யாழ் ஜீனோமிக்ஸ் YAAZH XENOMICS

METAGENOMICS ANALYSIS REPORT

16S rRNA Amplicon Sequencing Report

Demo: 16S rRNA microbiome profiling (4 samples)

Report ID	YX-16S-DEMO-001
Project	Demo: 16S rRNA microbiome profiling (4 samples)
Client	Demo report - public Illumina test data (nf-core)
Order reference	—
Analysis	16S rRNA gene amplicon sequencing (V4)
Samples	4
Report date	27 November 2026

Prepared by Yaazh Xenomics · ISO 9001:2015 certified genomics laboratory
TICEL Biopark, Maruthamalai Road, Coimbatore 641046, India · +91 99431 32020 ·
info@yaazhxenomics.com · www.yaazhxenomics.com

Confidential: prepared for the client named above. Results relate only to the samples received and analysed.

Summary

- **4** samples sequenced; **4,931** high-quality reads retained (49.3% of 10,000 raw read pairs).
- **370** amplicon sequence variants (ASVs) identified; 34.9% of reads classified to genus and 2.0% to species (exact match).
- Dominant phyla (mean relative abundance): Pseudomonadota 37.8%, Nanoarchaeota 17.4%, Patescibacteria 9.8%.
- Most abundant genera: Acinetobacter 7.4%, Candidatus Omnitrophus 6.2%, Pseudomonas 5.2%, Nitrosarchaeum 3.7%, Aquabacterium 1.8%.
- Community composition did not differ significantly between Treatment groups (PERMANOVA on bray: R^2 = 0.81, p = 0.333).
- Caution: group(s) Group A, Group B have fewer than 3 samples. With so few replicates, PERMANOVA cannot reach significance and differential abundance results are unreliable; treat the statistics as exploratory.

Project	Demo: 16S rRNA microbiome profiling (4 samples)
Client	Demo report - public Illumina test data (nf-core)
Report date	27 November 2026
Target region	16S rRNA V4 (GTGYCAGCMGCCGCGGTAA / GGACTACNVGGGTWTCTAAT)
Samples	4
Raw read pairs	10,000
Final reads (after QC, merging, chimera removal)	4,931 (49.3%)
ASVs	370
Reads classified to genus	34.9%
Rarefaction depth for diversity	1,041 reads
Group comparison	Treatment (Group A: n=2, Group B: n=2)
Reference database	SILVA 138.2 (DADA2 formatted, Zenodo 10.5281/zenodo.14169026)
Pipeline version	Yaazh Xenomics 16S pipeline v1.0.0

1. Sequencing quality

Raw read quality per position, pooled across samples. Dashed lines mark Q30 (green) and Q20 (amber). Reads were truncated at 230 bp (R1) and 229 bp (R2) before denoising.

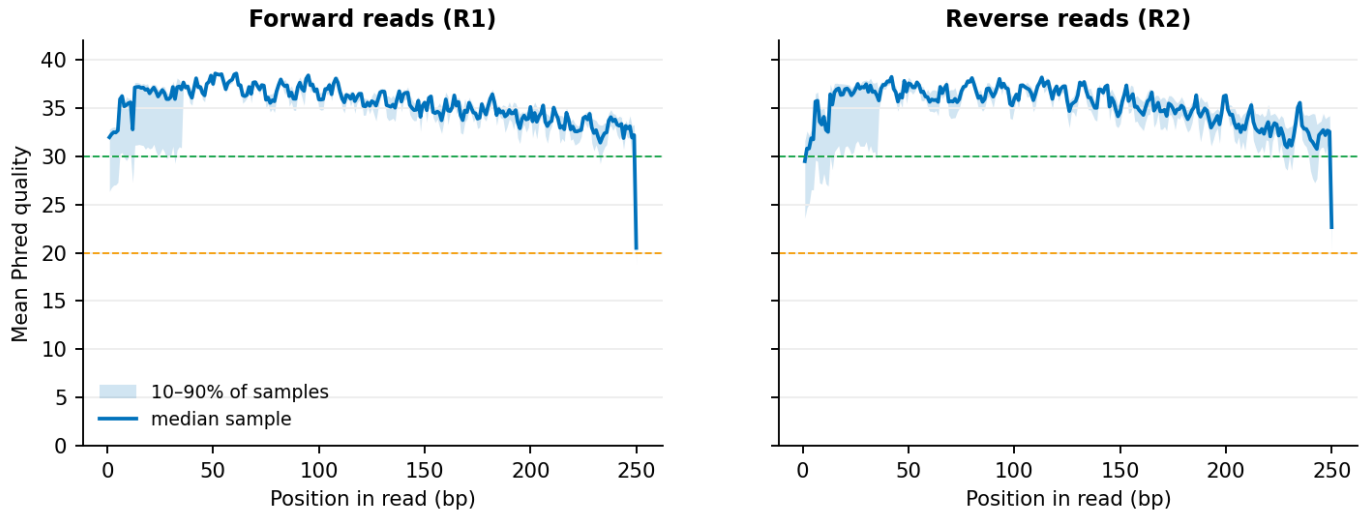


Figure 1. Read quality profiles

Sample	Read pairs	R1 mean length	R2 mean length	R1 ≥Q30 (%)	R2 ≥Q30 (%)	GC (%)	Primers found (%)
1	2,500	194.90	195.80	83.62	81.13	50.79	73.36
1a	2,500	243.00	243.10	92.94	93.07	52.71	95.88
2	2,500	242.90	242.90	88.88	86.49	52.86	95.00
2a	2,500	243.40	243.50	92.66	92.83	52.71	95.40

2. Read processing

Read pairs remaining after each step. Typical V3-V4 runs keep 40–80% of raw pairs; most losses come from quality filtering and chimera removal.

