



MASSIVE
BIOINFORMATICS

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

by Massive Bioinformatics

Version: 1.1.0

Sample Name: SE

EXECUTIVE SUMMARY

QC Verdict

Total Reads	131036
Reads After Filtering	131036
Classified	99.60%

Diversity Tier

Moderate

Top Species

Citrobacter freundii	36.95%
Citrobacter braakii	15.33%
Segatella copri	8.02%
Citrobacter murlinae	4.42%
Enterobacter sichuanensis	2.45%

Functional Highlights

- peptidoglycan maturation (meso-diaminopimelate ...
- (S)-propane-1,2-diol degradation
- glycolysis IV

Classification Confidence

1.1% of the community was classified with excellent confidence.

1. METHOD

1.1 Wet Lab Protocol

Library preparation was performed using the Native Barcoding Kit 24 (SQK-NBD114.24; Oxford Nanopore Technologies). Initially, samples underwent end-preparation, dA-tailing, and nick-repair in a final reaction volume of 10 μ l. Barcodes were subsequently ligated to each sample, followed by purification using MobioMx MagBeads and pooling. Sequencing adapters were ligated to the pooled samples, which were then purified and quantified spectrofluorimetrically. The final library was loaded onto a primed PromethION flow cell (FLO-PRO114M), and the sequencing run was initiated on a PromethION P2 Solo™ device (Oxford Nanopore Technologies) using the latest MinKNOW™ software. Sequencing proceeded until sufficient data yield was achieved or the maximum run time of 72 hours was reached.

1.2 Bioinformatics Analysis

Raw sequencing data in POD5 format was converted to FASTQ format using Dorado software for base-calling and demultiplexing. Quality control was performed using Fastplong with parameters: 0 bp trimmed from read starts, minimum Phred score 10, maximum 30% low-quality bases, minimum mean quality 15, and read length filter ≥ 1000 bp. Reads failing thresholds were excluded.

Reads were classified using the Massive Alignment Engine (MAE) with minimap2 alignment against ncbi 16s containing 16S ribosomal RNA gene sequences. MAE uses an EM algorithm to estimate species abundances from alignment probabilities. Taxa below 0.01% abundance were filtered.

Taxonomic profiles were converted to BIOM format. Taxonomy summaries were generated at all levels. Alpha diversity indices (Shannon, Simpson, Chao1, Fisher's alpha, McIntosh) were calculated using scikit-bio. Community complexity was assessed using Pielou's Evenness, effective species number, Berger-Parker dominance, and Good's coverage. Per-taxon classification confidence was derived from EM-algorithm read-assignment probabilities. Functional metabolic pathway prediction was performed using PICRUSt2.

Consensus marker gene sequences were reconstructed using the following strategies. All quality-filtered reads were oriented to a consistent strand direction using minimap2 self-alignment, then profiled using canonical 5-mer frequency composition, projected into a low-dimensional embedding with UMAP (Bray-Curtis metric), and clustered using adaptive HDBscan (first pass). Each cluster consensus was polished with 2 rounds of Racon error correction. Over-split clusters were then merged by comparing polished consensus sequences at 97% sequence identity (second pass). Chimeric sequences were detected and removed using VSEARCH uchime_denovo. Taxonomy was assigned to each cluster by majority vote over EM read assignments; clusters with no classified reads were flagged as unknown/novel sequences. All reads were remapped to the final consensus sequences for accurate abundance counts. Species-level representative sequences and feature tables were generated for QIIME2 diversity analysis and PICRUSt2 functional prediction. Taxa/clusters with fewer than 5 reads were excluded from consensus generation. Visualizations were created with R/ggplot2, matplotlib, and Krona.

2. GENERAL QUALITY STATISTICS

Summary

	Before Filtering	After Filtering
Total Reads	131036	131036
Read Mean Length	1584	1457
GC Content	53.85%	54.51%

Filtering Result

	Count	%
Passed Filter Reads	131,036	100.00%
Low Quality Reads	2	0.00%
Too Many N Reads	0	0.00%
Too Short Reads	32	0.02%
Too Long Reads	0	0.00%

Classification Report

	Count	%
Total Reads Used in Classification	131,035	
Unmapped	96	0.07%
Mapped Classified	130,510	99.60%
Mapped Unclassified	429	0.33%

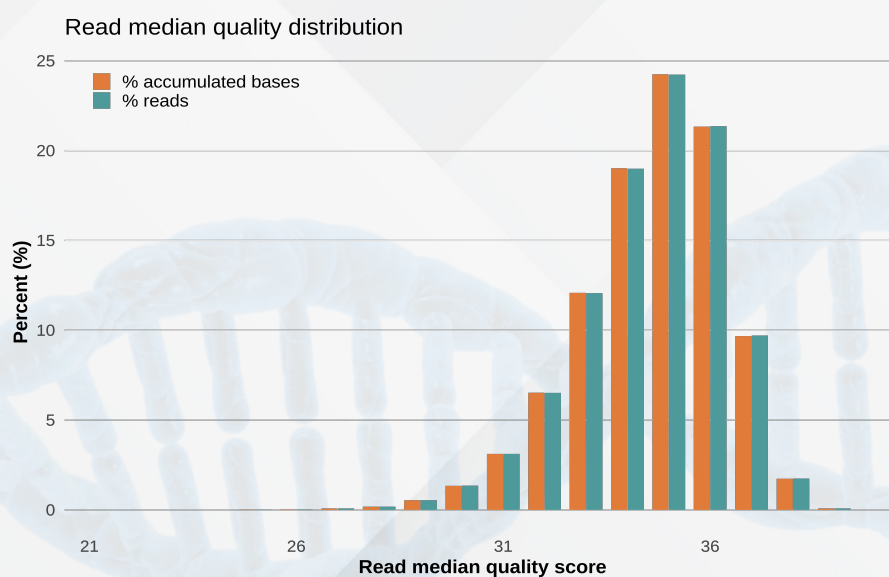


Figure 1. Per-Read Quality Distribution

3. BIOINFORMATICS ANALYSIS

3.1 Alpha Diversity

Alpha diversity measures the variety of species within a single sample. It combines two aspects: how many different species are present (richness) and how evenly they are distributed (evenness). Higher alpha diversity generally indicates a more stable and resilient microbial community. The table and radar chart below summarize the key diversity indices calculated for this sample.

