identity in 399bp to: 5 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.8%**	0.7%
s	**Acidobacterium sp. WY65** (2 OTUs with 90-92% identity in 399bp to: Acidobacterium sp. WY65)	**0.8%**	0.6%
s	**Woodsholea maritima** (2 OTUs with 95-98% identity in 399bp to: Woodsholea maritima)	**0.8%**	0.8%
g	**Thauera** (6 OTUs with 99-100% identity in 424bp to: 11 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**0.7%**	1.6%
g	**Geobacter** (3 OTUs with 88-90% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.7%**	1.5%
s	**Roseisolibacter agri** (2 OTUs with 91-93% identity in 416bp to: Roseisolibacter agri)	**0.7%**	0.5%
s	**Ignavibacterium album** (4 OTUs with 80-84% identity in 422-423bp to: Ignavibacterium album)	**0.7%**	1.2%
s	**Reyranella soli** (2 OTUs with 99% identity in 399bp to: Reyranella soli)	**0.7%**	1.5%
s	**Dongia sp. URHE0060** (2 OTUs with 94% identity in 399bp to: Dongia sp. URHE0060)	**0.6%**	1.6%
s	**Nitrospira calida** (2 OTUs with 96% identity in 415-417bp to: Nitrospira calida)	**0.6%**	0.4%
s	**Hyphomicrobium sp.** (3 OTUs with 97-99% identity in 399bp to: Hyphomicrobium sp.)	**0.5%**	0.5%
s	**Nitrospirillum amazonense** (3 OTUs with 89-97% identity in 399-400bp to: Nitrospirillum amazonense)	**0.5%**	1.3%
s	**Micavibrio aeruginosavorus** (1 OTU with 91% identity in 402bp to: Micavibrio aeruginosavorus)	**0.5%**	0.8%
o	**Fragilariales** (1 OTU with 98% identity in 402bp to: Fragilaria pinnata, Nanofrustulum shiloi)	**0.5%**	0.3%
s	**Pandoraea sp.** (3 OTUs with 92-94% identity in 424bp to: Pandoraea sp.)	**0.4%**	1.7%
s	**Vermiphilus pyriformis** (1 OTU with 90% identity in 427bp to: Vermiphilus pyriformis)	**0.4%**	0.7%
s	**Parvularcula sp. P33** (4 OTUs with 94-95% identity in 399bp to: Parvularcula sp. P33)	**0.4%**	0.6%
s	**Thiobacter sp. SL-1** (1 OTU with 92% identity in 425bp to: Thiobacter sp. SL-1)	**0.4%**	0.9%
s	**Desulfovibrio cavernae** (1 OTU with 89% identity in 429bp to: Desulfovibrio cavernae)	**0.4%**	0.6%
o	**Rhodocyclales** (1 OTU with 91% identity in 424bp to: 3 unclassified Azoarcus strains, Aromatoleum aromaticum, Aromatoleum buckelii, Azoarcus olearius)	**0.4%**	0.8%
s	**Pelobacter carbinolicus** (1 OTU with 87% identity in 425bp to: Pelobacter carbinolicus)	**0.4%**	0.3%
s	**Enhydrobacter sp.** (4 OTUs with 91-94% identity in 399-400bp to: Enhydrobacter sp.)	**0.4%**	0.8%
s	**Conexibacter sp.** (1 OTU with 92% identity in 424bp to: Conexibacter sp.)	**0.3%**	0.2%
g	**Phenylobacterium** (1 OTU with 89% identity in 403bp to: Phenylobacterium koreense, Phenylobacterium sp. S140)	**0.3%**	0.2%
s	**Endomicrobium sp.** (1 OTU with 80% identity in 420bp to: Endomicrobium sp.)	**0.3%**	0.2%

g	**Bradyrhizobium** (1 OTU with 100% identity in 399bp to: 4 unclassified Bradyrhizobium strains, Bradyrhizobium elkanii, Bradyrhizobium jicamae, Bradyrhizobium paxllaeri)	**0.3%**	0.3%
s	**Bauldia consociata** (2 OTUs with 96-99% identity in 401bp to: Bauldia consociata)	**0.3%**	0.5%
g	**Lewinella** (1 OTU with 87% identity in 424bp to: Lewinella marina, Lewinella sp. SRO_346)	**0.3%**	0.5%
s	**Desulfuromonas sp. AOP6** (2 OTUs with 87% identity in 427-429bp to: Desulfuromonas sp. AOP6)	**0.3%**	0.2%
s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**0.3%**	0.2%
s	**Blastochloris viridis** (1 OTU with 95% identity in 399bp to: Blastochloris viridis)	**0.3%**	0.3%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.3%**	0.7%
s	**Mesorhizobium sp. PETBA10** (2 OTUs with 98-100% identity in 399bp to: Mesorhizobium sp. PETBA10)	**0.3%**	0.3%
s	**Nitrospira sp. Ecomares 2.1** (1 OTU with 95% identity in 417bp to: Nitrospira sp. Ecomares 2.1)	**0.3%**	0.2%
s	**Hyphomicrobium sp. Pd-S-(I)-e-D-8(2)** (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. Pd-S-(I)-e-D-8(2))	**0.3%**	0.3%
s	**Blastochloris sp. AT2101** (3 OTUs with 95% identity in 400bp to: Blastochloris sp. AT2101)	**0.3%**	0.3%
g	**Gaiella** (2 OTUs with 96-97% identity in 423bp to: 2 unclassified Gaiella strains)	**0.3%**	0.2%
c	**Betaproteobacteria** (1 OTU with 92% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Azospira sp. AH, Glaciimonas alpina)	**0.3%**	0.4%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.3%**	0.3%
g	**Pseudomonas** (3 OTUs with 99-100% identity in 424bp to: 10 unclassified Pseudomonas strains, Pseudomonas fluorescens, Pseudomonas hunanensis, Pseudomonas juntendi, Pseudomonas monteilii, Pseudomonas mosselii, Pseudomonas plecoglossicida, Pseudomonas putida, Pseudomonas taiwanensis)	**0.3%**	0.8%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.2%**	0.2%
g	**Rhodovibrio** (1 OTU with 93% identity in 400bp to: 3 unclassified Rhodovibrio strains)	**0.2%**	0.2%
s	**Steroidobacter sp.** (1 OTU with 94% identity in 424bp to: Steroidobacter sp.)	**0.2%**	0.3%
o	**Rhizobiales** (2 OTUs with 95-100% identity in 399-400bp to: 10 unclassified Bradyrhizobium strains, 13 unclassified Shinella strains, 2 unclassified Mycoplana strains, 2 unclassified Rhizobium strains, Bradyrhizobium genosp. Y, Bradyrhizobium vignae, Methylocapsa aurea, Rhizobium herbae, Shinella kummerowiae, Shinella zoogloeoides)	**0.2%**	0.3%
s	**Microcystis sp. SAG 43.90** (1 OTU with 91% identity in 416bp to: Microcystis sp. SAG 43.90)	**0.2%**	0.2%
s	**Fodinicurvata sp.** (1 OTU with 92% identity in 399bp to: Fodinicurvata sp.)	**0.2%**	0.5%
s	**Nitrosomonas sp. Is343** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp. Is343)	**0.2%**	0.1%
o	**Cytophagales** (2 OTUs with 89% identity in 421bp to: Chryseolinea soli, Chryseotalea sanaruensis, Flexibacter flexilis)	**0.2%**	0.3%
g	**Mesorhizobium** (1 OTU with 100% identity in 399bp to: 2 unclassified Mesorhizobium strains, alpha proteobacterium WG1)	**0.2%**	0.2%
s	**Ilumatobacter fluminis** (1 OTU with 97% identity in 400bp to: Ilumatobacter fluminis)	**0.2%**	0.2%
s	**Gaiella occulta** (1 OTU with 97% identity in 423bp to: Gaiella occulta)	**0.2%**	0.1%
s	**Asprobacter aquaticus** (1 OTU with 99% identity in 399bp to: Asprobacter aquaticus)	**0.2%**	0.2%
s	**Aquicella siphonis** (1 OTU with 93% identity in 425bp to: Aquicella siphonis)	**0.2%**	0.1%
s	**Hyphomicrobium sp. WG6** (1 OTU with 98% identity in 400bp to: Hyphomicrobium sp. WG6)	**0.2%**	0.2%
g	**Chondromyces** (2 OTUs with 86% identity in 425bp to: Chondromyces apiculatus, Chondromyces pediculatus, Chondromyces robustus)	**0.2%**	0.4%
s	**Haematospirillum sp.** (1 OTU with 96% identity in 371bp to: Haematospirillum sp.)	**0.2%**	0.4%
s	**Aquihabitans daechungensis** (1 OTU with 95% identity in 401bp to: Aquihabitans daechungensis)	**0.1%**	0.2%
s	**Pseudorhodoplanes sinuspersici** (1 OTU with 99% identity in 399bp to: Pseudorhodoplanes sinuspersici)	**0.1%**	0.2%
f	**Sphingomonadaceae** (1 OTU with 98% identity in 399bp to: 3 unclassified Sphingomonas strains, Sphingosinicella cucumeris)	**0.1%**	0.2%
s	**Sphingomonas sp.** (1 OTU with 97% identity in 399bp to: Sphingomonas sp.)	**0.1%**	0.2%
s	**Marimicrobium arenosum** (1 OTU with 95% identity in 424bp to: Marimicrobium arenosum)	**0.1%**	0.2%

