



@nanopore

# **SHOTGUN METAGENOMICS WORKFLOW USING OXFORD NANOPORE PLATFORM**

Presenter: PhD Student Quynh Nhu Nguyen, Heidelberg University Hospital, Germany

Phu Quy Ho - Can Tho University, Vietnam

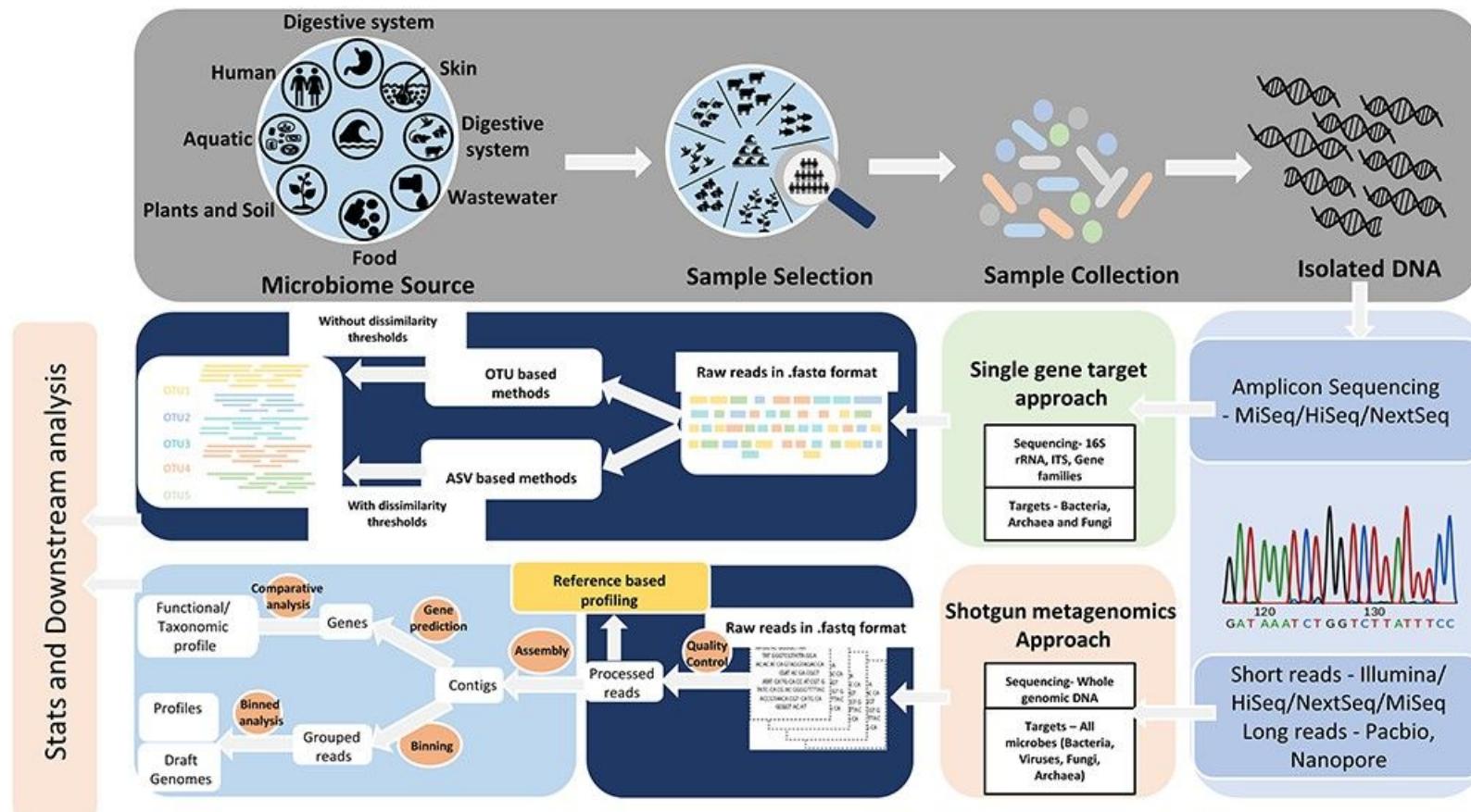
13.10.2024

# CONTENTS

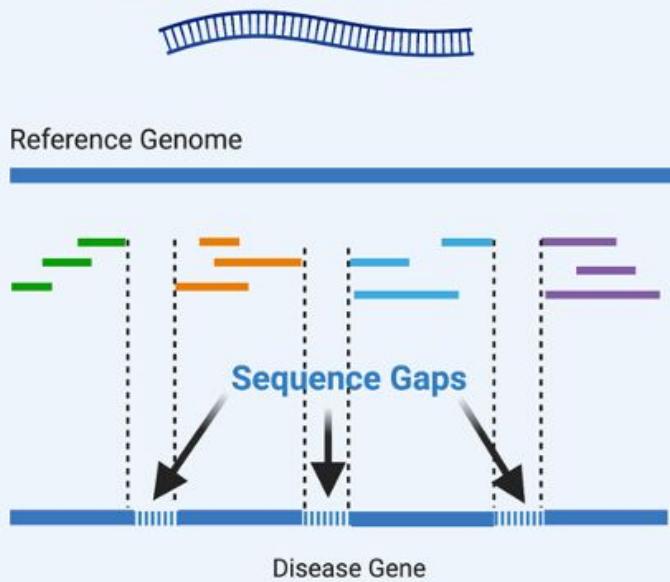
1. Introduction
2. Applications
3. Challenges
4. A Typical Article
5. Workflow
6. A Simulation Pipeline

# **INTRODUCTION**

# Targeted amplicon and metagenomic sequencing

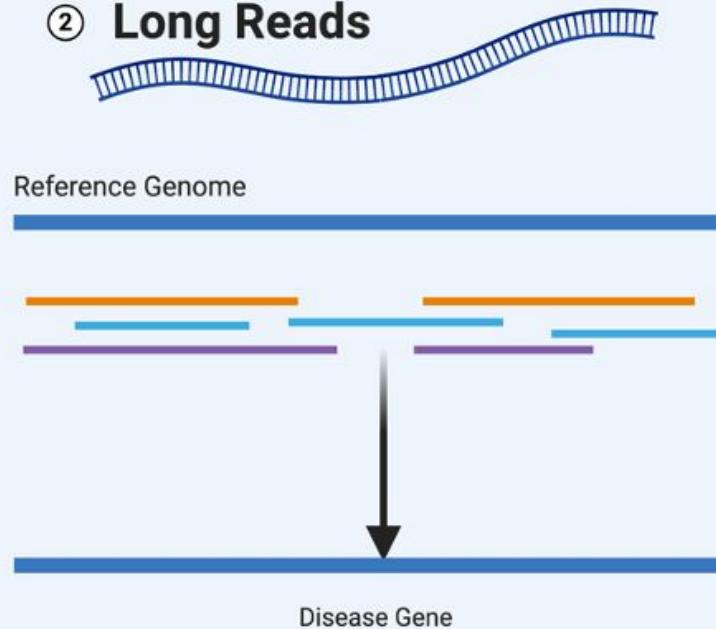


## ① Short Reads



Missing sequence data leads to gaps  
in genome coverage and limits variant detection

## ② Long Reads



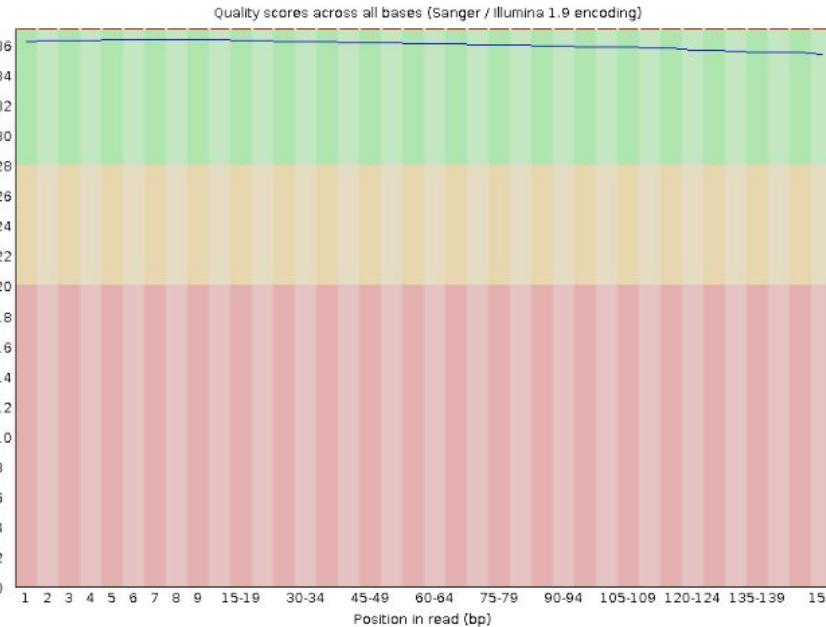
Long reads map uniquely and  
span large variants providing  
comprehensive variant detection

## Basic Statistics

Measure	Value
Filename	TD_92_run0_raw_1.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	19513006
Total Bases	2.9 Gbp
Sequences flagged as poor quality	0
Sequence length	19-150
%GC	46

## Short reads

### ✓ Per base sequence quality



## Quality control after trimming

## Basic Statistics

Measure	Value
Filename	TD_92_long.unmapped.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	584180
Total Bases	2.8 Gbp
Sequences flagged as poor quality	0
Sequence length	1000-45642
%GC	43

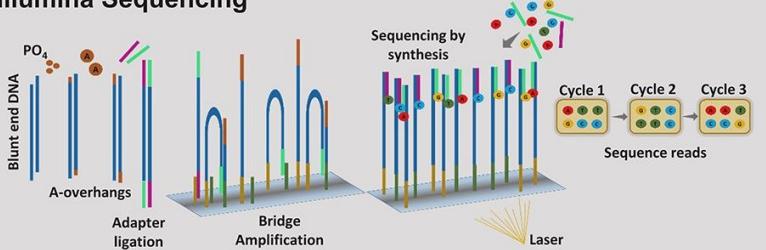
## Long reads

### ✗ Per base sequence quality

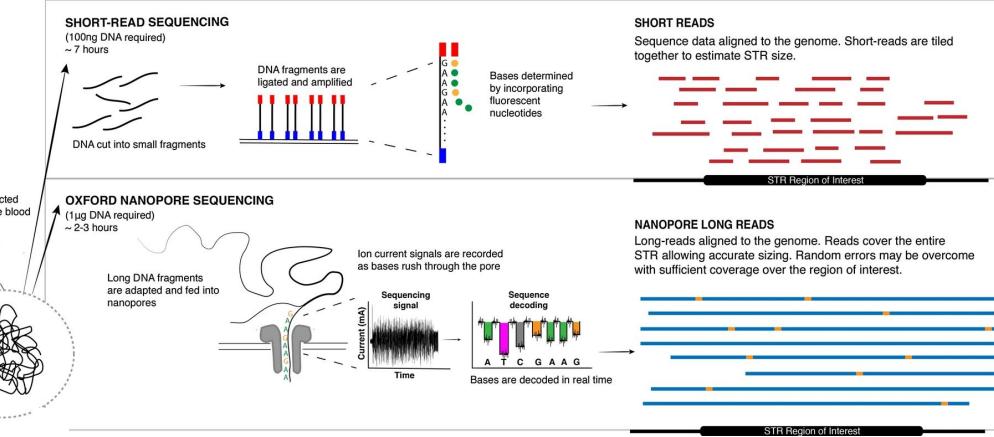
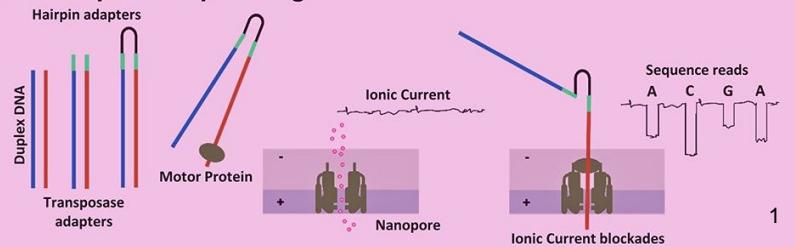


# Illumina Sequencing and Nanopore Sequencing

## A. Illumina Sequencing



## C. Nanopore Sequencing



2

1.Richa Bharti, Dominik G Grimm, Current challenges and best-practice protocols for microbiome analysis, *Briefings in Bioinformatics*, Volume 22, Issue 1, January 2021, Pages 178–193, <https://doi.org/10.1093/bib/bbz155>

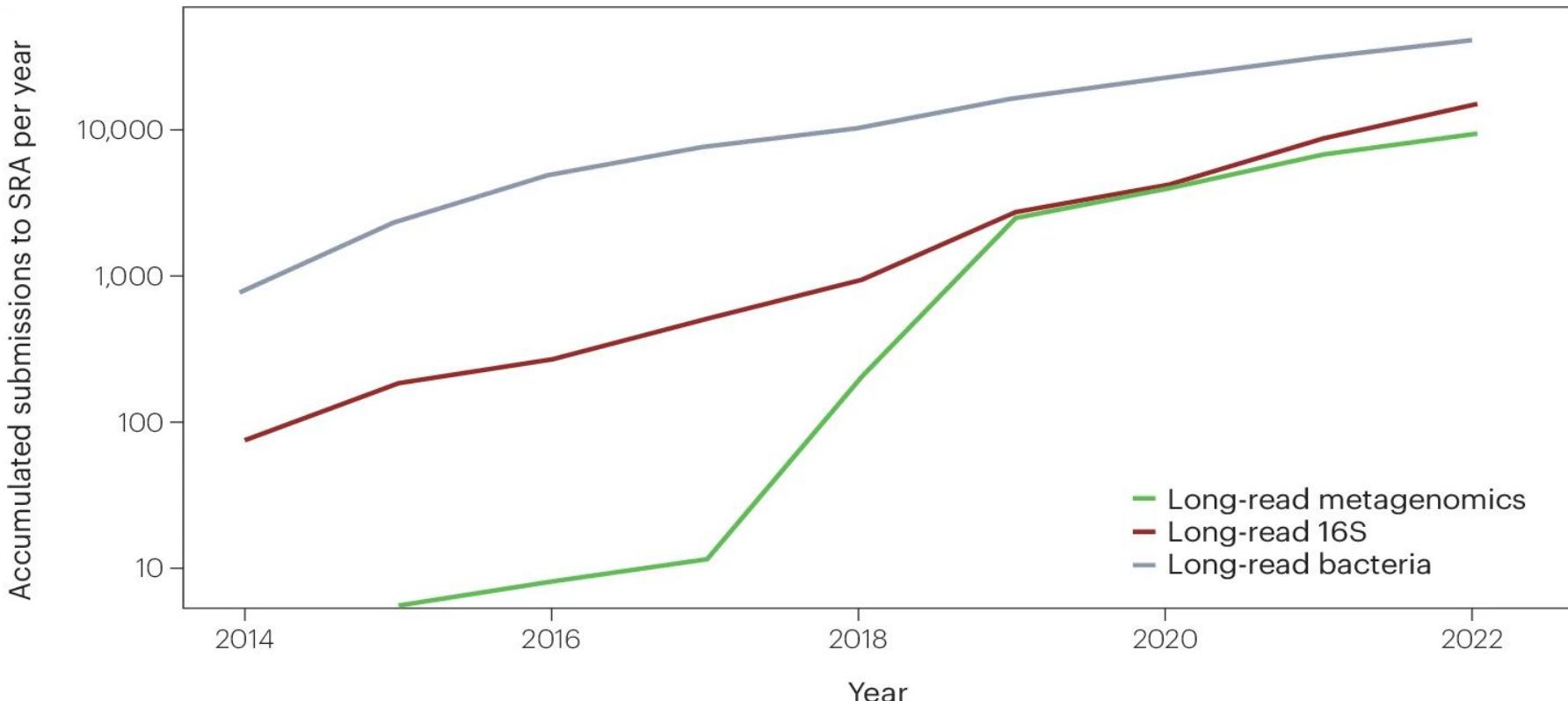
2.Chintalaphani, S.R., Pineda, S.S., Deveson, I.W. et al. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *acta neuropathol commun* 9, 98 (2021). <https://doi.org/10.1186/s40478-021-01201-x>