Diversity Indices

Shannon	Simpson	Chao1	Fisher Alpha	Enspie	McIntosh E	McIntosh D
2.818	0.828	201.0	23.289	5.813	0.415	0.587

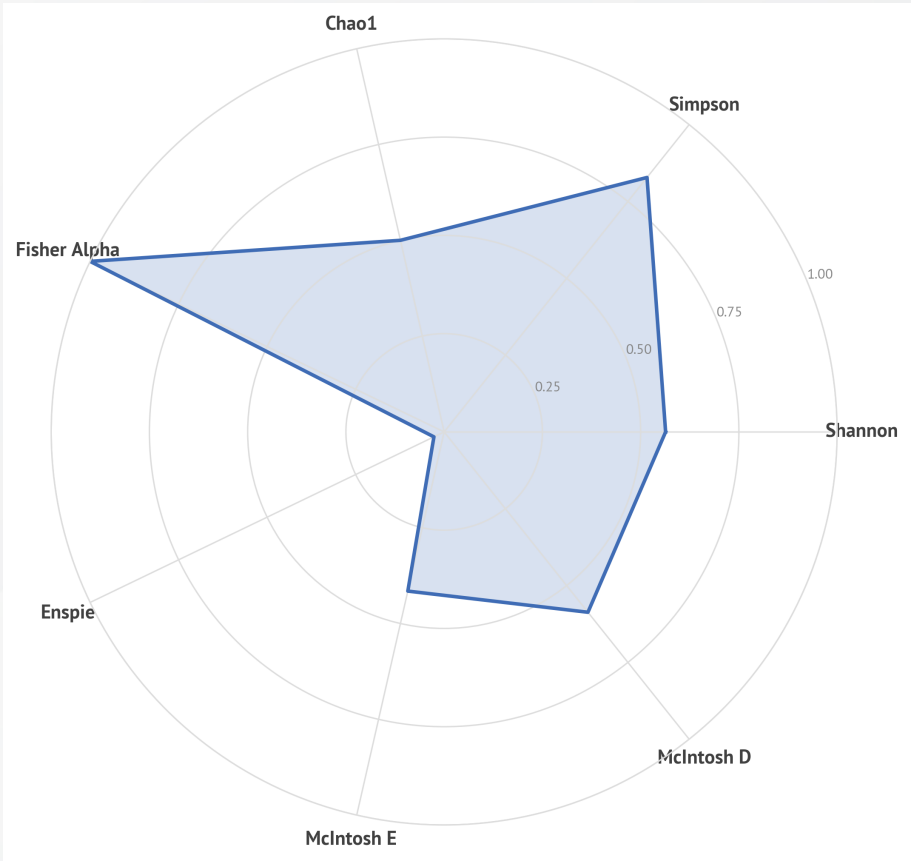


Figure 2. Alpha Diversity Radar Chart

This sample exhibits **moderate** overall diversity (Shannon $H' = 2.818$). The estimated species richness (Chao1 = 201) matches the observed count of 201 species, suggesting sampling has captured the full community. Species evenness (McIntosh E = 0.415) indicates a moderately balanced community. Together, the Shannon and Simpson (0.828) indices suggest a moderately diverse microbial community.

Index Definitions

Shannon: Combines species count and their relative abundances into a single entropy-based score. Higher values indicate greater diversity. Typical values for microbial communities range from 1.5 (low diversity) to over 3.5 (high diversity).

Simpson: Reflects the probability that two randomly chosen individuals belong to different species. Values range from 0 to 1, where higher values indicate lower dominance by any single species and thus greater effective diversity.

Chao1: Estimates total species richness, including undetected rare species, based on the frequency of singletons and doubletons. When Chao1 is close to the observed species count, sampling is considered sufficient; a much higher estimate suggests additional rare species remain undetected.

Fisher Alpha: Estimates species richness from the relationship between the number of species and the number of individuals. It is relatively robust to sample size differences and is widely used for comparing richness across samples of unequal depth.

Evenness (Enspie): Measures how equally individuals are distributed among species. Values range from 0 to 1, where values near 1 indicate that all species are present at similar abundances, while low values indicate dominance by one or a few species.

McIntosh E: An evenness index derived from McIntosh D, focusing on how uniformly species are represented. Values range from 0 to 1, where 1 represents a perfectly even community.

McIntosh D: A distance-based diversity measure that balances species richness and evenness. It quantifies how far the community's abundance distribution is from a perfectly even distribution, with higher values indicating greater diversity.

3.2 Sample Complexity Assessment

Community complexity describes how diverse and balanced the microbial population in the sample is. A **low-complexity** community is dominated by one or a few organisms, typical of acute infections or highly selective environments, while a **high-complexity** community has many species at similar abundances, indicating a rich and balanced ecosystem.

The gauge bar below shows **Pielou's Evenness** (J), the primary complexity indicator, on a 0–1 scale. Values near 0 indicate strong dominance by few species; values near 1 indicate even distribution. The metric cards summarize key ecological indices used in the assessment.

Note: This section focuses on how evenly species are distributed (Pielou's J), whereas the diversity assessment in Section 3.1 also takes into account how many species are present and how dominant they are (Shannon + Simpson). These complementary perspectives may yield different tier labels for the same sample.

