

Introduction to Pacbio Long reads Full-Length 16S HiFi Sequencing

Presenters: Mạnh Hùng, Phương Thảo

20 Oct 2024

Table of contents

1. Quick overview of Pacbio Hifi sequencing technology
2. 16S full-length rRNA sequencing for accurate species-level taxonomic classification
3. Case study: Comparison of 16S rRNA Pacbio Long reads vs Illumina Short reads sequencing
4. Google Colab code

Pacbio Hifi sequencing produce highly accurate long reads

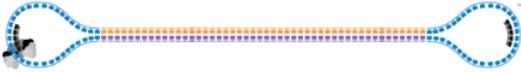
Start with high-quality double stranded DNA



Prepare SMRTbell libraries



Anneal primers and bind DNA polymerase



Circularized DNA is sequenced in repeated passes

> 10kb reads

The polymerase reads are trimmed of adapters to yield subreads

Consensus and methylation status are called from subreads



HiFi read
(99.9% accuracy)



Long reads

- Tens of kilobases
- Sequence from 500 bp to >20,000 bp inserts



High accuracy

- Free of systematic errors
- Achieves >99.999% (Q50) consensus accuracy

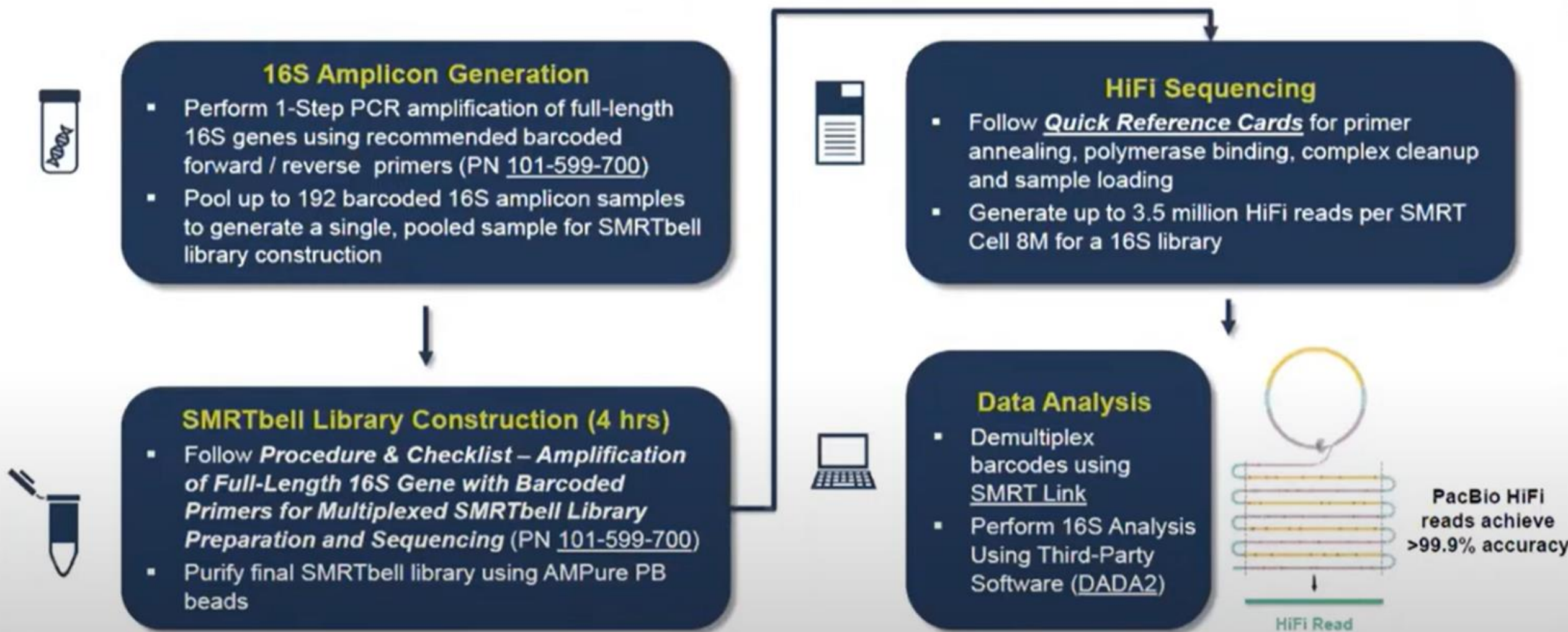


Single-molecule resolution

- Sequence DNA or RNA
- Long reads with $\geq Q30$ (>99.9%) single-molecule accuracy

16S LIBRARY SAMPLE PREPARATION & SEQUENCING WORKFLOW OVERVIEW

Workflow summary for constructing SMRTbell libraries suitable for HiFi sequencing on the Sequel, Sequel II and Sequel IIe Systems for 16S metagenomics applications

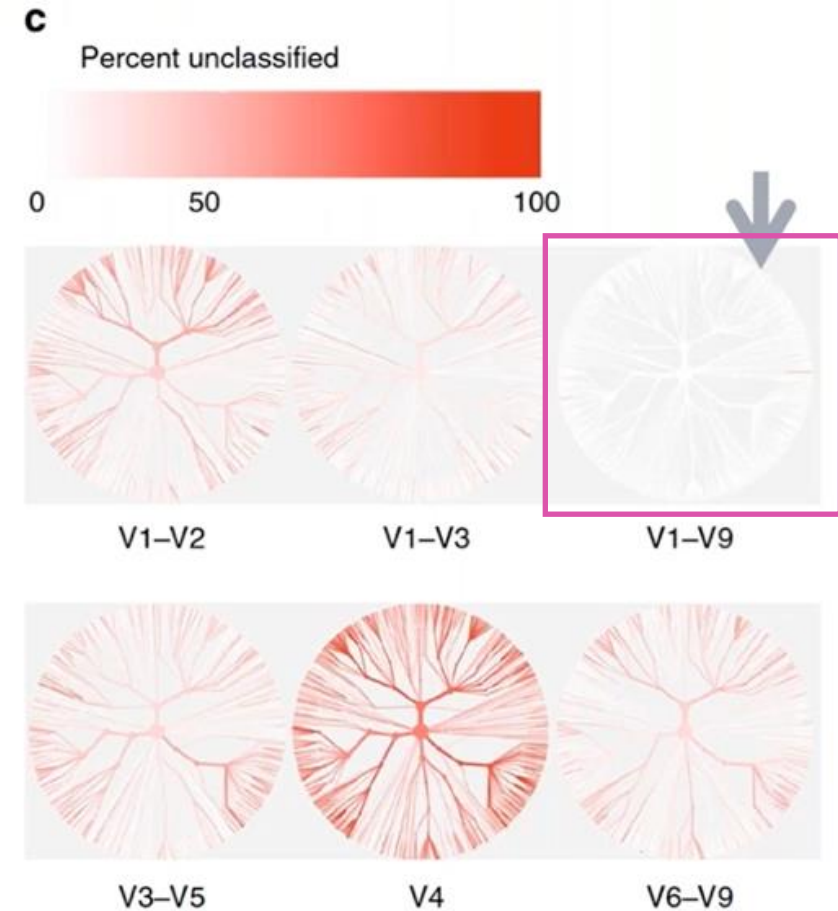


Full-length 16S sequencing is the only way to resolve all the clades seen in the human gut, without bias

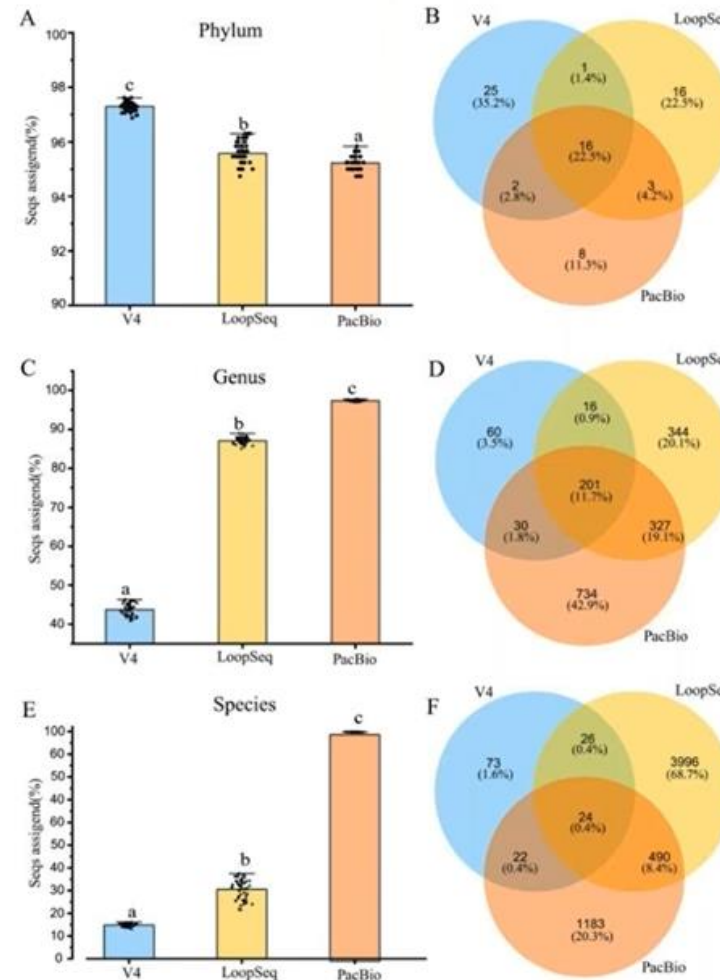
Question: What proportion of data can be re-identified at the species level using different 16S amplicons?