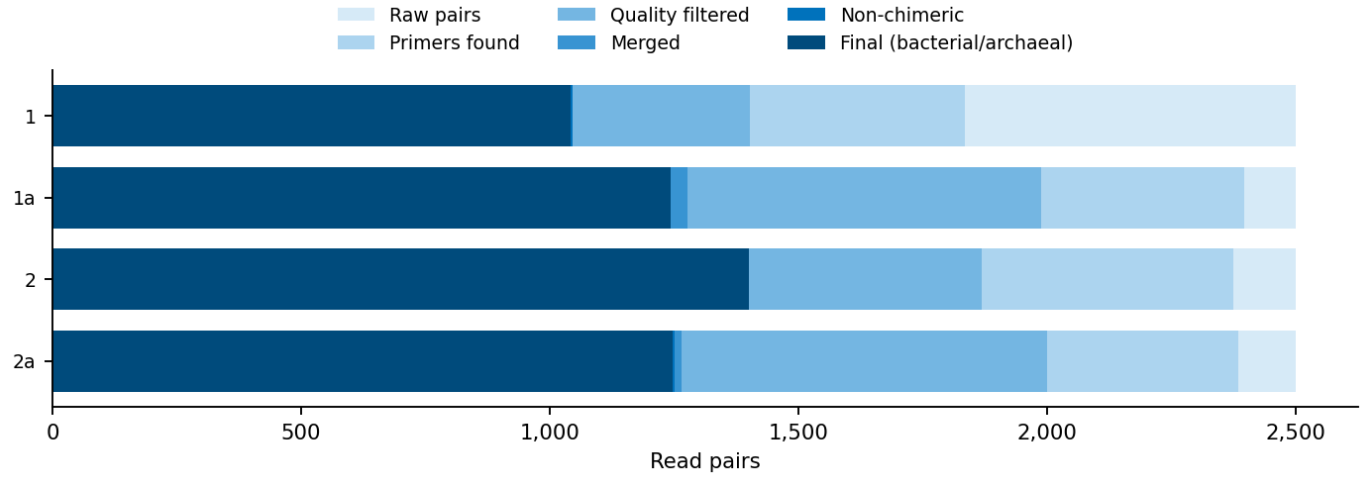


Figure 2. Reads retained at each processing step

Sample	Raw pairs	With primers	Filtered	Denoised F	Denoised R	Merged	Length OK	Non-chimeric	Final	Retained (%)
1	2,500	1,834	1,403	1,170	1,149	1,045	1,045	1,045	1,041	41.60
1a	2,500	2,397	1,989	1,381	1,378	1,277	1,243	1,243	1,243	49.70
2	2,500	2,375	1,868	1,509	1,489	1,401	1,401	1,401	1,401	56.00

Sample	Raw pairs	With primers	Filtered	Denoised F	Denoised R	Merged	Length OK	Non-chimeric	Final	Retained (%)
2a	2,500	2,385	2,001	1,414	1,430	1,265	1,250	1,250	1,246	49.80

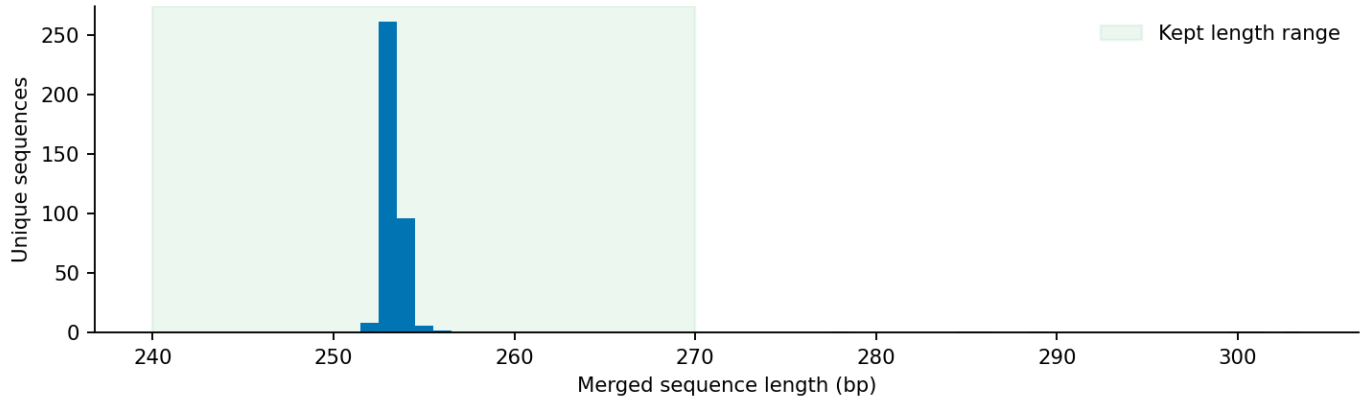


Figure 3. Length of merged sequences

Removed as non-target: Chloroplast (4 reads, 1 ASVs), Mitochondria (2 reads, 1 ASVs), Unassigned kingdom (2 reads, 1 ASVs).

3. Sequencing depth

Rarefaction curves show how many ASVs are found as more reads are sampled. Curves that level off indicate the sample was sequenced deeply enough. For diversity analyses all samples were randomly subsampled to 1,041 reads.