7

# Comparison between short-read and long-read technologies

Platform	MiSeq	NovaSeq 6000	Sequel II	Revio	Flongle	MinION	GridION	PromethION
<b>Company</b>	Illumina	Illumina	PacBio	PacBio	Nanopore	Nanopore	Nanopore	Nanopore
<b>Read length (average)<sup>a</sup></b>	Up to 2 × 300 bp	Up to 2 × 250 bp <sup>b</sup>	~13.5–20 kb <sup>52</sup>	15–18 kb	20 kb	20 kb	20 kb	20 kb
<b>Yield per cell (Gb)</b>	Up to 15	~350	25 (HiFi)	90	1–1.5	15–20	15–20	~120
<b>Runtime (h)</b>	4–56	13–44	30	24	16–24	72	72	72
<b>Read accuracy (Q20+)</b>	99.75% <sup>128</sup>	99.75% <sup>128</sup>	99.18–99.8% <sup>52</sup>	>99.5% <sup>129</sup>	97–99% <sup>12</sup>	97–99% <sup>12</sup>	97–99% <sup>12</sup>	97–99% <sup>12</sup>
<b>Can perform Direct RNA-seq</b>	No	No	No	No	Yes	Yes	Yes	Yes
<b>DNA input needed</b>	1–500 ng <sup>18</sup>	1–500 ng <sup>18</sup>	150 ng–1 µg <sup>7</sup> <sup>c</sup>	150 ng–1 µg <sup>7</sup> <sup>c</sup>	150 ng–1 µg <sup>7</sup>	150 ng–1 µg <sup>7</sup>	150 ng–1 µg <sup>7</sup>	150 ng–1 µg <sup>7</sup>
<b>Estimated sequencing costs per Gb<sup>d</sup></b>	US\$178–1,705 <sup>130,131</sup>	US\$3.95 <sup>131</sup>	US\$30–43 <sup>131,132</sup>	US\$8–11 <sup>132</sup>	US\$118–437 <sup>131,132</sup>	US\$21–51 <sup>131,132</sup>	US\$29–51 <sup>132</sup>	US\$6–12 <sup>132</sup>

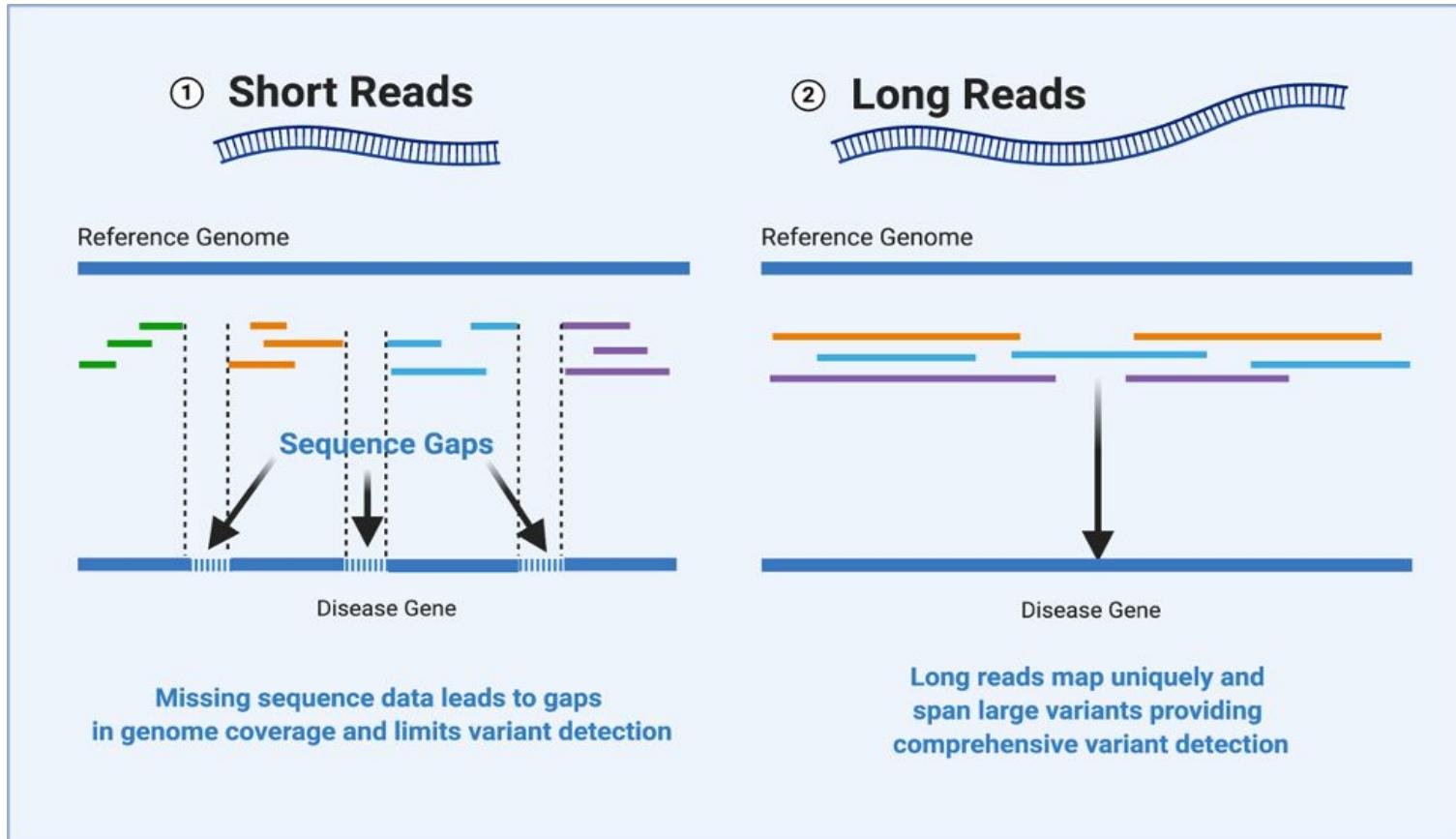
# The growth of long-read-related submission to the Sequence Read Archive (SRA) in recent years



# APPLICATION

1. De novo genome assembly and Metagenomics
2. Structural variant detection
3. Epigenetics studies
4. Portable and real-time sequencing device

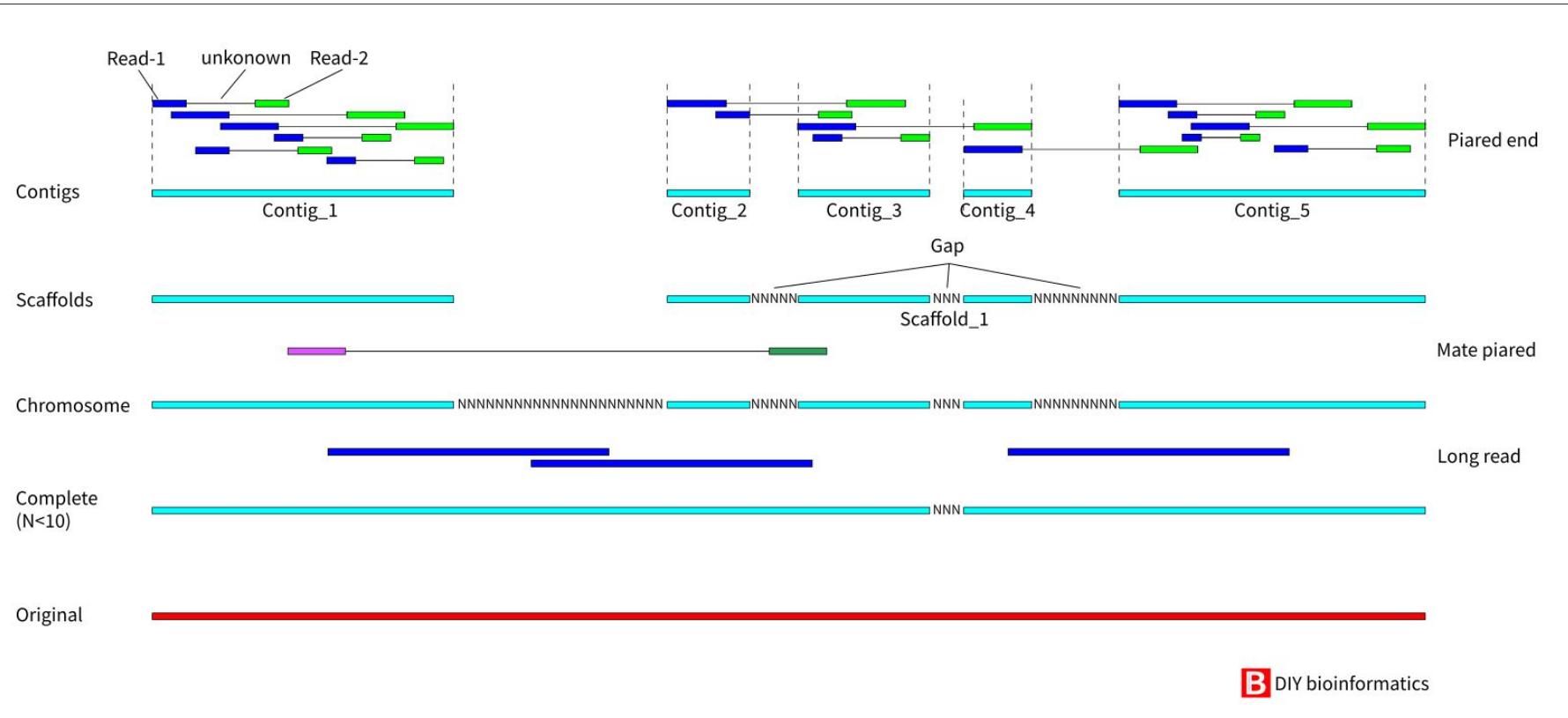
# 1. DE NOVO GENOME ASSEMBLY AND METAGENOMICS

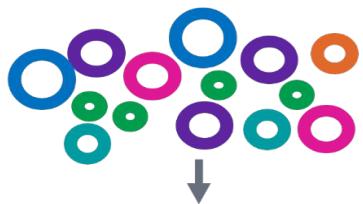


<https://hudsonalpha.org/wp-content/uploads/2021/02/short-read-2.png>

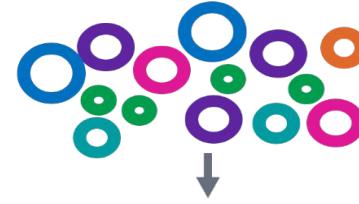
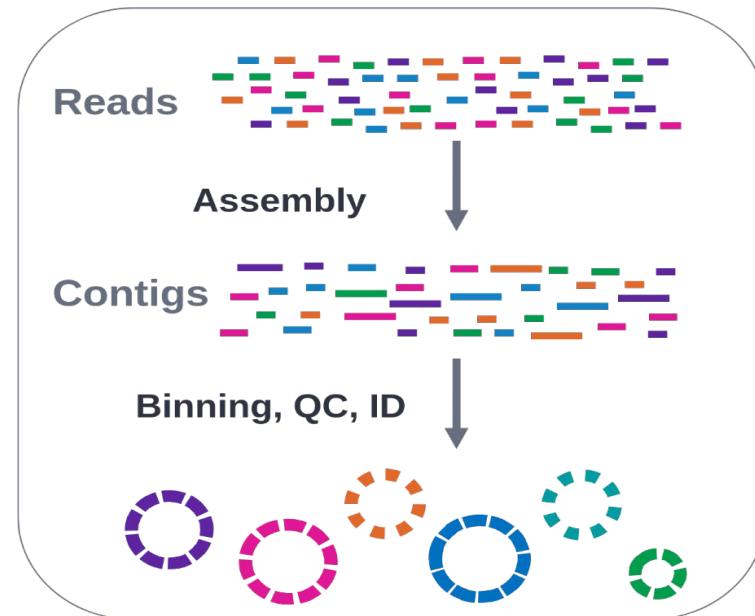
<https://www.youtube.com/watch?v=LmUQ4f2f5nc>

# Contigs, Scaffolds

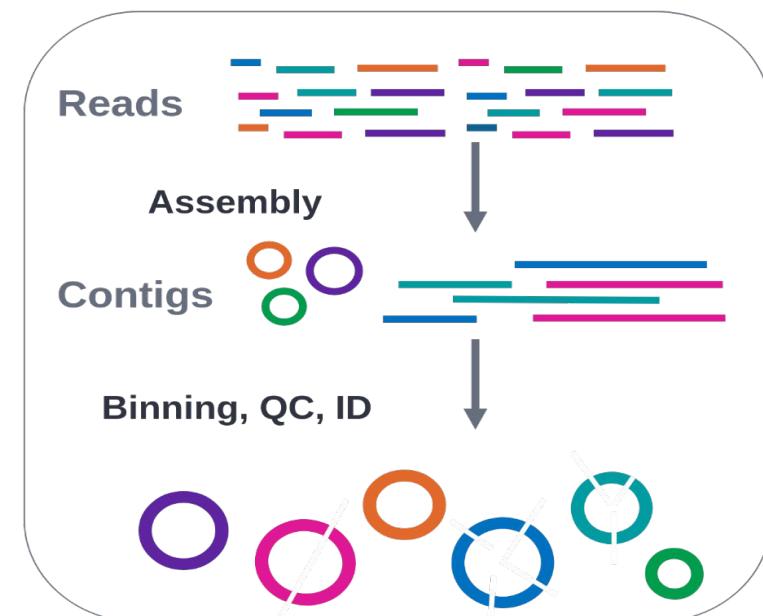




### Short-read metagenomics



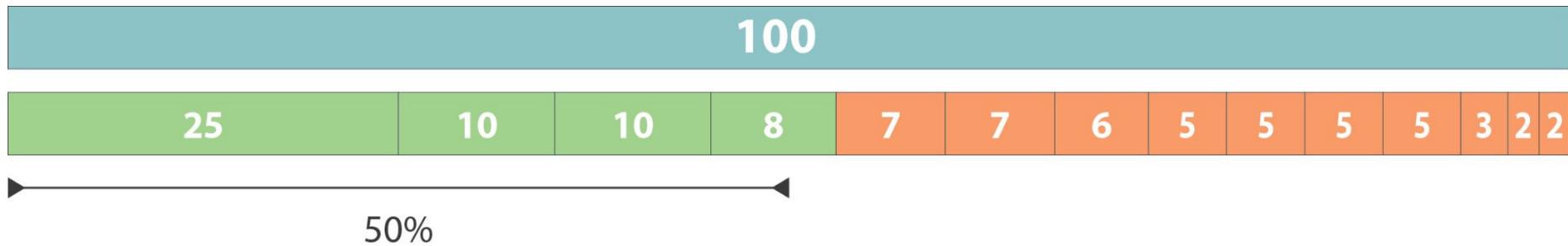
### Long-read metagenomics



Draft-quality MAGs

High-quality MAGs

# N50 and L50



Genome size = 100

Sequence sorted by size list L = (25, 10, 10, 8 , 7, 7 , 6 , 5, 5, 5, 3, 2, 2 ) = 100