Problem: George Weinstock, leader of the *Human Microbiome Project*, showed that different amplicons show bias in the taxa they can identify

- V4: Consistently poor performance
- V1–V2: poor for *Proteobacteria*
- V3–V5: poor for *Actinobacteria*
- V1–V3: good results for *Escherichia / Shigella*
- V3–V5: good results for *Klebsiella*
- V6–V9: good results for *Clostridium* and *Staphylococcus*



More reads are classified to species and genus level with PacBio full-length 16S sequencing compared to V4 and synthetic long-read full-length 16S



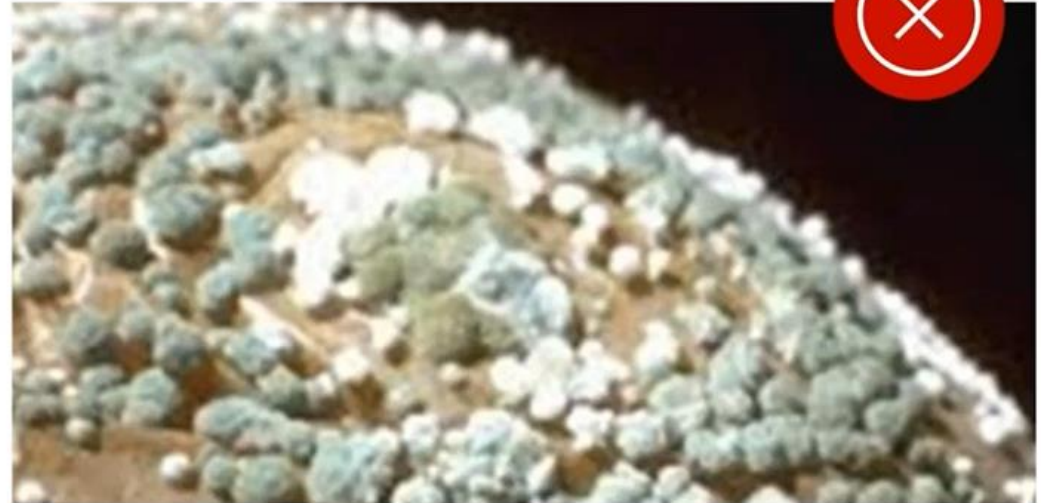
97% sequences assigned to **genus level** using PacBio vs LoopSeq at 87% and V4 at 44%

99.7% sequences assigned to **species level** using PacBio vs LoopSeq at ~31% and V4 at ~15%

Small [species] differences really matter as it relates to health and disease



Penicillium roqueforti



Penicillium verrucosum

Being able to call and identify these subtle differences can make something taste good vs something that can be deadly

► [Medicine \(Baltimore\)](#). 2022 Jul 22;101(29):e29826. doi: [10.1097/MD.00000000000029826](https://doi.org/10.1097/MD.00000000000029826) [↗](#)

Role of gut microbiota in postoperative complications and prognosis of gastrointestinal surgery: A narrative review

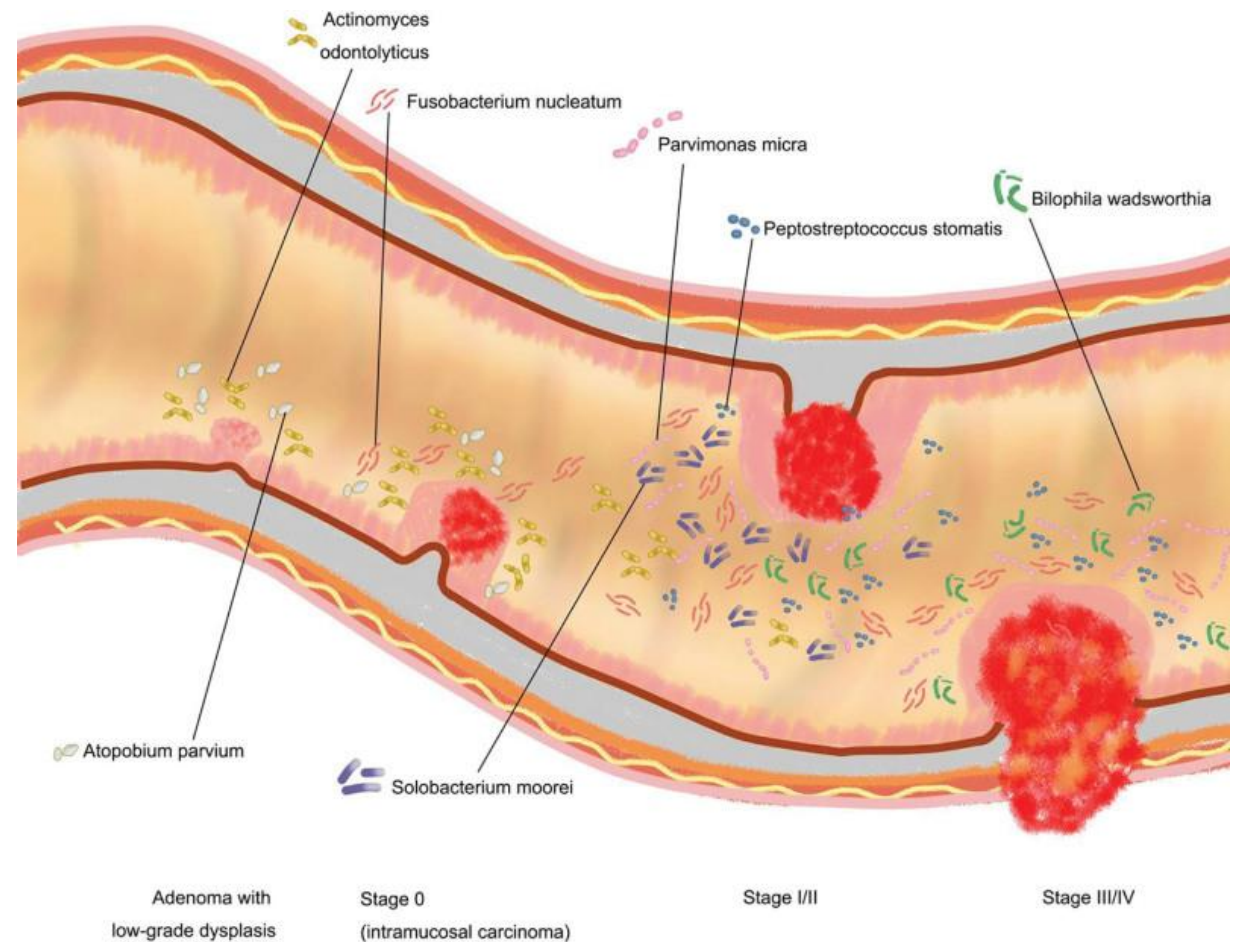
[Yong Shi](#)^a, [Huxiao Cui](#)^a, [Fangjie Wang](#)^a, [Yanxia Zhang](#)^a, [Qingbin Xu](#)^a, [Dan Liu](#)^a, [Kunhui Wang](#)^a, [Sen Hou](#)^{a,*}

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#)

PMCID: PMC9302249 PMID: [35866808](#)

Abstract

Gastrointestinal surgery is often challenging because of unexpected postoperative complications such as pouchitis, malabsorption, anastomotic leak, diarrhea, inflammatory responses, and life-threatening infections. Moreover, the gut microbiota has been shown to be associated with the complications described above. Major intestinal reconstruction, such as Roux-en-Y gastric bypass (RYGB) and ileal pouch-anal anastomosis surgery, could result in altered gut microbiota, which might lead to some of the benefits of these procedures but could also contribute to the development of postsurgical complications. Moreover, postsurgical reestablishment of the gut microbiota population is still poorly understood. Here, we review evidence outlining the role of gut microbiota in complications of gastrointestinal surgery, especially malabsorption, anastomotic leak, pouchitis, and infections. In addition, this review will evaluate the risks and benefits of live biotherapeutics in the complications of gastrointestinal surgery.



Major change of microbial species during multistep CRC progression.

Case study

Comparison of taxonomy resolution

16S rRNA Pacbio Long reads vs Illumina Short reads sequencing

Link github: <https://github.com/thaocaoHPzbook/16S-Long-Reads-vs-Short-Reads>

Research | [Open access](#) | Published: 25 March 2024

Full-length 16S rRNA gene sequencing by PacBio improves taxonomic resolution in human microbiome samples

Elena Buetas, [Marta Jordán-López](#), [Andrés López-Roldán](#), [Giuseppe D'Auria](#), [Lucia Martínez-Priego](#), [Griselda De Marco](#), [Miguel Carda-Diéguez](#)  & [Alex Mira](#)

BMC Genomics 25, Article number: 310 (2024) | [Cite this article](#)