p	**Proteobacteria** (3 OTUs with 90-100% identity in 399-427bp to: 3 unclassified Methylosinus strains, 3 unclassified Pseudomonas strains, Comamonas sp., Comamonas testosteroni, Comamonas thiooxydans, Methylomonas sp. GYJ3, Methylosinus trichosporium, Paracoccus sp. 36, Pseudomonas fluvialis, Pseudomonas guguanensis, Pseudomonas oleovorans, Pseudomonas pharmafabricae, Sphingobium sp.)	**0.1%**	0.3%
s	**Thauera phenylacetica** (1 OTU with 97% identity in 425bp to: Thauera phenylacetica)	**0.1%**	0.3%
s	**Bdellovibrio sp. S1** (1 OTU with 89% identity in 405bp to: Bdellovibrio sp. S1)	**0.1%**	0.3%
g	**Sphingomonas** (2 OTUs with 90-91% identity in 399bp to: 4 unclassified Sphingomonas strains, Sphingomonas hankyongensis, Sphingomonas sediminicola)	**0.1%**	0.2%
s	**Curvibacter sp. R-36930** (1 OTU with 99% identity in 424bp to: Curvibacter sp. R-36930)	**0.1%**	0.2%
g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.1%**	0.3%
g	**Rurimicrobium** (1 OTU with 82% identity in 421bp to: Rurimicrobium arvi, Rurimicrobium sp.)	**0.1%**	0.2%
s	**Georgfuchsia toluolica** (1 OTU with 93% identity in 424bp to: Georgfuchsia toluolica)	**0.1%**	0.1%
s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.1%**	0.2%
s	**Enhygromyxa sp. SHK-11** (2 OTUs with 93-94% identity in 420bp to: Enhygromyxa sp. SHK-11)	**0.1%**	0.2%
f	**Parachlamydiaceae** (1 OTU with 87% identity in 427bp to: Candidatus Metachlamydia lacustris, Parachlamydia acanthamoebae)	**0.1%**	0.1%
s	**Pseudolabrys sp. W2.09-308** (1 OTU with 97% identity in 399bp to: Pseudolabrys sp. W2.09-308)	**0.1%**	0.2%
s	**Roseomonas sp.** (1 OTU with 95% identity in 399bp to: Roseomonas sp.)	**0.1%**	0.2%
f	**Myxococcaceae** (1 OTU with 87% identity in 424bp to: Corallococcus macrosporus, Corallococcus sp. xj73, Myxococcus sp. 004-95, Myxococcus virescens)	**0.1%**	0.2%
s	**Parachlamydia acanthamoebae** (1 OTU with 84% identity in 430bp to: Parachlamydia acanthamoebae)	**0.1%**	0.1%
s	**Arenimonas subflava** (1 OTU with 96% identity in 424bp to: Arenimonas subflava)	**0.1%**	0.2%
	Other	**2.4%**	4.7%
	Unclassified (0 reads)		
	Filtered (0 reads)		

R1.Sumpf.V3V4a (38 869 reads)

g	**Nitrospira** (17 OTUs with 98-100% identity in 417bp to: 3 unclassified Nitrospira strains, Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira defluvii, Nitrospira marina)	**31.9%**	23.0%
s	**Nitrospira japonica** (8 OTUs with 97-99% identity in 419bp to: Nitrospira japonica)	**12.1%**	8.7%
g	**Nitrosomonas** (7 OTUs with 99-100% identity in 424bp to: 3 unclassified Nitrosomonas strains, Nitrosomonas mobilis, Nitrosomonas ureae)	**8.3%**	6.0%
s	**Nitrosomonas sp. Nm86** (1 OTU with 98% identity in 424bp to: Nitrosomonas sp. Nm86)	**3.1%**	2.2%
g	**Nitrosospira** (2 OTUs with 92% identity in 424bp to: 6 unclassified Nitrosospira strains)	**2.5%**	1.8%
s	**Chryseolinea serpens** (9 OTUs with 90-95% identity in 421-423bp to: Chryseolinea serpens)	**2.1%**	3.9%
s	**Halochromatium sp. MTK6IM088** (2 OTUs with 88-89% identity in 424-426bp to: Halochromatium sp. MTK6IM088)	**2.0%**	2.6%
s	**Paludibaculum sp.** (6 OTUs with 90-94% identity in 399bp to: Paludibaculum sp.)	**1.9%**	1.6%
g	**Pedomicrobium** (2 OTUs with 96-97% identity in 399-400bp to: 3 unclassified Pedomicrobium strains, Pedomicrobium americanum, Pedomicrobium australicum)	**1.8%**	1.8%
s	**Dongia sp. URHE0060** (4 OTUs with 94% identity in 399-413bp to: Dongia sp. URHE0060)	**1.6%**	4.0%
s	**Nitrospira moscoviensis** (3 OTUs with 98-99% identity in 417-419bp to: Nitrospira moscoviensis)	**1.5%**	1.1%
s	**Holophaga sp. WY42** (4 OTUs with 89-97% identity in 423-425bp to: Holophaga sp. WY42)	**1.5%**	1.1%
s	**Gemmatimonas aurantiaca** (5 OTUs with 89-95% identity in 402bp to: Gemmatimonas aurantiaca)	**1.4%**	1.0%
g	**Hyphomicrobium** (2 OTUs with 99% identity in 399bp to: 6 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**1.3%**	1.3%
s	**Hypericibacter terrae** (4 OTUs with 94% identity in 399bp to: Hypericibacter terrae)	**1.3%**	3.4%
s	**Roseisolibacter agri** (2 OTUs with 91-93% identity in 416bp to: Roseisolibacter agri)	**1.2%**	0.9%
s	**Hyphomicrobium sp.** (3 OTUs with 97-99% identity in 399bp to: Hyphomicrobium sp.)	**1.1%**	1.1%
s	**Woodsholea maritima** (6 OTUs with 93-99% identity in 399bp to: Woodsholea maritima)	**1.0%**	1.1%
s	**Candidatus Amoebophilus asiaticus** (1 OTU with 93% identity in 421bp to: Candidatus Amoebophilus asiaticus)	**1.0%**	1.9%