50% of the total length is contained within sequences of at least 8bp:  $25 + 10 + 10 + 8 \geq 50$

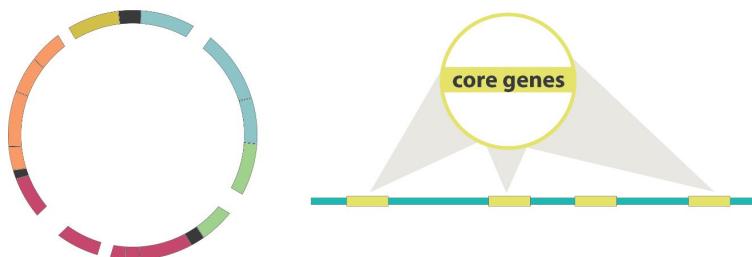
**N50 = 8 and L50 = 4**

# Metrics considered for assessing the performance of the different sequencing strategies

Metric	Description
Assembly length	The total number of nucleotides of the unfiltered (no minimum contig size threshold applied) metagenomic assembly.
Assembly contiguity	The N50 metric of assembly contiguity, defined as the length of the shortest contig for which longer and equal-length contigs cover at least 50% of the assembly.
Assembly coverage	Percentage of short-read sequences (a random subset of 20 Gbp) mapping to the assembly.
MAG number	Number of metagenome-assembled genomes with >70% CheckM completeness and <10% redundancy recovered from the sequencing data.
MAG contiguity	Average number of contigs per MAG.
MAG completeness	Average completeness of the MAGs in terms of CheckM assessment based on the presence of single-copy core genes.
MAG contamination	Inverse to the average contamination of the MAGs in terms of CheckM assessment based on the redundancy of single-copy core genes.

1

## Contiguity      Completeness      Correctness

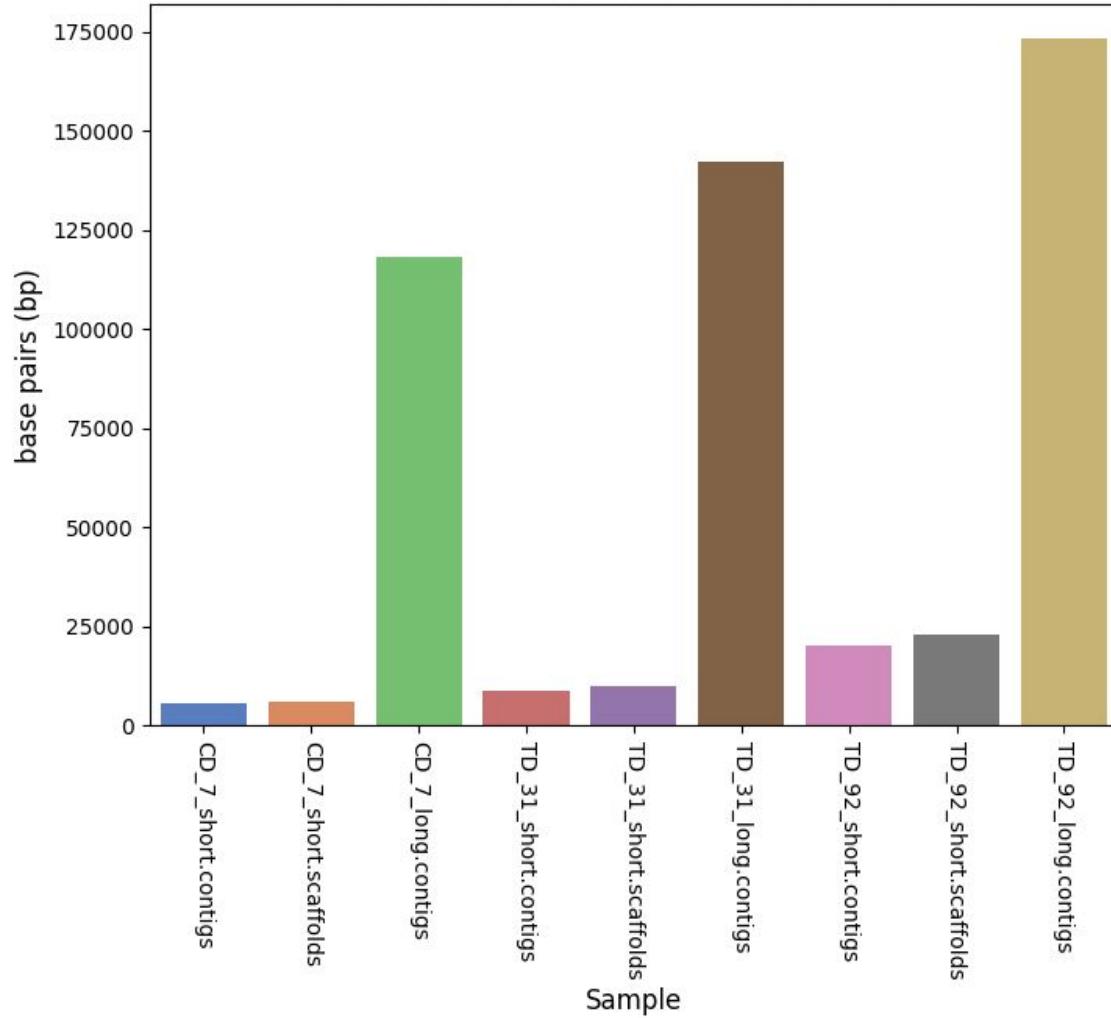


2

Real  
T G C A G T A A C G A T T  
T G C A A T A A C T A T T  
Assembly

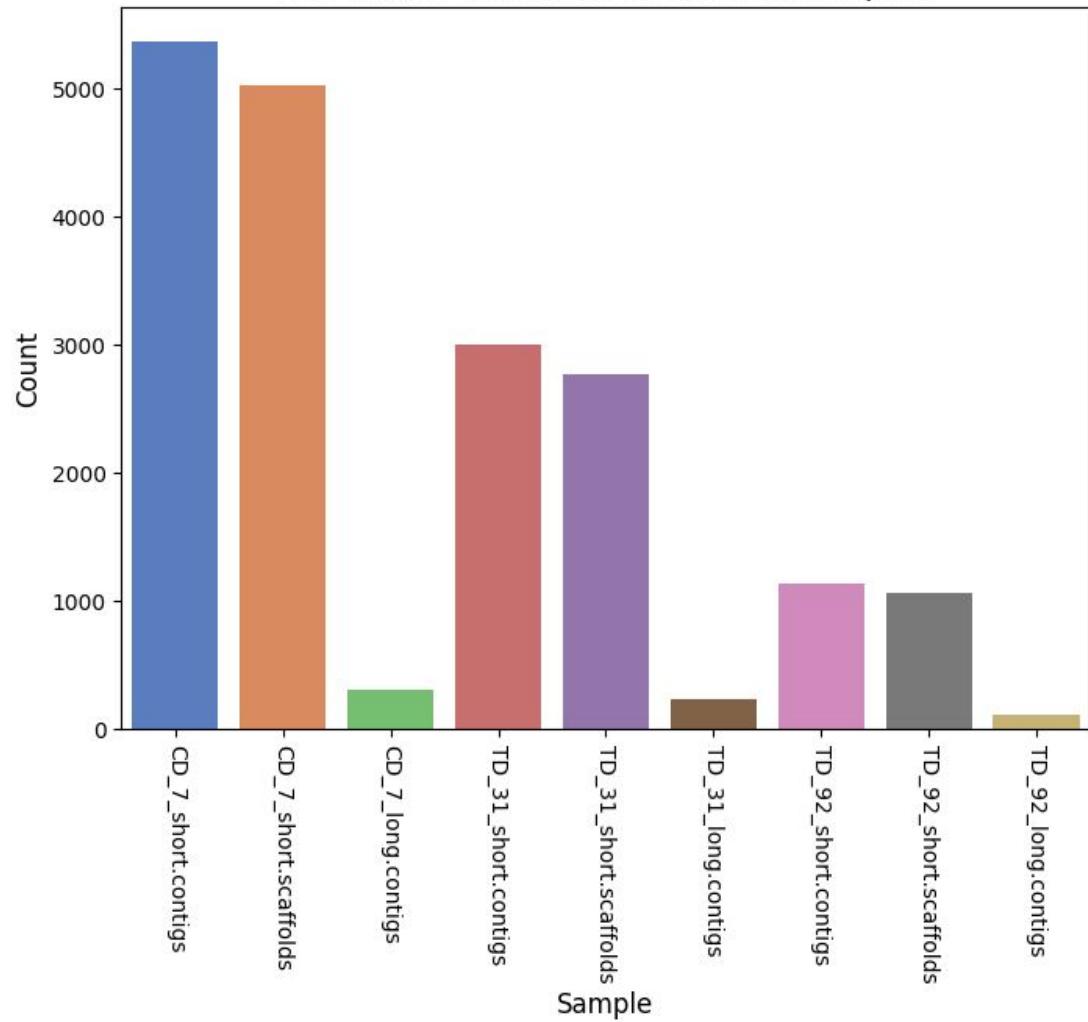
15

## N50 value distribution across samples



Long-read sequencing provides a **higher N50** value, which indicates longer contiguous sequences in genome assemblies.

L50 value distribution across samples



Long-read sequencing offers a **lower L50** value, meaning fewer contigs are required to cover 50% of the genome, contributing to more continuous assemblies.