3946 Accesses | 8 Citations | 6 Altmetric | [Metrics](#)

Abstract

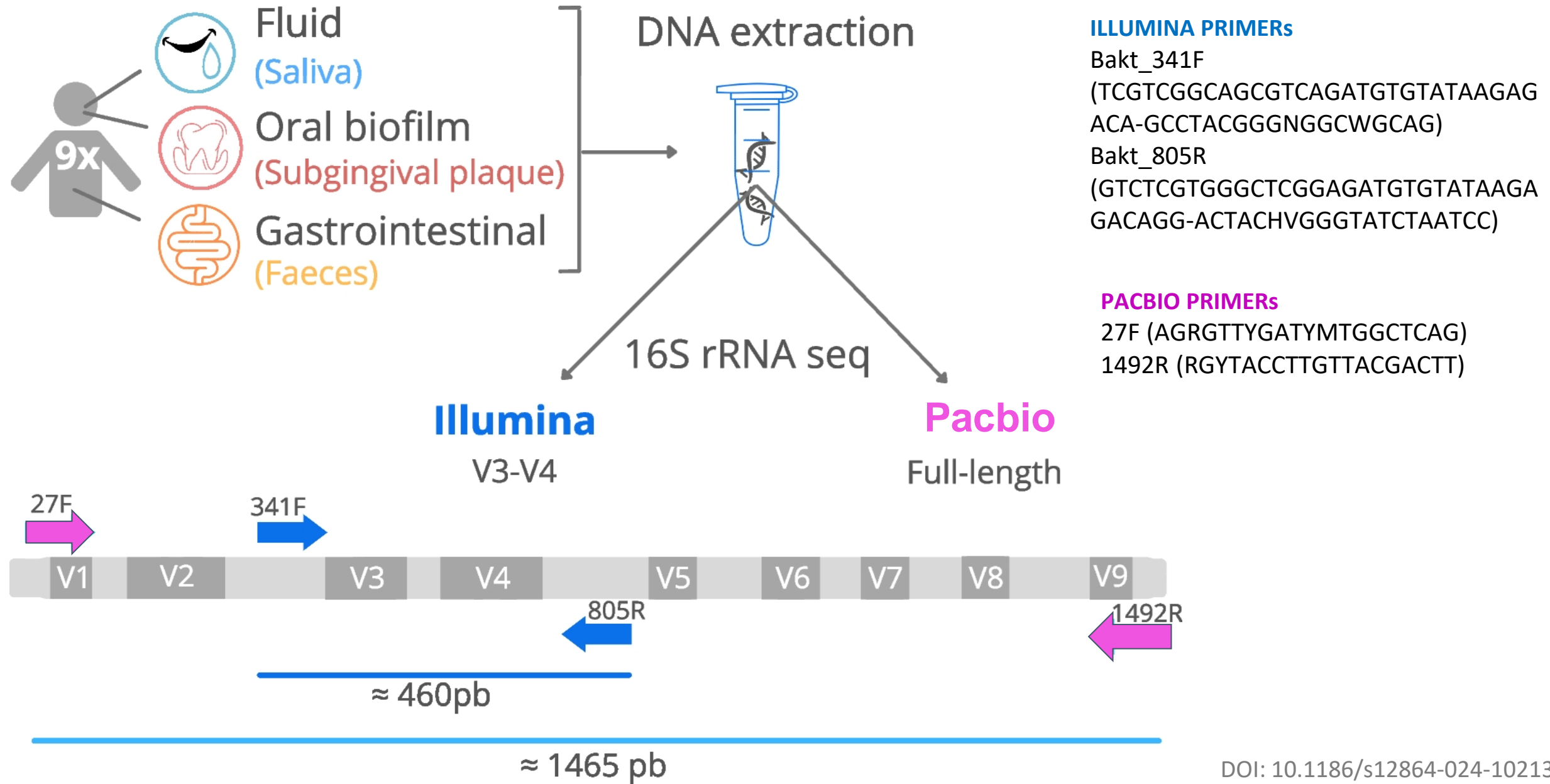
Background

Sequencing variable regions of the 16S rRNA gene (≈ 300 bp) with Illumina technology is commonly used to study the composition of human microbiota. Unfortunately, short reads are unable to differentiate between highly similar species. Considering that species from the same genus can be associated with health or disease it is important to identify them at the lowest possible taxonomic rank. Third-generation sequencing platforms such as PacBio SMRT, increase read lengths allowing to sequence the whole gene with the maximum taxonomic resolution. Despite its potential, full length 16S rRNA gene sequencing is not widely used yet. The aim of the current study was to compare the sequencing output and taxonomic annotation performance of the two approaches (Illumina short read sequencing and PacBio long read sequencing of 16S rRNA gene) in different human microbiome samples. DNA from saliva, oral biofilms (subgingival plaque) and faeces of 9 volunteers was isolated. Regions V3-V4 and V1-V9 were amplified and sequenced by Illumina Miseq and by PacBio Sequel II sequencers, respectively.

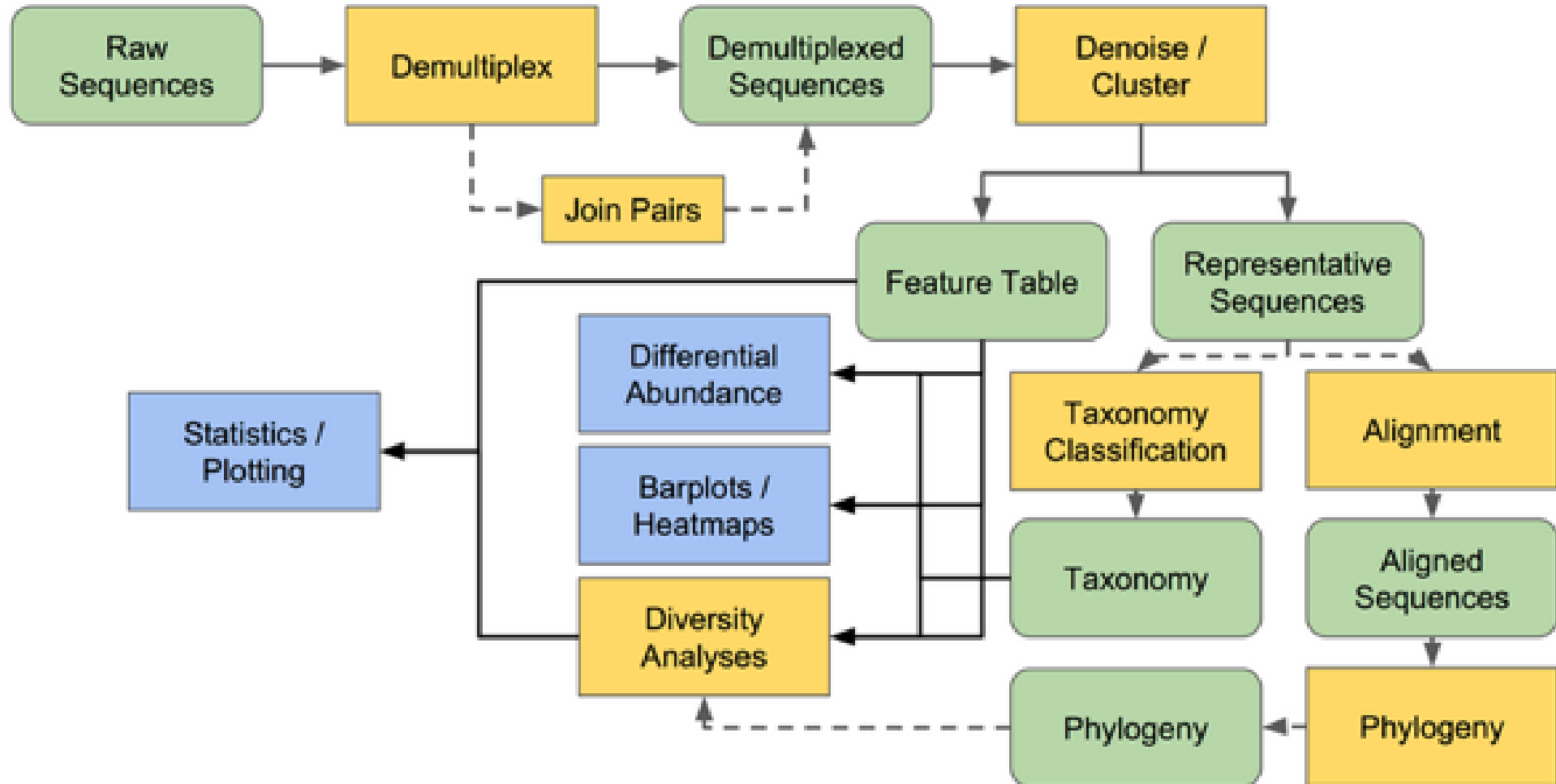
DOI: [10.1186/s12864-024-10213-5](https://doi.org/10.1186/s12864-024-10213-5)

NCBI SRA under accession number [PRJNA933120](#)

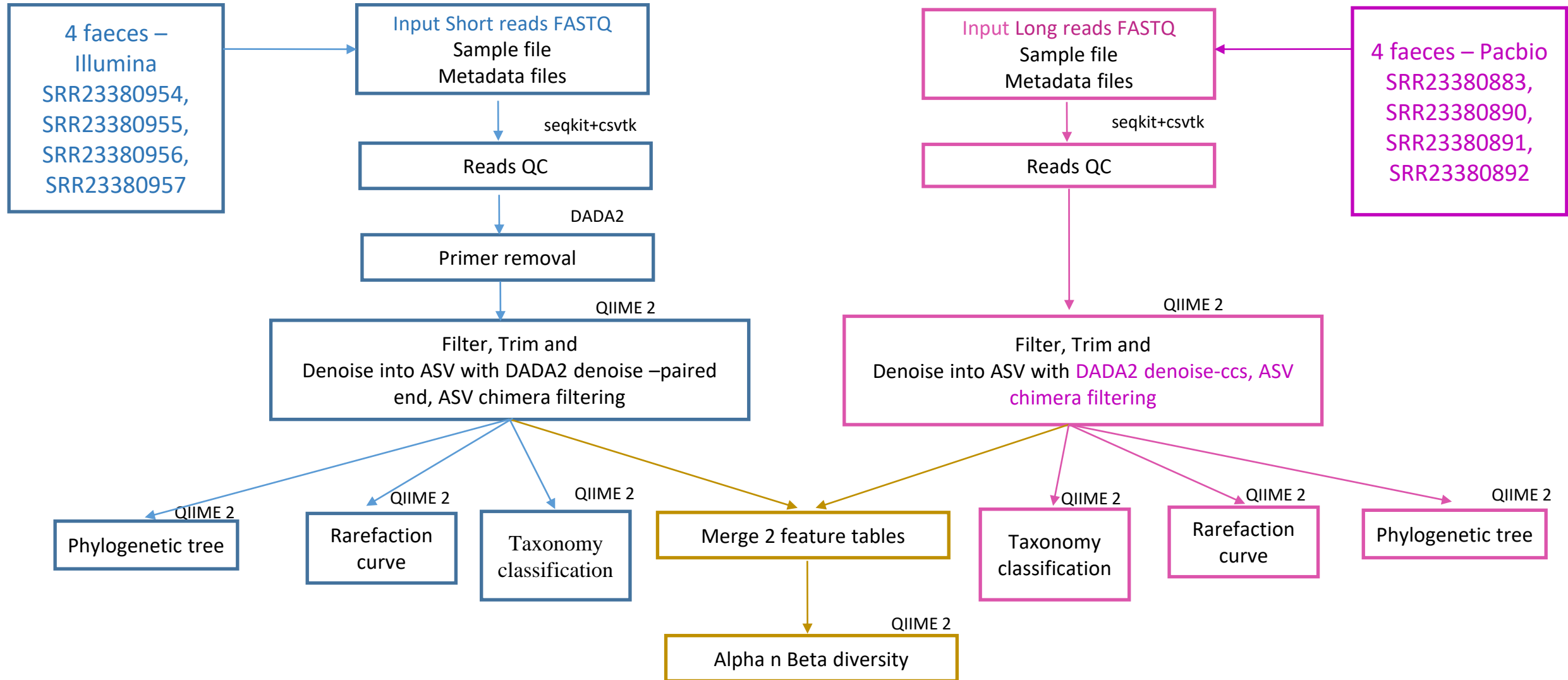
Summary of study design



Qiime 2 pipeline



Summary of sequencing pipeline



Feature table

SHORT READS

Overview	Interactive Sample Detail	Feature Detail
Table summary		
Summary Statistic		Value
Number of samples		4
Number of features		809
Total frequency		412,745