s	**Gemmatimonas phototrophica** (2 OTUs with 91-92% identity in 416bp to: Gemmatimonas phototrophica)	**0.9%**	0.7%
s	**Nitrosomonas ureae** (1 OTU with 100% identity in 424bp to: Nitrosomonas ureae)	**0.9%**	0.7%
s	**Nitrospira calida** (2 OTUs with 96% identity in 415-417bp to: Nitrospira calida)	**0.9%**	0.6%
g	**Bradyrhizobium** (2 OTUs with 100% identity in 399bp to: 6 unclassified Bradyrhizobium strains, Bradyrhizobium centrolobii, Bradyrhizobium cytisi, Bradyrhizobium elkanii, Bradyrhizobium guangdongense, Bradyrhizobium guangzhouense, Bradyrhizobium jicamae, Bradyrhizobium paxllaeri, Bradyrhizobium rifense)	**0.7%**	0.6%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.7%**	0.5%
s	**Acidobacterium sp. WY65** (1 OTU with 92% identity in 399bp to: Acidobacterium sp. WY65)	**0.6%**	0.5%
s	**Mesorhizobium sp. PETBA10** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp. PETBA10)	**0.5%**	0.7%
s	**Parvularcula sp. P33** (5 OTUs with 94-95% identity in 399bp to: Parvularcula sp. P33)	**0.5%**	0.8%
o	**Rhodocyclales** (1 OTU with 91% identity in 424bp to: 3 unclassified Azoarcus strains, Aromatoleum aromaticum, Aromatoleum buckelii, Azoarcus olearius)	**0.5%**	1.1%
s	**Pandoraea sp.** (4 OTUs with 92-94% identity in 424-425bp to: Pandoraea sp.)	**0.5%**	2.0%
s	**Desulfovibrio cavernae** (1 OTU with 89% identity in 429bp to: Desulfovibrio cavernae)	**0.5%**	0.9%
s	**Micavibrio aeruginosavorus** (1 OTU with 91% identity in 402bp to: Micavibrio aeruginosavorus)	**0.5%**	0.7%
g	**Thauera** (5 OTUs with 99-100% identity in 424bp to: 8 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica)	**0.5%**	1.0%
s	**Elioraea tepidiphila** (1 OTU with 100% identity in 399bp to: Elioraea tepidiphila)	**0.4%**	1.0%
o	**Fragilariales** (1 OTU with 98% identity in 402bp to: Fragilaria pinnata, Nanofrustulum shiloi)	**0.4%**	0.3%
s	**Hyphomicrobium sp. Pd-S-(I)-e-D-8(2)** (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. Pd-S-(I)-e-D-8(2))	**0.4%**	0.4%
s	**Nitrospirillum amazonense** (2 OTUs with 89-97% identity in 399-400bp to: Nitrospirillum amazonense)	**0.4%**	1.0%
s	**Tabrizicola piscis** (1 OTU with 100% identity in 399bp to: Tabrizicola piscis)	**0.4%**	0.7%
s	**Reyranella soli** (1 OTU with 99% identity in 399bp to: Reyranella soli)	**0.4%**	0.8%
s	**Nitrosomonas sp.** (1 OTU with 95% identity in 425bp to: Nitrosomonas sp.)	**0.3%**	0.3%
g	**Geobacter** (2 OTUs with 88-90% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.3%**	0.7%
s	**Thiobacter sp. SL-1** (1 OTU with 92% identity in 425bp to: Thiobacter sp. SL-1)	**0.3%**	0.8%
s	**Endomicrobium sp.** (1 OTU with 80% identity in 420bp to: Endomicrobium sp.)	**0.3%**	0.2%
s	**Desulfuromonas sp. AOP6** (2 OTUs with 87% identity in 427-429bp to: Desulfuromonas sp. AOP6)	**0.3%**	0.2%
s	**Steroidobacter sp.** (1 OTU with 94% identity in 424bp to: Steroidobacter sp.)	**0.3%**	0.4%
s	**Vermiphilus pyriformis** (1 OTU with 90% identity in 427bp to: Vermiphilus pyriformis)	**0.3%**	0.4%
s	**Blastochloris sp. AT2101** (1 OTU with 95% identity in 400bp to: Blastochloris sp. AT2101)	**0.3%**	0.3%
s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**0.2%**	0.2%
g	**Phenylobacterium** (1 OTU with 89% identity in 403bp to: Phenylobacterium koreense, Phenylobacterium sp. S140)	**0.2%**	0.2%
s	**Ignavibacterium album** (2 OTUs with 82-84% identity in 422-423bp to: Ignavibacterium album)	**0.2%**	0.4%
g	**Pseudomonas** (3 OTUs with 98-100% identity in 424bp to: 12 unclassified Pseudomonas strains, Pseudomonas fluorescens, Pseudomonas hunanensis, Pseudomonas juntendi, Pseudomonas monteilii, Pseudomonas mosselii, Pseudomonas plecoglossicida, Pseudomonas putida, Pseudomonas taiwanensis)	**0.2%**	0.8%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.2%**	0.6%
s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.2%**	0.5%
s	**Enhydrobacter sp.** (4 OTUs with 91-94% identity in 399-400bp to: Enhydrobacter sp.)	**0.2%**	0.5%
s	**Blastochloris viridis** (1 OTU with 95% identity in 399bp to: Blastochloris viridis)	**0.2%**	0.2%
s	**Candidatus Nitrosarchaeum limnium** (1 OTU with 94% identity in 402bp to: Candidatus Nitrosarchaeum limnium)	**0.2%**	0.2%
s	**Microcystis sp. SAG 43.90** (1 OTU with 91% identity in 416bp to: Microcystis sp. SAG 43.90)	**0.2%**	0.2%
g	**Gaiella** (1 OTU with 96% identity in 423bp to: 2 unclassified Gaiella strains)	**0.2%**	0.1%

s	**Aquihabitans daechungensis** (1 OTU with 95% identity in 401bp to: Aquihabitans daechungensis)	**0.2%**	0.3%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.2%**	0.2%
s	**Hyphomicrobium sp. WG6** (1 OTU with 98% identity in 400bp to: Hyphomicrobium sp. WG6)	**0.2%**	0.2%
s	**Nitrosomonas sp. Is343** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp. Is343)	**0.2%**	0.1%
s	**Asprobacter aquaticus** (1 OTU with 99% identity in 399bp to: Asprobacter aquaticus)	**0.2%**	0.2%
s	**Pelobacter carbinolicus** (1 OTU with 87% identity in 425bp to: Pelobacter carbinolicus)	**0.2%**	0.1%
s	**Ilumatobacter fluminis** (1 OTU with 97% identity in 400bp to: Ilumatobacter fluminis)	**0.2%**	0.2%
c	**Betaproteobacteria** (1 OTU with 92% identity in 424bp to: 2 unclassified Glaciimonas strains, 2 unclassified Herbaspirillum strains, Azospira sp. AH, Glaciimonas alpina)	**0.2%**	0.3%
s	**Sterolibacterium sp. TKU1** (1 OTU with 100% identity in 424bp to: Sterolibacterium sp. TKU1)	**0.2%**	0.1%
s	**Nitrospira sp. Ecomares 2.1** (1 OTU with 95% identity in 417bp to: Nitrospira sp. Ecomares 2.1)	**0.1%**	0.1%
o	**Cytophagales** (2 OTUs with 89% identity in 421bp to: Chryseolinea soli, Chryseotalea sanaruensis, Flexibacter flexilis)	**0.1%**	0.3%
s	**Pseudorhodoplanes sinuspersici** (1 OTU with 99% identity in 399bp to: Pseudorhodoplanes sinuspersici)	**0.1%**	0.2%
s	**Enhygromyxa sp. SHK-11** (2 OTUs with 93-94% identity in 420bp to: Enhygromyxa sp. SHK-11)	**0.1%**	0.2%
s	**Bdellovibrio sp. S1** (1 OTU with 89% identity in 405bp to: Bdellovibrio sp. S1)	**0.1%**	0.3%
s	**Bauldia consociata** (1 OTU with 96% identity in 401bp to: Bauldia consociata)	**0.1%**	0.2%
s	**Nordella sp.** (1 OTU with 97% identity in 399bp to: Nordella sp.)	**0.1%**	0.2%
s	**Roseomonas sp.** (1 OTU with 95% identity in 399bp to: Roseomonas sp.)	**0.1%**	0.3%
c	**Gammaproteobacteria** (1 OTU with 90% identity in 424bp to: Acinetobacter radioresistens, Halochromatium roseum, Halochromatium sp. MTK2IM039, Thioalkalivibrio denitrificans)	**0.1%**	0.2%
o	**Rhizobiales** (1 OTU with 95% identity in 400bp to: 10 unclassified Bradyrhizobium strains, Bradyrhizobium genosp. Y, Bradyrhizobium vignae, Methylocapsa aurea)	**0.1%**	0.2%
g	**Mesorhizobium** (1 OTU with 100% identity in 399bp to: 2 unclassified Mesorhizobium strains, alpha proteobacterium WG1)	**0.1%**	0.1%
f	**Myxococcaceae** (1 OTU with 87% identity in 424bp to: Corallococcus macrosporus, Corallococcus sp. xj73, Myxococcus sp. 004-95, Myxococcus virescens)	**0.1%**	0.2%
g	**Lewinella** (1 OTU with 87% identity in 424bp to: Lewinella marina, Lewinella sp. SRO_346)	**0.1%**	0.2%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.1%**	0.1%
s	**Nitrosospira sp. REGAU** (1 OTU with 94% identity in 424bp to: Nitrosospira sp. REGAU)	**0.1%**	0.1%
s	**Haematospirillum sp.** (1 OTU with 97% identity in 371bp to: Haematospirillum sp.)	**0.1%**	0.3%
g	**Frigoribacterium** (1 OTU with 100% identity in 404bp to: 5 unclassified Frigoribacterium strains)	**0.1%**	0.2%
s	**Magnetospira sp. QH-2** (1 OTU with 92% identity in 399bp to: Magnetospira sp. QH-2)	**0.1%**	0.3%
f	**Bradyrhizobiaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Bradyrhizobium strains, 3 unclassified Afipia strains, Afipia massiliensis)	**0.1%**	0.1%
	Other	**1.9%**	3.9%
	Unclassified (76 reads)		
	Filtered (0 reads)		