# Assembly visualization

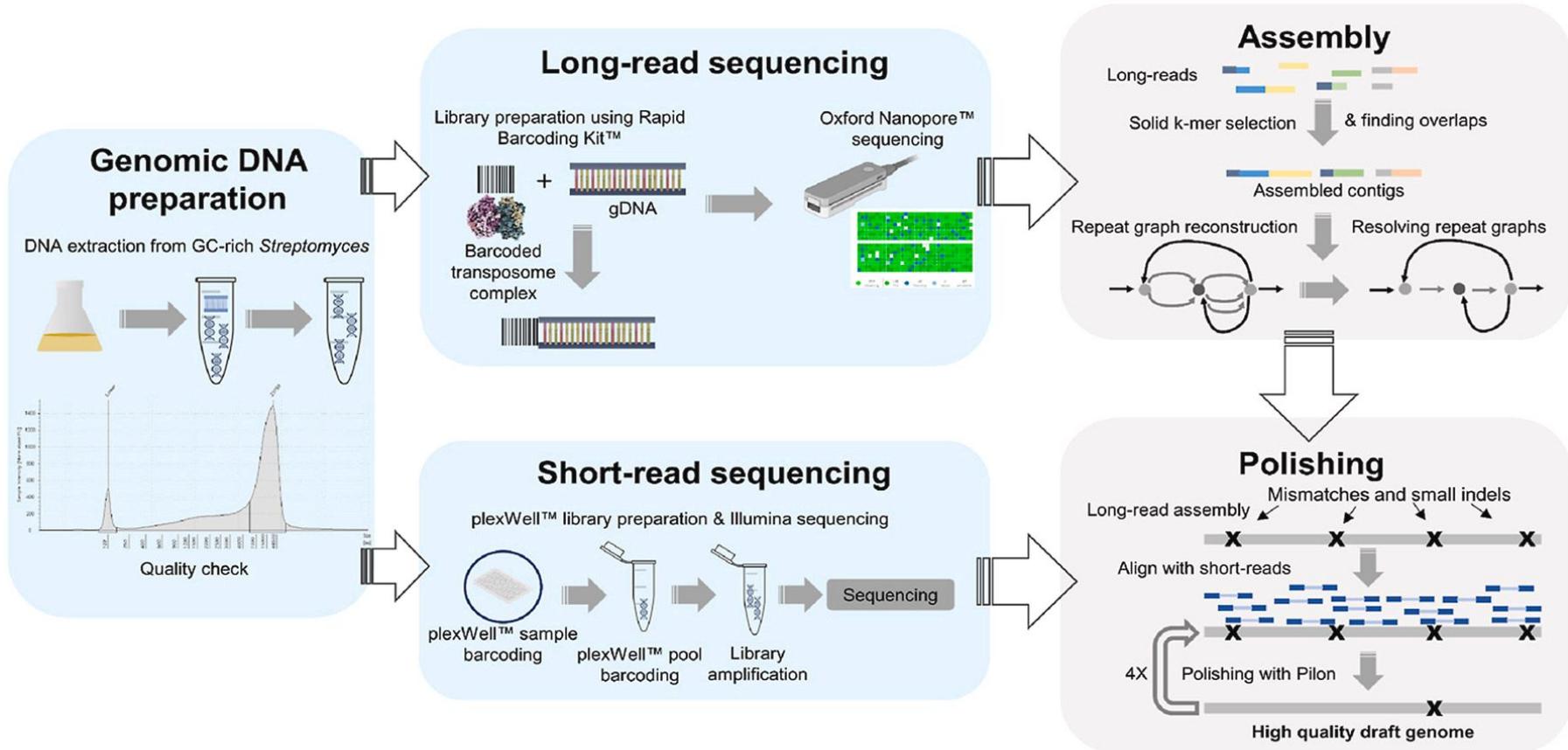
short reads TD\_92



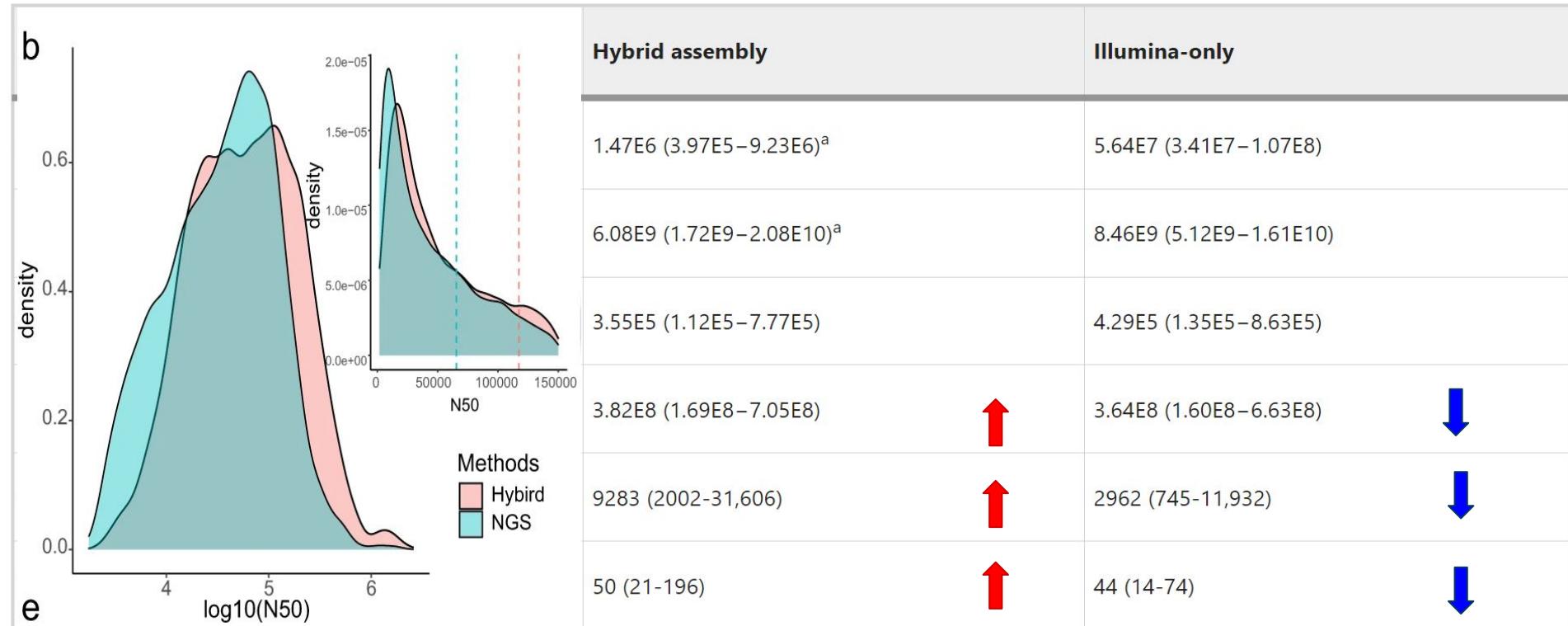
long reads TD\_92

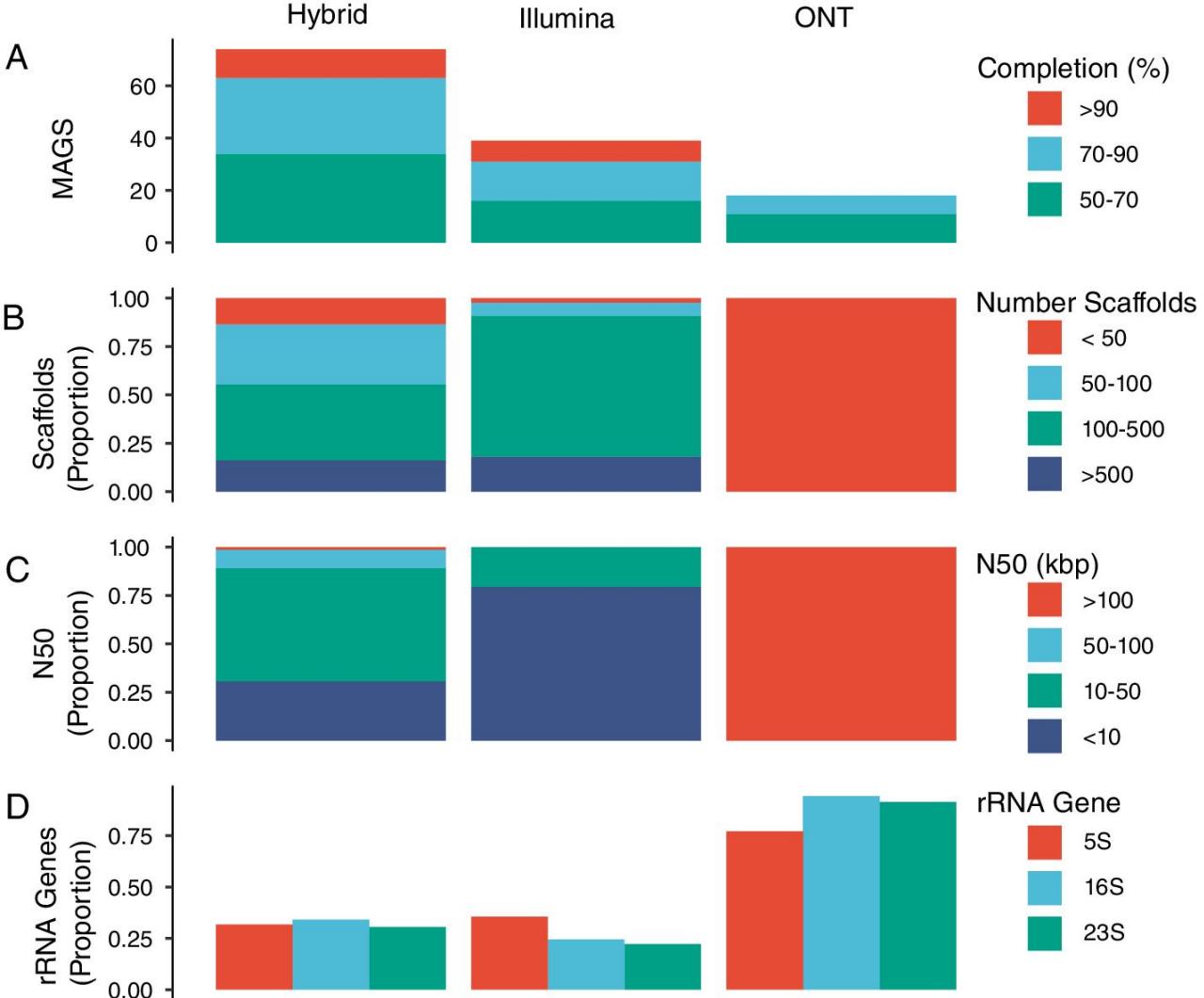


# Hybrid sequencing



# Hybrid sequencing improves the quality of metagenome assembly

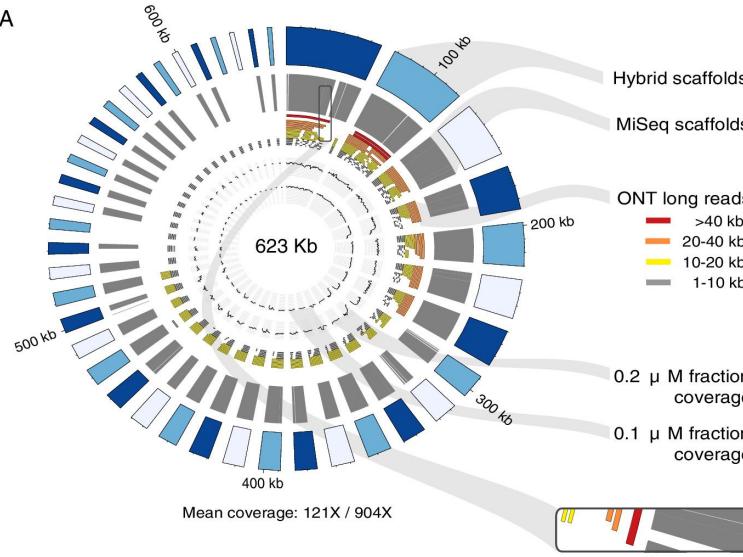




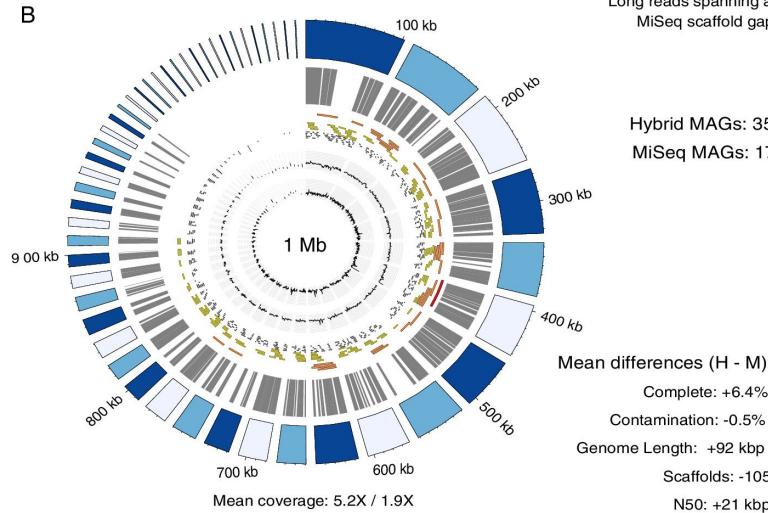
Hybrid sequencing improves the quality of metagenome assembly

Overholt WA, Hölzer M, Geesink P, Diezel C, Marz M, Küsel K. Inclusion of Oxford Nanopore long reads improves all microbial and viral metagenome-assembled genomes from a complex aquifer system. *Environ Microbiol*. 2020;22(9):4000-4013.  
doi:10.1111/1462-2920.15186

A



B



Genome circos plots for the most (A) and least (B) covered Patescibacteria MAGs retrieved by the hybrid and Illumina-only assembly methods

Overholt WA, Hölzer M, Geesink P, Diezel C, Marz M, Küsel K. Inclusion of Oxford Nanopore long reads improves all microbial and viral metagenome-assembled genomes from a complex aquifer system. *Environ Microbiol*. 2020;22(9):4000-4013. doi:10.1111/1462-2920.15186

## **2. STRUCTURAL VARIANT DETECTION**

**Structural variants (SVs)** are large-scale events (>50 bp) where entire sections of genetic material have changed.

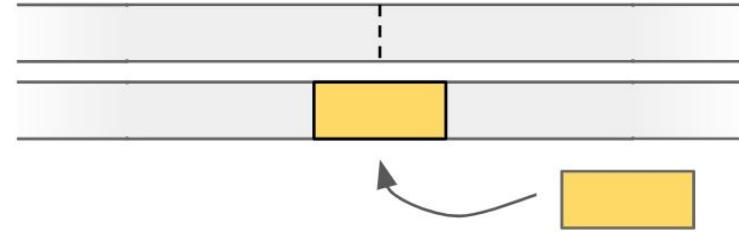
## Sequence Variants

### SNV (Single Nucleotide Variant)

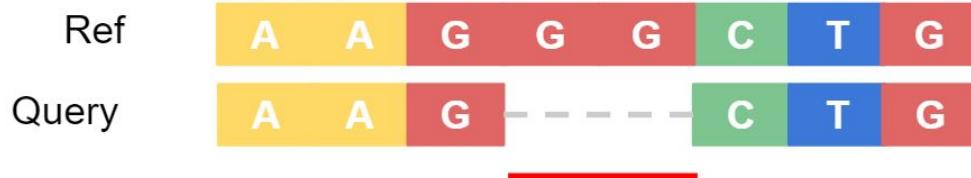


## Structural Variants

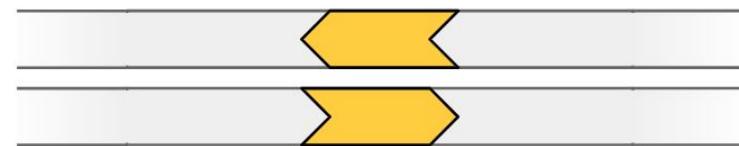
### Insertion



### INDEL (Insertion or Deletion)



### Inversion



SVs are difficult to detect due to their association with repeat elements, but **long reads improve mappability, allowing for more accurate detection.**

Tandem Repeat

Reference



Isolate



Short read origin



Equal best alignment locations



Long read origin



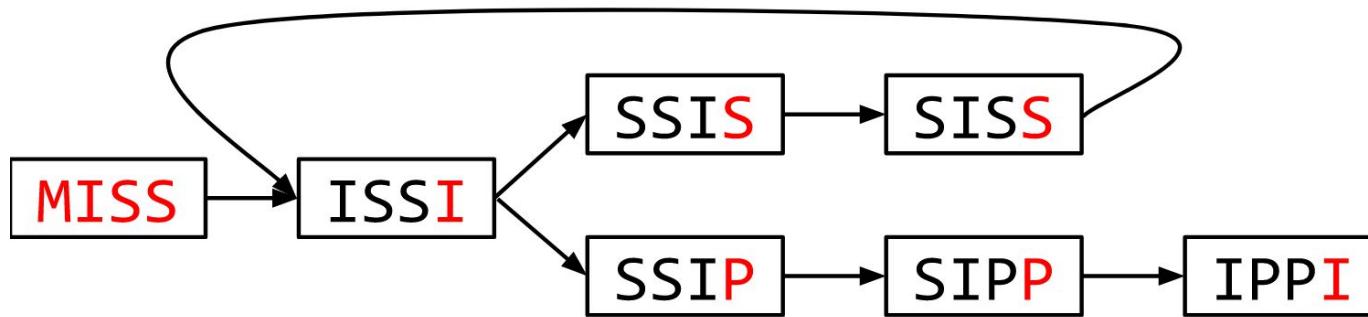
Equal best alignment locations



# Reference: MISSISSIPPI

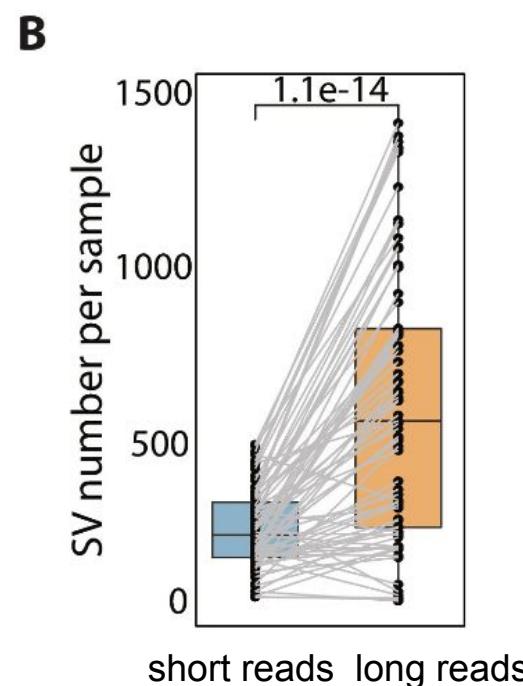
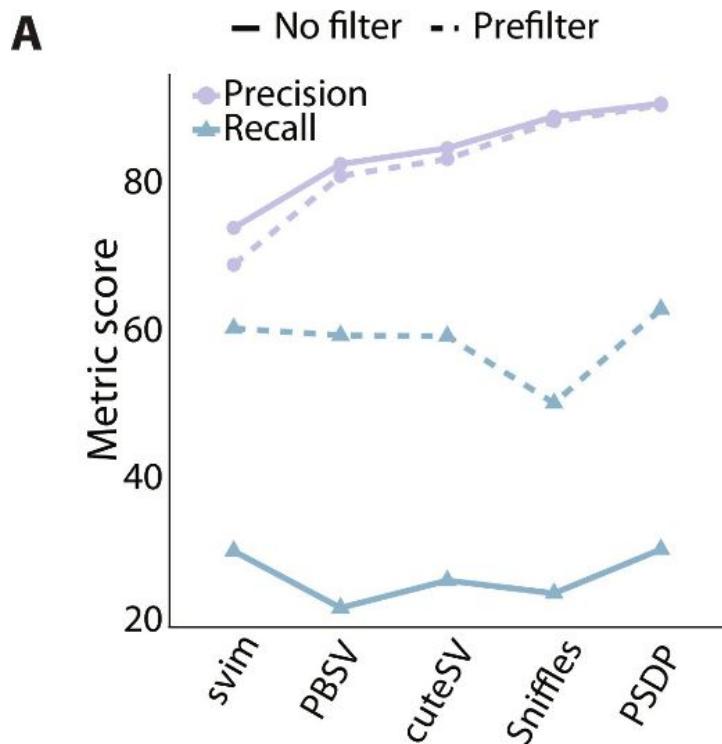
Reads:

MISS ISSI SSIS SISS SSIP SIPP IPPI

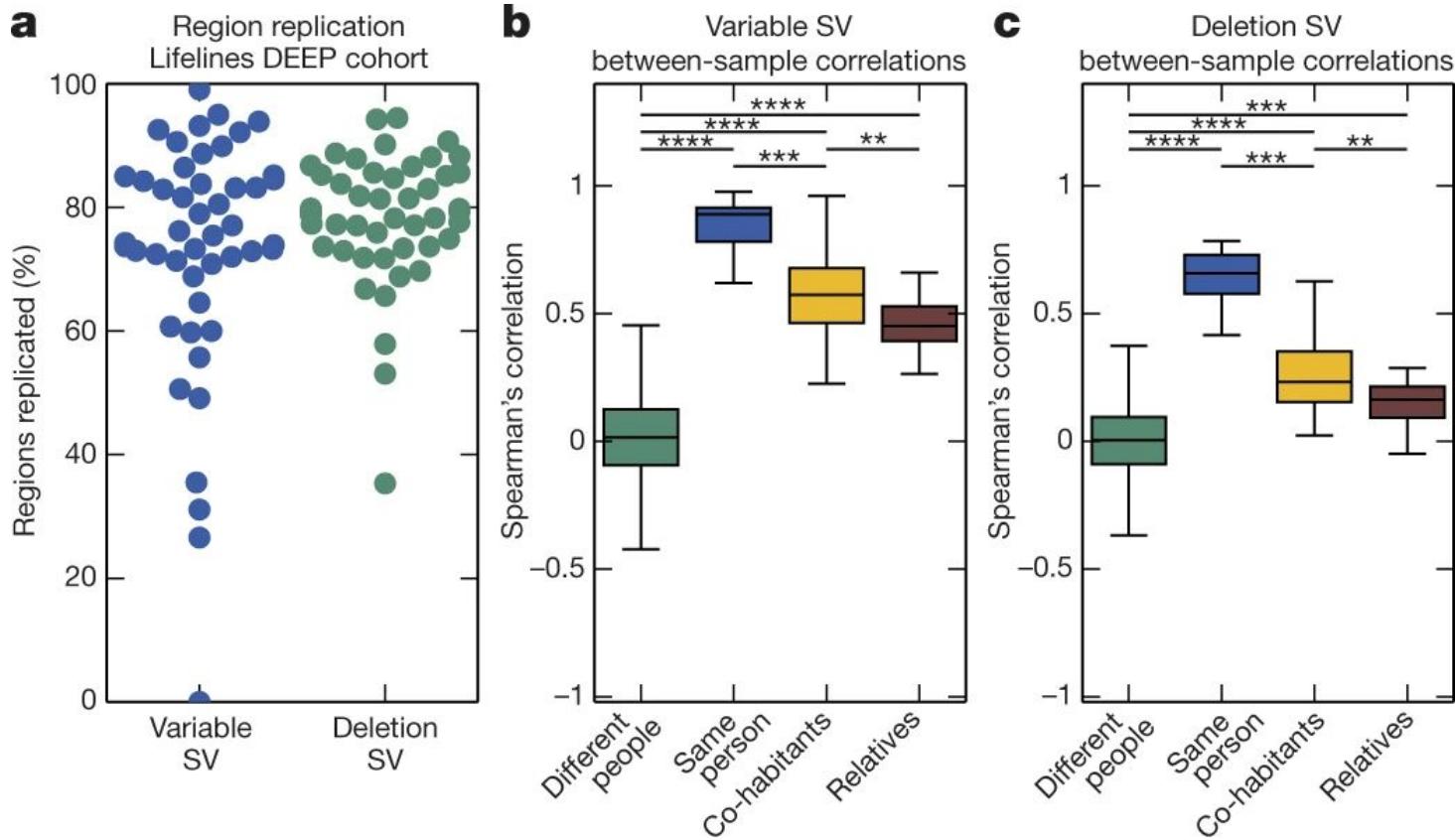


MISSISSIPPI or MISSISSISSISSIPPI or ...

Long reads enhances the accuracy and resolution in identifying large insertions, deletions, inversions, and translocations.

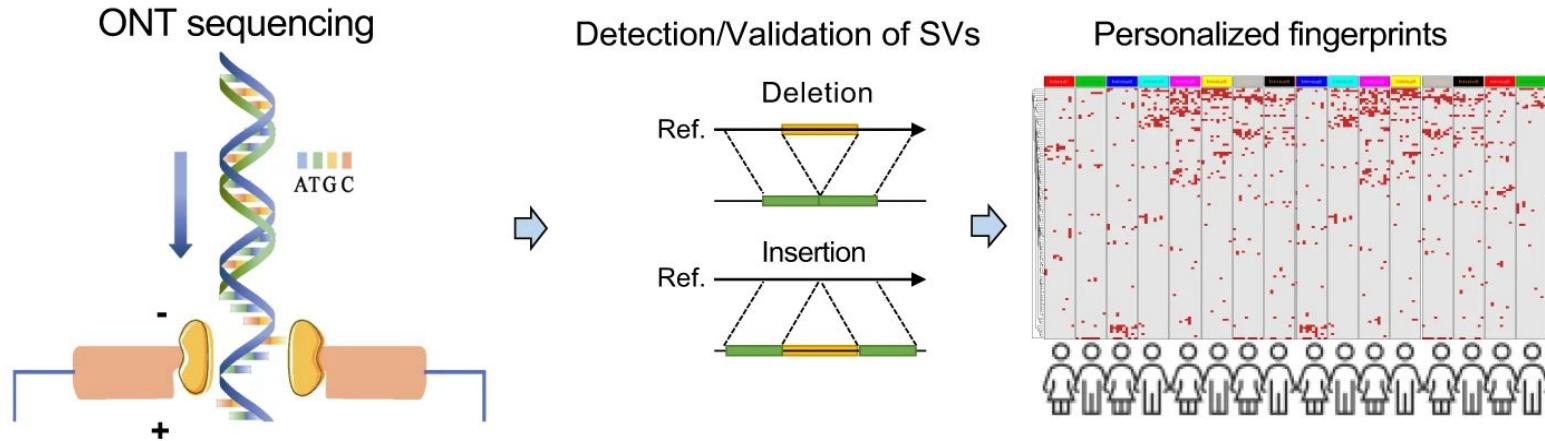


# SVs are prevalent across distinct populations



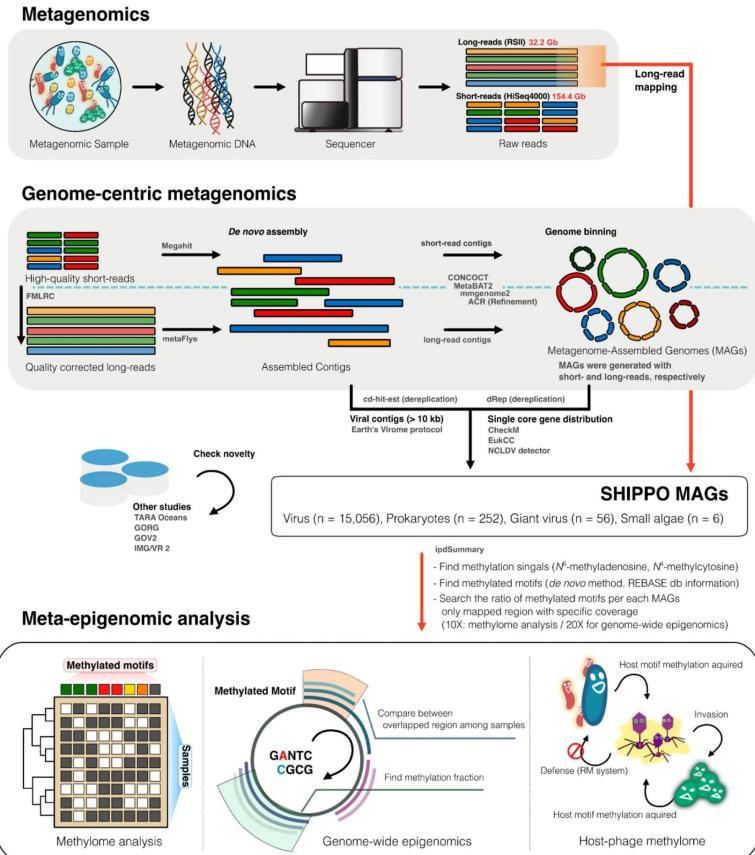
# Hybrid sequencing expands SVs like insertions and inversions that represent personalized microbial signatures, complicating correlations with metabolites and host health indicators

a



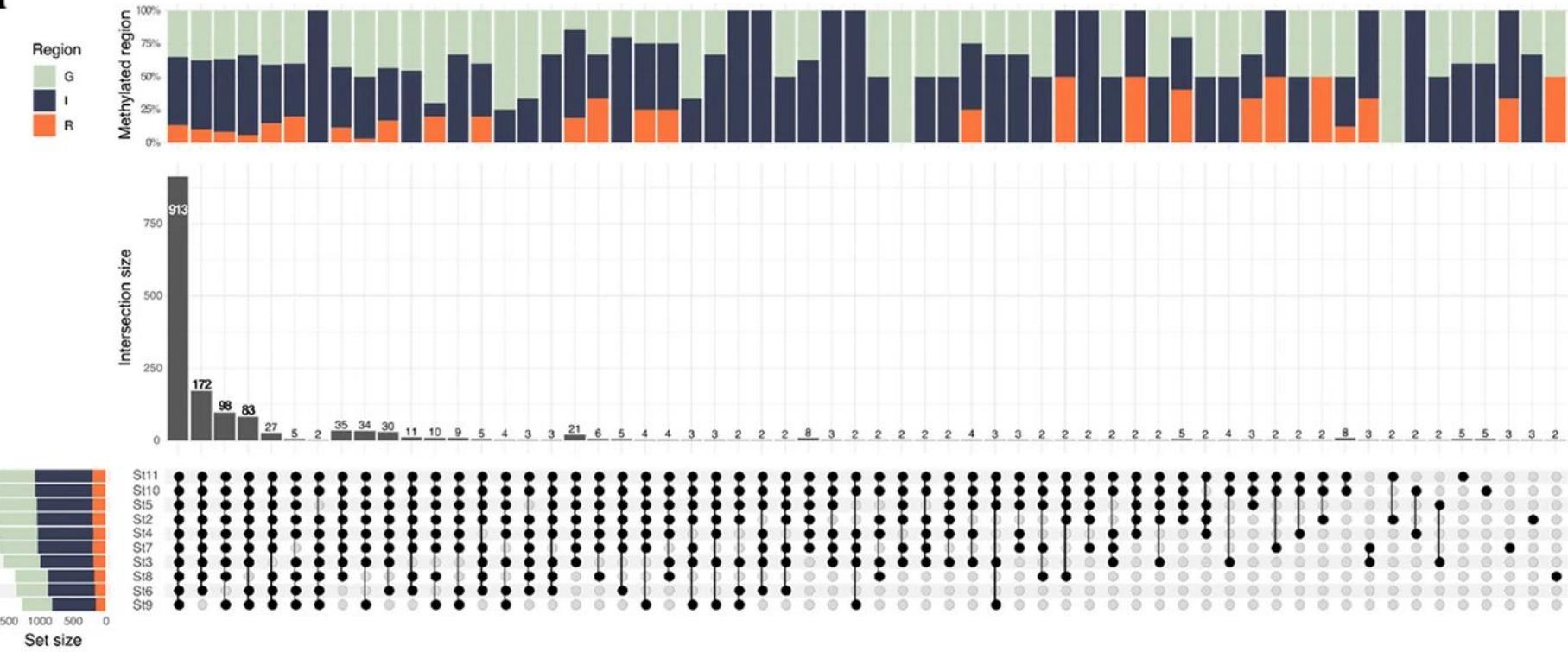
### **3. EPIGENETICS STUDIES**

Meta-epigenome analysis benefits from long-read sequencing by offering comprehensive insights into DNA methylation patterns and structural variations across diverse microbial communities, even at **sub-strain levels**.



## Compares methylation at each genome position on a Pelagibacter MAG

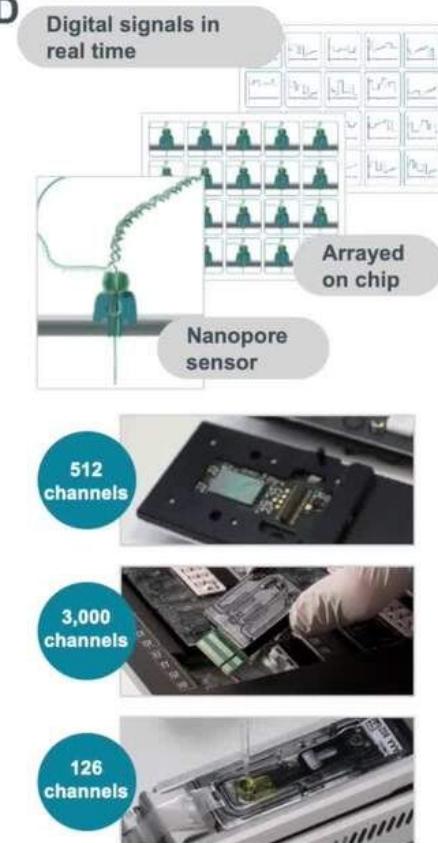
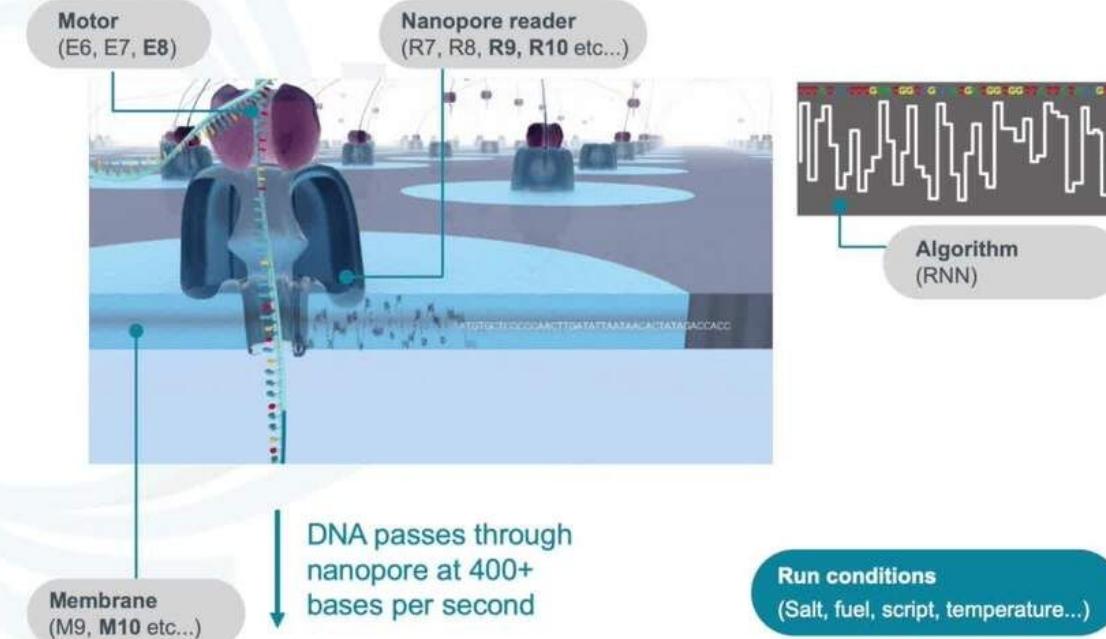
a