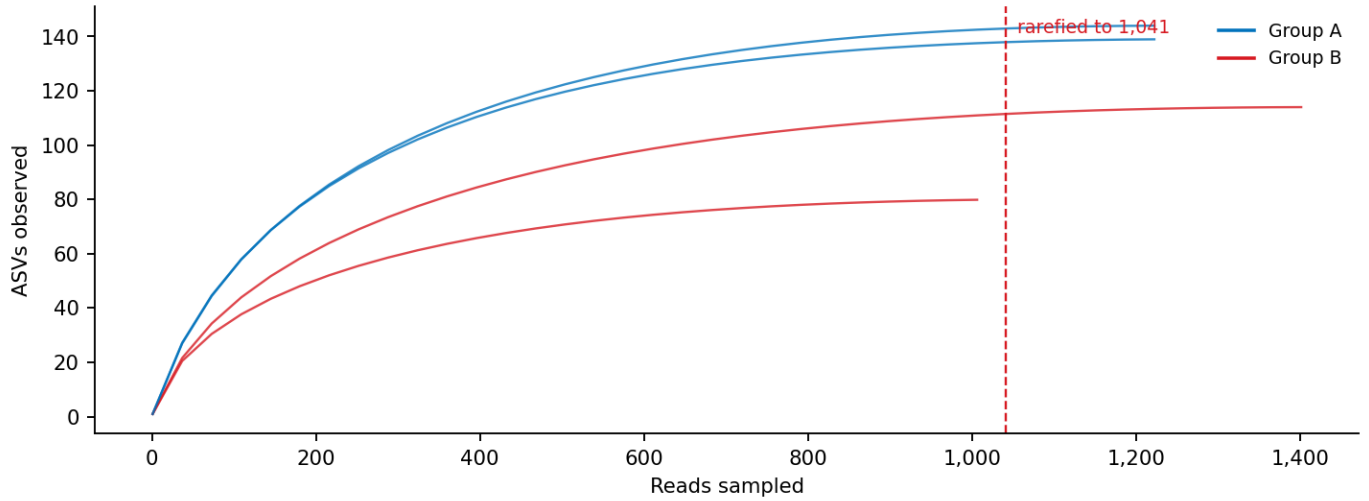


Figure 4. Rarefaction curves

4. Taxonomic composition

Classification against SILVA 138.2: 91.9% of reads assigned to phylum, 52.3% to family, 34.9% to genus and 2.0% to species. Species names come from exact matches only and should be treated as indicative for short amplicons.

