FULL - LENGTH 16S OXFORD NANOPORE TECHNOLOGIES

Hoang Kim

06-10-2024

1 Introduction

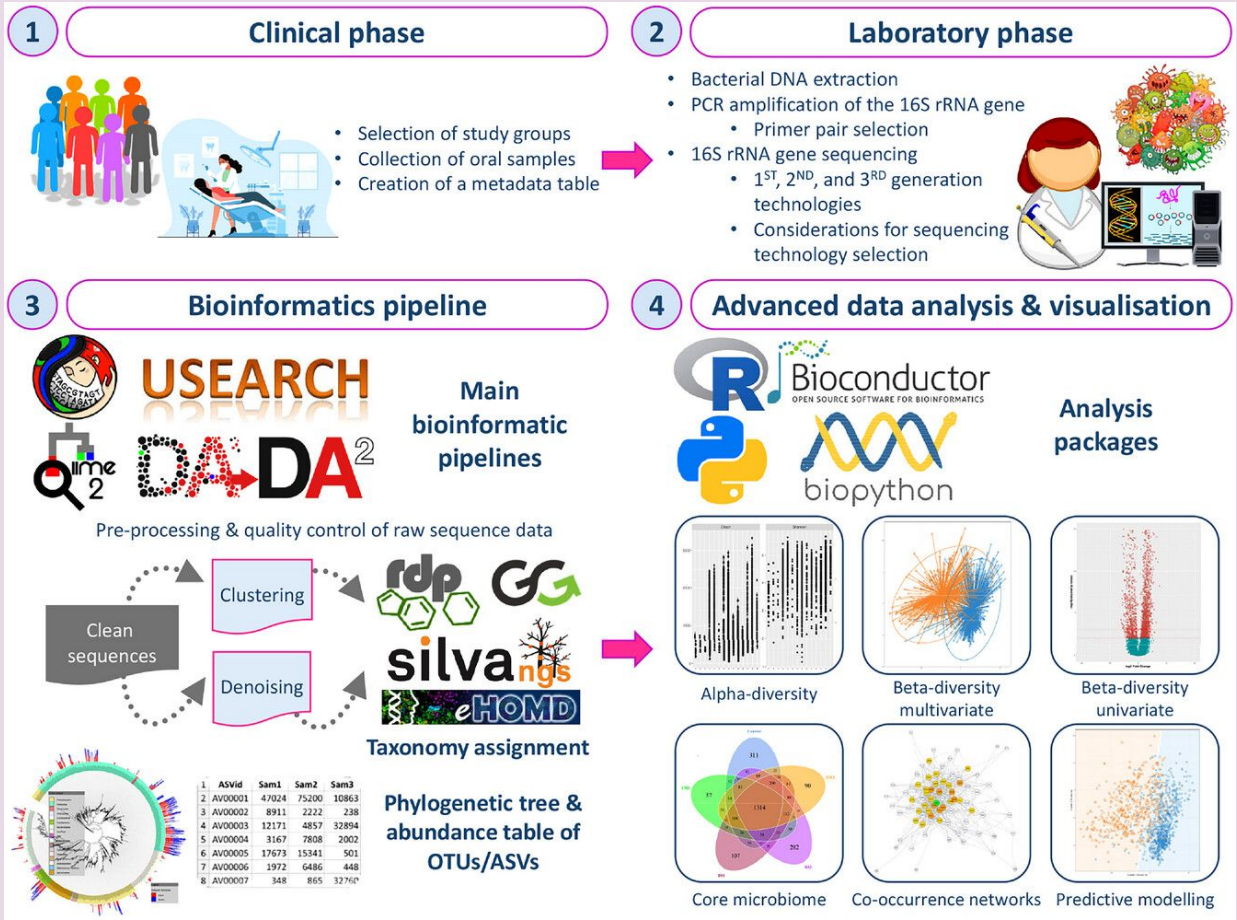
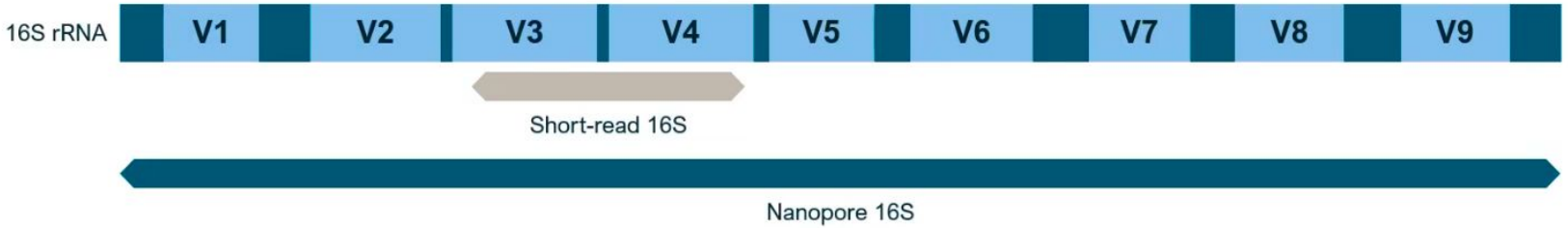


Figure. Summarises the workflow followed in 16S rRNA metabarcoding oral microbiome studies

https://onlinelibrary.wiley.com/doi/full/10.1111/omi.12434

Why implement nanopore sequencing for 16S analysis?

Full-length 16S rRNA sequencing improves taxonomic resolution of microbial communities



Key features	Sanger	Short-read	Oxford Nanopore
Read length	~500 bp	~150 – 300 bp	~1,500 bp
Resolves polymicrobial samples	✗	✓	✓
High taxonomic resolution[1]	✗	✗	✓
Rapid library preparation	✗	✗	✓
Real-time sequencing	✗	✗	✓

1. Zhang, et al. "The newest Oxford Nanopore R10. 4.1 full-length 16S rRNA sequencing enables the accurate resolution of species-level microbial community profiling." Applied and Environ Microbio 89.10 (2023): e00605-23.