Frequency per sample

		Frequency
Minimum frequency		27,021
1st quartile		85,401
Median frequency		120,829
3rd quartile		138,614.2
Maximum frequency		144,066
Mean frequency		103,186.2

Frequency per feature

		Frequency
Minimum frequency		2
1st quartile		16
Median frequency		60
3rd quartile		256
Maximum frequency		15,771
Mean frequency		510.2

LONG READS

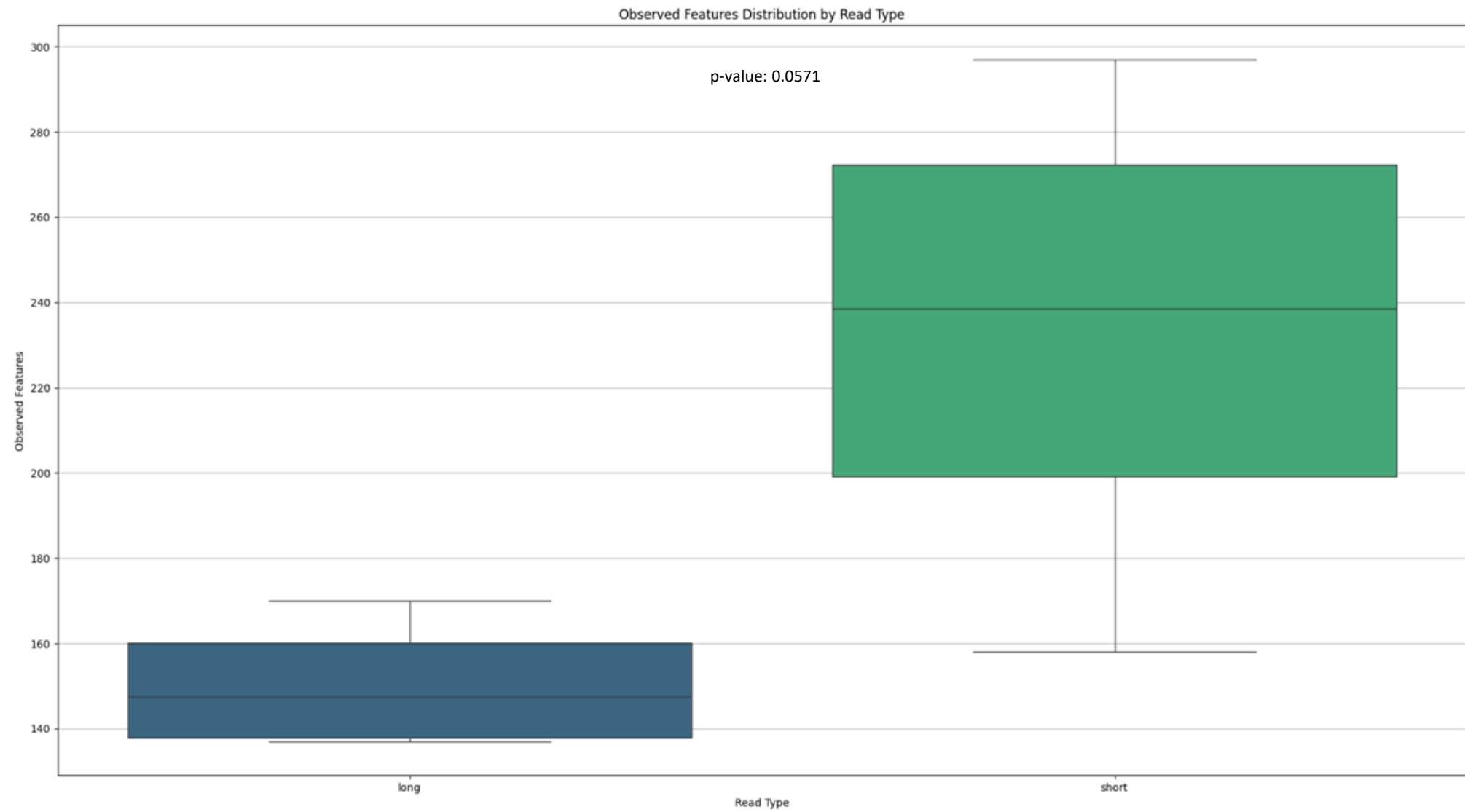
Overview	Interactive Sample Detail	Feature Detail
Table summary		
Summary Statistic		Value
Number of samples		4
Number of features		616
Total frequency		32,889

Frequency per sample

		Frequency
Minimum frequency		7,718
1st quartile		8,027.8
Median frequency		8,233.5
3rd quartile		8,428
Maximum frequency		8,704
Mean frequency		8,222.2

Frequency per feature

		Frequency
Minimum frequency		2
1st quartile		11
Median frequency		23
3rd quartile		47
Maximum frequency		1,472
Mean frequency		53.4



Taxa barplots

SHORT READS

Download

SVG (bars)

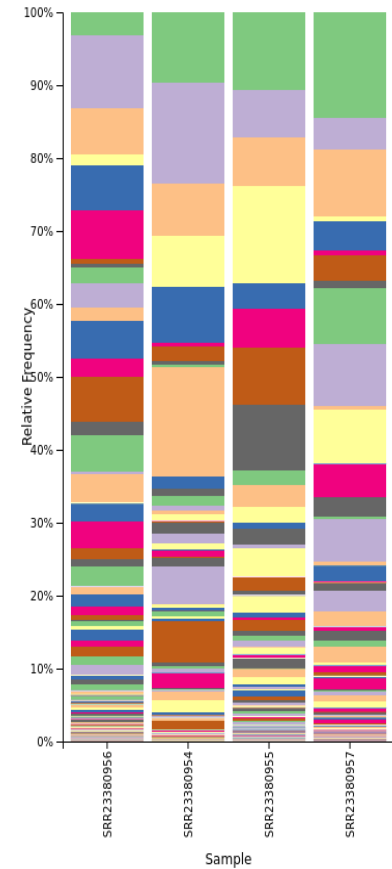
SVG (legend)

CSV

Bar Width



Hover over the plot to learn more



- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Prevotellaceae;g_Prevotella;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;__
- d_Bacteria;p_Firmicutes;c_Negativicutes;o_Veillonellales-Selenomonadales;f_Veillonellaceae;g_Dialister;__
- d_Bacteria;p_Firmicutes;c_Negativicutes;o_Acidaminococcales;f_Acidaminococcaceae;g_Phascolartobacterium;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Erysipelotrichales;f_Erysipelotrichaceae;g_Holdemania;__
- d_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Aeromonadales;f_Succinivibrionaceae;g_Succinivibrio;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Dorea;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_uncultured;g_gut_metagenome
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Coproccoccus;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Eubacterium]_ruminantium_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Fusicatibacter;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Erysipelotrichales;f_Erysipelatoclostridiaceae;g_Catenibacterium;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Christensenellales;f_Christensenellaceae;g_;;
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Ruminococcus]_torques_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_UCG-005;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Christensenellales;f_Christensenellaceae;g_Christensenellaceae_R-7_group;s_uncultured_spirochrome
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Monoglobales;f_Monoglobaceae;g_Monoglobus;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Prevotellaceae;g_Alloprevotella;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Erysipelotrichales;f_Erysipelatoclostridiaceae;g_Erysipelotrichaceae_UCG-003;__
- d_Bacteria;p_Firmicutes;c_Clostridia;_;;
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Roseburia;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Rikenellaceae;g_Alistipes;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Ruminococcus;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Christensenellales;f_Christensenellaceae;g_Christensenellaceae_R-7_group;s_uncultured_prokaryote
- d_Bacteria;p_Verrucomicrobiota;c_Verrucomicrobiae;o_Verrucomicrobiales;f_Akkermansiaceae;g_Akkermansia;__
- d_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae;g_;;
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Eubacterium]_eligans_group;__
- d_Bacteria;p_Actinobacteriota;c_Coribacteriia;o_Coribacteriales;f_Coribacteriaceae;g_Collinsella;__
- d_Bacteria;p_Desulfobacterota;c_Desulfovibrionia;o_Desulfovibrionales;f_Desulfovibrionaceae;g_Desulfovibrio;s_metagenome
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_NK44214_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Eubacterium]_hallii_group;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Tannerellaceae;g_Parabacteroides;__
- d_Bacteria;p_Firmicutes;c_Negativicutes;o_Veillonellales-Selenomonadales;f_Selenomonadaceae;g_Mitsuokella;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Anaerostipes;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_;;
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g_Lactobacillus;s_Lactobacillus_ruminis
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Ruminococcus_gut_metagenome

LONG READS

Download

SVG (bars)

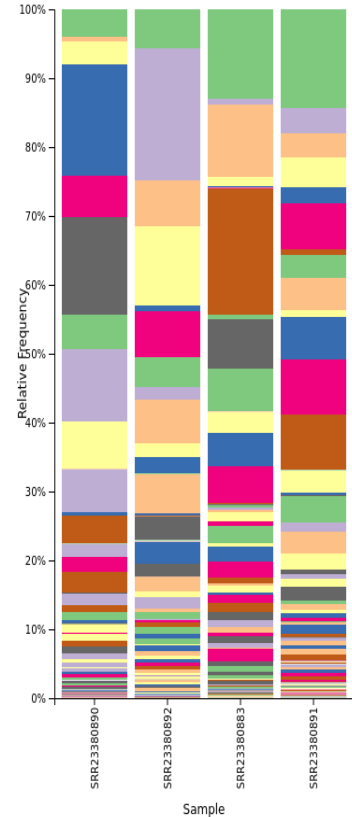
SVG (legend)

