Metagenomics processing

E5444 Analysis of Sequencing Data

Vojtěch Bartoň, 2023

Metagenomics

- **Microbial Community Genetics:** Metagenomics studies the genetics of entire microbial communities in environmental samples.
- **Genomic Snapshot:** It doesn't require isolating or culturing individual microorganisms, providing a holistic view of diverse microbes.
- **Health Insights:** It helps us study the human microbiome and its impact on health without traditional culturing methods
- **Applications in Ecology and Biotech:** Metagenomics informs ecosystem understanding and aids in biotech discoveries like enzymes and antibiotics

Metagenomics: WMGS x 16S

WMGS	16S
Captures the entire genetic content of all microorganisms.	Targets a specific 16S gene marker in bacteria and archaea.
High-resolution data for functional analysis and taxonomic identification.	Lower resolution, primarily for taxonomic identification.
Suitable for complex ecosystems and novel gene discovery.	Commonly used for microbial community profiling and diversity studies.
Used in environmental and clinical metagenomics.	Cost-effective method for taxonomic analysis.

MetaPhlAn

- **Meta**genomic **Ph**y**l**ogenetic **An**alysis
- **Taxonomic Profiling:** MetaPhlAn identifies and quantifies microbes in samples.
- **Marker Gene Approach:** It uses unique genetic markers for speedy and accurate identification.
- **Efficiency:** Known for fast analysis of large datasets.
- **Applications:** Widely used in microbiome research and clinical metagenomics.

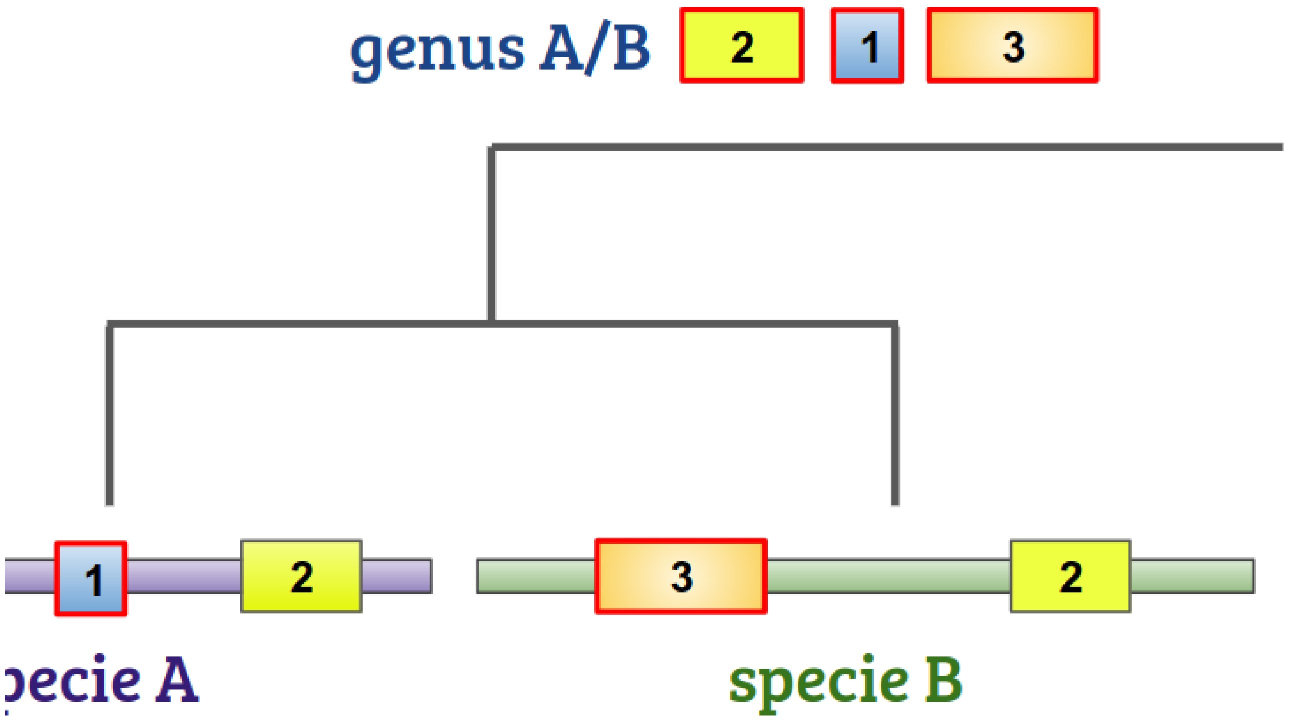
- https://github.com/biobakery/MetaPhlAn