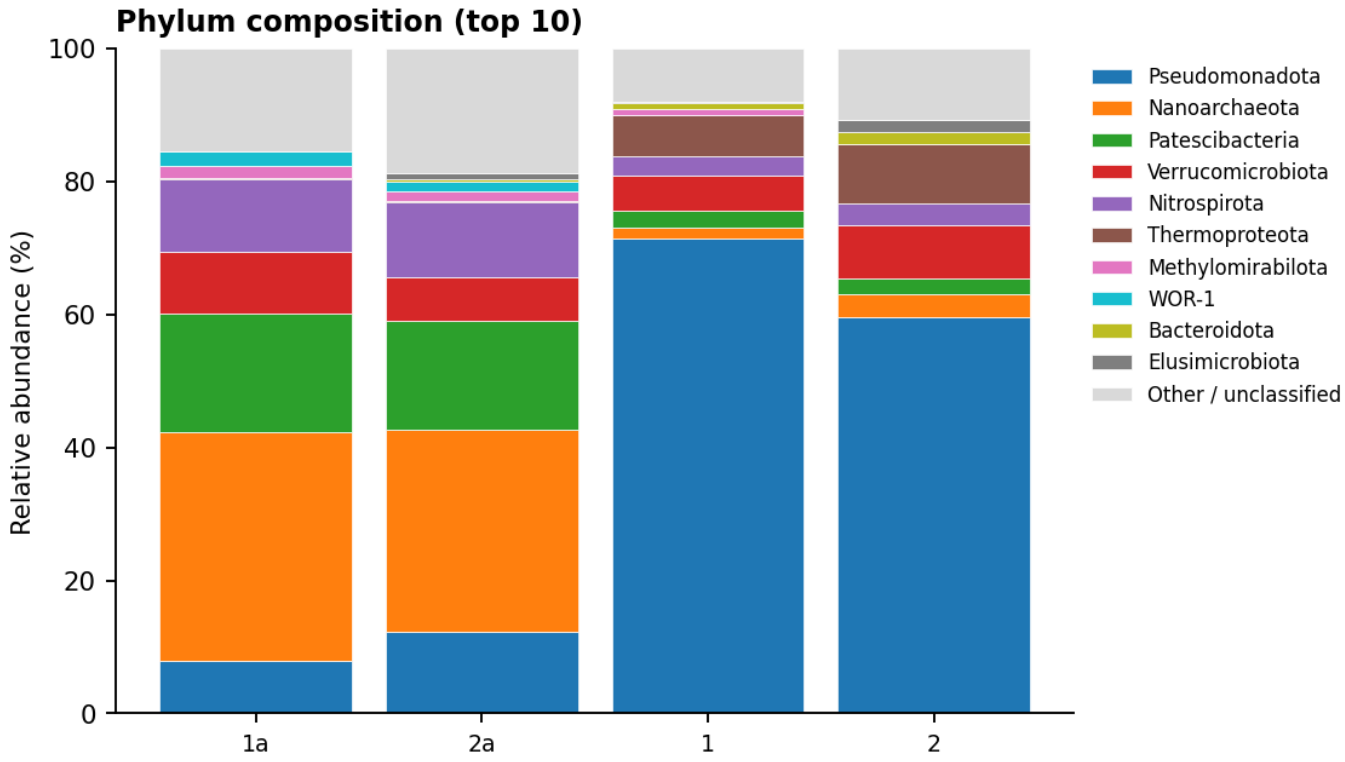


Figure 5. Phylum-level composition

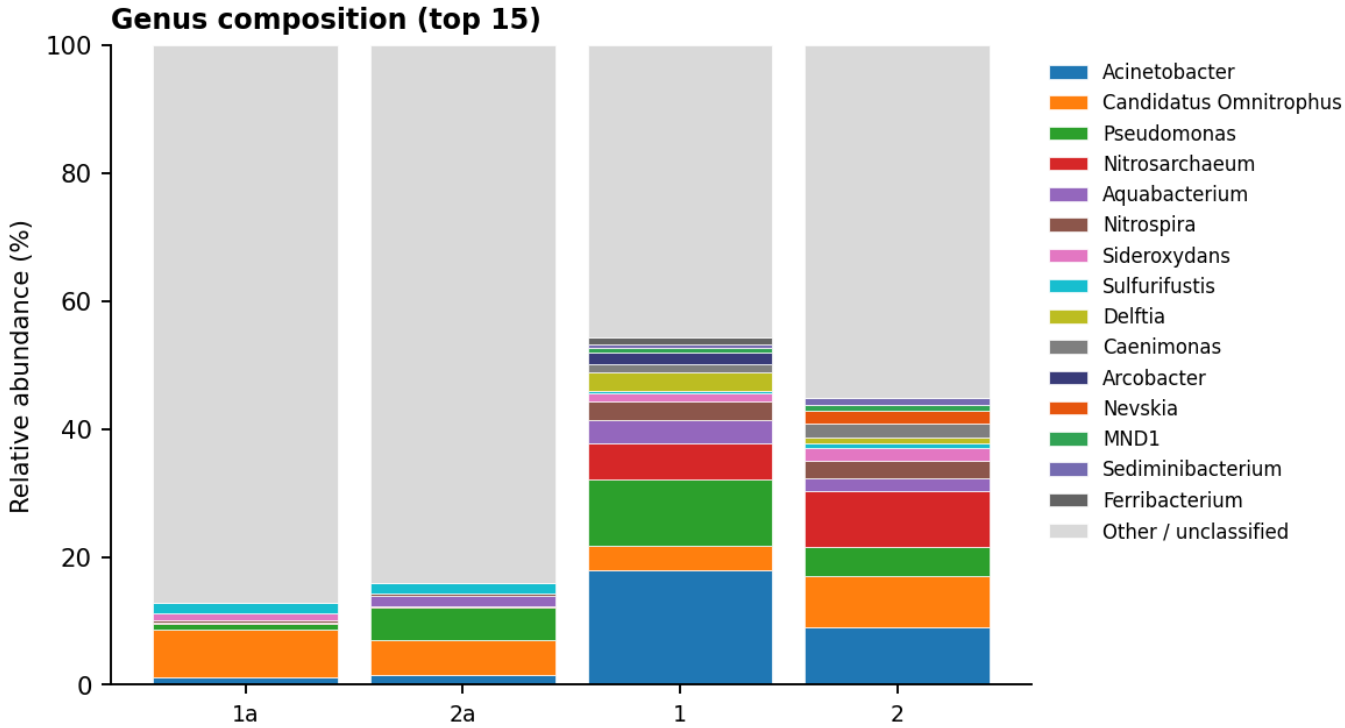


Figure 6. Genus-level composition

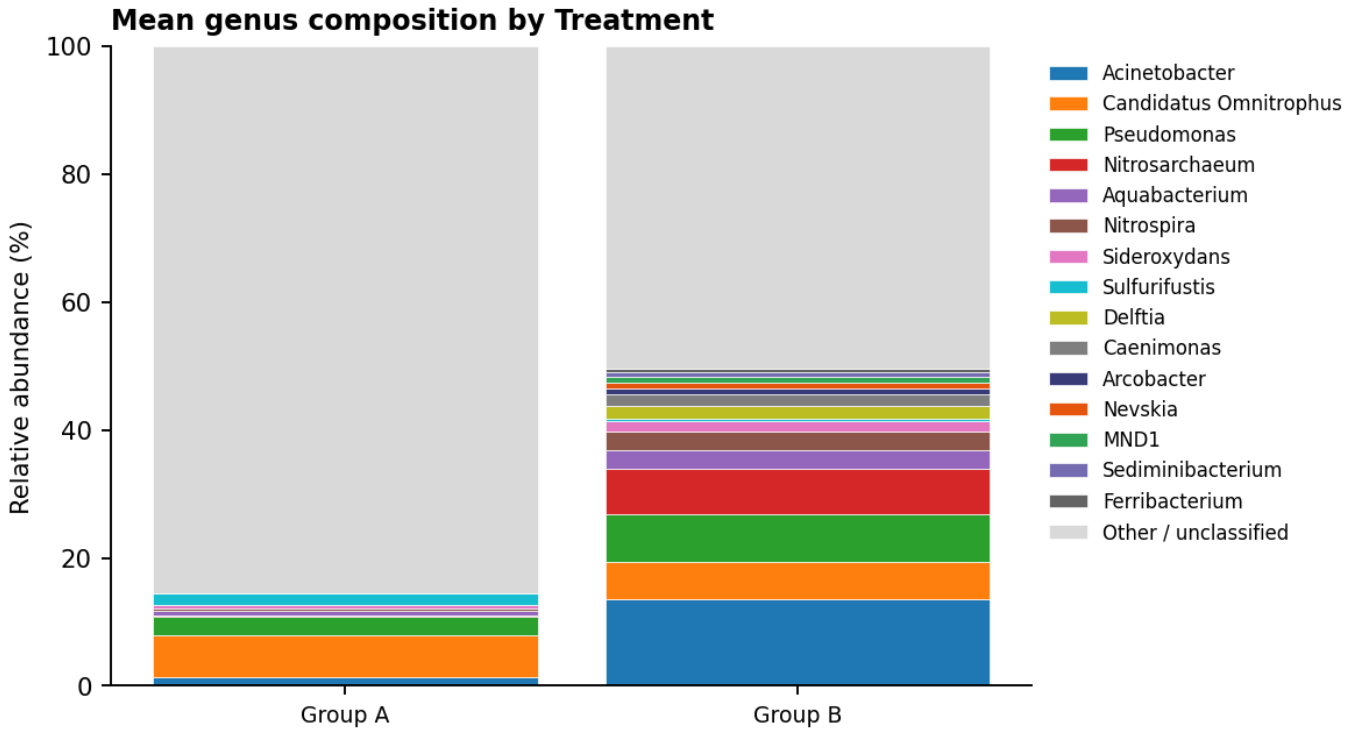


Figure 6b. Mean genus composition by group

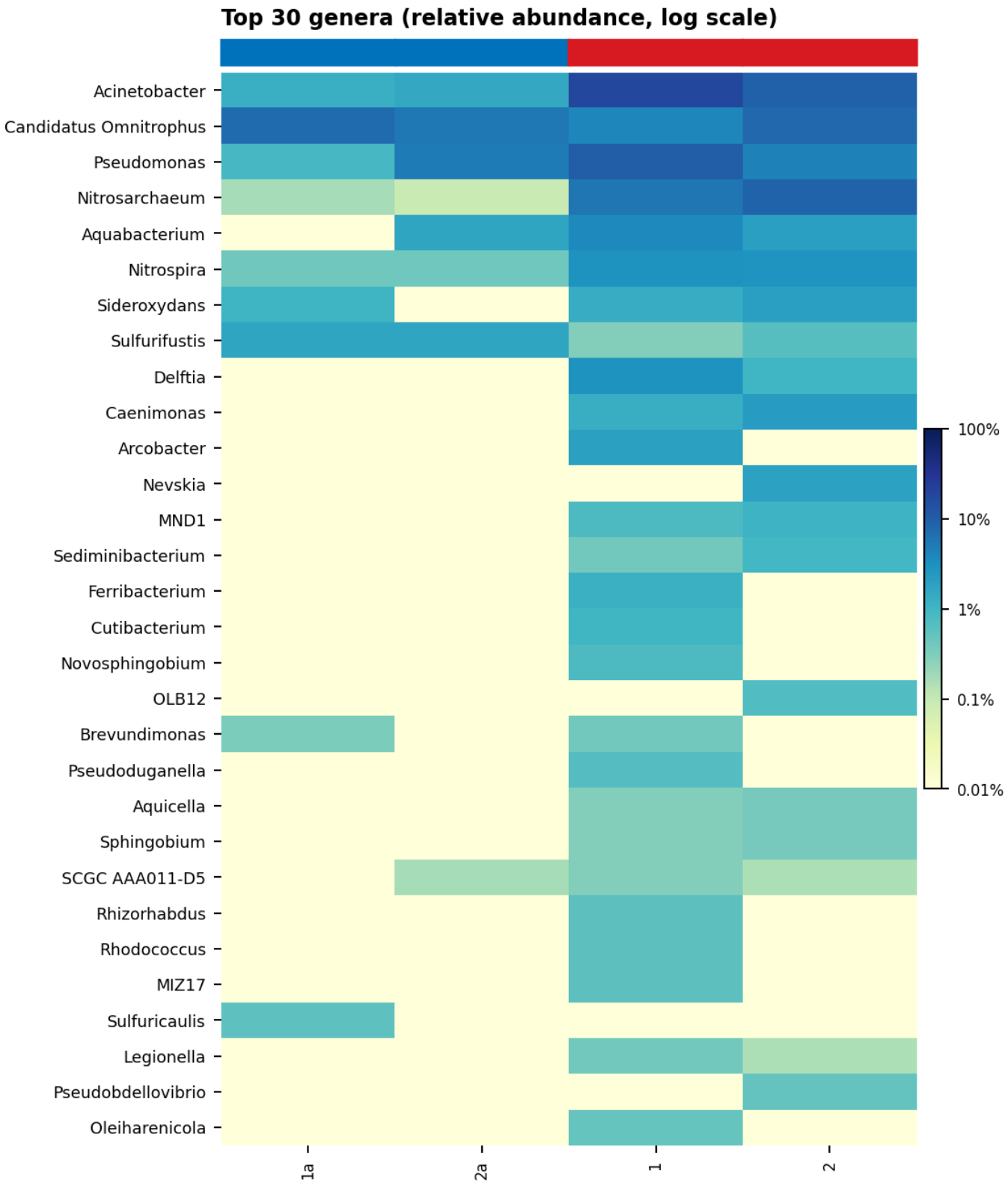


Figure 7. Heatmap of the 30 most abundant genera

Top 25 genera (relative abundance, %)

Genus	Mean %	1a	2a	1	2
Acinetobacter	7.42	1.21	1.52	17.87	9.06
Candidatus Omnitrophus	6.22	7.40	5.54	3.94	7.99
Pseudomonas	5.16	0.88	5.06	10.28	4.43

Genus	Mean %	1a	2a	1	2
Nitrosarchaeum	3.67	0.16	0.08	5.67	8.78
Aquabacterium	1.83	0.00	1.69	3.65	2.00
Nitrospira	1.62	0.40	0.40	2.88	2.78
Sideroxydans	1.10	1.05	0.00	1.34	2.00
Sulfurifustis	1.08	1.69	1.69	0.29	0.64
Delftia	0.97	0.00	0.00	2.88	1.00