R2.Ende.V3V4a (22 213 reads)

g	**Thauera** (50 OTUs with 99-100% identity in 283-425bp to: 14 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**51.1%**	60.1%
s	**Denitratisoma oestradiolicum** (5 OTUs with 95-96% identity in 424bp to: Denitratisoma oestradiolicum)	**9.4%**	3.7%
s	**Gemmatimonas aurantiaca** (4 OTUs with 89-93% identity in 402bp to: Gemmatimonas aurantiaca)	**3.7%**	1.5%
s	**Ignavibacterium album** (5 OTUs with 82-95% identity in 422-425bp to: Ignavibacterium album)	**3.5%**	3.3%
s	**Desulforegula conservatrix** (3 OTUs with 81% identity in 424bp to: Desulforegula conservatrix)	**2.8%**	3.9%
f	**Sterolibacteriaceae** (1 OTU with 96% identity in 424bp to: Denitratisoma sp. DHT3, Sulfuritalea hydrogenivorans)	**2.8%**	1.1%
s	**Chryseolinea serpens** (2 OTUs with 91-95% identity in 421-423bp to: Chryseolinea serpens)	**2.2%**	2.2%
s	**Holophaga sp. WY42** (3 OTUs with 88-90% identity in 424-425bp to: Holophaga sp. WY42)	**1.7%**	0.7%

s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**1.7%**	0.7%
s	**Thauera sp. JPB-3.02** (1 OTU with 100% identity in 424bp to: Thauera sp. JPB-3.02)	**1.3%**	1.5%
s	**Candidatus Aquirestis calciphila** (1 OTU with 88% identity in 422bp to: Candidatus Aquirestis calciphila)	**1.2%**	1.3%
s	**Desulfocarbo indianensis** (1 OTU with 89% identity in 430bp to: Desulfocarbo indianensis)	**1.1%**	1.0%
g	**Mesorhizobium** (3 OTUs with 99-100% identity in 399bp to: 4 unclassified Mesorhizobium strains, Mesorhizobium carbonis, Mesorhizobium sangaii, alpha proteobacterium WG1)	**1.0%**	0.7%
s	**Hypericibacter terrae** (1 OTU with 94% identity in 399bp to: Hypericibacter terrae)	**1.0%**	1.4%
g	**Thiocapsa** (1 OTU with 99% identity in 424bp to: 3 unclassified Thiocapsa strains)	**0.9%**	0.8%
s	**Parvibaculum sp.** (1 OTU with 93% identity in 399bp to: Parvibaculum sp.)	**0.8%**	0.7%
s	**Aliifodinibius sp.** (1 OTU with 81% identity in 423bp to: Aliifodinibius sp.)	**0.8%**	0.8%
s	**Pandoraea sp.** (3 OTUs with 93-94% identity in 424bp to: Pandoraea sp.)	**0.8%**	1.7%
c	**Betaproteobacteria** (3 OTUs with 98-100% identity in 424bp to: 17 unclassified Simplicispira strains, 2 unclassified Aquaspirillum strains, Simplicispira limi)	**0.8%**	0.7%
f	**Bacillaceae** (2 OTUs with 83% identity in 430bp to: Bacillus sp. G3DM-87, Terribacillus aidingensis, Terribacillus halophilus, Terribacillus saccharophilus, Terribacillus sp.)	**0.7%**	2.6%
s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.7%**	0.8%
s	**Denitratimonas tolerans** (1 OTU with 100% identity in 424bp to: Denitratimonas tolerans)	**0.7%**	0.8%
s	**Nitrosomonas sp.** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp.)	**0.5%**	0.2%
g	**Desulfovibrio** (1 OTU with 84% identity in 425bp to: 4 unclassified Desulfovibrio strains, Desulfovibrio fairfieldensis)	**0.5%**	0.6%
g	**Geobacter** (3 OTUs with 90-100% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.5%**	0.5%
s	**Lewinella cohaerens** (2 OTUs with 88-92% identity in 422-426bp to: Lewinella cohaerens)	**0.5%**	0.5%
s	**Desulfoprunum benzoelyticum** (2 OTUs with 98-99% identity in 428bp to: Desulfoprunum benzoelyticum)	**0.5%**	0.6%
s	**Geobacter sulfurreducens** (2 OTUs with 99-100% identity in 425bp to: Geobacter sulfurreducens)	**0.5%**	0.4%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.5%**	0.2%
s	**Paludibaculum sp.** (1 OTU with 93% identity in 399bp to: Paludibaculum sp.)	**0.4%**	0.2%
g	**Simplicispira** (2 OTUs with 100% identity in 424bp to: 3 unclassified Simplicispira strains, Simplicispira limi, Simplicispira piscis)	**0.4%**	0.3%
s	**Halochromatium sp. MTK6IM088** (1 OTU with 88% identity in 426bp to: Halochromatium sp. MTK6IM088)	**0.4%**	0.2%
s	**Pelobacter carbinolicus** (1 OTU with 82% identity in 426bp to: Pelobacter carbinolicus)	**0.3%**	0.1%
g	**Nitrospira** (1 OTU with 98% identity in 417bp to: Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira sp.)	**0.3%**	0.1%
g	**Hyphomicrobium** (1 OTU with 99% identity in 399bp to: 4 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.3%**	0.2%
s	**Thiobacter sp. SL-1** (2 OTUs with 92-95% identity in 424-425bp to: Thiobacter sp. SL-1)	**0.3%**	0.4%
g	**Lewinella** (1 OTU with 86% identity in 423bp to: Lewinella litorea, Lewinella lutea, Lewinella marina, Lewinella maritima, Lewinella sp. SRO_346)	**0.3%**	0.3%
g	**Cystobacter** (1 OTU with 87% identity in 424bp to: Cystobacter ferrugineus, Cystobacter fuscus, Cystobacter sp. MCy9104)	**0.3%**	0.2%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.2%**	0.1%
o	**Rhizobiales** (1 OTU with 99% identity in 399bp to: 13 unclassified Sinorhizobium strains, 2 unclassified Bradyrhizobium strains, 39 unclassified Ensifer strains, Ensifer glycinis, Ensifer mexicanus, Ensifer shofinae, Mesorhizobium sp., Sinorhizobium americanum, Sinorhizobium chiapanecum, Sinorhizobium fredii, Sinorhizobium terangae)	**0.2%**	0.2%
f	**Bradyrhizobiaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Rhodopseudomonas strains, Nitrobacter sp. NS5-5, Rhodopseudomonas faecalis)	**0.2%**	0.1%
s	**Melioribacter roseus** (1 OTU with 89% identity in 427bp to: Melioribacter roseus)	**0.2%**	0.2%
s	**Dokdonella sp.** (1 OTU with 97% identity in 424bp to: Dokdonella sp.)	**0.2%**	0.2%
s	**Aquincola sp.** (1 OTU with 99% identity in 424bp to: Aquincola sp.)	**0.2%**	0.2%
s	**Thiothrix unzii** (1 OTU with 100% identity in 402bp to: Thiothrix unzii)	**0.2%**	0.1%

s	**Lentimicrobium saccharophilum** (1 OTU with 96% identity in 423bp to: Lentimicrobium saccharophilum)	**0.2%**	0.2%
s	**Mesorhizobium sp.** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp.)	**0.1%**	0.1%
g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.1%**	0.2%
s	**Maribellus luteus** (1 OTU with 95% identity in 422bp to: Maribellus luteus)	**0.1%**	0.2%
f	**Xanthomonadaceae** (1 OTU with 100% identity in 424bp to: 20 unclassified Pseudoxanthomonas strains, Pseudoxanthomonas mexicana, Xanthomonas campestris)	**0.1%**	0.2%
f	**Prolixibacteraceae** (1 OTU with 93% identity in 422bp to: Maribellus luteus, Tangfeifania diversioriginum)	**0.1%**	0.1%
s	**Thauera mechernichensis** (1 OTU with 100% identity in 424bp to: Thauera mechernichensis)	**0.1%**	0.1%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.1%**	0.1%
	Other	**0.8%**	1.0%
	Unclassified (690 reads)		
	Filtered (0 reads)		