Oxford NANOPORE Technologies

2

Advantages

Sample Data

Archives of Microbiology (2024) 206:248
https://doi.org/10.1007/s00203-024-03985-7

ORIGINAL PAPER

A comparison between full-length 16S rRNA Oxford nanopore sequencing and Illumina V3-V4 16S rRNA sequencing in head and neck cancer tissues

Kenny Yeo[1,2] · James Connell[1,2] · George Bouras[1,2] · Eric Smith[1,3] · William Murphy[1,2] · John-Charles Hodge[1,4] · Suren Krishnan[1,4] · Peter-John Wormald[5] · Rowan Valentine[2] · Alkis James Psaltis[1,2] · Sarah Vreugde[1,2] · Kevin Aaron Fenix[1,2]

Received: 13 March 2024 / Accepted: 28 April 2024 / Published online: 7 May 2024

Abstract
Describing the microbial community within the tumour has been a key aspect in understanding the pathophysiology of the tumour microenvironment. In head and neck cancer (HNC), most studies on tissue samples have only performed 16S rRNA short-read sequencing (SRS) on V3-V5 region. SRS is mostly limited to genus level identification. In this study, we compared full-length 16S rRNA long-read sequencing (FL-ONT) from Oxford Nanopore Technology (ONT) to V3-V4 Illumina SRS (V3V4-Illumina) in 26 HNC tumour tissues. Further validation was also performed using culture-based methods in 16 bacterial isolates obtained from 4 patients using MALDI-TOF MS. We observed similar alpha diversity indexes between FL-ONT and V3V4-Illumina. However, beta-diversity was significantly different between techniques (PERMANOVA - $R^2 = 0.131$, $p < 0.0001$). At higher taxonomic levels (Phylum to Family), all metrics were more similar among sequencing techniques, while lower taxonomy displayed more discrepancies. At higher taxonomic levels, correlation in relative abundance from FL-ONT and V3V4-Illumina were higher, while this correlation decreased at lower levels. Finally, FL-ONT was able to identify more isolates at the species level that were identified using MALDI-TOF MS (75% vs. 18.8%). FL-ONT was able to identify lower taxonomic levels at a better resolution as compared to V3V4-Illumina 16S rRNA sequencing.

Keywords Microbiome · 16S ribosomal RNA · Long read sequencing · Head and neck cancer

NIH National Library of Medicine
National Center for Biotechnology Information

BioProject | BioProject | PRJNA1087430
Create alert | Advanced | Browse by Project attributes

Display Settings: | Send to:

Comparing Full-length Nanopore 16S rRNA sequencing to Illumina V3-V4 16S rRNA sequencing in head and neck cancer tissues
Accession: PRJNA1087430 ID: 1087430

In this study, we compared full-length 16S rRNA long-read sequencing (FL-ONT) from Oxford Nanopore Technology (ONT) to V3-V4 Illumina SRS (V3V4-Illumina). More...

Accession	PRJNA1087430
Data Type	Raw sequence reads
Scope	Multispecies
Submission	Registration date: 13-Mar-2024 - University of Adelaide - ENT BHI
Relevance	16S rRNA on HNC tissue samples

Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
SRA Experiments	52
OTHER DATASETS	
BioSample	52

SRA Data Details

Parameter	Value
Data volume, Gbases	3
Data volume, Mbytes	2494

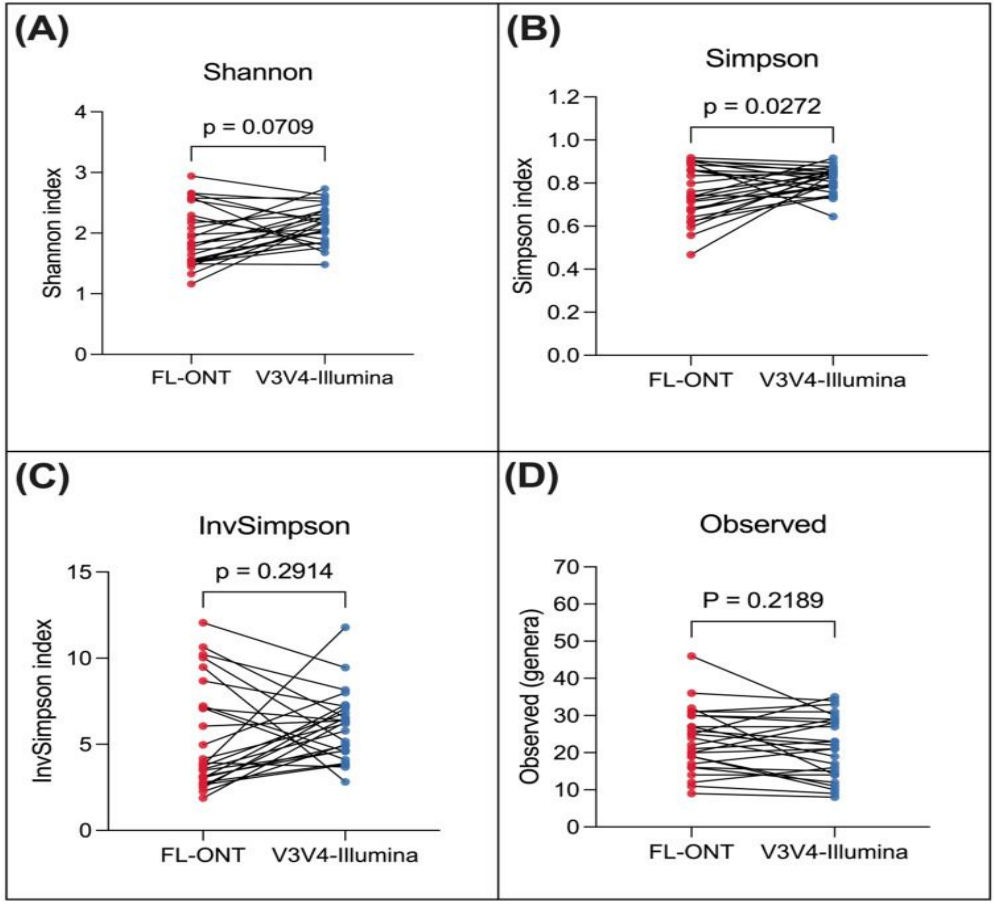
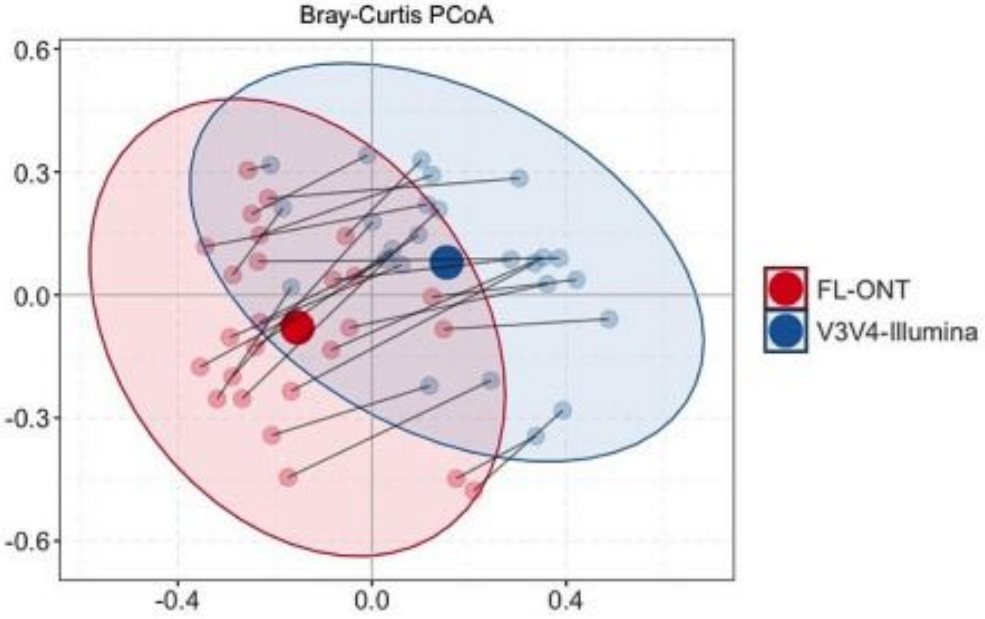