Caenimonas	0.87	0.00	0.00	1.25	2.21
Arcobacter	0.48	0.00	0.00	1.92	0.00
Nevskia	0.46	0.00	0.00	0.00	1.86
MND1	0.46	0.00	0.00	0.77	1.07
Sediminibacterium	0.33	0.00	0.00	0.38	0.93
Ferribacterium	0.29	0.00	0.00	1.15	0.00
Cutibacterium	0.24	0.00	0.00	0.96	0.00
Novosphingobium	0.19	0.00	0.00	0.77	0.00
OLB12	0.18	0.00	0.00	0.00	0.71
Brevundimonas	0.18	0.32	0.00	0.38	0.00
Pseudoduganella	0.17	0.00	0.00	0.67	0.00
Aquicella	0.16	0.00	0.00	0.29	0.36
Sphingobium	0.16	0.00	0.00	0.29	0.36
SCGC AAA011-D5	0.15	0.00	0.16	0.29	0.14
Rhizorhabdus	0.14	0.00	0.00	0.58	0.00
Rhodococcus	0.14	0.00	0.00	0.58	0.00

5. Alpha diversity

Alpha diversity describes diversity within each sample: richness (observed ASVs, Chao1), richness and evenness combined (Shannon, Gini–Simpson) and phylogenetic breadth (Faith's PD).