CSV

Bar Width

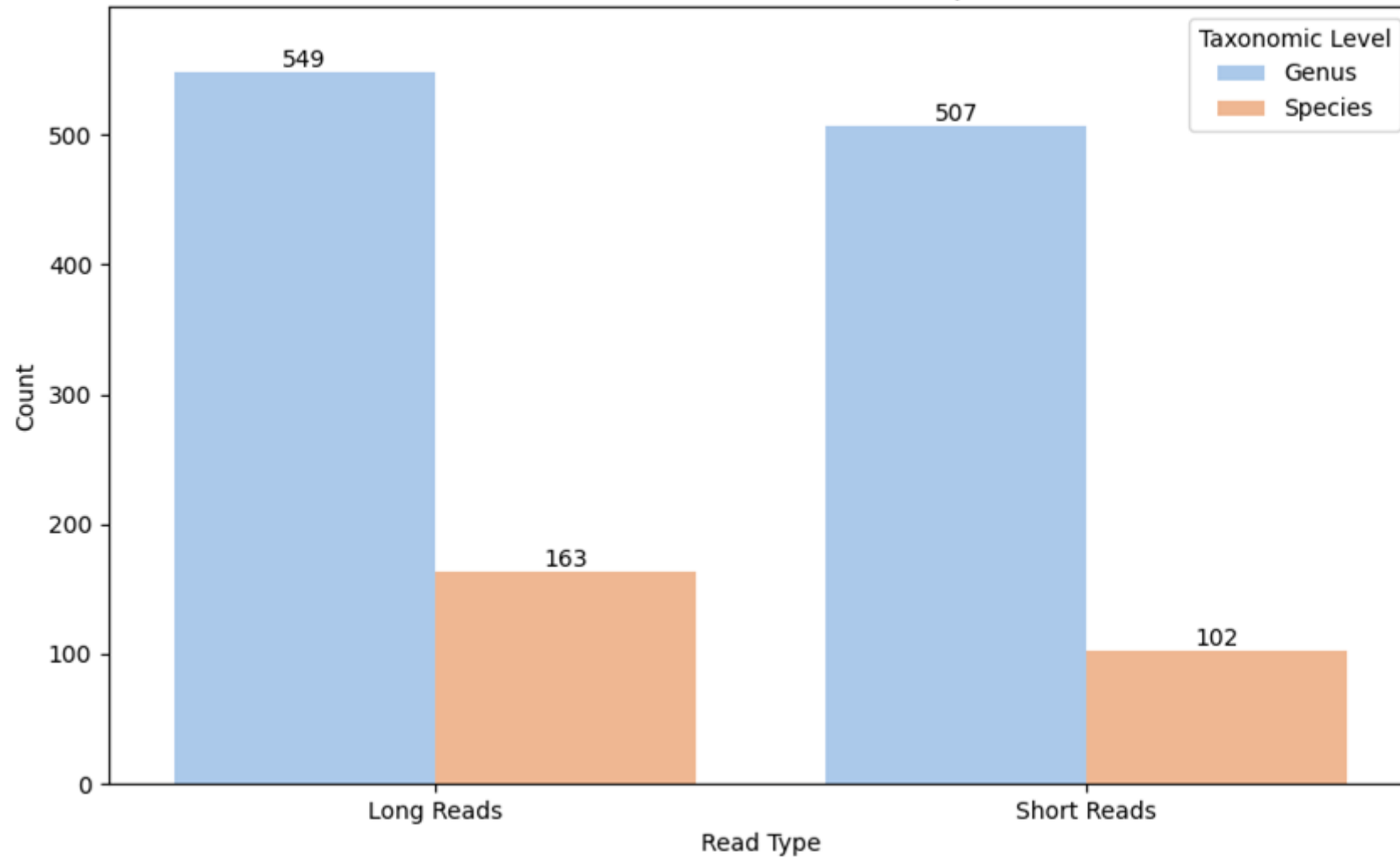


Hover over the plot to learn more



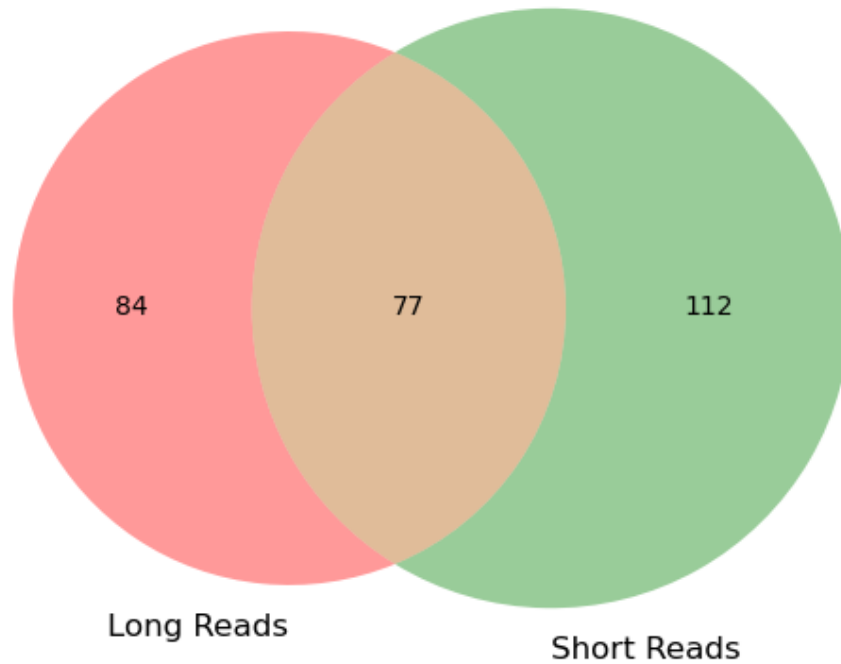
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Faecalibacterium;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Subdoligranulum;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridia_UCG-014;f_Clostridia_UCG-014;g_Clostridia_UCG-014;__
- d_Bacteria;p_Firmicutes;c_Negativicutes;o_Acidaminococcales;f_Acidaminococcaceae;g_Phascolartobacterium;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Agathobacter;__
- d_Bacteria;p_Actinobacteriota;c_Actinobacteria;o_Bifidobacteriales;f_Bifidobacteriaceae;g_Bifidobacterium;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_dorei
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Ruminococcus]_gnavus_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Dorea;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_[Eubacterium]_coprostanoligenes_group;g_[Eubacterium]_coprostanoligenes_group;s_gut_metagenome
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Rikenellaceae;g_Alistipes;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Blautia;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Erysipelotrichales;f_Erysipelotrichaceae;g_Holdemania;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_massiliensis
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_vulgatus
- d_Bacteria;p_Verrucomicrobiota;c_Verrucomicrobiae;o_Verrucomicrobiales;f_Akkermansiaceae;g_Akkermansia;s_uncultured_organism
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Oscillospiraceae;g_UCG-002;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Blautia;s_Blautia_hansenii
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Rikenellaceae;g_Alistipes;s_Alistipes_finegoldii
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_uniformis
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_;;
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Ruminococcus;s_gut_metagenome
- d_Bacteria;p_Actinobacteriota;c_Coribacteriia;o_Coribacteriales;f_Coribacteriaceae;g_Collinsella;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Christensenellales;f_Christensenellaceae;g_Christensenellaceae_R-7_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Ruminococcus]_torques_group;__
- d_Bacteria;p_Firmicutes;c_Bacilli;o_Erysipelotrichales;f_Erysipelatoclostridiaceae;g_Erysipelotrichaceae_UCG-003;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Eubacterium]_ruminantium_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_Ruminococcaceae;g_Ruminococcus;s_Ruminococcus_bicirculans
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Oscillospirales;f_[Eubacterium]_coprostanoligenes_group;g_[Eubacterium]_coprostanoligenes_group;s_human_gut
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_;;
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Lachnoclostridium;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_Anaerostipes;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides;s_Bacteroides_intestinalis
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Peptostreptococcales-Tissierellales;f_Anaerovoracaceae;g_Family_XIII_AD3011_group;s_uncultured_organism
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Prevotellaceae;g_Prevotellaceae_NK3B31_group;__
- d_Bacteria;p_Firmicutes;c_Clostridia;o_Lachnospirales;f_Lachnospiraceae;g_[Eubacterium]_eligans_group;__
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Tannerellaceae;g_Parabacteroides;s_Parabacteroides_merdae
- d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Bacteroidales;f_Desulfotomaculum;g_Paracaulobacter

Number of classified taxon at Genus and species levels of 2 difference sequencing methods

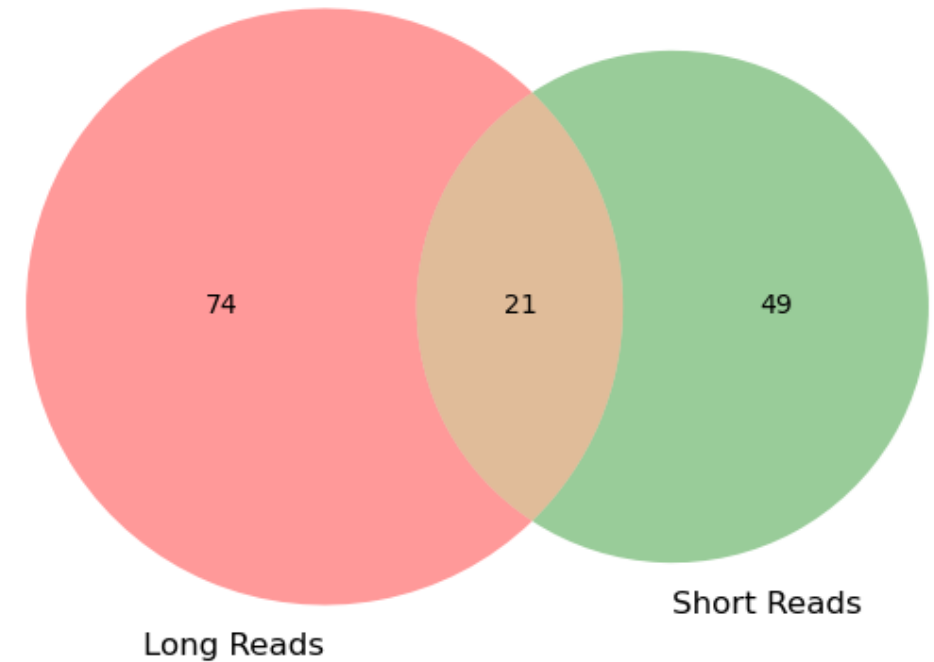


Number of identified taxon at Genus and species levels of 2 difference sequencing methods