Fig. Paired alpha diversity analysis of FL-ONT and V3V4-Illumina at the genus level

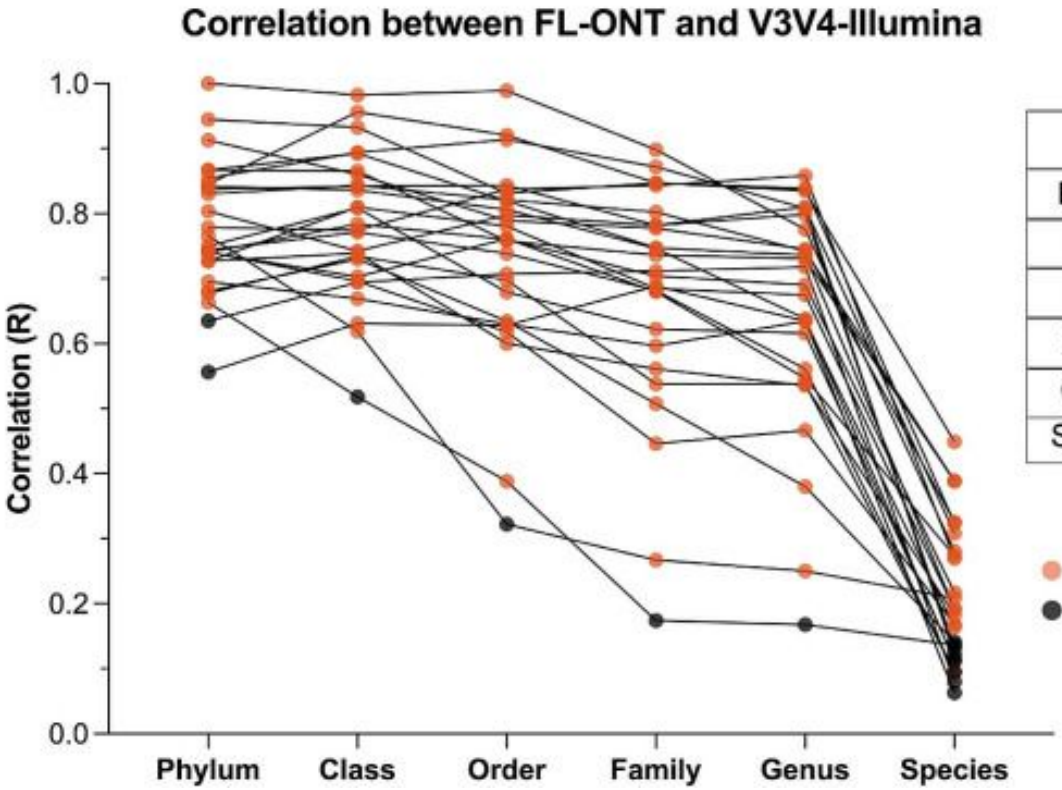
The results indicate that both sequencing methods are relatively similar in terms of alpha diversity.



PERMANOVA	R^2	
	0.131	$p < 0.0001$
ANOSIM	R	
	0.332	$p < 0.0001$
W_d	F	
	7.57	$p < 0.0001$

Fig. Paired beta diversity analysis of paired FL-ONT and V3V4-Illumina 16S rRNA sequencing on tissue samples at the genus level.