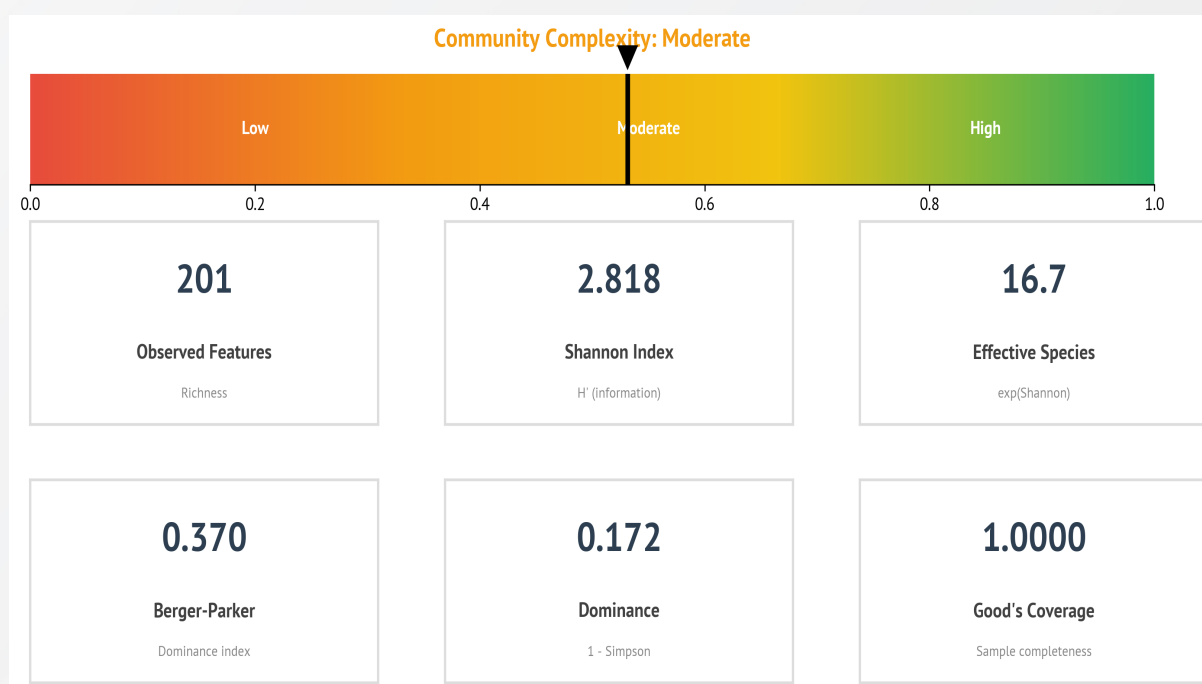


Figure 3. Sample Complexity Assessment Panel

This sample shows **moderate** community complexity (Pielou's J = 0.531). Out of 201 detected species, only 16.7 contribute meaningfully to the community profile; the remainder are present at very low abundances. The dominant phylum is **Pseudomonadota** (72.5%), and 100.0% of reads could be classified down to genus level. Good's coverage (100.0%) confirms that the sequencing depth was sufficient to capture the microbial community present in the sample.

1.1% of the community was classified with excellent confidence.

3.3 Alpha Rarefaction

Alpha rarefaction estimates whether the sequencing run produced enough data to reliably detect all species present in the sample. The curve below plots the number of distinct species detected (Y-axis) against the number of reads sampled (X-axis). At low read counts, new species are discovered rapidly; as more reads are added, the rate of new discoveries slows.

How to read the curve: If the curve flattens into a plateau toward the right side, the sequencing depth was sufficient, most species in the sample have been captured and the diversity metrics reported elsewhere in this document are reliable. If the curve is still rising steeply at the rightmost point, additional sequencing could reveal more species, and the reported diversity may underestimate the true community composition.

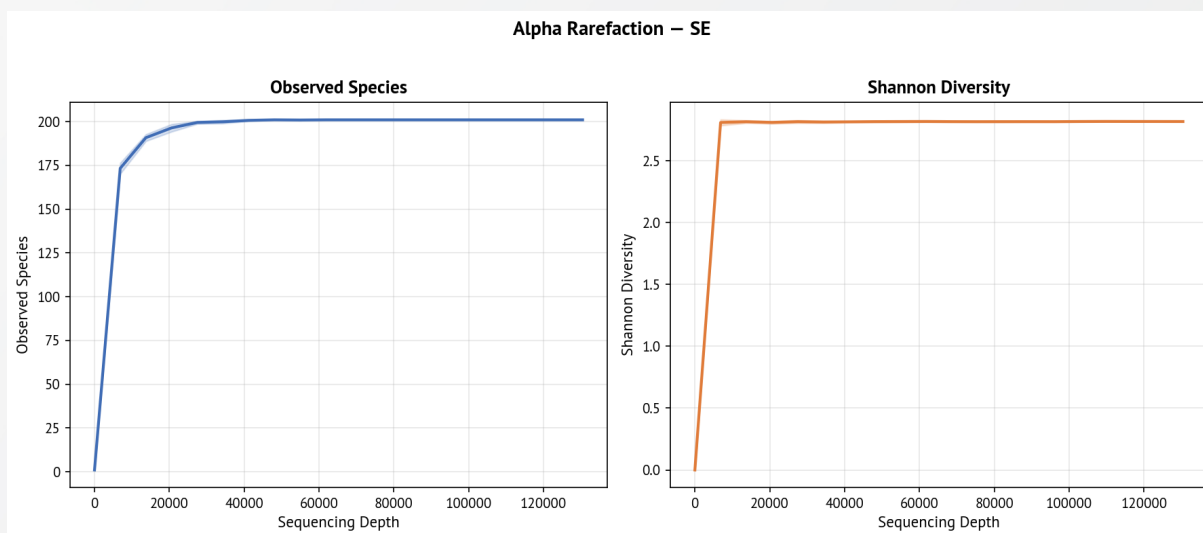


Figure 4. Alpha Rarefaction Curve

3.4 Abundance Information According to Taxonomic Levels

The taxonomic composition of bacterial and archaeal organisms in the sample was examined at different taxonomic levels. The relative abundances at each level are shown in the bar charts below.