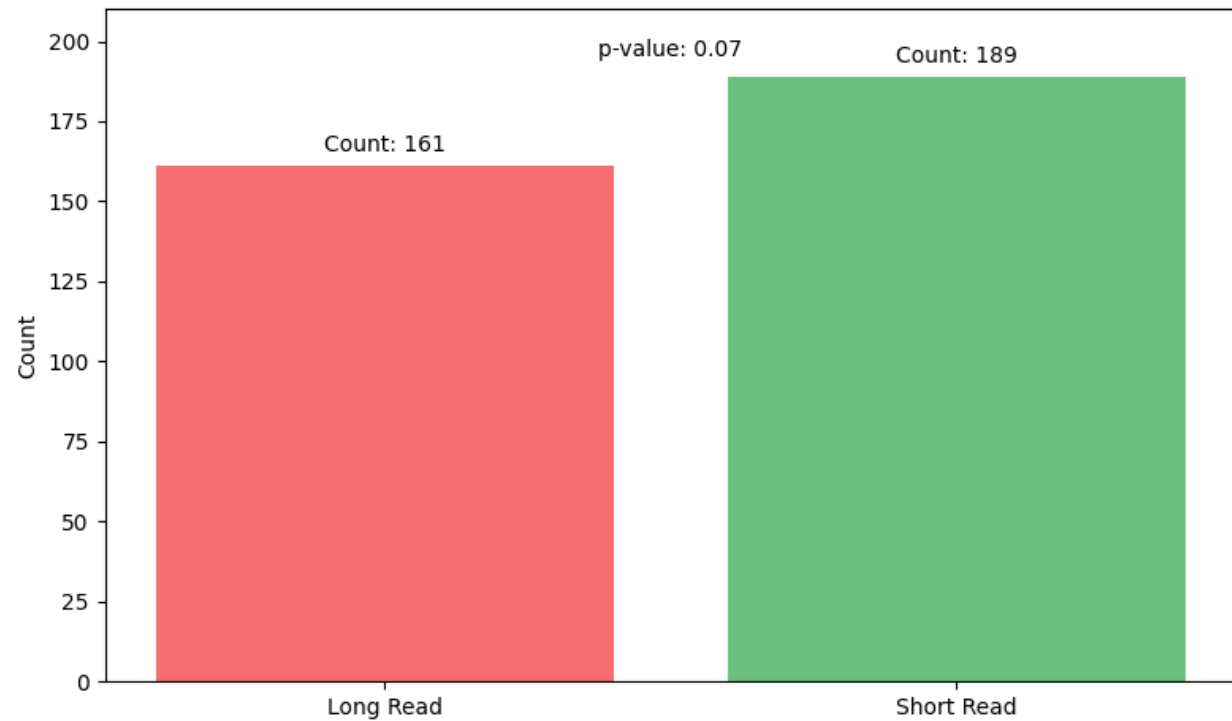
Venn Diagram of Genus



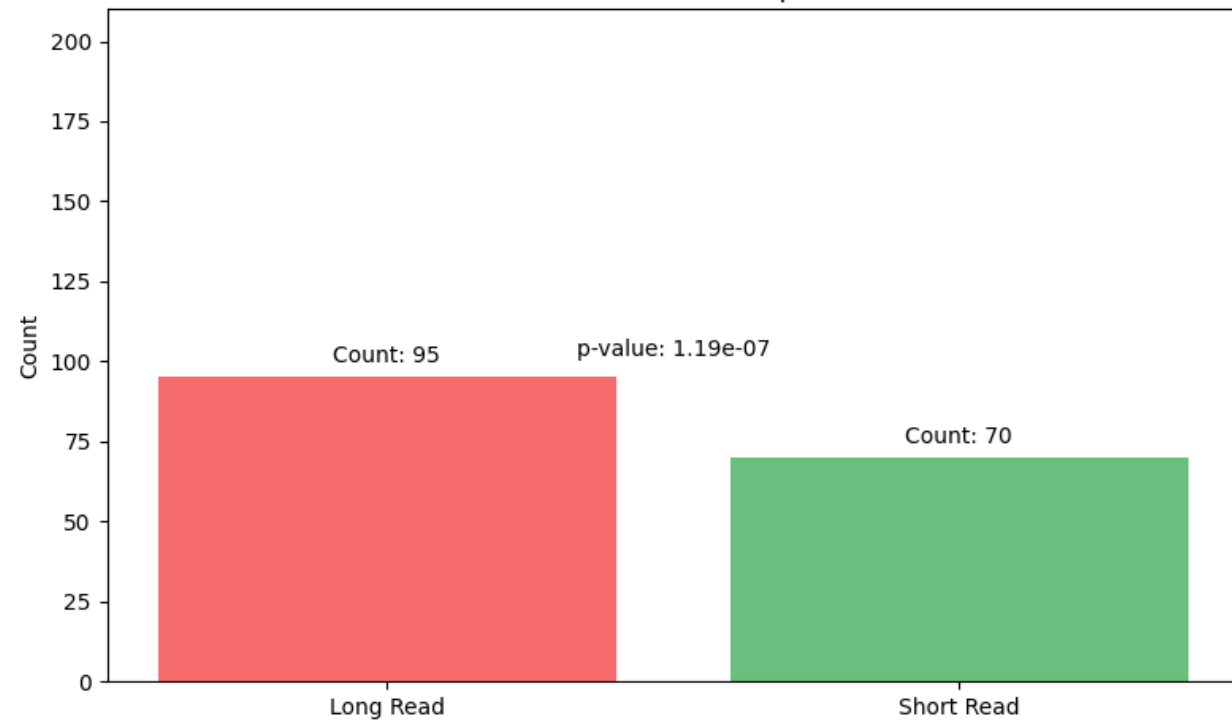
Venn Diagram of Species



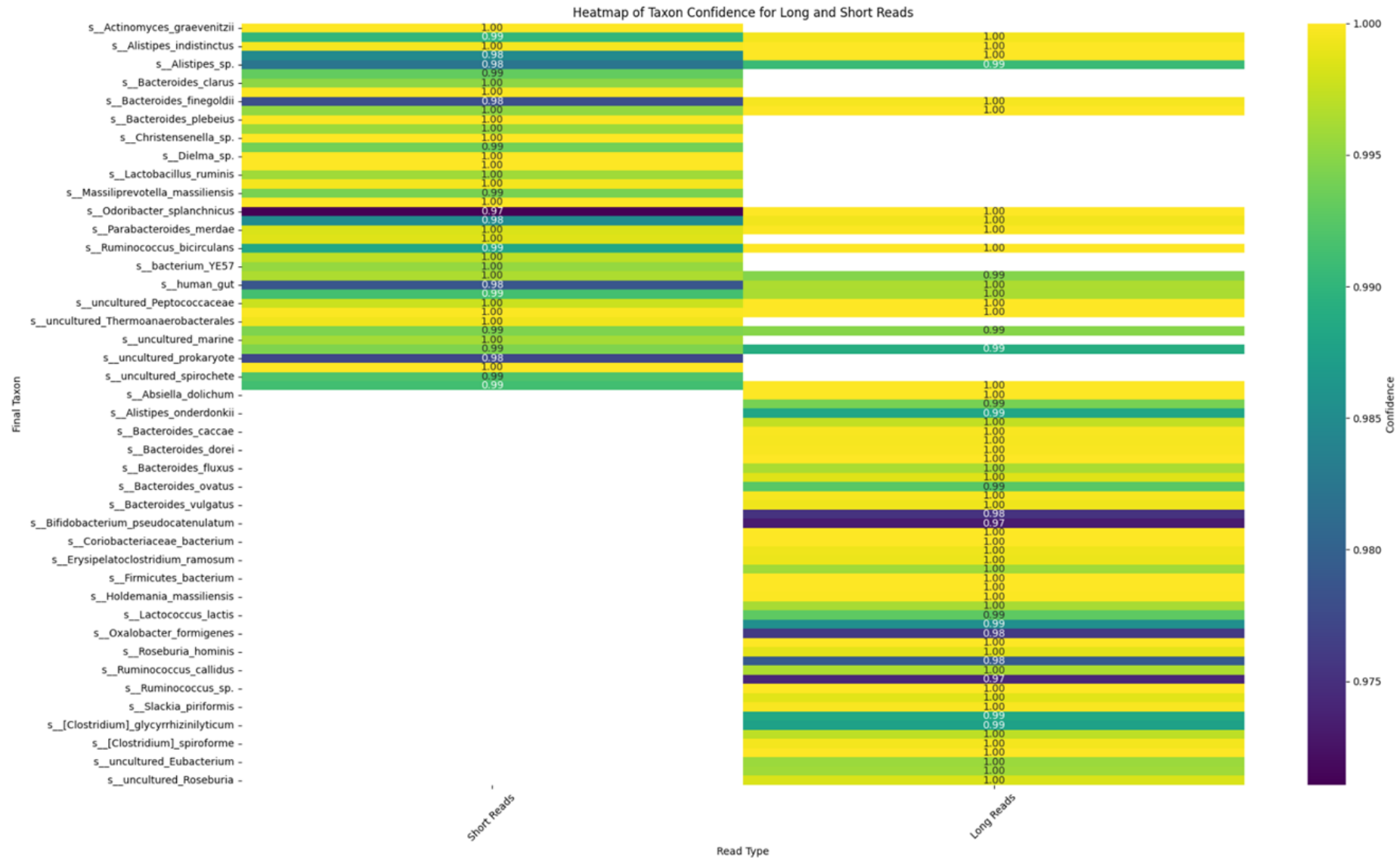
Number of Identified Genus



Number of Identified Species

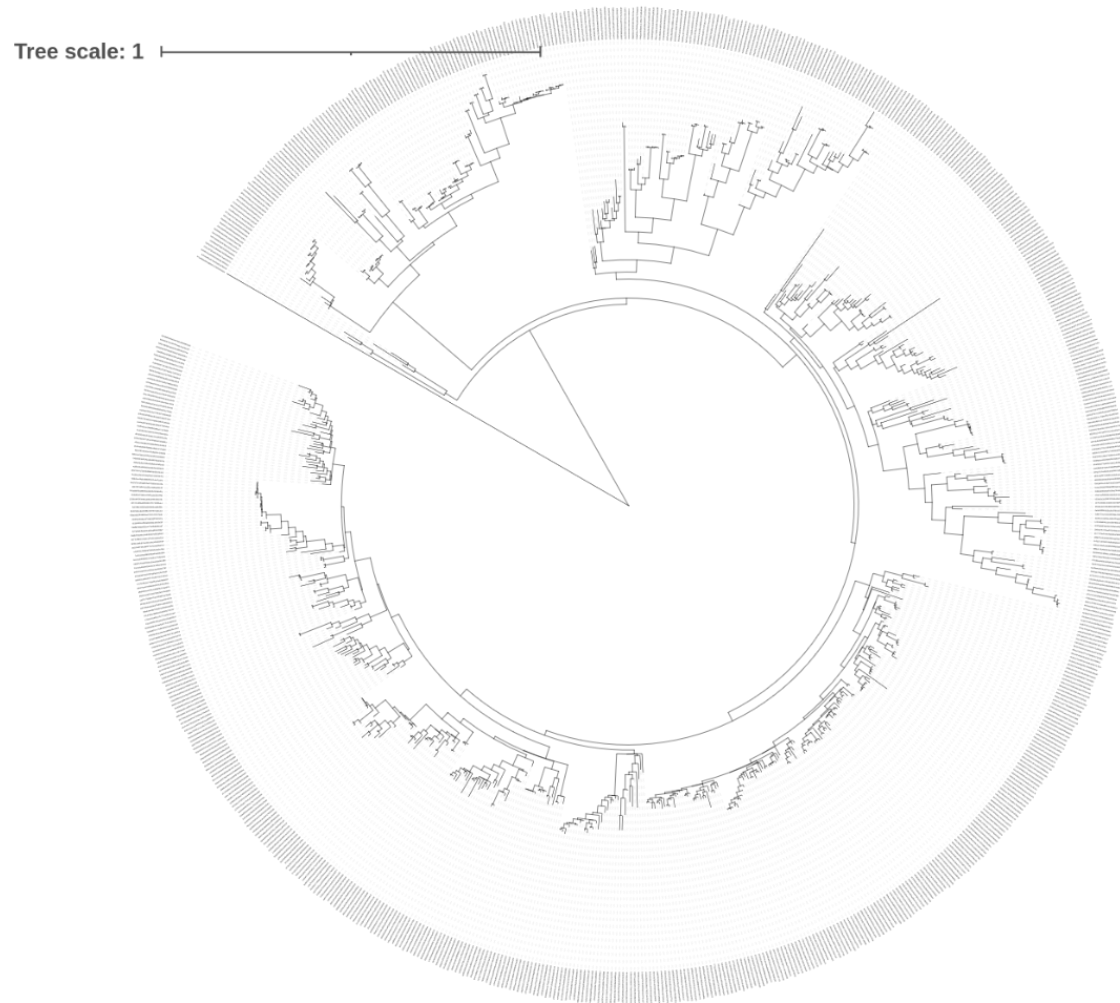


Confidence scores of final identified taxon



Phylogenetic trees

SHORT READS