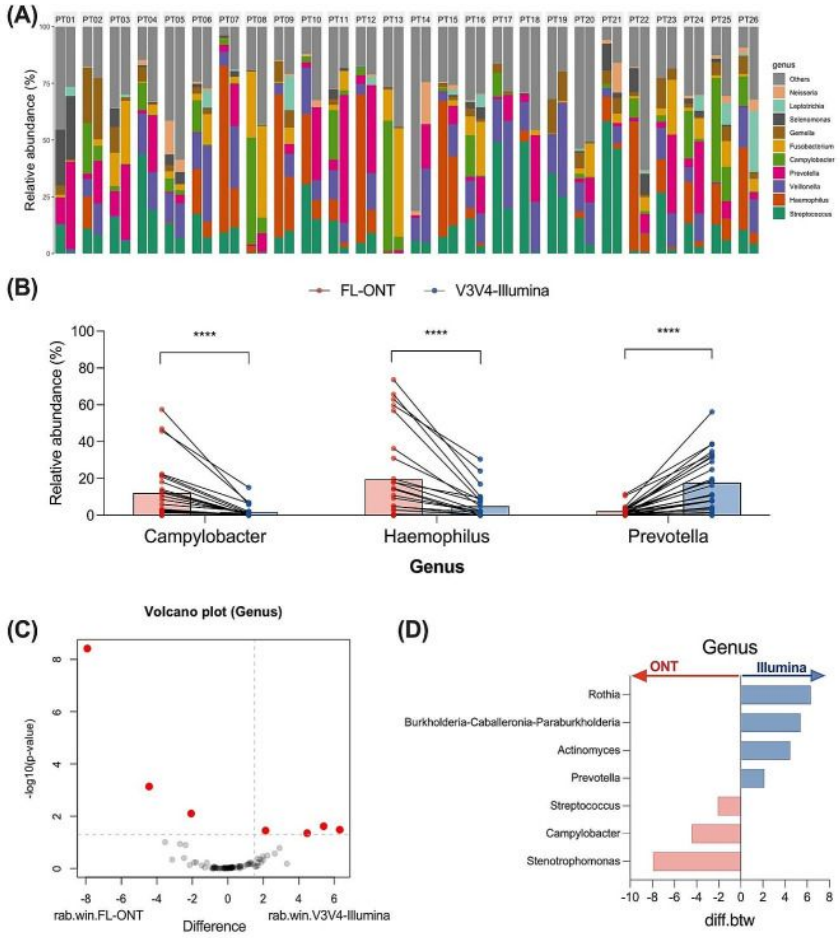
β-diversity differs between FL-ONT and V3V4-Illumina 16S rRNA sequencing at the genus level.



	Median R (95% CI)
Phylum	0.758 (0.727-0.842)
Class	0.779 (0.730-0.843)
Order	0.761 (0.679-0.820)
Family	0.708 (0.622-0.782)
Genus	0.683 (0.561-0.745)
Species	0.187 (0.132-0.274)

Fig. Correlation in bacterial relative abundance (%) at every taxonomic level between FL-ONT and V3V4-Illumina 16S rRNA sequencing.

Results: At higher taxonomic levels (phylum, class, order, family), the correlation in abundance ratios is high (R > 0.7). However, the correlation significantly decreases at the species level (R = 0.187).



The two sequencing methods, FL-ONT and V3V4-Illumina, detected bacterial genera with different abundances, with some genera showing significant differences (>10%) in richness between the two techniques.

Fig. Comparison of abundance between FL-ONT and V3V4-Illumina 16S rRNA sequencing at the Genus level.

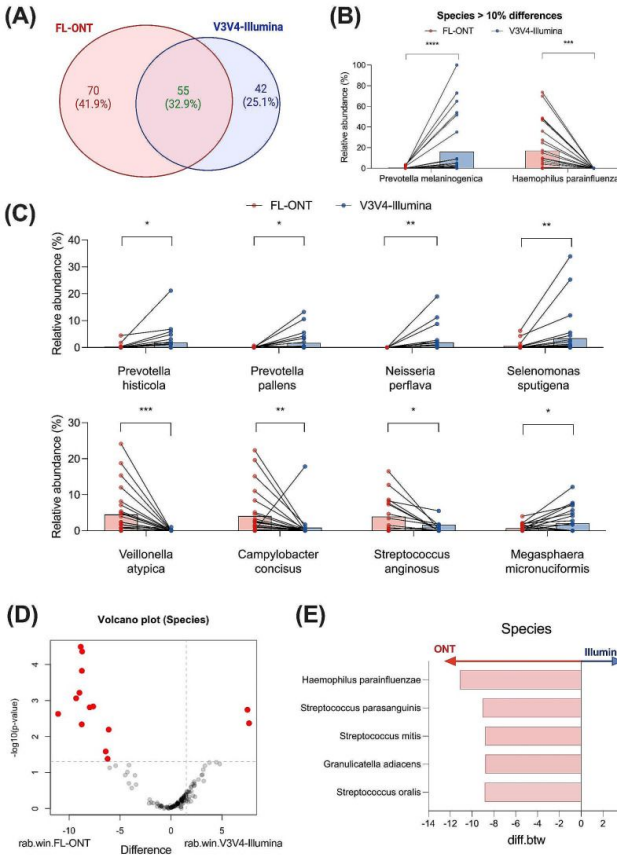


Fig. Comparison of abundance between FL-ONT and V3V4-Illumina 16S rRNA sequencing at the species level. After agglomerating to species level, a total of 167 species were identified.

Comparing the species between FL-ONT and V3V4-Illumina shows a significant difference in the relative abundance of the species.

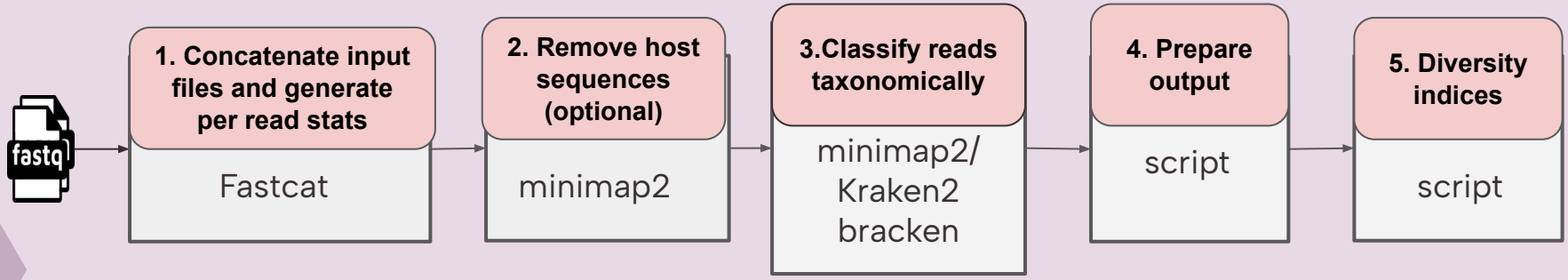
Summary

	FL-ONT (Oxford Nanopore Technologies)	V3V4 Illumina (Short Read Sequencing)
Application	• There is a need to identify bacteria at the species level, especially when dealing with many similar or hard-to-differentiate species. • Researching diverse and complex microbial communities. • High accuracy is required for bacterial species identification. • Analyzing samples with low or rare abundance.	• Research requires high accuracy but does not need species-level identification. • Rapid and cost-effective analysis of a large number of samples is needed. • Studying microbial communities at higher taxonomic levels (phylum, class, order).
Advantages	• Whole Genome Sequencing • High Detection Capability	• Lower Cost • High Accuracy at Higher Levels • Large Data Output
Disadvantages	• Higher Cost • Uneven Read Quality	• Limitations in Classification • Inability to Detect Unique Species