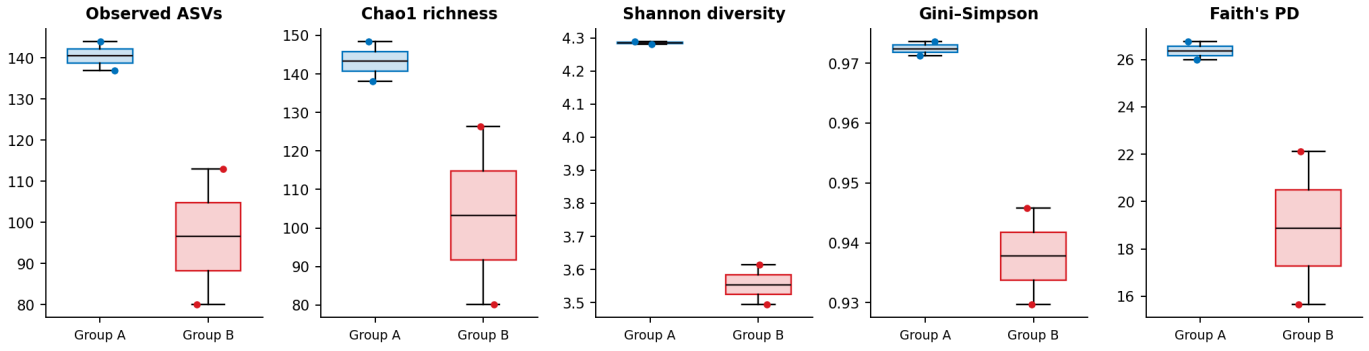


Figure 8. Alpha diversity

Group tests

metric	test	statistic	p_value	p_adj_BH
Observed	Wilcoxon rank-sum	4	0.333	0.333
Chao1	Wilcoxon rank-sum	4	0.333	0.333
Shannon	Wilcoxon rank-sum	4	0.333	0.333

metric	test	statistic	p_value	p_adj_BH
Simpson	Wilcoxon rank-sum	4	0.333	0.333
FaithPD	Wilcoxon rank-sum	4	0.333	0.333

Sample	Treatment	Reads (before rarefying)	Observed ASVs	Chao1	Shannon	Gini-Simpson	InvSimpson	Pielou	Faith PD
1	Group B	1,041	80	80.16	3.49	0.95	18.46	0.80	15.65
1a	Group A	1,243	144	148.39	4.28	0.97	34.71	0.86	26.76
2	Group B	1,401	113	126.32	3.61	0.93	14.22	0.76	22.11
2a	Group A	1,246	137	138.12	4.29	0.97	37.94	0.87	25.98

6. Beta diversity

Beta diversity compares community composition between samples. Each point is a sample; points closer together have more similar communities. Bray–Curtis and weighted UniFrac weigh abundant taxa; Jaccard and unweighted UniFrac consider presence or absence only.

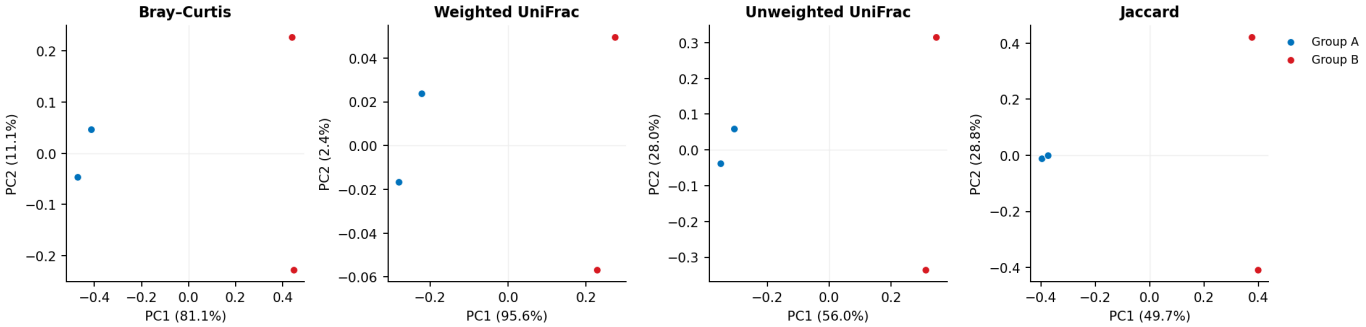


Figure 9. Principal coordinates analysis (PCoA)

PERMANOVA

Tests whether community composition differs between groups. A significant dispersion p-value means within-group spread differs, which can also make PERMANOVA significant.

distance	F	R2	p_value	dispersion p
bray	8.49	0.81	0.333	0.042
jaccard	1.98	0.50	0.333	0.042
unweighted_unifrac	2.54	0.56	0.333	0.042
weighted_unifrac	34.92	0.95	0.333	0.042

7. Differential abundance

28 genus–comparison pairs were significantly different (DESeq2, adjusted p < 0.05). With fewer than 3 samples in a group these results are exploratory only.