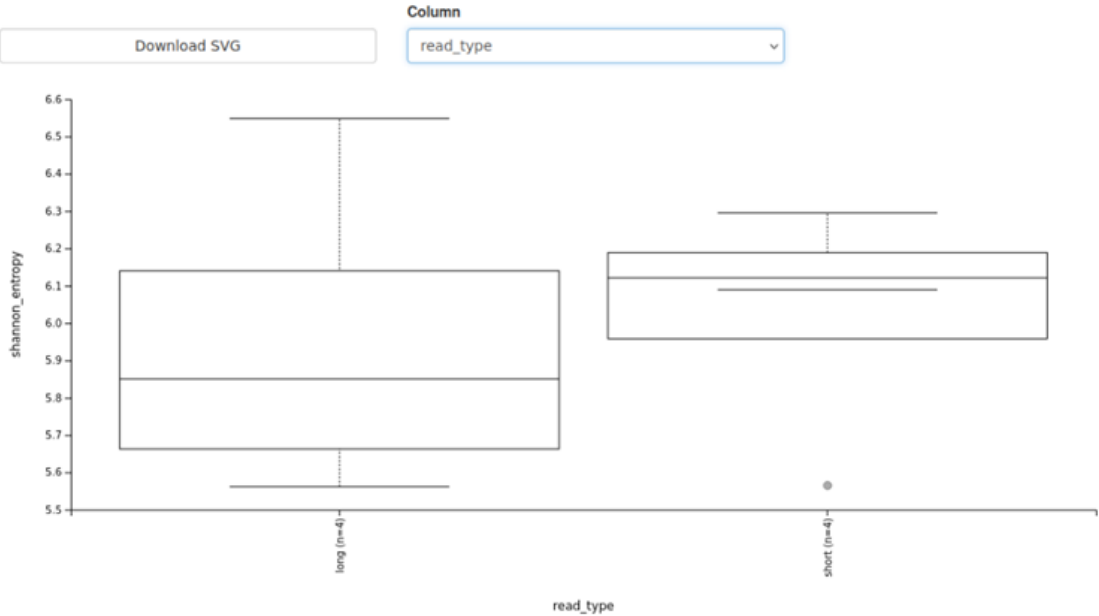


LONG READS



Alpha diversity

Alpha Diversity Boxplots



[Download raw data as TSV](#)

Kruskal-Wallis (all groups)

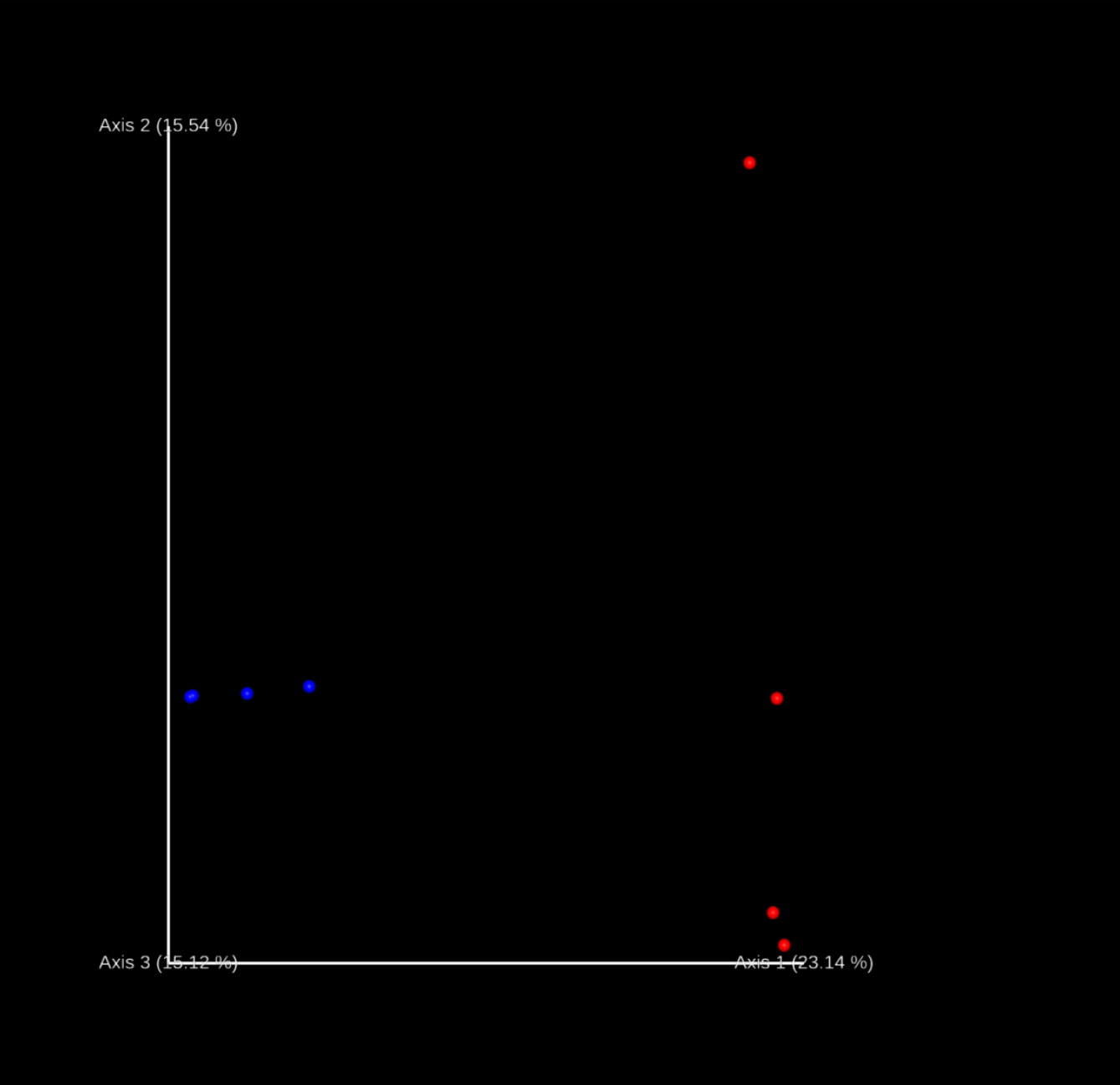
	Result
H	0.33333333333333215
p-value	0.5637028616507738

Kruskal-Wallis (pairwise)

[Download CSV](#)

		H	p-value	q-value
Group 1	Group 2			
long (n=4)	short (n=4)	0.333333	0.563703	0.563703