3

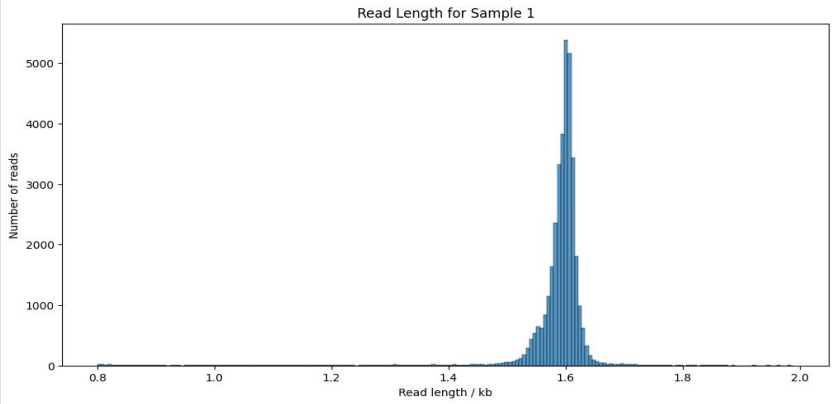
EPI2ME: pipeline wf-16S

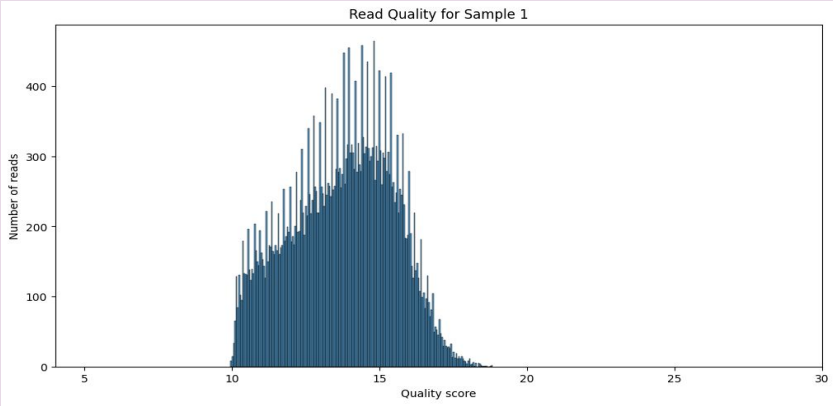
pipeline wf-16S

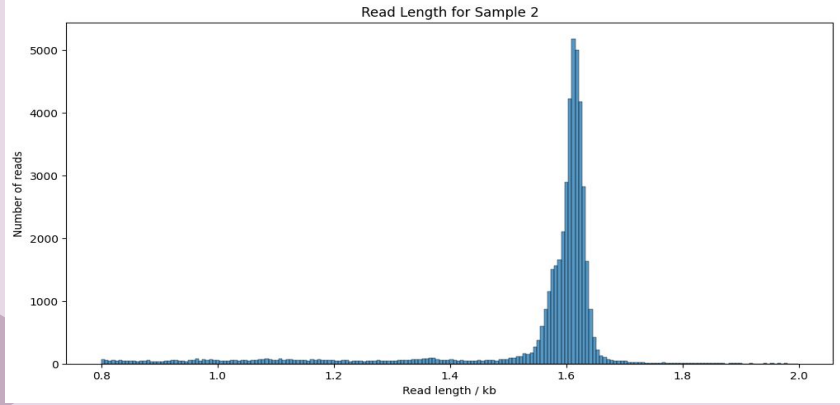


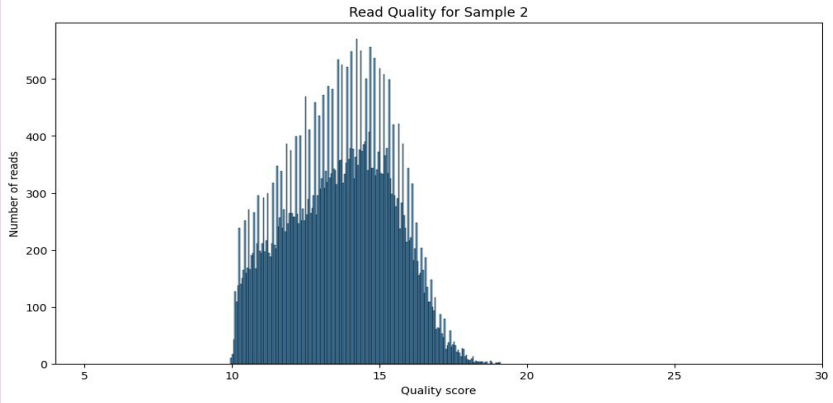
01	Concatenate input files and generate per read statsadipiscing	• Input: FASTQ files • Output: Concatenated file and per-read statistics (length & quality his)
02	Remove host sequences (optional)	• Input: Host reference • Output: BAM file of non-host reads
03	Classify reads taxonomically	**Kraken2**: • Input: Concatenated FASTQ, Kraken2 database • Output: Classification file (TSV), Report file (TXT) **Bracken**: • Input: Classification file from Kraken2 • Output: Bracken report file
04	Prepare output	• Input: Classification results • Output: Final report and additional summaries
05	Diversity indices	• Input: Taxonomic data • Output: Diversity metrics and curves