# 4. PORTABLE AND REAL-TIME SEQUENCING DEVICE

## NANOPORE DNA/RNA SEQUENCING: A NOVEL, SCALABLE METHOD

DNA/RNA strand passes through pore → signal interpreted into sequence data



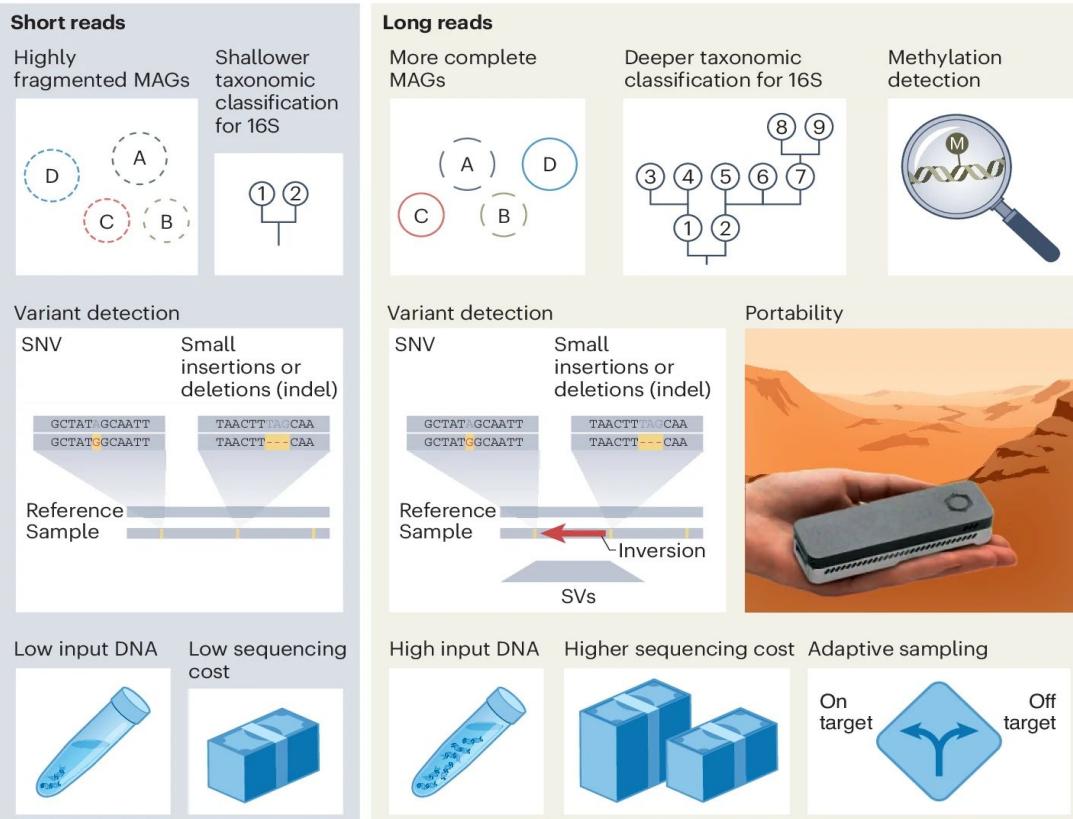
<https://www.youtube.com/watch?v=GUb1TzvMWsw>

5 | © Copyright 2020 Oxford Nanopore Technologies. Oxford Nanopore Technologies products are currently for Research Use Only

# CHALLENGES

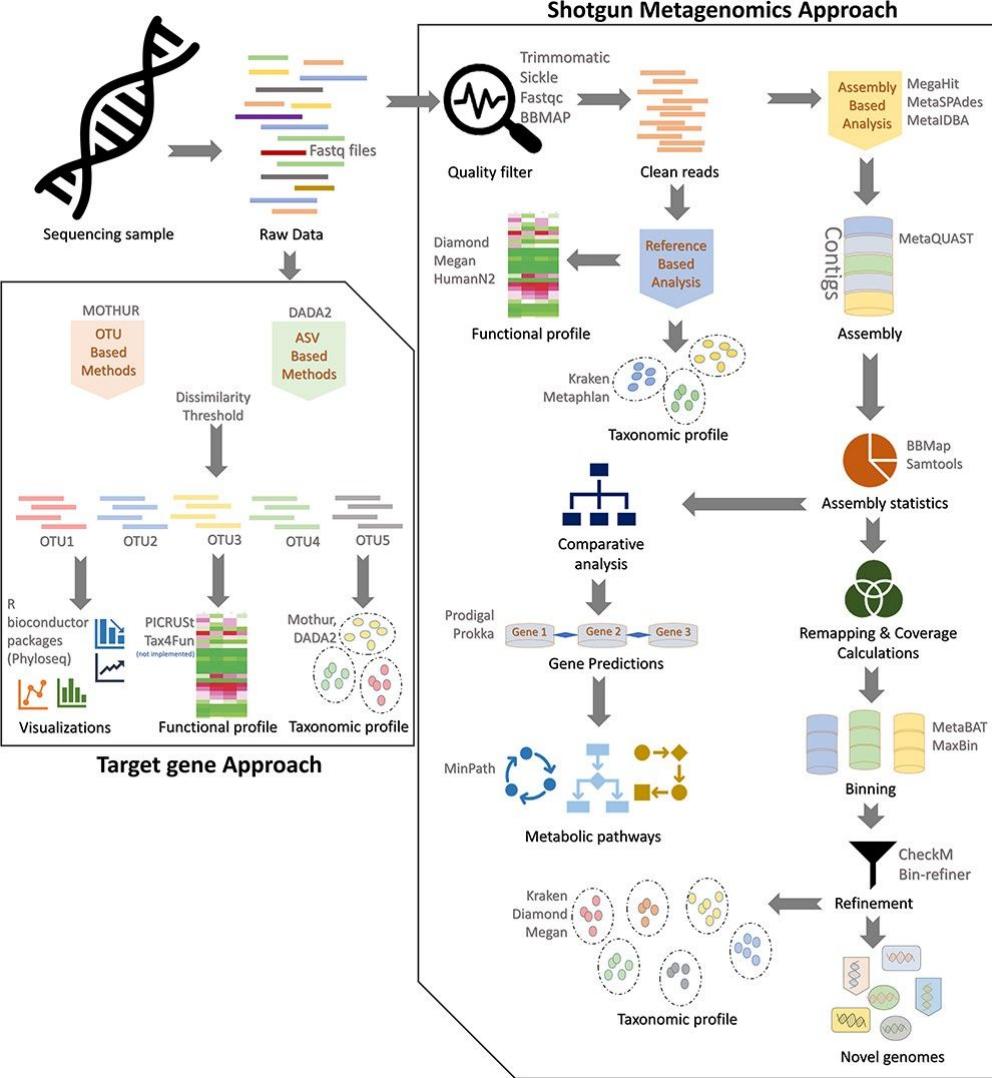
- Accurate and unbiased species identification
- Lack of universally accepted benchmark samples
- Difficulty in discerning sub-strain variations
- Computer resources

# SUMMARY



Long reads can

- generate less fragmented genome assemblies
- lower-level (species/strain) taxonomic characterization
- DNA/RNA methylation pattern identification
- large SV detection
- highly portable sequencers and realtime



# WORKFLOW

# Examples of Software

	Short reads	Long reads	Hybrid	EPI2ME
Adapter / quality trimming	fastp	porechop	fastp, porechop filtlong	No
Host read removal	Bowtie2	Minimap2	Bowtie2, Minimap2	Minimap2
Quality control	FastQC	Fastcat, NanoPlot, FastQC	FastQC, NanoPlot	Fastcat
Taxonomic classification	Reads: Kraken2, Bracken, Centrifuge Contigs, MAGs: CAT, GTDB-Tk			Kraken2, Bracken
Assembly	SPAdes, MEGAHIT QUAST (for evaluation)	Canu, MetaFlye, hifiasm-meta and pipelines such as Lathe QUAST (for evaluation)	SPADES Hybrid	No
Binning	MetaBAT2, MaxBin2, CONCONCT, MyCC, GroopM, MetaWRAP, Anvi'o, SemiBin QUAST (for evaluation)	MetaBAT2, MaxBin2, CONCONCT, MetaBCC-LR, LRBinner QUAST (for evaluation)	MetaBAT2, MaxBin2, CONCONCT, MyCC, GroopM, MetaWRAP, Anvi'o, SemiBin QUAST (for evaluation)	No
Binning refinement	DAS Tool, CheckM			No
Statistics	R, python packages			

# ONT Metagenomics workflow

## epi2me-labs/wf-metagenomics



Metagenomic classification of long-read sequencing data

8  
Contributors

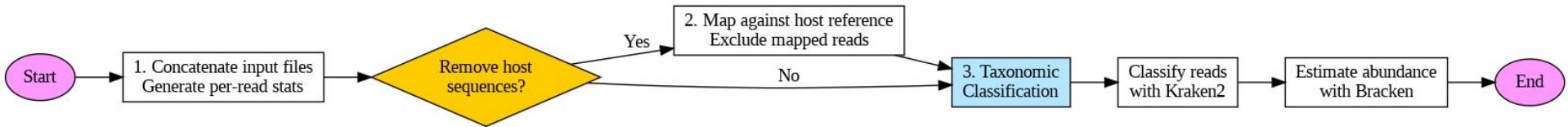
2  
Issues

55  
Stars

23  
Forks



# ONT Metagenomics workflow



## 1. Concatenate fastq(.gz) files whilst creating a summary of the sequences

- Use fastcat to concatenate FASTQ files
- Output per-read stats (lengths, average qualities)

## 2. Remove host sequences

- Use Minimap2 to map against provided host reference
- Exclude mapped reads from further analysis

## 3. Taxonomic Classification by kraken2 (faster method)

- Classify reads taxonomically
- Use Bracken for species abundance estimation
- Supports real-time analysis

# A SIMULATION PIPELINE

# Metadata

- BioProject: PRJNA820119
- Strategy: SHOTGUN
- Host: *Homo sapiens*
- Isolation source: fecal samples
- Collection date: November, 2019
- Geographic location: China: Beijing
- latitude and longitude: 39.3 N 116.194 E
- Instrument: Illumina NovaSeq 6000 and PromethION

# Metadata

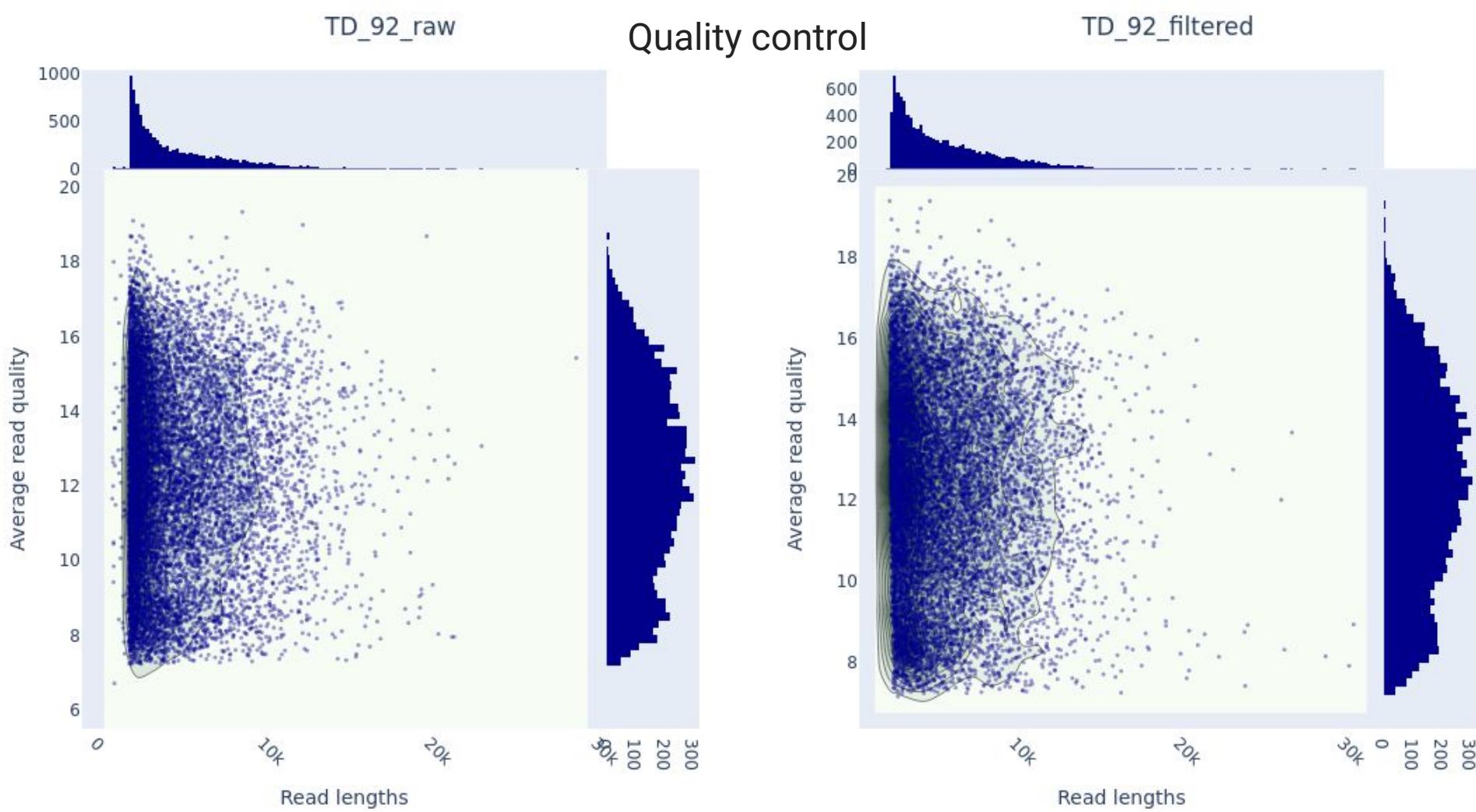
SRR18491279	CD_7_short	(19.6M spots, 5.9G bases, 1.8Gb) x 2
SRR18491202	CD_7_long	511,893 spots, 3.5G bases, 2.9Gb
SRR18491059	TD_31_short	(18.8M spots, 5.6G bases, 1.7Gb) x 2
SRR18490949	TD_31_long	476,473 spots, 2.9G bases, 2.5Gb
SRR18491219	TD_92_short	(19.7M spots, 5.9G bases, 1.8Gb) X 2
SRR18491134	TD_92_long	582,734 spots, 2.9G bases, 2.5Gb