Beta diversity



- Long reads
- Short reads

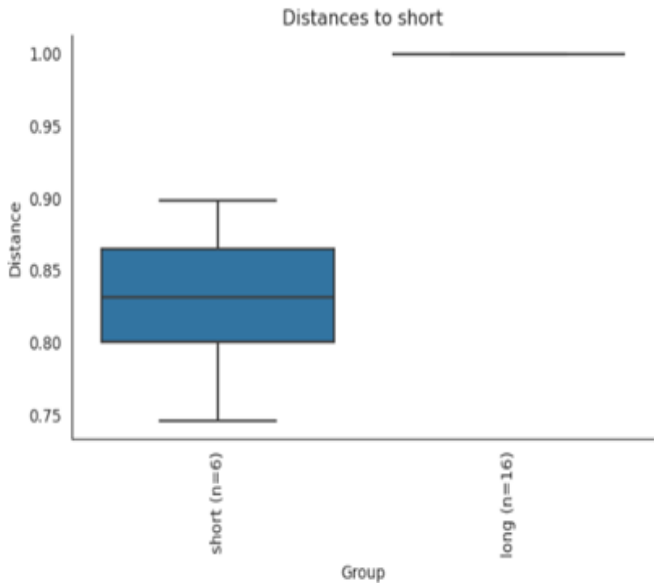
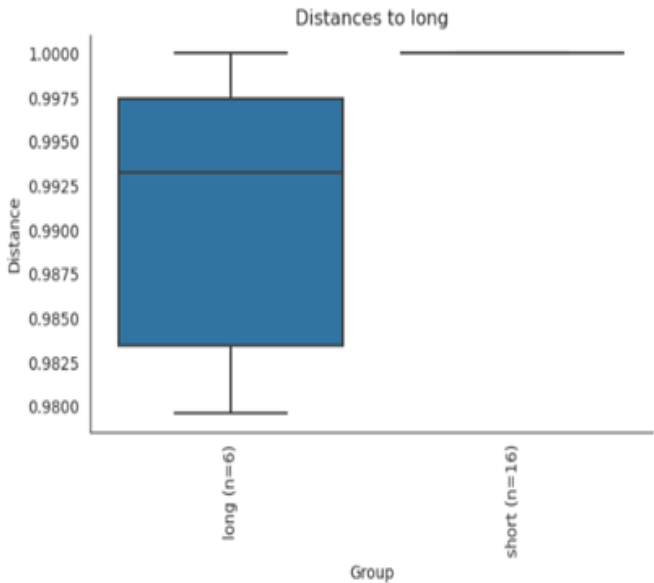
Beta diversity

Overview

PERMANOVA results	
method name	PERMANOVA
test statistic name	pseudo-F
sample size	8
number of groups	2
test statistic	1.786113
p-value	0.028
number of permutations	999

Group significance plots

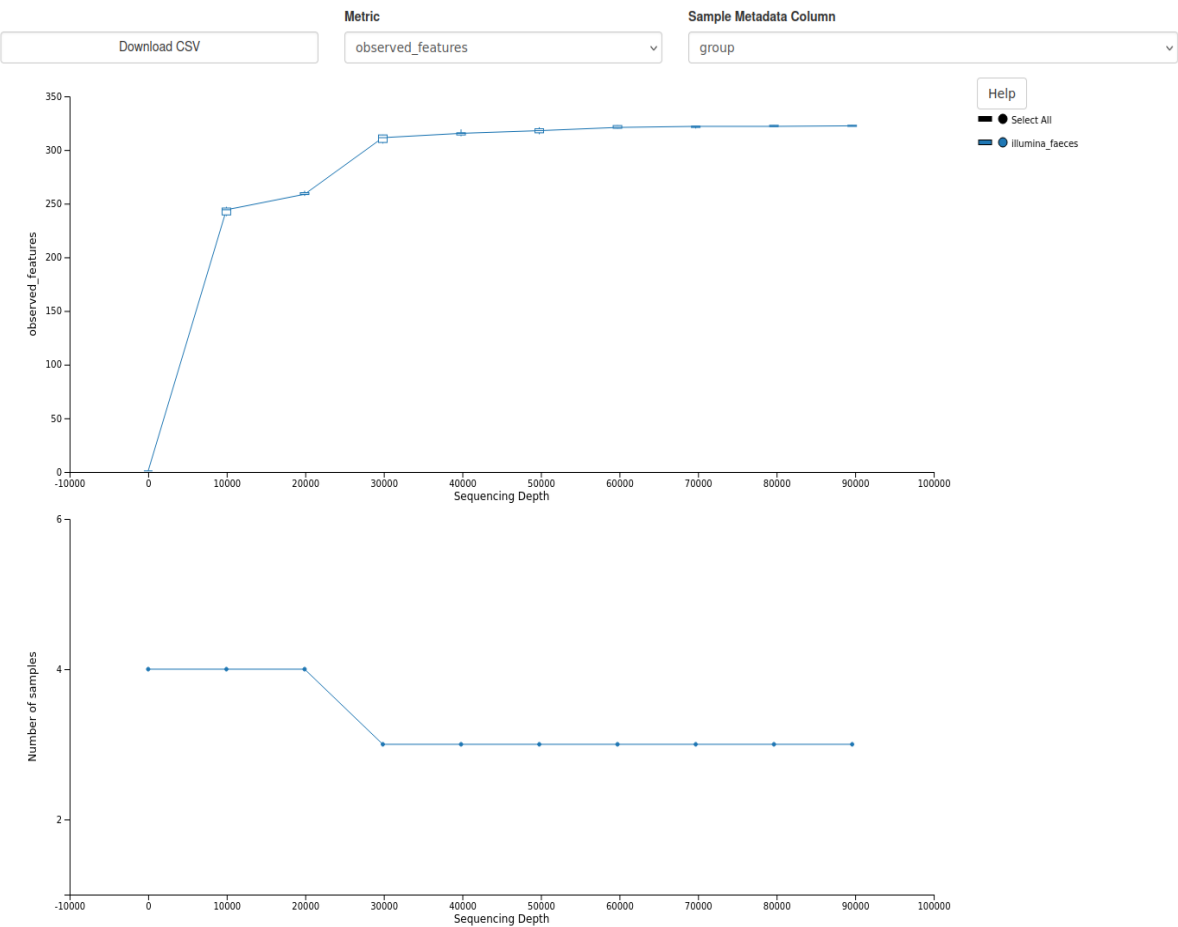
[Download raw data as TSV](#)



Rarefaction curves

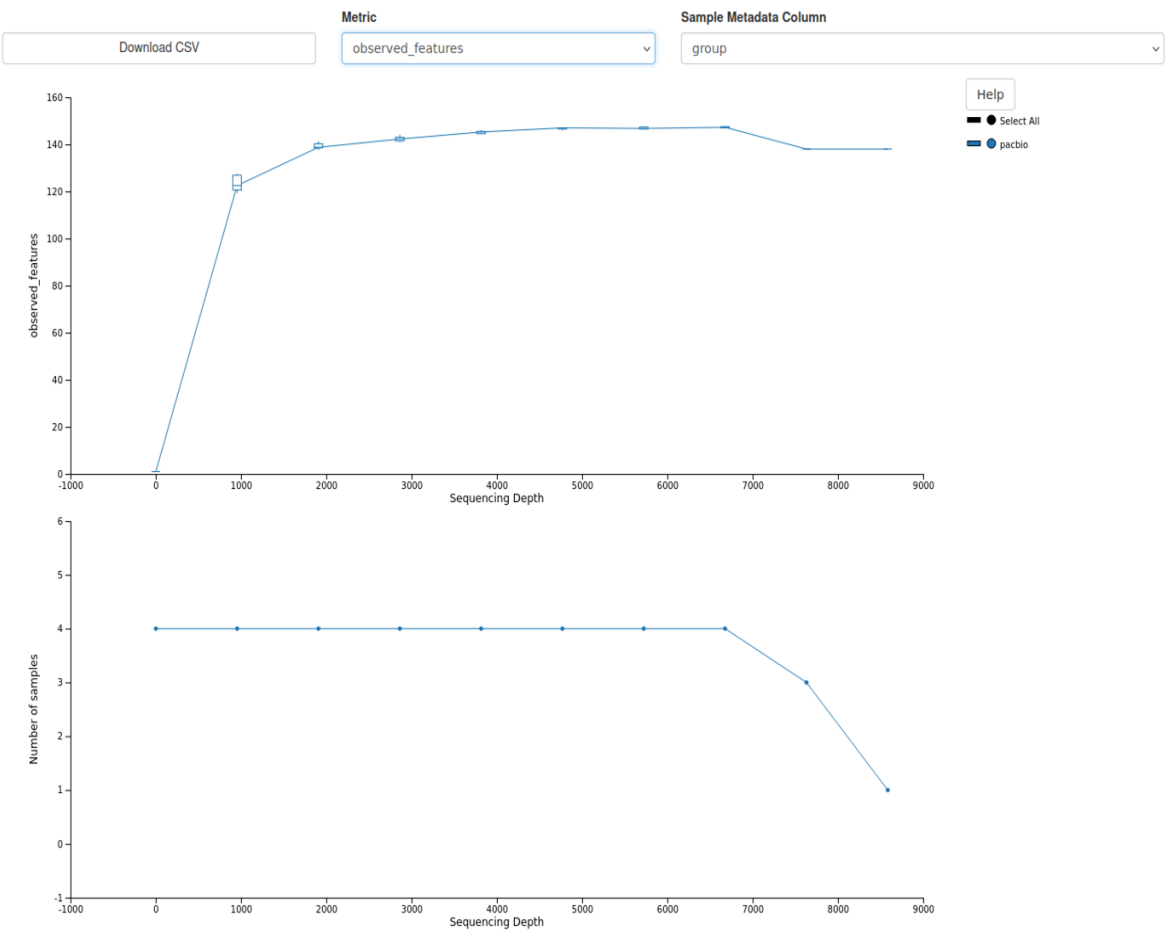
SHORT READS

Alpha rarefaction



LONG READS

Alpha rarefaction



Thanks for your attention!