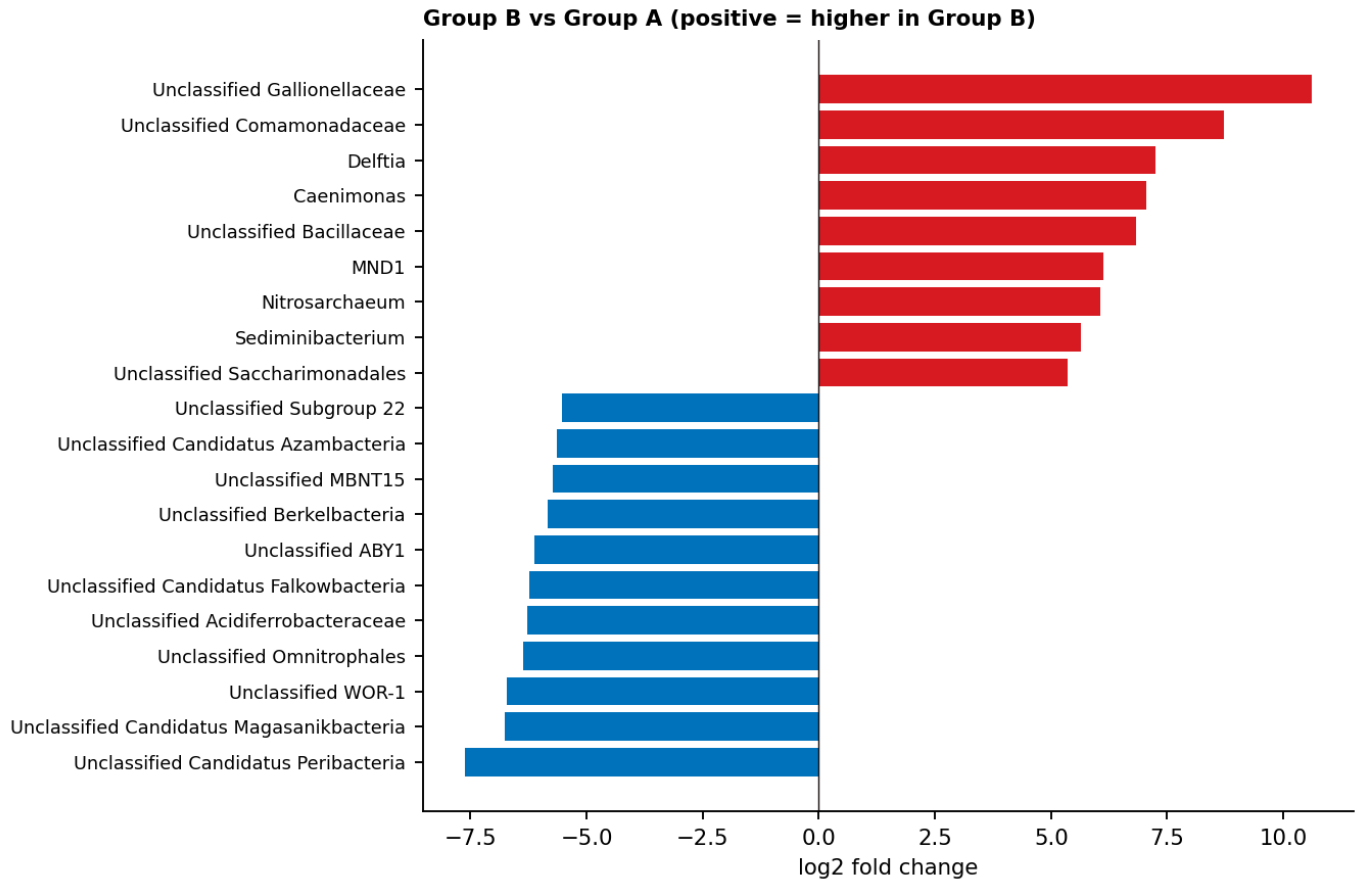


Figure 10. Differentially abundant genera

contrast	Genus	baseMean	log2FoldChange	p_adj_BH
Group B vs Group A	Unclassified Woesearchaeales	201.4	-3.74	<0.001
Group B vs Group A	Unclassified Gallionellaceae	132.7	10.61	<0.001
Group B vs Group A	Acinetobacter	97.6	3.52	<0.001
Group B vs Group A	Nitrosarchaeum	47.4	6.06	<0.001
Group B vs Group A	Unclassified Bacteria	93.5	-2.31	<0.001
Group B vs Group A	Unclassified Thermodesulfovibrionia	60.0	-4.77	<0.001
Group B vs Group A	Unclassified Comamonadaceae	35.6	8.71	<0.001
Group B vs Group A	Unclassified Candidatus Peribacteria	19.6	-7.62	<0.001
Group B vs Group A	Caenimonas	11.1	7.04	<0.001
Group B vs Group A	Delftia	13.1	7.25	<0.001
Group B vs Group A	Unclassified Bacillaceae	9.6	6.82	<0.001
Group B vs Group A	Unclassified Candidatus Magasanikbacteria	10.9	-6.77	<0.001
Group B vs Group A	Unclassified WOR-1	10.5	-6.72	<0.001
Group B vs Group A	Unclassified Acidiferrobacteraceae	7.7	-6.27	0.001
Group B vs Group A	Unclassified Omnitrophales	8.2	-6.36	0.001
Group B vs Group A	Unclassified Candidatus Falkowbacteria	7.4	-6.23	0.001
Group B vs Group A	Unclassified ABY1	7.0	-6.13	0.002
Group B vs Group A	MND1	6.0	6.13	0.002
Group B vs Group A	Unclassified Candidatus Azambacteria	5.0	-5.64	0.004

contrast	Genus	baseMean	log2FoldChange	p_adj_BH
Group B vs Group A	Sediminibacterium	4.2	5.64	0.004
Group B vs Group A	Unclassified Berkelbacteria	5.7	-5.85	0.004
Group B vs Group A	Unclassified MBNT15	5.3	-5.72	0.004
Group B vs Group A	Nitrospira	20.9	3.00	0.005
Group B vs Group A	Unclassified Subgroup 22	4.5	-5.52	0.005
Group B vs Group A	Unclassified Candidatus Peregrinibacteria	3.8	-5.26	0.006
Group B vs Group A	Unclassified Saccharimonadales	3.4	5.35	0.007
Group B vs Group A	Unclassified Candidatus Kerfeldbacteria	3.2	-5.01	0.010
Group B vs Group A	Unclassified Rokubacteriales	9.6	-3.28	0.018

Methods

Raw paired-end reads were assessed with fastp v0.23.4. Primer sequences (GTGYCAGCMGCCGCGGTAA forward, GGACTACNVGGGTWTCTAAT reverse) were removed with cutadapt v5.2 (maximum error rate 0.15), and read pairs in which either primer was not found were discarded.