# WORKFLOW

	Short reads	Long reads
Adapter / quality trimming	fastp	porechop
Host read removal	Bowtie2	Minimap2
Quality control	FastQC	NanoPlot
Taxonomic classification		Kraken2, Bracken
Assembly	SPAdes QUAST (for evaluation)	MetaFlye QUAST (for evaluation)
Binning		MetaWRAP - MetaBAT2
Binning evaluation, refinement and annotation		MetaWRAP - CheckM
Statistics: relative abundance, alpha diversity, beta diversity		R, python packages

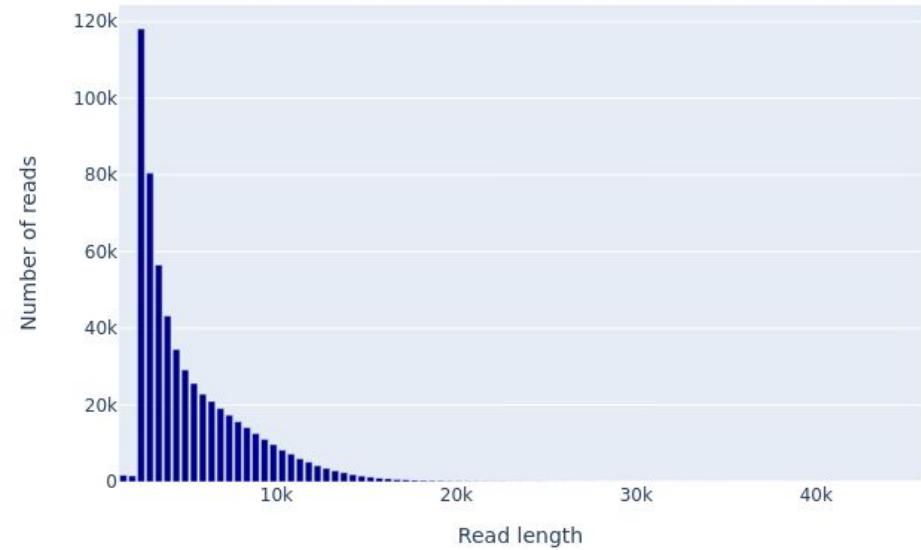
# 1. QUALITY CONTROL

- Quality Assessment
  - Before Trimming: Assess raw read quality (e.g., Phred scores).
  - After Trimming: Re-evaluate trimmed reads for quality improvements.
- Trimming and Filtering
  - Remove Adapters: Trim adapter sequences from reads.
  - Filter Low-Quality Bases: Remove bases with low-quality scores.
  - Filter Short Reads: Discard reads below a specified length.
- Decontamination
  - Remove Host DNA: Eliminate host DNA sequences.
  - Correct Sequencing Errors: Improve read accuracy by correcting miscalled bases.

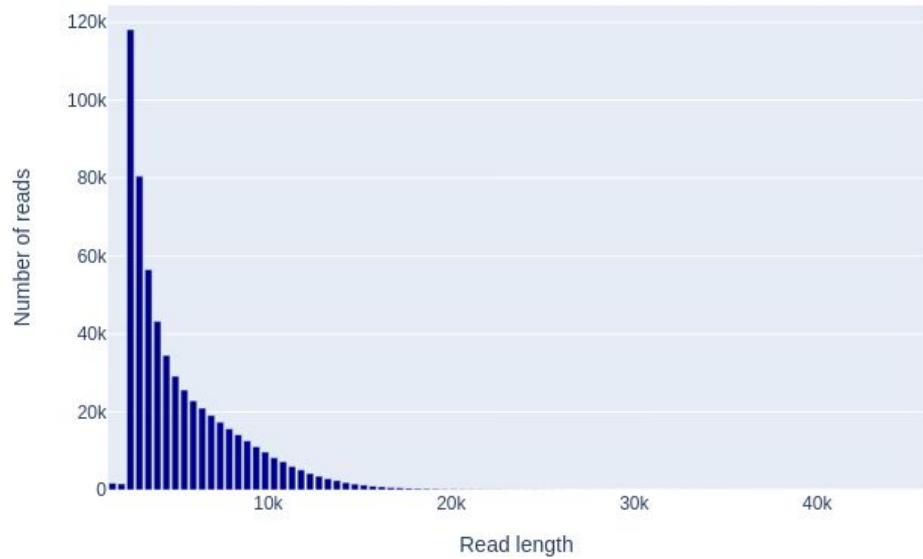


# Quality control

TD\_92\_raw

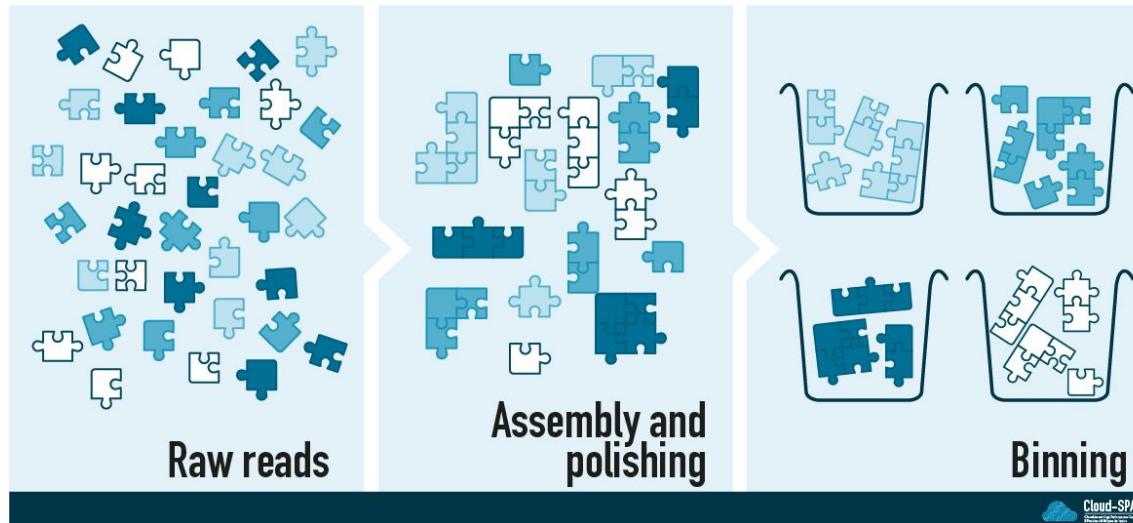


TD\_92\_filtered



## 2. ASSEMBLY: Reconstruct longer genomic fragments

- Challenges in Metagenomic Assembly:
  - Uneven Coverage: Variability in coverage across different species; conserved regions complicate assembly.
  - Strain-Level Variations: Differences between strains of the same species.
  - Repetitive Elements: Presence of repetitive sequences can hinder accurate assembly.
- Assembly Strategies
  - De Novo Assembly
  - Assembly Quality Assessment
  - Binning: Group contigs into putative genomes



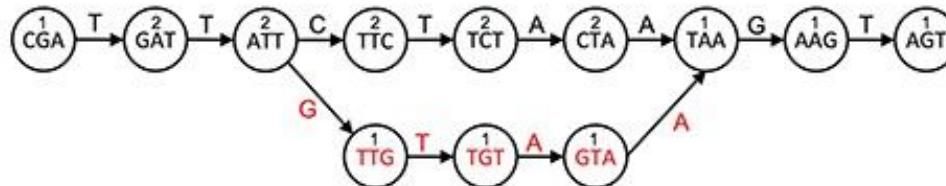
# SPAdes Genome Assembler

## (b) De Bruijn graph assembly

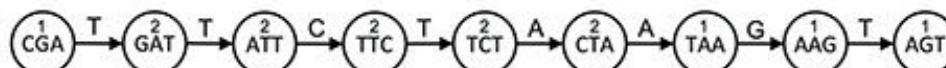
### (i) Make kmers

Read1: TTCTAAAGT	Read2: CGATTCTA	Read3: GATTGTAA
Kmers: TTC	Kmers: CGA	Kmers: GAT
TCT	GAT	ATT
CTA	ATT	TTG
TAA	TTC	TGT
AAG	TCT	GTA
AGT	CTA	TAA

### (ii) Build graph



### (iii) Walk graph and output contigs



**CGATTCTAAAGT**

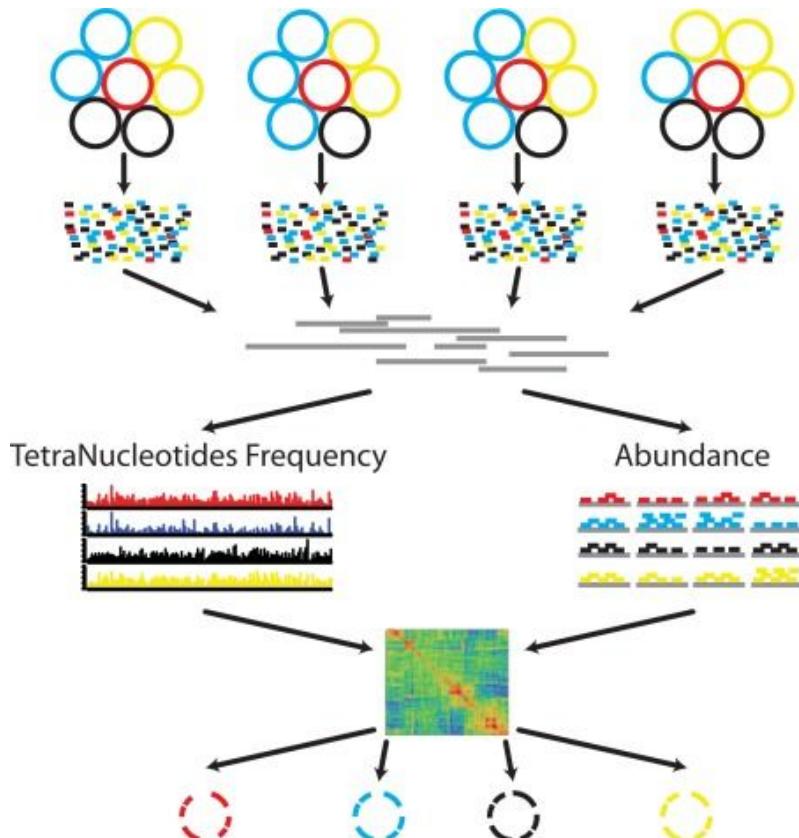
Martin Ayling, Matthew D Clark, Richard M Leggett, New approaches for metagenome assembly with short reads, *Briefings in Bioinformatics*, Volume 21, Issue 2, March 2020, Pages 584–594, <https://doi.org/10.1093/bib/bbz020>

Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19(5):455-477. doi:10.1089/cmb.2012.0021

# QUAST for contigs

Assembly	CD_7_short contigs	CD_7_short scaffolds	CD_7_long	TD_31_short	TD_31_short scaffolds	TD_31_long	TD_92_short	TD_92_short scaffolds	TD_92_long
# contigs (>= 0 bp)	352497	348091	3430	229053	225253	2996	136094	133897	1750
# contigs (>= 1000 bp)	35521	34893	3408	28789	28175	2977	19763	19094	1742
# contigs (>= 5000 bp)	5869	5924	3320	5413	5360	2839	3879	3841	1651
# contigs (>= 10000 bp)	2600	2661	2934	2660	2716	2407	2048	2084	1391
# contigs (>= 25000 bp)	787	835	1738	910	939	1293	940	968	798
# contigs (>= 50000 bp)	276	293	877	351	368	666	490	503	443
Total length (>= 0 bp)	265543681	265641333	194584883	220188265	220277443	159179728	169341946	169411224	106053939
Total length (>= 1000 bp)	155166798	156498690	194571080	147726318	148887068	159166732	128183543	128793310	106048669
Total length (>= 5000 bp)	97131210	99950938	194245324	102441551	104838575	158686661	96413741	98505521	105725045
Total length (>= 10000 bp)	74532639	77396318	191274569	83385709	86547617	155381648	83799568	86358980	103729294
Total length (>= 25000 bp)	46818895	49422186	171171254	56895322	59488372	137283925	66570189	68911734	93837056
Total length (>= 50000 bp)	29151419	30713587	140563593	37653832	39851956	115167055	50485398	52341712	81276403
# contigs	85334	84489	3426	65996	65090	2995	39368	38498	1750
Largest contig	529830	755834	2584321	565327	816179	2342557	1314428	1314428	2229045
Total length	189046883	190225872	194583122	173202607	174179833	159179245	141689831	142166149	106053939
GC (%)	46.63	46.64	45.46	45.92	45.92	45.11	43.9	43.9	42.74
N50	5405	5788	118255	8863	9820	142056	20197	22907	173207
N90	715	718	22293	787	794	19835	1035	1045	22330
auN	33120.8	35692.2	368224.9	41572.5	47205.6	321501.9	65353.1	68282.9	402776.2
L50	5367	5024	307	3002	2771	237	1131	1060	114
L90	53364	52378	1905	38006	37035	1562	19112	18269	867
# N's per 100 kbp	0	51.28	0	0	51.05	0	0	49.25	0
# predicted rRNA genes	38 + 17 part	39 + 18 part	732 + 25 part	31 + 26 part	31 + 26 part	456 + 42 part	36 + 20 part	36 + 20 part	319 + 26 part

# MetaBAT



## Preprocessing

- 1 Samples from multiple sites or times
- 2 Metagenome libraries
- 3 Initial de-novo assembly using the combined library

## MetaBAT

- 4 Calculate TNF for each contig
- 5 Calculate Abundance per library for each contig
- 6 Calculate the pairwise distance matrix using pre-trained probabilistic models
- 7 Forming genome bins iteratively

# CheckM - assessing the quality of genome bins

TD\_92\_long/binning/bin\_refinement/binsA.stats

bin	completeness	contamination	GC	lineage	N50	size	binner	
bin.10	0.0	0.0	0.268	root	37144	248280	binsA	
bin.11	73.66	0.0	0.464	Bacteroidales	1031723	3054580	binsA	
bin.12	97.60	0.227	0.593	Bifidobacteriaceae		2229045	2229045	
				binsA				
bin.1	33.01	0.0	0.416	Clostridiales	1104936	1647912	binsA	
bin.2	79.95	2.230	0.435	Bacteroidales	1008497	4241802	binsA	
bin.3	74.58	4.474	0.373	Clostridiales	654883	2711616	binsA	
bin.4	65.82	0.158	0.412	Clostridiales	660293	2538921	binsA	
bin.5	43.10	8.620	0.429	Bacteria	467667	3921431	binsA	
bin.6	81.89	102.5	0.419	Bacteria	152868	15164013	binsA	
bin.7	0.0	0.0	0.422	root	371188	395268	binsA	
bin.8	0.0	0.0	0.437	Bacteria	306292	354129	binsA	
bin.9	90.75	0.309	0.462	Bacteroidales	604225	3654929	binsA	
bin.unbinned		5.199	0.458	0.418	Bacteroides	41611	1037190	binsA

### 3. TAXONOMIC CLASSIFICATION

**Kraken:** K-mer-based taxonomic classification that is fast and highly accurate.

- Collects k-mers from reference genomes and associates them with taxonomic IDs (LCA).
- Breaks query sequences into k-mers and looks them up in the database.
- Uses a weighted voting scheme for taxonomy assignment.
- Output: Provides taxonomic classifications and a summary report of taxonomic composition.