R2.Mitte.V3V4a (22 428 reads)

g	**Thauera** (40 OTUs with 99-100% identity in 283-437bp to: 14 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**46.9%**	56.4%
s	**Denitratisoma oestradiolicum** (4 OTUs with 95-96% identity in 424bp to: Denitratisoma oestradiolicum)	**10.7%**	4.3%
s	**Ignavibacterium album** (6 OTUs with 82-95% identity in 422-425bp to: Ignavibacterium album)	**3.5%**	3.4%
f	**Sterolibacteriaceae** (1 OTU with 96% identity in 424bp to: Denitratisoma sp. DHT3, Sulfuritalea hydrogenivorans)	**3.2%**	1.3%
s	**Gemmatimonas aurantiaca** (5 OTUs with 89-93% identity in 402bp to: Gemmatimonas aurantiaca)	**3.1%**	1.3%
s	**Desulforegula conservatrix** (3 OTUs with 81% identity in 424bp to: Desulforegula conservatrix)	**2.9%**	4.1%
s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**2.1%**	0.8%
s	**Chryseolinea serpens** (2 OTUs with 91-95% identity in 421-423bp to: Chryseolinea serpens)	**2.0%**	2.1%
s	**Holophaga sp. WY42** (4 OTUs with 88-90% identity in 424-425bp to: Holophaga sp. WY42)	**1.8%**	0.7%
s	**Desulfocarbo indianensis** (1 OTU with 89% identity in 430bp to: Desulfocarbo indianensis)	**1.5%**	1.4%
s	**Aliifodinibius sp.** (1 OTU with 81% identity in 423bp to: Aliifodinibius sp.)	**1.2%**	1.2%
g	**Thiocapsa** (1 OTU with 99% identity in 424bp to: 3 unclassified Thiocapsa strains)	**1.1%**	0.9%
s	**Candidatus Aquirestis calciphila** (1 OTU with 88% identity in 422bp to: Candidatus Aquirestis calciphila)	**1.1%**	1.1%
g	**Mesorhizobium** (3 OTUs with 99-100% identity in 399bp to: 4 unclassified Mesorhizobium strains, Mesorhizobium carbonis, Mesorhizobium sangaii, alpha proteobacterium WG1)	**1.0%**	0.7%
s	**Hypericibacter terrae** (1 OTU with 94% identity in 399bp to: Hypericibacter terrae)	**1.0%**	1.4%
s	**Thauera sp. JPB-3.02** (1 OTU with 100% identity in 424bp to: Thauera sp. JPB-3.02)	**1.0%**	1.2%
c	**Betaproteobacteria** (3 OTUs with 98-100% identity in 424bp to: 17 unclassified Simplicispira strains, 2 unclassified Aquaspirillum strains, Simplicispira limi)	**0.9%**	0.8%
s	**Parvibaculum sp.** (1 OTU with 93% identity in 399bp to: Parvibaculum sp.)	**0.8%**	0.7%
f	**Bacillaceae** (2 OTUs with 83% identity in 430bp to: Bacillus sp. G3DM-87, Terribacillus aidingensis, Terribacillus halophilus, Terribacillus saccharophilus, Terribacillus sp.)	**0.8%**	2.7%
s	**Nitrosomonas sp.** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp.)	**0.8%**	0.3%
s	**Pandoraea sp.** (3 OTUs with 93-94% identity in 424bp to: Pandoraea sp.)	**0.7%**	1.4%
g	**Desulfovibrio** (1 OTU with 84% identity in 425bp to: 4 unclassified Desulfovibrio strains, Desulfovibrio fairfieldensis)	**0.7%**	0.9%
s	**Desulfoprunum benzoelyticum** (2 OTUs with 98-99% identity in 428bp to: Desulfoprunum benzoelyticum)	**0.6%**	0.8%
g	**Geobacter** (3 OTUs with 90-100% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.6%**	0.7%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.6%**	0.3%
s	**Denitratimonas tolerans** (1 OTU with 100% identity in 424bp to: Denitratimonas tolerans)	**0.6%**	0.7%
s	**Geobacter sulfurreducens** (2 OTUs with 99-100% identity in 425bp to: Geobacter sulfurreducens)	**0.6%**	0.5%
s	**Paludibaculum sp.** (1 OTU with 93% identity in 399bp to: Paludibaculum sp.)	**0.5%**	0.2%

s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.5%**	0.6%
g	**Nitrospira** (1 OTU with 98% identity in 417bp to: Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira sp.)	**0.5%**	0.2%
s	**Lewinella cohaerens** (2 OTUs with 88-92% identity in 422-426bp to: Lewinella cohaerens)	**0.5%**	0.5%
g	**Lewinella** (1 OTU with 86% identity in 423bp to: Lewinella litorea, Lewinella lutea, Lewinella marina, Lewinella maritima, Lewinella sp. SRO_346)	**0.4%**	0.4%
s	**Thiobacter sp. SL-1** (2 OTUs with 92-95% identity in 424-425bp to: Thiobacter sp. SL-1)	**0.4%**	0.5%
o	**Rhizobiales** (1 OTU with 99% identity in 399bp to: 13 unclassified Sinorhizobium strains, 2 unclassified Bradyrhizobium strains, 39 unclassified Ensifer strains, Ensifer glycinis, Ensifer mexicanus, Ensifer shofinae, Mesorhizobium sp., Sinorhizobium americanum, Sinorhizobium chiapanecum, Sinorhizobium fredii, Sinorhizobium terangae)	**0.3%**	0.3%
s	**Halochromatium sp. MTK6IM088** (1 OTU with 88% identity in 426bp to: Halochromatium sp. MTK6IM088)	**0.3%**	0.2%
g	**Simplicispira** (2 OTUs with 100% identity in 424bp to: 3 unclassified Simplicispira strains, Simplicispira limi, Simplicispira piscis)	**0.3%**	0.2%
g	**Hyphomicrobium** (1 OTU with 99% identity in 399bp to: 4 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.3%**	0.2%
s	**Pelobacter carbinolicus** (1 OTU with 82% identity in 426bp to: Pelobacter carbinolicus)	**0.3%**	0.1%
s	**Thiothrix unzii** (1 OTU with 100% identity in 402bp to: Thiothrix unzii)	**0.2%**	0.2%
s	**Rhodoplanes sp. Z2-YC6860** (1 OTU with 97% identity in 400bp to: Rhodoplanes sp. Z2-YC6860)	**0.2%**	0.1%
s	**Elioraea tepidiphila** (1 OTU with 100% identity in 399bp to: Elioraea tepidiphila)	**0.2%**	0.2%
s	**Aquincola sp.** (1 OTU with 99% identity in 424bp to: Aquincola sp.)	**0.2%**	0.2%
f	**Bradyrhizobiaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Rhodopseudomonas strains, Nitrobacter sp. NS5-5, Rhodopseudomonas faecalis)	**0.2%**	0.1%
g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.2%**	0.2%
s	**Acidobacterium sp. WY65** (1 OTU with 92% identity in 399bp to: Acidobacterium sp. WY65)	**0.2%**	0.1%
s	**Mesorhizobium sp.** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp.)	**0.2%**	0.1%
g	**Cystobacter** (1 OTU with 87% identity in 424bp to: Cystobacter ferrugineus, Cystobacter fuscus, Cystobacter sp. MCy9104)	**0.2%**	0.1%
s	**Lentimicrobium saccharophilum** (1 OTU with 96% identity in 423bp to: Lentimicrobium saccharophilum)	**0.1%**	0.2%
s	**Chitinophaga arvensicola** (1 OTU with 92% identity in 422bp to: Chitinophaga arvensicola)	**0.1%**	0.1%
s	**Melioribacter roseus** (1 OTU with 89% identity in 427bp to: Melioribacter roseus)	**0.1%**	0.1%
s	**Dongia sp. URHE0060** (1 OTU with 94% identity in 399bp to: Dongia sp. URHE0060)	**0.1%**	0.2%
f	**Xanthomonadaceae** (1 OTU with 100% identity in 424bp to: 20 unclassified Pseudoxanthomonas strains, Pseudoxanthomonas mexicana, Xanthomonas campestris)	**0.1%**	0.1%
s	**Proteiniphilum sp. I28** (1 OTU with 80% identity in 425bp to: Proteiniphilum sp. I28)	**0.1%**	0.1%
f	**Prolixibacteraceae** (1 OTU with 93% identity in 422bp to: Maribellus luteus, Tangfeifania diversioriginum)	**0.1%**	0.1%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.1%**	0.1%
s	**Thauera mechernichensis** (1 OTU with 100% identity in 424bp to: Thauera mechernichensis)	**0.1%**	0.1%
o	**Desulfuromonadales** (1 OTU with 85% identity in 429bp to: Desulfuromonas carbonis, Desulfuromonas sp. DDH964, Geoalkalibacter subterraneus)	**0.1%**	0.1%
s	**Maribellus luteus** (1 OTU with 95% identity in 422bp to: Maribellus luteus)	**0.1%**	0.1%
g	**Halochromatium** (1 OTU with 80% identity in 428bp to: 4 unclassified Halochromatium strains)	**0.1%**	0.1%
	Other	**1.1%**	1.4%
	Unclassified (557 reads)		
	Filtered (0 reads)		