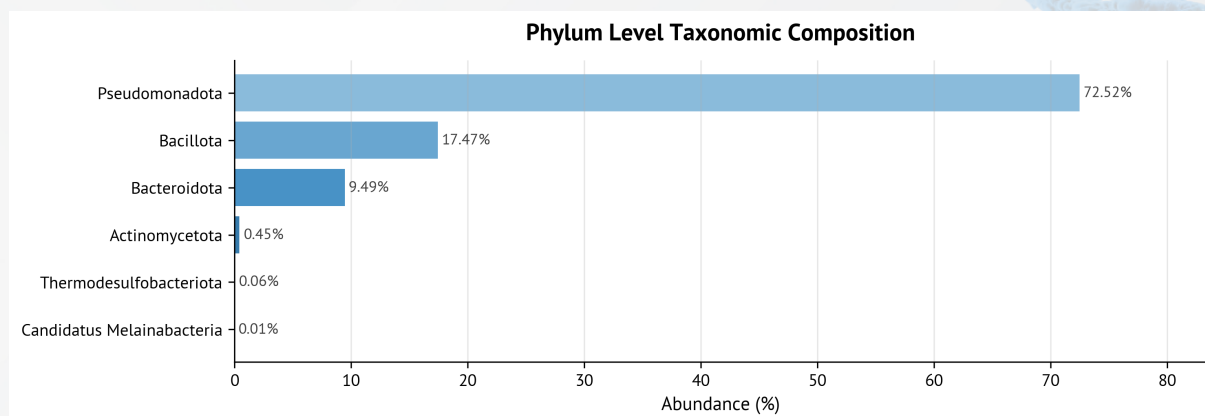


Figure 5. Bar Chart at Phylum Level

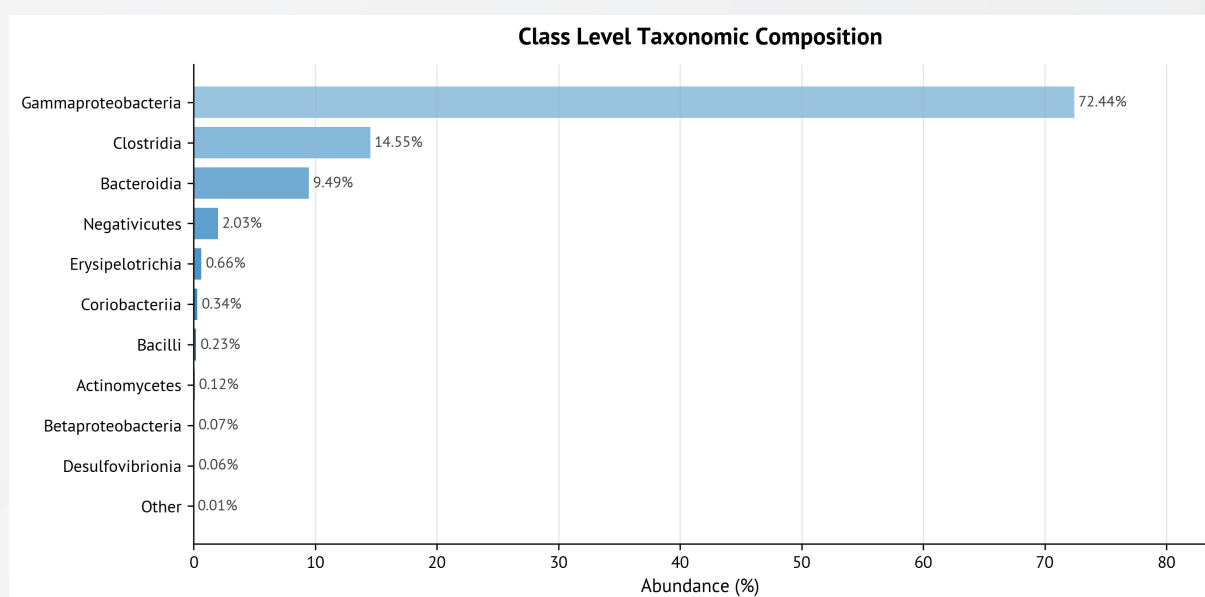


Figure 6. Bar Chart at Class Level

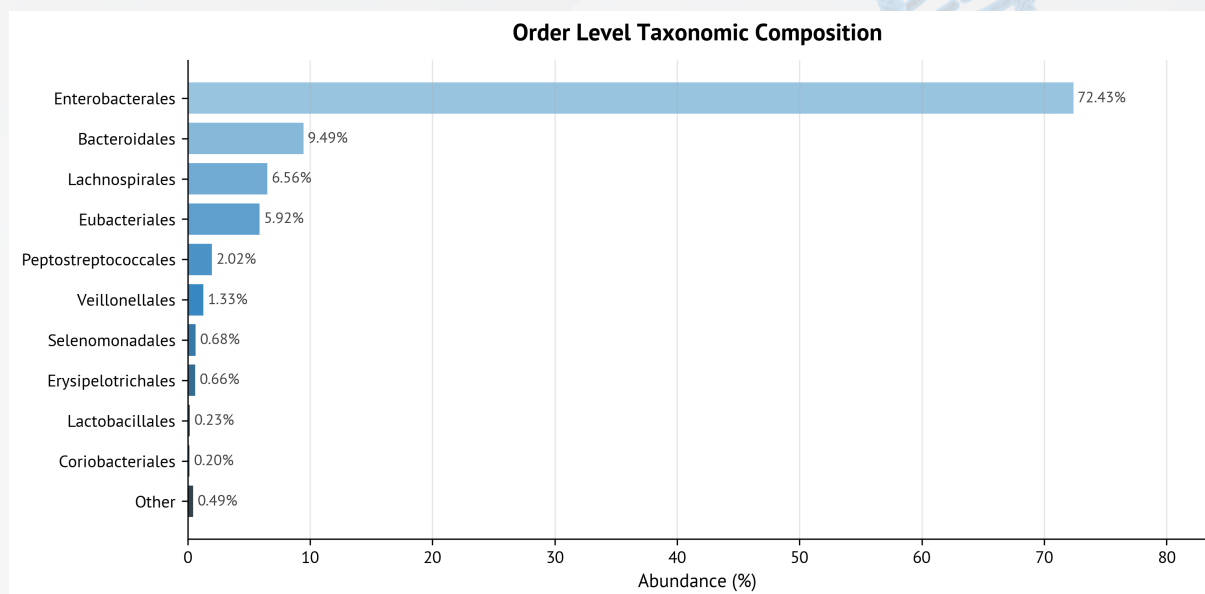


Figure 7. Bar Chart at Order Level

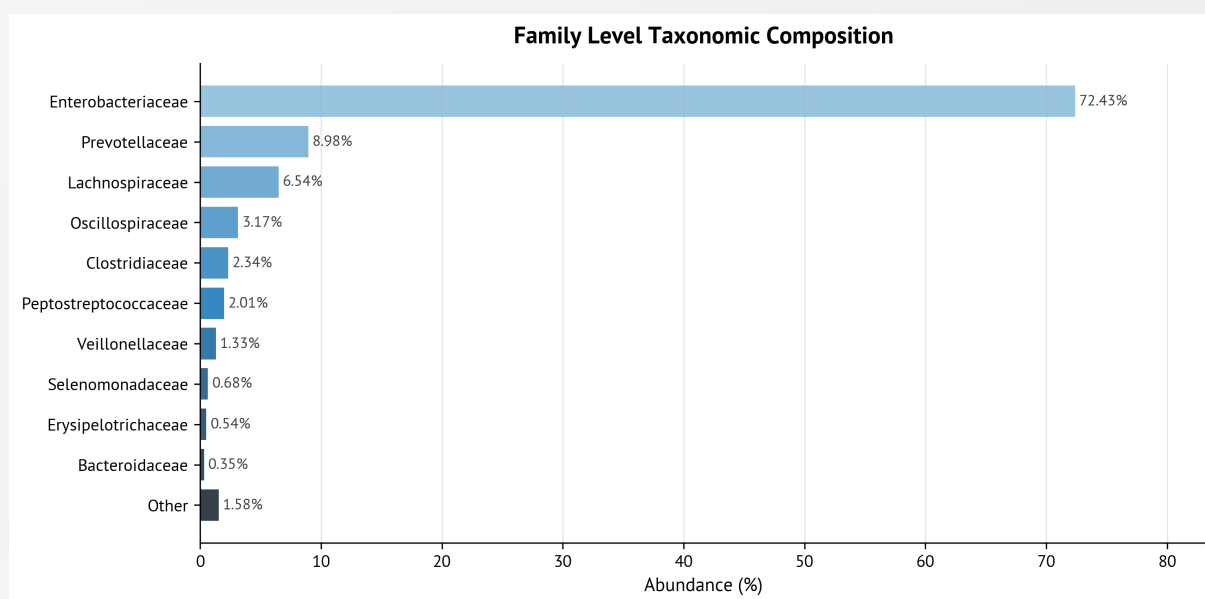


Figure 8. Bar Chart at Family Level

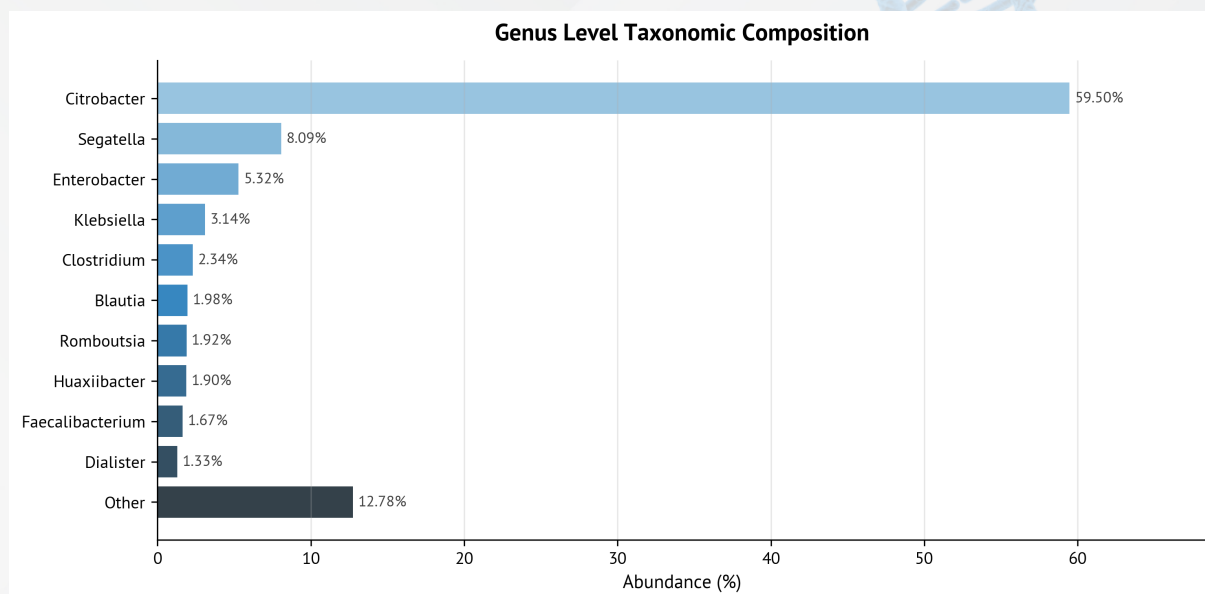


Figure 9. Bar Chart at Genus Level

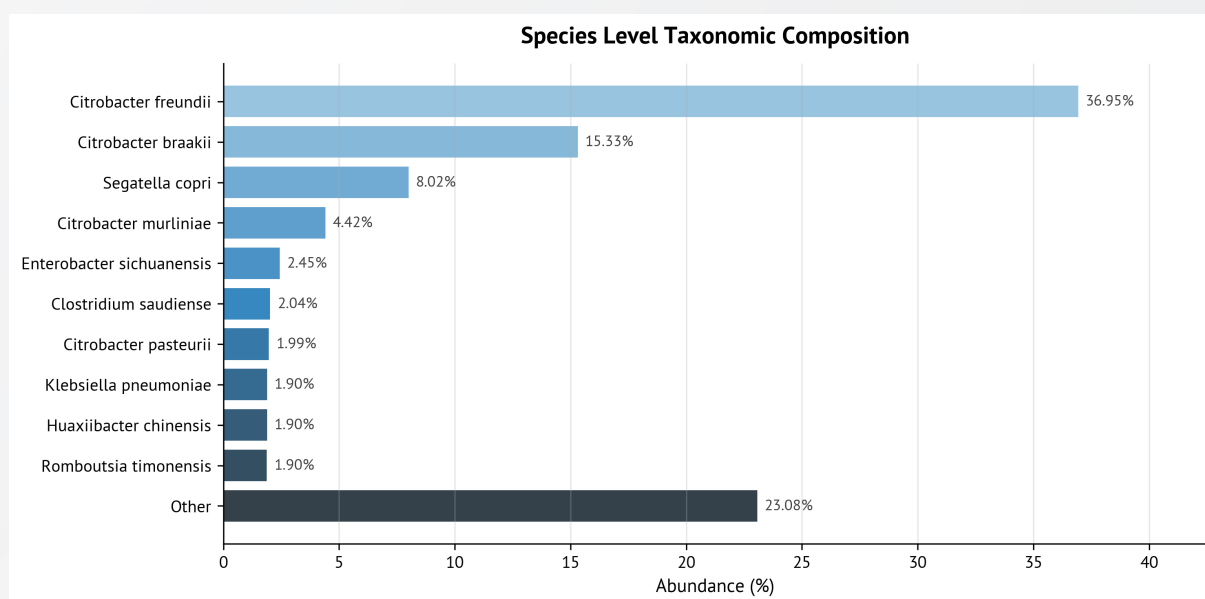


Figure 10. Bar Chart at Species Level

3.5 Taxonomic Flow Across Ranks

The Sankey diagram traces how reads distribute across the eight taxonomic ranks, from Domain on the left to Species on the right. Each coloured ribbon is one lineage; its width is proportional to the number of reads it carries, and its colour is held constant across every rank so a species can be followed visually from its root to its leaf. Ranks are stacked by descending read count and separated by ribbons that merge and split as taxa aggregate into common ancestors. Low-abundance species below the read-share threshold are rolled into a single grey "Other" ribbon; labels are drawn only on strata large enough to remain legible.