**Bracken:** Bayesian Reestimation of Abundance with Kraken

- Estimates species abundance in metagenomic samples, addressing Kraken's limitations.
- Improves species-level abundance estimates using a Bayesian algorithm to redistribute reads.
- Fast and efficient, works directly with Kraken output.

Collection	Contains	Date	Archive size (GB)	Index size (GB)	HTTPS URL	Inspect	Library	MD5
Viral	RefSeq viral	6/5/2024	0.5	0.6	.tar.gz	.txt	.tsv	.md5
MinusB	RefSeq archaea, viral, plasmid, human <sup>1</sup> , UniVec_Core	6/5/2024	7.1	10.2	.tar.gz	.txt	.tsv	.md5
Standard	RefSeq archaea, bacteria, viral, plasmid, human <sup>1</sup> , UniVec_Core	6/5/2024	60	78	.tar.gz	.txt	.tsv	.md5
Standard-8	Standard with DB capped at 8 GB	6/5/2024	5.5	7.5	.tar.gz	.txt	.tsv	.md5
Standard-16	Standard with DB capped at 16 GB	6/5/2024	11	15	.tar.gz	.txt	.tsv	.md5
PlusPF	Standard plus RefSeq protozoa & fungi	6/5/2024	64	83	.tar.gz	.txt	.tsv	.md5
PlusPF-8	PlusPF with DB capped at 8 GB	6/5/2024	5.5	7.5	.tar.gz	.txt	.tsv	.md5
PlusPF-16	PlusPF with DB capped at 16 GB	6/5/2024	11	15	.tar.gz	.txt	.tsv	.md5
PlusPFP	Standard plus RefSeq protozoa, fungi & plant	6/5/2024	135	182	.tar.gz	.txt	.tsv	.md5
PlusPFP-8	PlusPFP with DB capped at 8 GB	6/5/2024	5.1	7.5	.tar.gz	.txt	.tsv	.md5
PlusPFP-16	PlusPFP with DB capped at 16 GB	6/5/2024	11	15	.tar.gz	.txt	.tsv	.md5
EuPathDB46 <sup>2</sup>	Eukaryotic pathogen genomes with contaminants removed	4/18/2023	8.4	11	.tar.gz	.txt	N/A	N/A
nt Database	Very large collection, inclusive of GenBank, RefSeq, TPA and PDB	5/30/2024	684	889	.tar.gz	.txt	.tsv	.md5

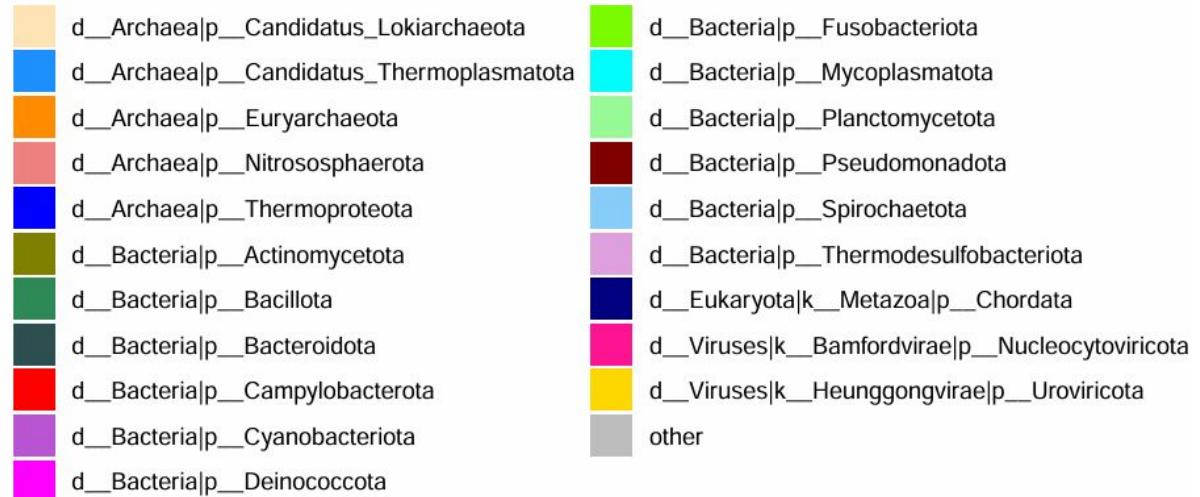
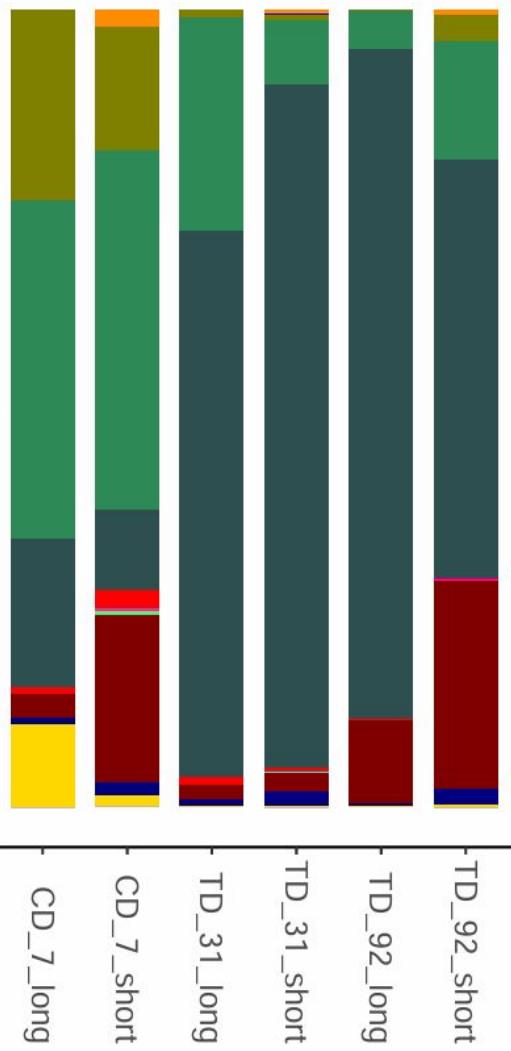
# Kraken 2 / Bracken database

<https://benlangmead.github.io/aws-indexes/k2>

Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, Salzberg SL, Steinegger M. (2022) Metagenome analysis using the Kraken software suite.. Nature Protocols doi: 10.1038/s41596-022-00738-y

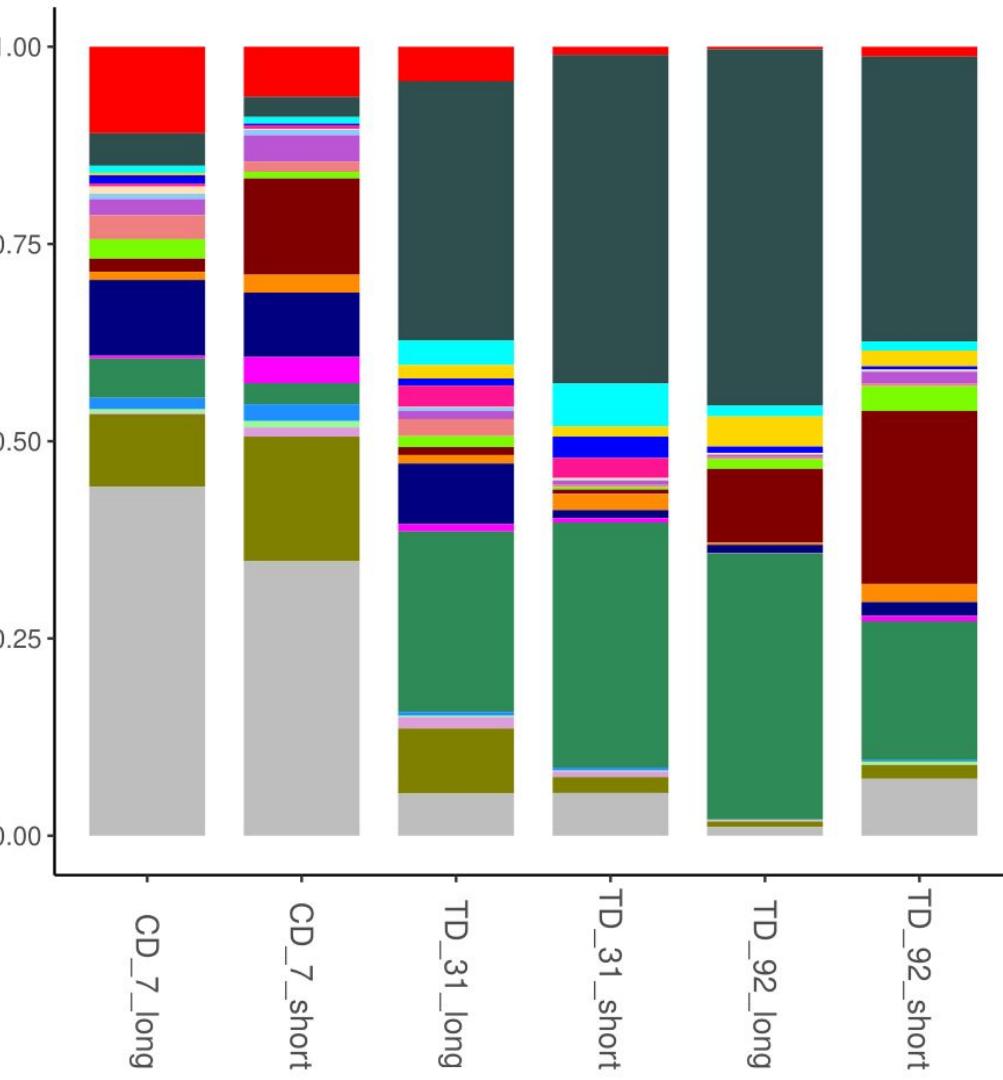
# Relative abundance by Phylum

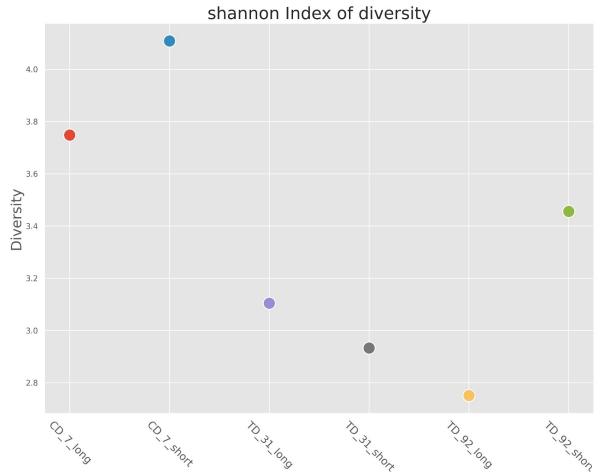
Relative Abundance



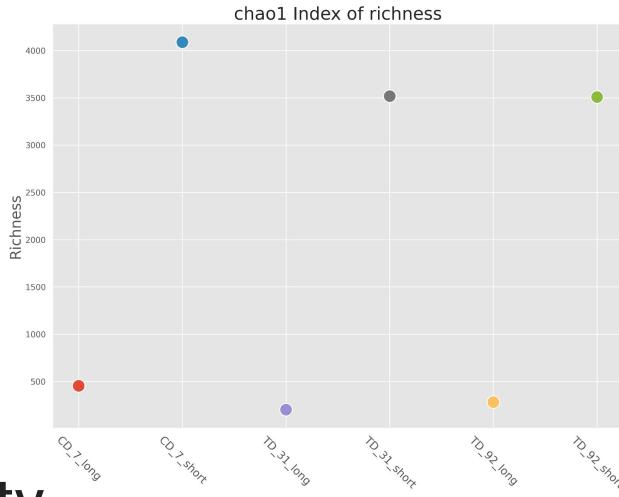
# Relative abundance by Species

Relative Abundance



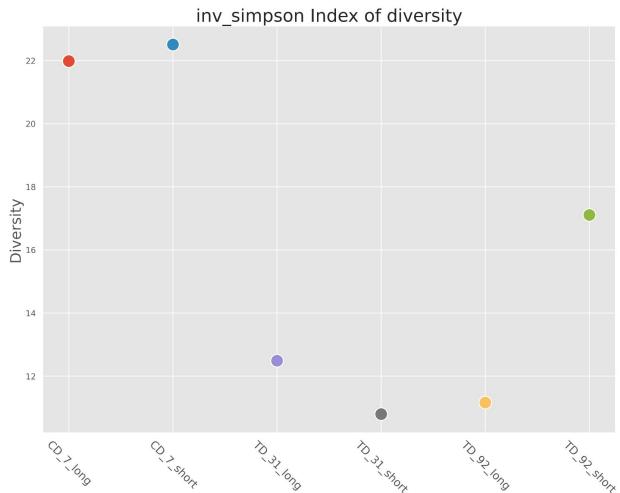


Group  
 CD\_7\_long  
 CD\_7\_short  
 TD\_31\_long  
 TD\_31\_short  
 TD\_92\_long  
 TD\_92\_short

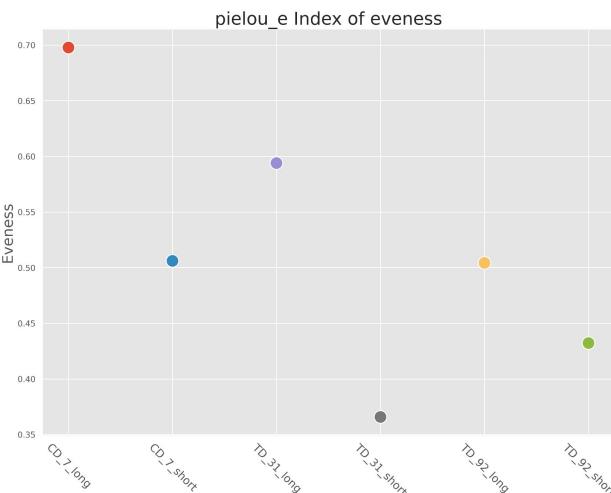


Group  
 CD\_7\_long  
 CD\_7\_short  
 TD\_31\_long  
 TD\_31\_short  
 TD\_92\_long  
 TD\_92\_short

# Alpha diversity



Group  
 CD\_7\_long  
 CD\_7\_short  
 TD\_31\_long  
 TD\_31\_short  
 TD\_92\_long  
 TD\_92\_short



Group  
 CD\_7\_long  
 CD\_7\_short  
 TD\_31\_long  
 TD\_31\_short  
 TD\_92\_long  
 TD\_92\_short

# Beta diversity