Amplicon sequence variants (ASVs) were inferred with DADA2 v1.30.0 in R v4.3.3. Reads were truncated to 230 bp (forward) and 229 bp (reverse) (chosen automatically from the quality profiles so that pairs still overlap), and reads with ambiguous bases or more than 2/2 expected errors were removed. Error models were learned from the data, samples were denoised with pseudo-pooling, read pairs were merged (minimum overlap 12 bp, no mismatches), sequences outside 240–270 bp were removed and chimeras were removed with the consensus method.

Taxonomy was assigned with the DADA2 naive Bayesian classifier against SILVA 138.2 (DADA2 formatted, Zenodo 10.5281/zenodo.14169026) (minimum bootstrap confidence 80), and species were added by exact sequence matching. ASVs classified as chloroplast, mitochondria, eukaryote or unassigned at kingdom level were removed.

ASVs were aligned with MAFFT v7.505 and a phylogeny was built with FastTree v2.1.11 (GTR+CAT, Gamma20) and midpoint-rooted. Diversity analyses used phyloseq v1.46.0 and vegan v2.6.4 on counts rarefied to 1,041 reads per sample. Alpha diversity was measured as observed ASVs, Chao1, Shannon, Gini–Simpson and Faith's phylogenetic diversity; beta diversity as Bray–Curtis, Jaccard and weighted and unweighted UniFrac distances, visualised by principal coordinates analysis.

Samples were compared by 'Treatment'. Alpha diversity was tested with the Wilcoxon rank-sum (Benjamini–Hochberg correction), community composition with PERMANOVA (adonis2, 999 permutations) and homogeneity of dispersion with betadisper. Differentially abundant genera were identified with DESeq2 v1.42.0 (poscounts size factors; Benjamini–Hochberg adjusted $p < 0.05$) on unrarefied genus counts.

Software versions

fastp	0.23.4
cutadapt	5.2
MAFFT	7.505

FastTree	2.1.11
dada2	1.30.0
phyloseq	1.46.0
vegan	2.6.4
DESeq2	1.42.0
ape	5.7.1
phangorn	2.11.1
R	4.3.3

Files delivered

Demo_16S_rRNA_microbiome_profiling_4_samples_report.html / .pdf	This report
Demo_16S_rRNA_microbiome_profiling_4_samples_results.xlsx	All tables: QC, read tracking, ASV table, taxonomy at every rank (% and counts), diversity and statistics
data/asv_sequences.fasta	ASV sequences
data/asv_table_with_taxonomy.tsv, asv_counts.tsv	ASV × sample counts with taxonomy
data/feature_table.biom, qiime2_taxonomy.tsv	Import into QIIME 2 (BIOMV100Format and HeaderlessTSVTaxonomyFormat)
data/asv_tree.nwk, asv_alignment.fasta	ASV phylogeny (Newick) and alignment
data/phyloseq.rds	R phyloseq object (counts, taxonomy, metadata and tree)
data/taxonomy_by_rank/	Relative abundance and counts from phylum to species
data/diversity/, data/statistics/	Alpha diversity, distance matrices, PCoA coordinates and test results
data/qc/	Per-sample fastp QC reports, read tracking and DADA2 error models
figures/	Every figure as PNG and SVG

About Yaazh Xenomics

Yaazh Xenomics is an ISO 9001:2015 certified DNA sequencing and genomics laboratory founded in 2012, with its head office and laboratory in Coimbatore and sample collection across India (Coimbatore, Madurai, Hyderabad, Visakhapatnam, Bengaluru and Mumbai). We provide Sanger sequencing, next-generation sequencing (16S/ITS amplicon and shotgun metagenomics, whole genome, exome and transcriptome sequencing) and bioinformatics analysis for research institutions, universities, hospitals and industry.

Head office & laboratory	Module No. 103, TICEL Biopark Phase III, 1st Floor, Maruthamalai Road, Coimbatore 641046, Tamil Nadu, India
Phone / WhatsApp	+91 99431 32020
Bioinformatics support	+91 95002 45454
Email	info@yaazhxenomics.com

Website	www.yaazhxenomics.com

For questions about this report, quote the report ID shown on the cover.

References

1. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.
2. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.
3. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12.
4. Quast C, Pruesse E, Yilmaz P, et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.
5. McLaren MR, Callahan BJ (2024) Silva 138.2 prokaryotic SSU taxonomic training data formatted for DADA2. Zenodo. doi:10.5281/zenodo.14169026.
6. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267.
7. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7. Mol Biol Evol 30:772–780.
8. Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490.
9. McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217.
10. Oksanen J, et al. vegan: Community Ecology Package. R package. https://CRAN.R-project.org/package=vegan
11. Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71:8228–8235.
12. Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46.
13. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
14. Klindworth A, Pruesse E, Schweer T, et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1.

Yaazh Xenomics · ISO 9001:2015 certified genomics laboratory · Coimbatore, India · www.yaazhxenomics.com · +91 99431 32020. Generated by the Yaazh Xenomics 16S pipeline v1.0.0 on 27 Nov 2026 14:29. Results relate only to the samples received.