Taxonomic flow across ranks

Top 20 species – each colour tracks one lineage from Domain to Species

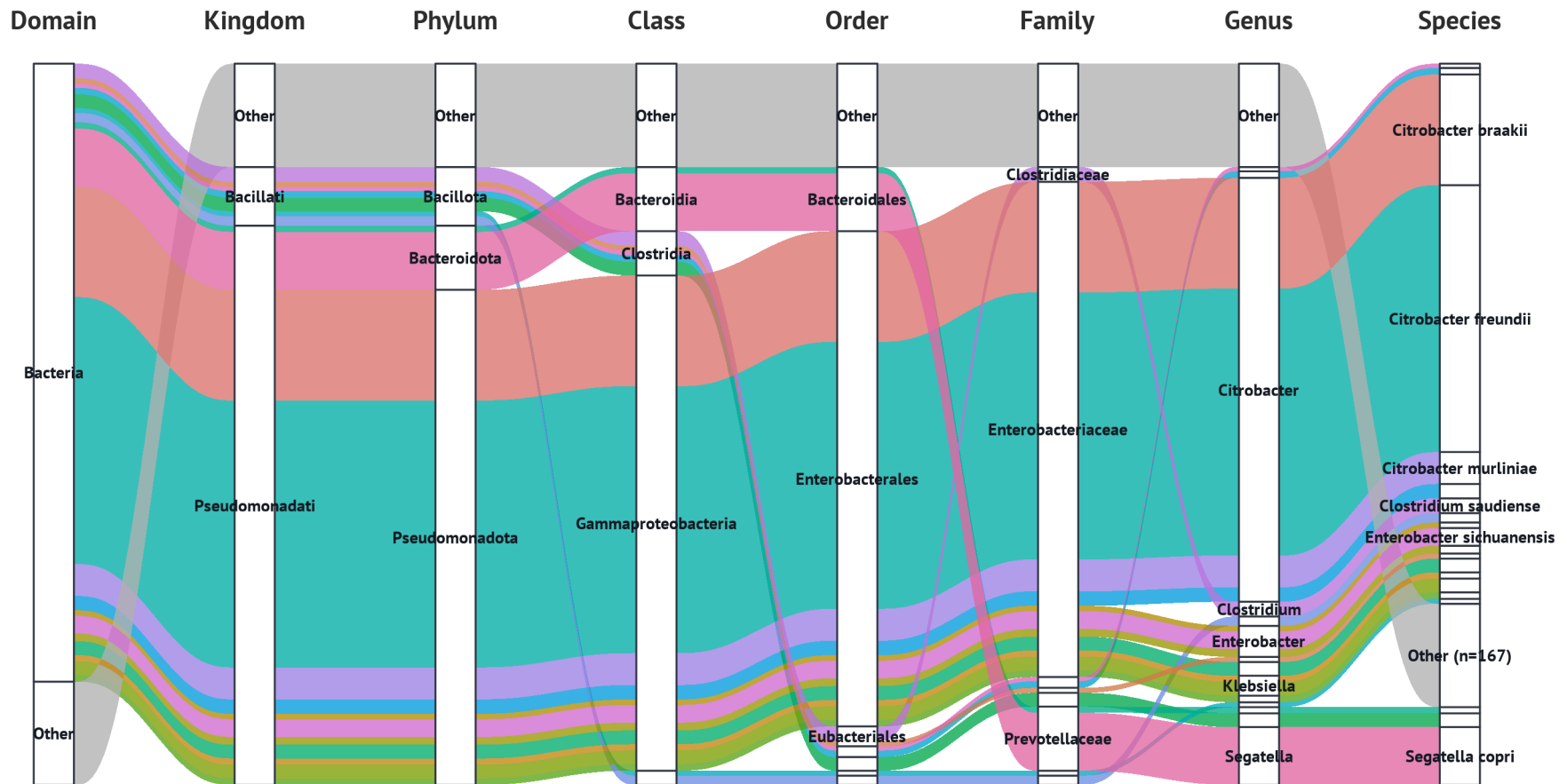


Figure 11. Taxonomic Sankey Diagram (top species by read count)

3.6 Functional Prediction

PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) predicts the functional potential of microbial communities based on 16S rRNA marker gene sequences. The predicted metabolic pathways represent the inferred functional capacity of the community.

Table 1. Top Predicted Metabolic Pathways (PICRUSt2)

Pathway ID	Description	Rel. Abundance (%)
PWY0-1586	peptidoglycan maturation (meso-diaminopimelate containing)	1.71
PWY-7013	(S)-propane-1,2-diol degradation	1.06
PWY-1042	glycolysis IV	1.05
PWY0-1296	purine ribonucleosides degradation	0.86
PWY0-1477	ethanolamine utilization	0.80
PWY-5695	inosine 5'-phosphate degradation	0.79
PWY-6609	adenine and adenosine salvage III	0.76
PWY-8178	pentose phosphate pathway (non-oxidative branch) II	0.72
PWY0-301	L-ascorbate degradation I (bacterial, anaerobic)	0.71
PWY4FS-8	phosphatidylglycerol biosynthesis II (non-plastidic)	0.68
PWY4FS-7	phosphatidylglycerol biosynthesis I (plastidic)	0.68
PWY0-1319	CDP-diacylglycerol biosynthesis II	0.67
PWY-5667	CDP-diacylglycerol biosynthesis I	0.67
PWY-7663	gondoate biosynthesis (anaerobic)	0.66
FERMENTATION-PWY	mixed acid fermentation	0.66
PWY-5659	GDP-mannose biosynthesis	0.65
NONOXIPENT-PWY	pentose phosphate pathway (non-oxidative branch) I	0.64
GOLPDLCAT-PWY	superpathway of glycerol degradation to 1,3-propanediol	0.62
PHOSLIPSYN-PWY	superpathway of phospholipid biosynthesis I (bacteria)	0.58
PWY-5136	fatty acid & beta;-oxidation II (plant peroxisome)	0.56

Table 2. Pathway Category Summary (PICRUSt2)

Category	Rel. Abundance (%)
Biosynthesis	51.29
Degradation	20.87
TCA/Energy Cycles	4.43
Superpathway	2.17
Fermentation	1.88
Other	19.35

3.7 ABUNDANCE TABLE

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
<i>Citrobacter freundii</i>	Species	48191	36.95	Fair
<i>Citrobacter braakii</i>	Species	19994	15.33	Fair
<i>Segatella copri</i>	Species	10465	8.024	Excellent
<i>Citrobacter murlinae</i>	Species	5770	4.424	Fair
<i>Enterobacter sichuanensis</i>	Species	3200	2.454	Fair
<i>Clostridium saudiense</i>	Species	2663	2.042	Good
<i>Citrobacter pasteurii</i>	Species	2592	1.987	Fair
<i>Klebsiella pneumoniae</i>	Species	2483	1.904	Fair
<i>Huaxiibacter chinensis</i>	Species	2481	1.902	Fair
<i>Romboutsia timonensis</i>	Species	2477	1.899	Excellent
<i>Dialister succinatiphilus</i>	Species	1694	1.299	Excellent
<i>Enterobacter wuhouensis</i>	Species	1407	1.079	Fair
<i>Kluyvera ascorbata</i>	Species	1183	0.907	Fair
<i>Blautia luti</i>	Species	1166	0.894	Good
<i>Klebsiella aerogenes</i>	Species	1123	0.861	Fair
<i>Prevotellamassilia timonensis</i>	Species	1117	0.856	Excellent
<i>Enterobacter quasihormaechei</i>	Species	999	0.766	Fair
<i>Faecalibacterium hattorii</i>	Species	920	0.705	Fair
<i>Mitsuokella jalaludinii</i>	Species	885	0.679	Excellent
<i>Agathobacter rectalis</i>	Species	782	0.6	Good
<i>Anaerobutyricum soehngenii</i>	Species	716	0.549	Good
<i>Holdemanella porci</i>	Species	699	0.536	Excellent
<i>Anaerostipes hadrus</i>	Species	583	0.447	Good
<i>Faecalibacterium duncaniae</i>	Species	574	0.44	Fair
<i>Citrobacter gillenii</i>	Species	570	0.437	Fair
<i>Ruminococcoides bili</i>	Species	523	0.401	Good
<i>Lelliottia aquatilis</i>	Species	493	0.378	Fair
<i>Faecalibacterium langellae</i>	Species	474	0.363	Fair