R2.Start.V3V4a (23 648 reads)

g	**Thauera** (43 OTUs with 99-100% identity in 283-424bp to: 12 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**39.7%**	50.6%
s	**Denitratisoma oestradiolicum** (4 OTUs with 95-96% identity in 424bp to: Denitratisoma oestradiolicum)	**11.7%**	5.0%
f	**Sterolibacteriaceae** (3 OTUs with 95-96% identity in 424bp to: Denitratisoma sp. DHT3, Methyloversatilis sp. PL20, Sulfurisoma sediminicola, Sulfuritalea hydrogenivorans)	**5.0%**	2.1%

s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**4.3%**	1.8%
s	**Gemmatimonas aurantiaca** (5 OTUs with 89-93% identity in 402bp to: Gemmatimonas aurantiaca)	**2.8%**	1.2%
s	**Ignavibacterium album** (6 OTUs with 82-95% identity in 422-425bp to: Ignavibacterium album)	**2.7%**	2.8%
s	**Desulforegula conservatrix** (3 OTUs with 81% identity in 424bp to: Desulforegula conservatrix)	**2.4%**	3.6%
g	**Thiocapsa** (1 OTU with 99% identity in 424bp to: 3 unclassified Thiocapsa strains)	**2.2%**	1.9%
s	**Holophaga sp. WY42** (4 OTUs with 88-90% identity in 424-425bp to: Holophaga sp. WY42)	**2.0%**	0.9%
s	**Chryseolinea serpens** (3 OTUs with 90-95% identity in 421-423bp to: Chryseolinea serpens)	**1.7%**	2.0%
s	**Desulfocarbo indianensis** (1 OTU with 89% identity in 430bp to: Desulfocarbo indianensis)	**1.3%**	1.2%
s	**Aliifodinibius sp.** (1 OTU with 81% identity in 423bp to: Aliifodinibius sp.)	**1.2%**	1.2%
s	**Thauera sp. JPB-3.02** (1 OTU with 100% identity in 424bp to: Thauera sp. JPB-3.02)	**1.0%**	1.3%
s	**Geobacter sulfurreducens** (2 OTUs with 99-100% identity in 425bp to: Geobacter sulfurreducens)	**1.0%**	0.8%
g	**Nitrospira** (3 OTUs with 98-100% identity in 417bp to: 2 unclassified Nitrospira strains, Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira defluvii)	**1.0%**	0.4%
c	**Betaproteobacteria** (3 OTUs with 98-100% identity in 424bp to: 17 unclassified Simplicispira strains, 2 unclassified Aquaspirillum strains, Simplicispira limi)	**0.9%**	0.9%
s	**Candidatus Aquirestis calciphila** (1 OTU with 88% identity in 422bp to: Candidatus Aquirestis calciphila)	**0.9%**	1.0%
g	**Mesorhizobium** (3 OTUs with 99-100% identity in 399bp to: 4 unclassified Mesorhizobium strains, Mesorhizobium carbonis, Mesorhizobium sangaii, alpha proteobacterium WG1)	**0.9%**	0.7%
s	**Hypericibacter terrae** (2 OTUs with 94% identity in 399-413bp to: Hypericibacter terrae)	**0.9%**	1.4%
g	**Geobacter** (4 OTUs with 90-100% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.8%**	1.0%
s	**Paludibaculum sp.** (2 OTUs with 93-95% identity in 399bp to: Paludibaculum sp.)	**0.8%**	0.4%
f	**Bacillaceae** (3 OTUs with 82-83% identity in 430bp to: Bacillus sp. G3DM-87, Terribacillus aidingensis, Terribacillus halophilus, Terribacillus saccharophilus, Terribacillus sp.)	**0.7%**	2.8%
s	**Pandoraea sp.** (3 OTUs with 93-94% identity in 424bp to: Pandoraea sp.)	**0.7%**	1.6%
s	**Parvibaculum sp.** (1 OTU with 93% identity in 399bp to: Parvibaculum sp.)	**0.7%**	0.7%
s	**Thiobacter sp. SL-1** (2 OTUs with 92-95% identity in 424-425bp to: Thiobacter sp. SL-1)	**0.6%**	0.9%
g	**Desulfovibrio** (2 OTUs with 84% identity in 425bp to: 4 unclassified Desulfovibrio strains, Desulfovibrio fairfieldensis)	**0.6%**	0.8%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.6%**	0.3%
s	**Nitrosomonas sp.** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp.)	**0.5%**	0.2%
s	**Thiothrix unzii** (1 OTU with 100% identity in 402bp to: Thiothrix unzii)	**0.5%**	0.4%
g	**Lewinella** (1 OTU with 86% identity in 423bp to: Lewinella litorea, Lewinella lutea, Lewinella marina, Lewinella maritima, Lewinella sp. SRO_346)	**0.5%**	0.6%
s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.5%**	0.6%
f	**Bradyrhizobiaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Rhodopseudomonas strains, Nitrobacter sp. NS5-5, Rhodopseudomonas faecalis)	**0.4%**	0.3%
s	**Desulfoprunum benzoelyticum** (2 OTUs with 98-99% identity in 428bp to: Desulfoprunum benzoelyticum)	**0.4%**	0.6%
o	**Rhizobiales** (1 OTU with 99% identity in 399bp to: 13 unclassified Sinorhizobium strains, 2 unclassified Bradyrhizobium strains, 39 unclassified Ensifer strains, Ensifer glycinis, Ensifer mexicanus, Ensifer shofinae, Mesorhizobium sp., Sinorhizobium americanum, Sinorhizobium chiapanecum, Sinorhizobium fredii, Sinorhizobium terangae)	**0.4%**	0.4%
g	**Simplicispira** (2 OTUs with 100% identity in 424bp to: 3 unclassified Simplicispira strains, Simplicispira limi, Simplicispira piscis)	**0.4%**	0.3%
s	**Denitratimonas tolerans** (1 OTU with 100% identity in 424bp to: Denitratimonas tolerans)	**0.4%**	0.4%
s	**Halochromatium sp. MTK6IM088** (1 OTU with 88% identity in 426bp to: Halochromatium sp. MTK6IM088)	**0.3%**	0.2%
s	**Geobacter sp. Ply1** (1 OTU with 98% identity in 425bp to: Geobacter sp. Ply1)	**0.3%**	0.4%
s	**Lewinella cohaerens** (2 OTUs with 88-92% identity in 422-426bp to: Lewinella cohaerens)	**0.3%**	0.3%
s	**Pelobacter carbinolicus** (1 OTU with 82% identity in 426bp to: Pelobacter carbinolicus)	**0.3%**	0.1%