Read Summary

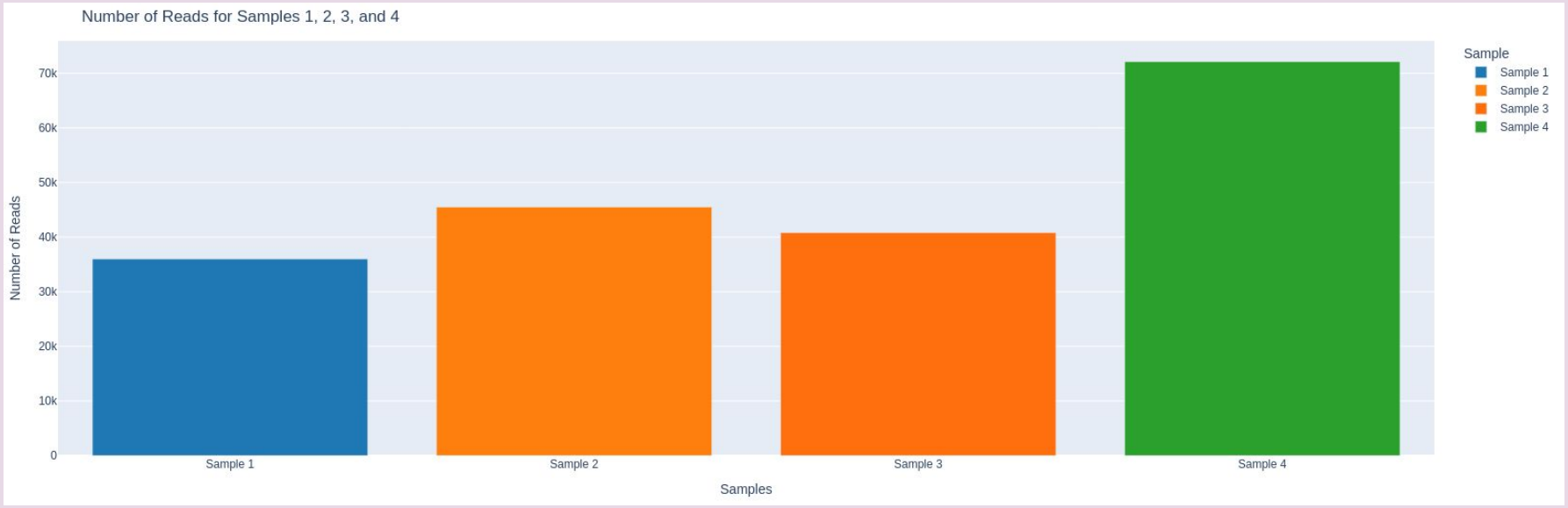




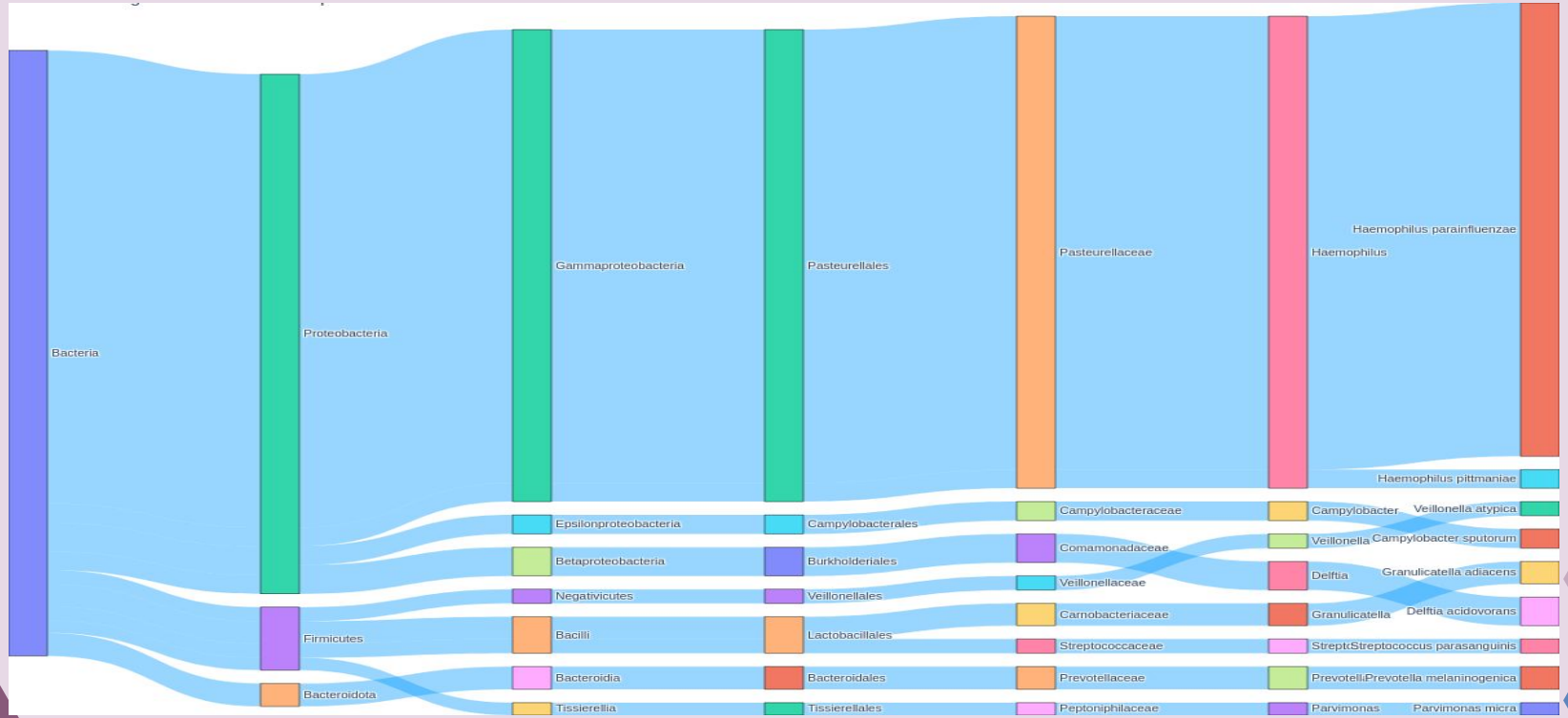




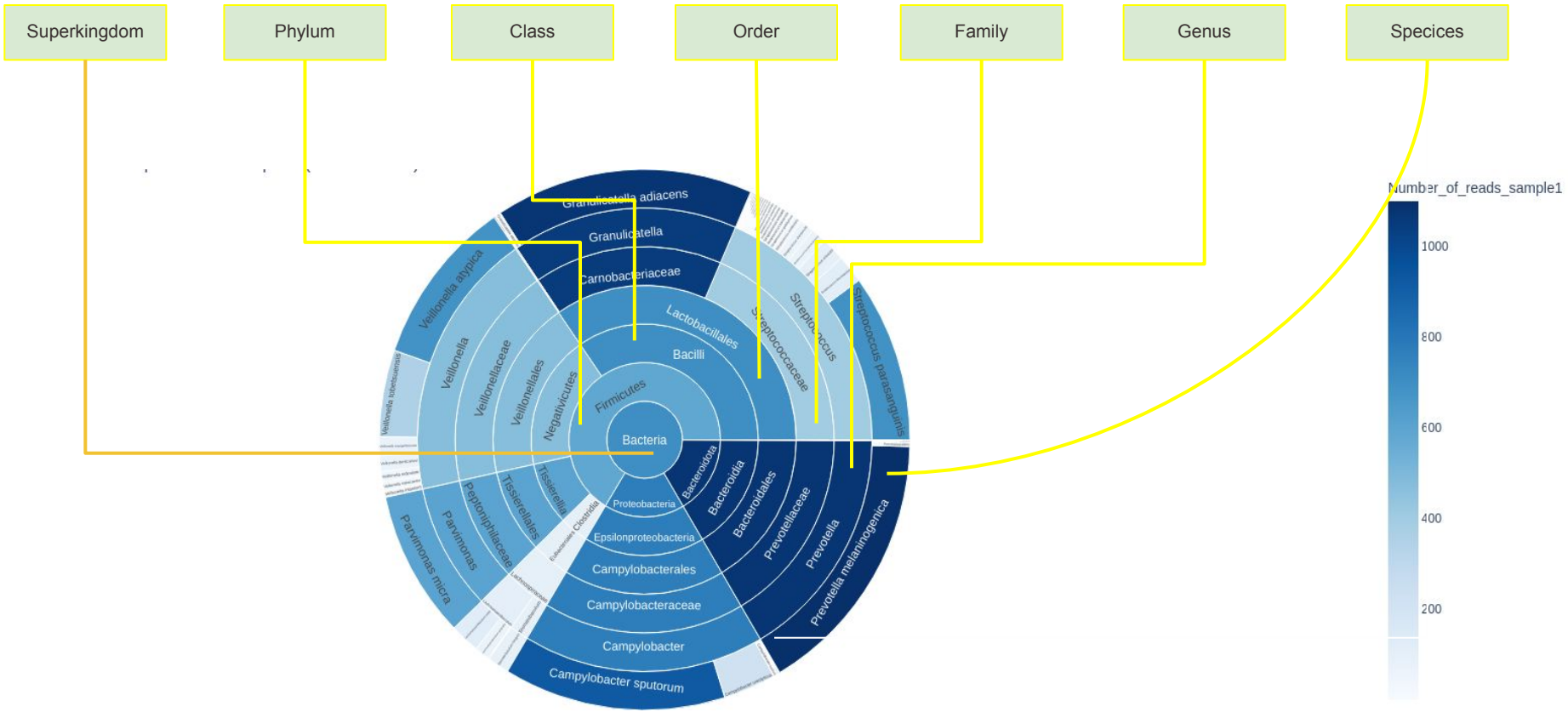
Sample Summary



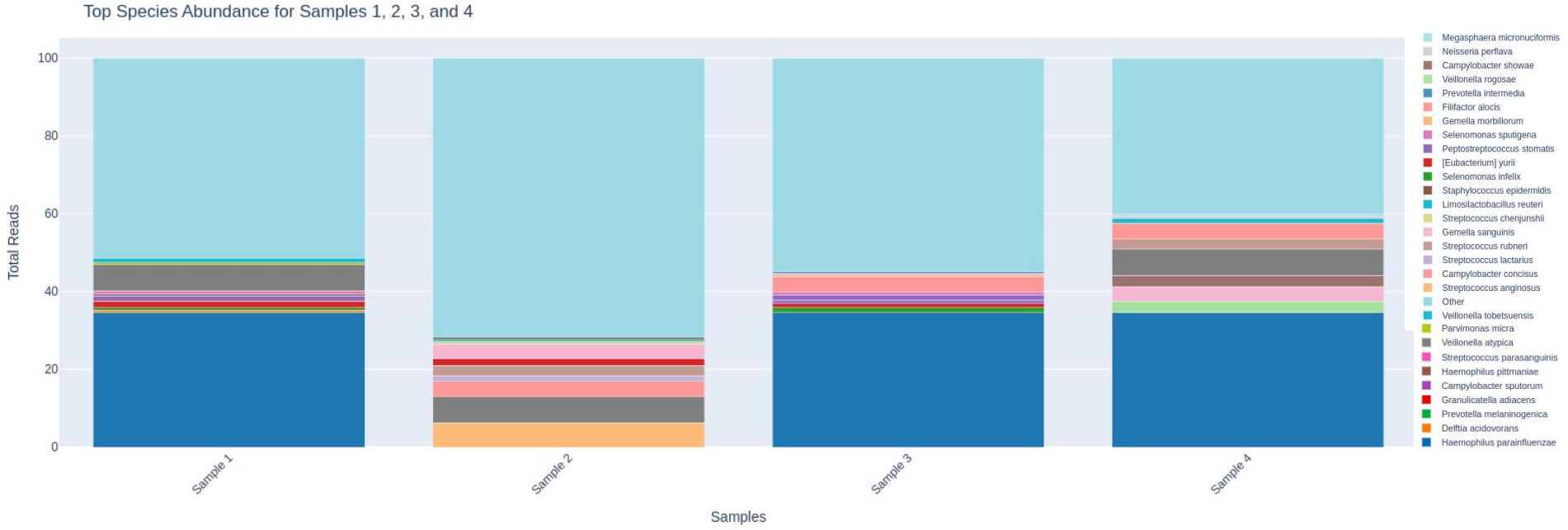
Lineages


Superkingdom
Phylum
Class
Order
Family
Genus
Specices
Bacteria
Proteobacteria
Firmicutes
Bacteroidota
Gammaproteobacteria
Epsilonproteobacteria
Betaproteobacteria
Negativicutes
Bacilli
Bacteroidia
Tissierellia
Pasteurellales
Campylobacterales
Burkholderiales
Veillonellales
Lactobacillales
Bacteroidales
Tissierellales
Pasteurellaceae
Campylobacteraceae
Comamonadaceae
Veillonellaceae
Carnobacteriaceae
Streptococcaceae
Prevotellaceae
Peptoniphilaceae
Haemophilus
Campylobacter
Veillonella
Delftia
Granulicatella
StreptStreptococcus parasanguinis
PrevotellPrevotella melaninogenica
Parvimonas
Haemophilus parainfluenzae
Haemophilus pittmaniae
Veillonella atypica
Campylobacter sputorum
Granulicatella adiacens
Delftia acidovorans