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
Klebsiella pasteurii	Species	473	0.363	Fair
Dorea phocaeensis	Species	470	0.36	Good
Ruminococcus bromii	Species	415	0.318	Fair
Fusicatenibacter saccharivorans	Species	410	0.314	Good
Enterobacter chuandaensis	Species	401	0.307	Fair
Mediterraneibacter faecis	Species	391	0.3	Good
Citrobacter youngae	Species	369	0.283	Fair
Kluyvera cryocrescens	Species	358	0.274	Fair
Enterobacter cloacae	Species	335	0.257	Fair
Scandinavium goeteborgense	Species	312	0.239	Fair
Eshraghiella crossota	Species	271	0.208	Excellent
Gemmiger formicilis	Species	265	0.203	Good
Escherichia fergusonii	Species	253	0.194	Fair
Alloprococcus comes	Species	237	0.182	Good
Enterobacter soli	Species	222	0.17	Fair
Cedecea lapagei	Species	215	0.165	Fair
Blautia glucerasea	Species	215	0.165	Fair
Coprococcus eutactus	Species	214	0.164	Good
Blautia intestinalis	Species	210	0.161	Fair
Blautia wexlerae	Species	197	0.151	Fair
Collinsella aerofaciens	Species	193	0.148	Excellent
Enterobacter quasiroggenkampii	Species	158	0.121	Fair
Blautia phocaeensis	Species	156	0.12	Fair
Blautia provencensis	Species	153	0.117	Fair
Faecalibacterium prausnitzii	Species	150	0.115	Fair
Corynebacterium dentalis	Species	150	0.115	Fair
Ruminococcus champanellensis	Species	148	0.113	Excellent
Lachnoclostridium edouardi	Species	143	0.11	Fair
Oliverpabstia intestinalis	Species	139	0.107	Fair
Blautia obeum	Species	137	0.105	Fair
Dorea longicatena	Species	135	0.104	Fair

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
Phocaeicola vulgatus	Species	134	0.103	Good
Eubacterium coprostanoligenes	Species	133	0.102	Excellent
Blautia stercoris	Species	131	0.1	Fair
Anaerobutyricum hallii	Species	122	0.094	Fair
Enterobacter chengduensis	Species	121	0.093	Fair
Roseburia faecis	Species	118	0.09	Fair
Clostridium tertium	Species	114	0.087	Fair
Marseillibacter massiliensis	Species	112	0.086	Good
Lactobacillus delbrueckii	Species	109	0.084	Excellent
Clostridium jeddahitimonense	Species	107	0.082	Fair
Blautia producta	Species	104	0.08	Fair
Hominisplanchenecus faecis	Species	103	0.079	Fair
Faecalibacillus intestinalis	Species	100	0.077	Excellent
Waltera intestinalis	Species	100	0.077	Good
Clostridium disporicum	Species	97	0.074	Fair
Segatella hominis	Species	89	0.068	Fair
Vescimonas coprocola	Species	89	0.068	Fair
Lachnospira eligens	Species	86	0.066	Fair
Vescimonas fastidiosa	Species	86	0.066	Fair
Adlercreutzia hattorii	Species	83	0.064	Fair
Roseburia inulinivorans	Species	83	0.064	Fair
Intestinibacter bartlettii	Species	82	0.063	Fair
Streptococcus thermophilus	Species	79	0.061	Fair
Roseburia hominis	Species	78	0.06	Fair
Escherichia marmotae	Species	77	0.059	Fair
Roseburia intestinalis	Species	76	0.058	Fair
Butyricoccus faecihominis	Species	73	0.056	Good
Bilophila wadsworthia	Species	72	0.055	Good
Bacteroides stercoris	Species	72	0.055	Excellent
Pseudocoproccoccus catus	Species	70	0.054	Good
Butyricoccus intestinalisimiae	Species	69	0.053	Excellent

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
Parabacteroides merdae	Species	68	0.052	Fair
Mediterraneibacter gnavus	Species	66	0.051	Fair
[Lactobacillus] rogosae	Species	66	0.051	Fair
Shigella flexneri	Species	65	0.05	Fair
Lacrimispora saccharolytica	Species	64	0.049	Fair
Yokenella regensburgei	Species	61	0.047	Fair
Phocaeicola plebeius	Species	61	0.047	Excellent
Enterocloster bolteae	Species	57	0.044	Fair
Parabacteroides distasonis	Species	57	0.044	Good
Enterobacter mori	Species	57	0.044	Fair
Streptococcus salivarius	Species	56	0.043	Fair
Citrobacter werkmanii	Species	56	0.043	Fair
Lacrimispora amygdalina	Species	56	0.043	Fair
Aristaeella hokkaidonensis	Species	52	0.04	Good
Bacteroides uniformis	Species	51	0.039	Excellent
Ruminococcus callidus	Species	51	0.039	Fair
Blautia faecicola	Species	51	0.039	Fair
Acetivibrio alkalicellulosi	Species	50	0.038	Excellent
Sutterella seckii	Species	49	0.038	Excellent
Pseudocitrobacter vendiensis	Species	49	0.038	Fair
Clostridium fessum	Species	47	0.036	Fair
Adlercreutzia equolifaciens	Species	47	0.036	Fair
Flavonifractor plautii	Species	45	0.035	Fair
Faecalimonas umbilicata	Species	45	0.035	Fair
Eubacterium ramulus	Species	44	0.034	Good
Phytobacter massiliensis	Species	44	0.034	Fair
Faecalibacterium longum	Species	44	0.034	Fair
Lacrimispora indolis	Species	43	0.033	Fair
Eggerthella guodeyinii	Species	42	0.032	Good
Pseudenterobacter timonensis	Species	41	0.031	Fair
Blautia pseudococcoides	Species	41	0.031	Fair