g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.3%**	0.4%
s	**Nitrospira japonica** (1 OTU with 98% identity in 419bp to: Nitrospira japonica)	**0.3%**	0.1%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.2%**	0.2%
s	**Roseisolibacter agri** (1 OTU with 91% identity in 416bp to: Roseisolibacter agri)	**0.2%**	0.1%
s	**Rhodoplanes sp. Z2-YC6860** (1 OTU with 97% identity in 400bp to: Rhodoplanes sp. Z2-YC6860)	**0.2%**	0.1%
g	**Hyphomicrobium** (1 OTU with 99% identity in 399bp to: 4 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.2%**	0.1%
s	**Elioraea tepidiphila** (1 OTU with 100% identity in 399bp to: Elioraea tepidiphila)	**0.2%**	0.3%
s	**Melioribacter roseus** (2 OTUs with 89-90% identity in 424-427bp to: Melioribacter roseus)	**0.2%**	0.2%
g	**Ferribacterium** (1 OTU with 100% identity in 424bp to: 2 unclassified Ferribacterium strains)	**0.2%**	0.2%
s	**Candidatus Babela massiliensis** (1 OTU with 87% identity in 428bp to: Candidatus Babela massiliensis)	**0.2%**	0.2%
s	**Candidatus Amoebophilus asiaticus** (1 OTU with 93% identity in 421bp to: Candidatus Amoebophilus asiaticus)	**0.2%**	0.2%
s	**Acidobacterium sp. WY65** (1 OTU with 92% identity in 399bp to: Acidobacterium sp. WY65)	**0.2%**	0.1%
s	**Aquincola sp.** (1 OTU with 99% identity in 424bp to: Aquincola sp.)	**0.2%**	0.2%
s	**Microcystis sp. SAG 43.90** (1 OTU with 92% identity in 416bp to: Microcystis sp. SAG 43.90)	**0.1%**	0.1%
s	**Lentimicrobium saccharophilum** (1 OTU with 96% identity in 423bp to: Lentimicrobium saccharophilum)	**0.1%**	0.2%
s	**Chitinophaga arvensicola** (1 OTU with 92% identity in 422bp to: Chitinophaga arvensicola)	**0.1%**	0.1%
s	**Mesorhizobium sp.** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp.)	**0.1%**	0.1%
s	**Candidatus Sphaeronema italicum** (1 OTU with 91% identity in 399bp to: Candidatus Sphaeronema italicum)	**0.1%**	0.1%
s	**Dongia sp. URHE0060** (1 OTU with 94% identity in 399bp to: Dongia sp. URHE0060)	**0.1%**	0.2%
f	**Sphingomonadaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Sphingosinicella strains, Sphingomonas sp. B9, Sphingosinicella microcystinivorans, Sphingosinicella xenopeptidilytica)	**0.1%**	0.1%
g	**Zoogloea** (1 OTU with 100% identity in 424bp to: Zoogloea caeni, Zoogloea sp. Chol3)	**0.1%**	0.2%
o	**Desulfuromonadales** (1 OTU with 85% identity in 429bp to: Desulfuromonas carbonis, Desulfuromonas sp. DDH964, Geoalkalibacter subterraneus)	**0.1%**	0.1%
s	**Raoultibacter timonensis** (1 OTU with 89% identity in 399bp to: Raoultibacter timonensis)	**0.1%**	0.1%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.1%**	0.2%
s	**Vicingus serpentipes** (1 OTU with 93% identity in 424bp to: Vicingus serpentipes)	**0.1%**	0.1%
g	**Streptomyces** (1 OTU with 80% identity in 422bp to: 3 unclassified Streptomyces strains, Streptomyces aurantiacus, Streptomyces caniferus)	**0.1%**	0.1%
s	**Maribellus luteus** (1 OTU with 95% identity in 422bp to: Maribellus luteus)	**0.1%**	0.1%
f	**Prolixibacteraceae** (1 OTU with 93% identity in 422bp to: Maribellus luteus, Tangfeifania diversioriginum)	**0.1%**	0.1%
s	**Limnobacter sp.** (1 OTU with 99% identity in 424bp to: Limnobacter sp.)	**0.1%**	0.2%
	Other	**0.9%**	1.1%
	Unclassified (365 reads)		
	Filtered (0 reads)		

R2.Sumpf.V3V4a (22 296 reads)

g	**Thauera** (41 OTUs with 99-100% identity in 283-424bp to: 14 unclassified Thauera strains, Thauera aminoaromatica, Thauera phenylacetica, Thauera selenatis)	**47.5%**	58.0%
s	**Denitratisoma oestradiolicum** (6 OTUs with 94-96% identity in 424bp to: Denitratisoma oestradiolicum)	**8.4%**	3.4%
f	**Sterolibacteriaceae** (2 OTUs with 95-96% identity in 424bp to: Denitratisoma sp. DHT3, Methyloversatilis sp. PL20, Sulfurisoma sediminicola, Sulfuritalea hydrogenivorans)	**3.6%**	1.5%
s	**Sulfuritalea hydrogenivorans** (1 OTU with 98% identity in 424bp to: Sulfuritalea hydrogenivorans)	**3.1%**	1.3%
s	**Gemmatimonas aurantiaca** (5 OTUs with 89-93% identity in 402bp to: Gemmatimonas aurantiaca)	**2.8%**	1.2%
s	**Ignavibacterium album** (5 OTUs with 82-95% identity in 422-425bp to: Ignavibacterium album)	**2.5%**	2.4%
s	**Holophaga sp. WY42** (4 OTUs with 88-90% identity in 424-425bp to: Holophaga sp. WY42)	**1.9%**	0.8%
s	**Thauera sp. JPB-3.02** (2 OTUs with 99-100% identity in 424bp to: Thauera sp. JPB-3.02)	**1.8%**	2.1%

s	**Desulforegula conservatrix** (2 OTUs with 81% identity in 424bp to: Desulforegula conservatrix)	**1.7%**	2.4%
g	**Thiocapsa** (1 OTU with 99% identity in 424bp to: 3 unclassified Thiocapsa strains)	**1.4%**	1.2%
s	**Paludibaculum sp.** (1 OTU with 93% identity in 399bp to: Paludibaculum sp.)	**1.3%**	0.6%
s	**Desulfocarbo indianensis** (1 OTU with 89% identity in 430bp to: Desulfocarbo indianensis)	**1.1%**	1.0%
s	**Hypericibacter terrae** (1 OTU with 94% identity in 399bp to: Hypericibacter terrae)	**1.1%**	1.6%
s	**Aliifodinibius sp.** (1 OTU with 81% identity in 423bp to: Aliifodinibius sp.)	**1.0%**	1.0%
s	**Thiothrix unzii** (1 OTU with 100% identity in 402bp to: Thiothrix unzii)	**0.9%**	0.6%
s	**Parvibaculum sp.** (1 OTU with 93% identity in 399bp to: Parvibaculum sp.)	**0.9%**	0.9%
g	**Mesorhizobium** (3 OTUs with 99-100% identity in 399bp to: 4 unclassified Mesorhizobium strains, Mesorhizobium carbonis, Mesorhizobium sangaii, alpha proteobacterium WG1)	**0.8%**	0.6%
s	**Pandoraea sp.** (3 OTUs with 93-94% identity in 424bp to: Pandoraea sp.)	**0.8%**	1.7%
c	**Betaproteobacteria** (2 OTUs with 98% identity in 424bp to: 3 unclassified Simplicispira strains, Aquaspirillum sp. R-22832)	**0.8%**	0.7%
s	**Thiobacter sp. SL-1** (2 OTUs with 92-95% identity in 424-425bp to: Thiobacter sp. SL-1)	**0.7%**	1.0%
o	**Rhizobiales** (2 OTUs with 95-99% identity in 399-400bp to: 12 unclassified Bradyrhizobium strains, 13 unclassified Sinorhizobium strains, 39 unclassified Ensifer strains, Bradyrhizobium genosp. Y, Bradyrhizobium vignae, Ensifer glycinis, Ensifer mexicanus, Ensifer shofinae, Mesorhizobium sp., Methylocapsa aurea, Sinorhizobium americanum, Sinorhizobium chiapanecum, Sinorhizobium fredii, Sinorhizobium terangae)	**0.7%**	0.6%
f	**Bradyrhizobiaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Rhodopseudomonas strains, Nitrobacter sp. NS5-5, Rhodopseudomonas faecalis)	**0.7%**	0.4%
s	**Nitrosomonas sp.** (1 OTU with 99% identity in 424bp to: Nitrosomonas sp.)	**0.7%**	0.3%
s	**Geobacter sulfurreducens** (2 OTUs with 99-100% identity in 425bp to: Geobacter sulfurreducens)	**0.7%**	0.5%
f	**Bacillaceae** (2 OTUs with 83% identity in 430bp to: Bacillus sp. G3DM-87, Terribacillus aidingensis, Terribacillus halophilus, Terribacillus saccharophilus, Terribacillus sp.)	**0.6%**	2.3%
g	**Geobacter** (3 OTUs with 90-100% identity in 425bp to: 4 unclassified Geobacter strains, Geobacter sulfurreducens)	**0.6%**	0.6%
s	**Chryseolinea serpens** (1 OTU with 95% identity in 423bp to: Chryseolinea serpens)	**0.5%**	0.6%
g	**Desulfovibrio** (2 OTUs with 84% identity in 425bp to: 4 unclassified Desulfovibrio strains, Desulfovibrio fairfieldensis)	**0.5%**	0.7%
g	**Nitrospira** (1 OTU with 98% identity in 417bp to: Nitrospira cf. moscoviensis SBR1015, Nitrospira cf. moscoviensis SBR1024, Nitrospira sp.)	**0.5%**	0.2%
s	**Rhodoplanes sp. Z2-YC6860** (1 OTU with 97% identity in 400bp to: Rhodoplanes sp. Z2-YC6860)	**0.5%**	0.3%
g	**Hyphomicrobium** (1 OTU with 99% identity in 399bp to: 4 unclassified Hyphomicrobium strains, Hyphomicrobium facile)	**0.5%**	0.3%
s	**Coxiella endosymbiont of Ornithodoros marocanus** (1 OTU with 92% identity in 425bp to: Coxiella endosymbiont of Ornithodoros marocanus)	**0.5%**	0.2%
s	**Denitromonas sp. D12-45.1** (1 OTU with 99% identity in 424bp to: Denitromonas sp. D12-45.1)	**0.4%**	0.5%
s	**Desulfoprunum benzoelyticum** (2 OTUs with 98-99% identity in 428bp to: Desulfoprunum benzoelyticum)	**0.4%**	0.5%
s	**Halochromatium sp. MTK6IM088** (1 OTU with 88% identity in 426bp to: Halochromatium sp. MTK6IM088)	**0.3%**	0.2%
s	**Pseudorhodoplanes sinuspersici** (1 OTU with 99% identity in 399bp to: Pseudorhodoplanes sinuspersici)	**0.3%**	0.3%
s	**Hyphomicrobium sp.** (1 OTU with 97% identity in 399bp to: Hyphomicrobium sp.)	**0.3%**	0.2%
s	**Candidatus Aquirestis calciphila** (1 OTU with 88% identity in 422bp to: Candidatus Aquirestis calciphila)	**0.3%**	0.3%
s	**Simkania negevensis** (1 OTU with 85% identity in 427bp to: Simkania negevensis)	**0.3%**	0.2%
g	**Simplicispira** (2 OTUs with 100% identity in 424bp to: 3 unclassified Simplicispira strains, Simplicispira limi, Simplicispira piscis)	**0.3%**	0.2%
s	**Raoultibacter timonensis** (1 OTU with 89% identity in 399bp to: Raoultibacter timonensis)	**0.3%**	0.1%
s	**Thioprofundum lithotrophicum** (1 OTU with 94% identity in 424bp to: Thioprofundum lithotrophicum)	**0.2%**	0.2%
s	**Denitratimonas tolerans** (1 OTU with 100% identity in 424bp to: Denitratimonas tolerans)	**0.2%**	0.3%
s	**Geothrix sp.** (1 OTU with 84% identity in 427bp to: Geothrix sp.)	**0.2%**	0.1%
g	**Pseudomonas** (1 OTU with 94% identity in 424bp to: 2 unclassified Pseudomonas strains)	**0.2%**	0.4%
s	**Candidatus Koribacter versatilis** (1 OTU with 84% identity in 411bp to: Candidatus Koribacter versatilis)	**0.2%**	0.1%