Parvimonas micra

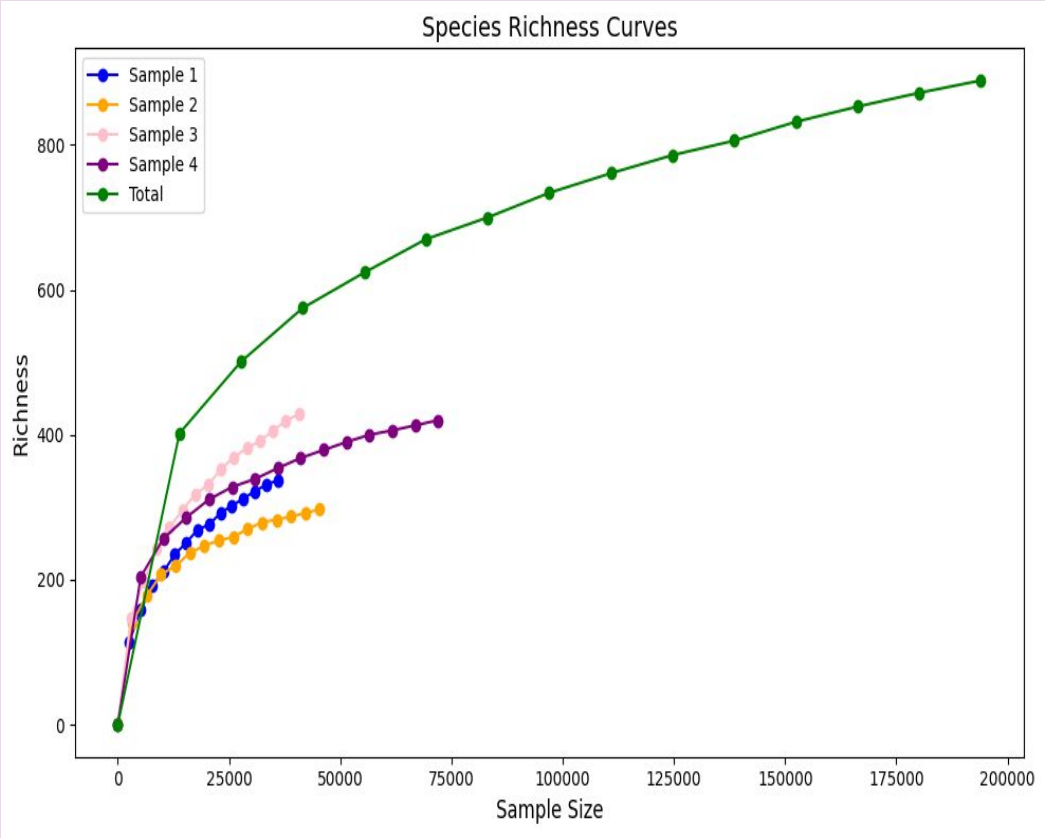
Sunburst


Superkingdom
Phylum
Class
Order
Family
Genus
Species
Number_of_reads_sample1
1000
800
600
400
200
Granulicatella adiacens
Granulicatella
Carnobacteriaceae
Lactobacillales
Bacilli
Streptococcaceae
Streptococcus
Streptococcus parasanguinis
Veillonella atypica
Veillonella
Veillonellaceae
Veillonellales
Negativicutes
Firmicutes
Bacteria
Bacteroidota
Bacteroidia
Bacteroidales
Prevotellaceae
Prevotella
Prevotella melaninogenica
Proteobacteria
Epsilonproteobacteria
Campylobacterales
Campylobacteraceae
Campylobacter
Campylobacter sputorum
Tissierella
Tissierellales
Peptoniphilaceae
Parvimonas
Parvimonas micra

Taxonomy


Top Species Abundance for Samples 1, 2, 3, and 4
Total Reads
100
80
60
40
20
0
Sample 1
Sample 2
Sample 3
Sample 4
Samples
Megasphaera micronuciformis
Neisseria perflava
Campylobacter showae
Veillonella rogosae
Prevotella intermedia
Filifactor alocis
Gemella morbillorum
Selenomonas sputigena
Peptostreptococcus stomatis
[Eubacterium] yurii
Selenomonas infelix
Staphylococcus epidermidis
Limosilactobacillus reuteri
Streptococcus chenjunshii
Gemella sanguinis
Streptococcus rubneri
Streptococcus lactarius
Campylobacter concisus
Streptococcus anginosus
Other
Veillonella tobetsuensis
Parvimonas micra
Veillonella atypica
Streptococcus parasanguinis
Haemophilus pittmaniae
Campylobacter sputorum
Granulicatella adiacens
Prevotella melaninogenica
Delftia acidovorans
Haemophilus parainfluenzae

Species Richness Curves



Tools

Name	Version	Link
fastcat	0.18.6	https://github.com/epi2me-labs/fastcat
minimap2	2.28-r1209	https://github.com/lh3/minimap2
Kraken2	2.1.2	https://github.com/DerrickWood/kraken2/archive/refs/tags/v2.1.2.tar.gz
Bracken	2.9	https://github.com/jenniferlu717/Bracken/archive/refs/tags/v2.9.tar.gz
Taxonkit	0.17.0	https://github.com/shenwei356/taxonkit/releases/tag/v0.17.0