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
Gemmiger gallinarum	Species	40	0.031	Fair
Duncaniella dubosii	Species	38	0.029	Good
Senegalimassilia faecalis	Species	37	0.028	Excellent
Phocaeicola coprocola	Species	36	0.028	Excellent
Dialister massiliensis	Species	36	0.028	Fair
Konateibacter massiliensis	Species	35	0.027	Fair
Faecalibacillus faecis	Species	34	0.026	Fair
Phocaeicola massiliensis	Species	33	0.025	Excellent
Romboutsia hominis	Species	32	0.025	Fair
Enterococcus faecium	Species	32	0.025	Excellent
Hominimerdicola intestinalis	Species	31	0.024	Good
Howardella ureilytica	Species	31	0.024	Excellent
Dryocola mayonis	Species	30	0.023	Fair
Leyella stercorea	Species	30	0.023	Good
Agathobaculum butyriciproducens	Species	29	0.022	Fair
Terrisporobacter mayombeii	Species	29	0.022	Fair
Duodenibacillus massiliensis	Species	29	0.022	Excellent
Salmonella enterica	Species	29	0.022	Fair
Faecalicatena contorta	Species	28	0.021	Fair
Phocaeicola dorei	Species	27	0.021	Fair
Citrobacter europaeus	Species	27	0.021	Fair
Citrobacter tructae	Species	27	0.021	Fair
Phascolarctobacterium faecium	Species	27	0.021	Excellent
Solibaculum mannosilyticum	Species	26	0.02	Fair
Catenibacterium mitsuokai	Species	26	0.02	Good
Leclercia pneumoniae	Species	25	0.019	Fair
Senegalimassilia anaerobia	Species	25	0.019	Fair
Intestinimonas gabonensis	Species	25	0.019	Good
[Eubacterium] siraeum	Species	25	0.019	Excellent
[Ruminococcus] lactaris	Species	24	0.018	Fair
Ruthenibacterium lactatiformans	Species	24	0.018	Excellent

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
Kluyvera sichuanensis	Species	23	0.018	Fair
Aristaeella lactis	Species	23	0.018	Fair
Ligilactobacillus ruminis	Species	23	0.018	Good
Anaerotignum faecicola	Species	22	0.017	Excellent
Faecalibacterium butyricigenerans	Species	22	0.017	Fair
Faecalicatena fissicatena	Species	20	0.015	Fair
Enterobacter ludwigii	Species	20	0.015	Fair
Intestinimonas timonensis	Species	20	0.015	Fair
Pluralibacter gergoviae	Species	20	0.015	Fair
Enterobacter bugandensis	Species	19	0.015	Fair
Clostridium phoceensis	Species	19	0.015	Fair
Dryocola sharpae	Species	19	0.015	Fair
Barnesiella intestinihominis	Species	18	0.014	Fair
Guopingia tenuis	Species	18	0.014	Excellent
Caenibaculum baiyandianus	Species	18	0.014	Fair
Ruminococcus sporogenes	Species	18	0.014	Fair
Bengtsoniella intestinalis	Species	18	0.014	Fair
Wujia chipingensis	Species	18	0.014	Fair
Vampirovibrio chlorellavorus	Species	17	0.013	Excellent
Lentihominibacter faecis	Species	17	0.013	Excellent
Bacteroides finegoldii	Species	17	0.013	Good
Eubacterium oxidoreducens	Species	16	0.012	Fair
Eggerthella timonensis	Species	16	0.012	Fair
Perlabentimonas gracilis	Species	16	0.012	Excellent
Lacrimispora sphenoides	Species	16	0.012	Fair
Robinsoniella peoriensis	Species	15	0.012	Fair
Cronobacter muytjensii	Species	15	0.012	Fair
Anaeromassilibacillus senegalensis	Species	15	0.012	Good
Blautia hansenii	Species	15	0.012	Fair
Hominimerdicola alba	Species	14	0.011	Fair
Papillibacter cinnamivorans	Species	14	0.011	Fair

SINGLE SAMPLE METAGENOMIC ANALYSIS REPORT

Taxon	Taxonomy	Read Count	Abundance (%)	Confidence
<i>Shigella dysenteriae</i>	Species	14	0.011	Fair
<i>Massiliprevotella massiliensis</i>	Species	14	0.011	Good
<i>Buttiauxella massiliensis</i>	Species	14	0.011	Fair
<i>Duncaniella muris</i>	Species	14	0.011	Fair
<i>Bacteroides ovatus</i>	Species	13	0.01	Good
<i>Cuneatibacter caecimuris</i>	Species	13	0.01	Fair
<i>Kosakonia quasisacchari</i>	Species	13	0.01	Fair
<i>Klebsiella oxytoca</i>	Species	13	0.01	Fair
<i>Mediterraneibacter torques</i>	Species	13	0.01	Fair
<i>Anaerosacchariphilus polymeriproducens</i>	Species	12	0.009	Fair
<i>Bacteroides caccae</i>	Species	12	0.009	Good
<i>Jutongia huaianensis</i>	Species	12	0.009	Good
<i>Faecalicatena orotica</i>	Species	12	0.009	Fair
<i>Clostridium tarantellae</i>	Species	10	0.008	Fair
<i>Blautia argi</i>	Species	10	0.008	Fair
<i>Sporobacter termitidis</i>	Species	10	0.008	Fair
<i>Succinivibrio dextrinosolvens</i>	Species	10	0.008	Excellent
<i>Haemophilus parainfluenzae</i>	Species	10	0.008	Good

REFERENCES & TOOLS

Version: 1.1.0

The following bioinformatics tools were used in this analysis workflow. Versions are extracted from the conda environment definitions.

Tool	Version	Description
Fastplong	0.4.1	Quality control filtering for long reads
minimap2	2.27	Long-read sequence alignment
QIIME2	2026.1.0	Microbiome analysis framework
PICRUSt2	2.5.3	Functional prediction from 16S marker genes
scikit-bio	0.7.0	Biological data science calculations
Krona	2.8	Interactive taxonomy visualization
seqtk	1.5	FASTQ/FASTA processing toolkit
samtools	-	SAM/BAM/CRAM file manipulation
matplotlib	3.10.6	Scientific data visualization
VSEARCH	2.22	Sequence clustering and chimera detection
pysam	0.15	SAM/BAM file parsing for consensus generation
racon	1.5.0	Consensus sequence polishing via iterative correction



MASSIVE
BIOINFORMATICS

Contact

massivebioinformatics.com

info@massivebioinformatics.com

+90 216 474 60 01



THANK YOU FOR CHOOSING US!

Website