g	**Lewinella** (1 OTU with 86% identity in 423bp to: Lewinella litorea, Lewinella lutea, Lewinella marina, Lewinella maritima, Lewinella sp. SRO_346)	**0.2%**	0.2%
s	**Hyphomicrobium sp. Pd-S-(l)-e-D-8(2)** (1 OTU with 99% identity in 399bp to: Hyphomicrobium sp. Pd-S-(l)-e-D-8(2))	**0.2%**	0.1%
s	**Desulfonatronum thiosulfatophilum** (1 OTU with 90% identity in 426bp to: Desulfonatronum thiosulfatophilum)	**0.2%**	0.2%
s	**Desulfomonile tiedjei** (1 OTU with 99% identity in 428bp to: Desulfomonile tiedjei)	**0.2%**	0.1%
g	**Aquimonas** (1 OTU with 92% identity in 424bp to: 4 unclassified Aquimonas strains)	**0.2%**	0.3%
s	**Elioraea tepidiphila** (1 OTU with 100% identity in 399bp to: Elioraea tepidiphila)	**0.2%**	0.2%
g	**Zoogloea** (2 OTUs with 99-100% identity in 424bp to: Zoogloea caeni, Zoogloea sp. Chol3)	**0.2%**	0.2%
s	**Geobacter sp. Ply1** (1 OTU with 98% identity in 425bp to: Geobacter sp. Ply1)	**0.2%**	0.2%
f	**Sphingomonadaceae** (1 OTU with 100% identity in 399bp to: 2 unclassified Sphingosinicella strains, Sphingomonas sp. B9, Sphingosinicella microcystinivorans, Sphingosinicella xenopeptidilytica)	**0.2%**	0.1%
g	**Gaiella** (1 OTU with 96% identity in 423bp to: 2 unclassified Gaiella strains)	**0.2%**	0.1%
s	**Aquicella siphonis** (1 OTU with 93% identity in 424bp to: Aquicella siphonis)	**0.1%**	0.1%
s	**Mesorhizobium sp.** (1 OTU with 100% identity in 399bp to: Mesorhizobium sp.)	**0.1%**	0.1%
s	**Limnobacter sp.** (1 OTU with 99% identity in 424bp to: Limnobacter sp.)	**0.1%**	0.3%
s	**Thauera mechernichensis** (1 OTU with 100% identity in 424bp to: Thauera mechernichensis)	**0.1%**	0.2%
s	**Pseudolabrys sp. W2.09-308** (1 OTU with 97% identity in 399bp to: Pseudolabrys sp. W2.09-308)	**0.1%**	0.1%
s	**Hypericibacter adhaerens** (1 OTU with 96% identity in 399bp to: Hypericibacter adhaerens)	**0.1%**	0.2%
s	**Candidatus Sphaeronema italicum** (1 OTU with 91% identity in 399bp to: Candidatus Sphaeronema italicum)	**0.1%**	0.1%
c	**Actinobacteria** (1 OTU with 100% identity in 404bp to: 23 unclassified Nocardia strains, Nocardia asteroides, Nocardia higoensis, Nocardia neocaledoniensis, Nocardia nova, Nocardia rhizosphaerihabitans, Streptomyces sp.)	**0.1%**	0.1%
s	**Roseisolibacter agri** (1 OTU with 91% identity in 416bp to: Roseisolibacter agri)	**0.1%**	0.1%
s	**Aquincola sp.** (1 OTU with 99% identity in 424bp to: Aquincola sp.)	**0.1%**	0.2%
s	**Vicingus serpentipes** (1 OTU with 93% identity in 424bp to: Vicingus serpentipes)	**0.1%**	0.1%
s	**Lewinella cohaerens** (1 OTU with 88% identity in 426bp to: Lewinella cohaerens)	**0.1%**	0.1%
s	**Lentimicrobium saccharophilum** (1 OTU with 96% identity in 423bp to: Lentimicrobium saccharophilum)	**0.1%**	0.1%
s	**Maribellus luteus** (1 OTU with 95% identity in 422bp to: Maribellus luteus)	**0.1%**	0.1%
	Other	**1.4%**	1.7%
	Unclassified (73 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 outperformes 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 resultion.** 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% accross 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/nt
02-03_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.
- **Alignment:** Caporaso JG et al. (2010) PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26, 266-267.
- **Alpha diversity:** Faith PD, Baker AM (2006) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. Evol Bioinform Online. 2006; 2: 121-128. Chao A (1984) Nonparametric estimation of the number of classes in a population. Scand J Stat. 1984;11:265-270.

- **Beta diversity:** Lozupone C, Knight R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71(12): 8228-8235. Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. (2013) Emperor: A tool for visualizing high-throughput microbial community data. Gigascience 2(1):16.
- **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 indentified taxonomic units for each sample. The first two columns are the sample name and the total number of reads that where 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 where 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 abundancy-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 to 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.

11_Clustering: This directory contains sample trees in Newick tree format. The trees were derived from the distance matrices created by the beta diversity analysis using UPGMA clustering. The trees can be viewed by any Newick tree viewer, e.g. by FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

12_Diversity_plots: This directory contains a file `index.html`. This file can be opened with the Chrome, Safari, or Firefox web browser. Please note that some versions of Safari do not support WebGL in Microsoft Windows environments. In such a case, please choose another browser, e.g. Firefox. The **index.html** serves as an entry point to view the results of the alpha and beta diversity analysis. Alpha diversity has been calculated using the Chao1 and Faith's Phylogenetic Diversity metrics. Beta diversity has been calculated using the unweighted and weighted UniFrac distance. Alpha and beta diversities were calculated using the tree given in file `rep_set.tre` in the `12_Diversity_plots` directory. Quantitative measures (e.g. weighted UniFrac) are ideally suited to revealing community differences that are due to changes in relative taxon abundance. Qualitative measures (e.g. unweighted UniFrac) are most informative when communities differ primarily by what can live in them, in part because abundance information can obscure significant patterns of variation.