Metaphlan: input data

- **Shotgun** Whole Metagenome Sequencing

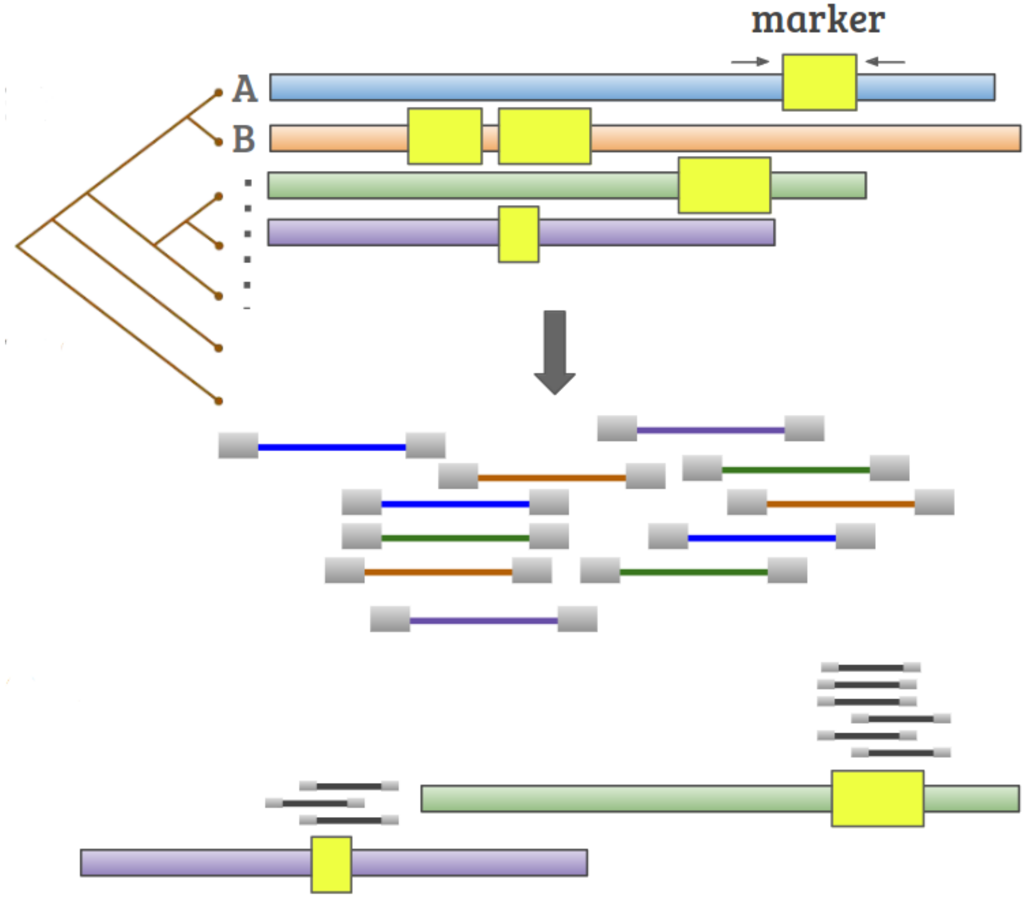
- Sequences (fasta, fastq)
- Database of taxonomically known sequences
  - Unique regions

Metaphlan: marker genes



- **Clades**
    - Groups of genomes (organisms) believed to have evolved from a common ancestor
- **Clade-specific marker-genes**
    - Strongly conserved within the clade's genomes
    - Not similar to any sequence in other clades (of the same level)
    - Unique markers change as the clade level grows
    - They also accumulate in a way (direct vs indirect)...

Metaphlan: Overview



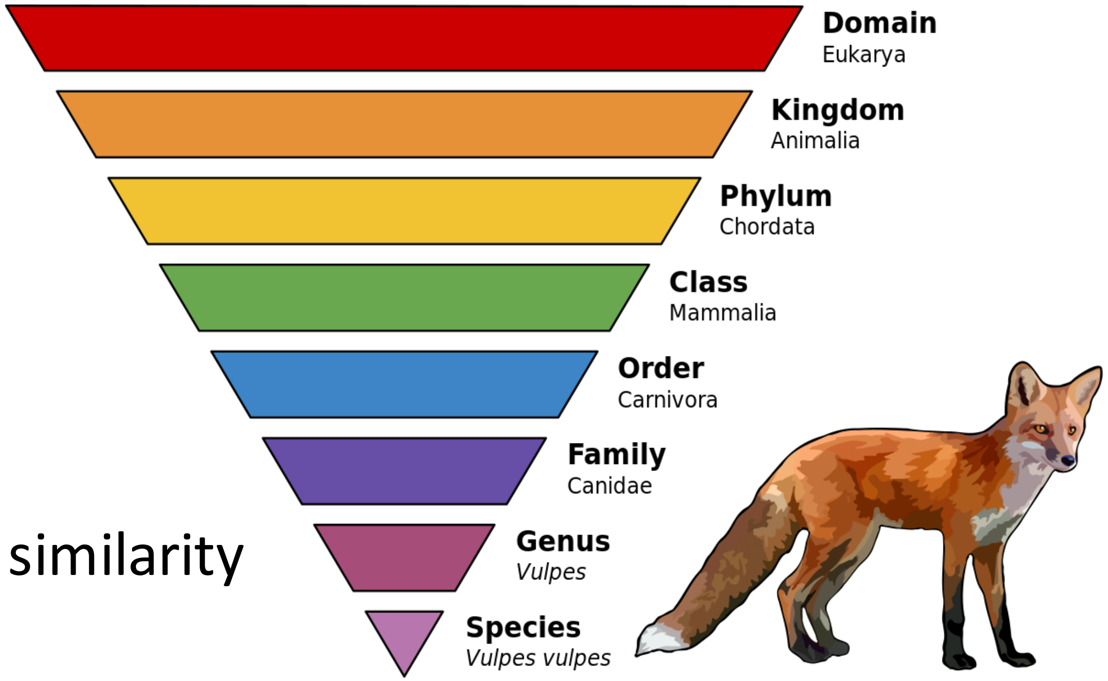
- Reference genomes and their taxonomy
- Find clade-specific marker genes

- Sequence your sample

- Map to marker genes
- Count taxonomic units

Metaphlan: Reference database

- **ChocoPhlAn**
- Acquire reference genomes
  - De novo assembly
  - Cultured species
  - **Uniprot core data**
- Acquire taxonomy
  - Hierarchical clustering tree based on similarity
  - **Ncbi taxonomy**

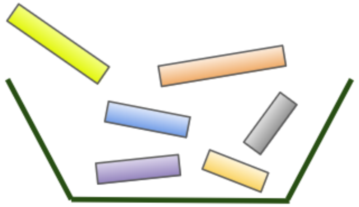


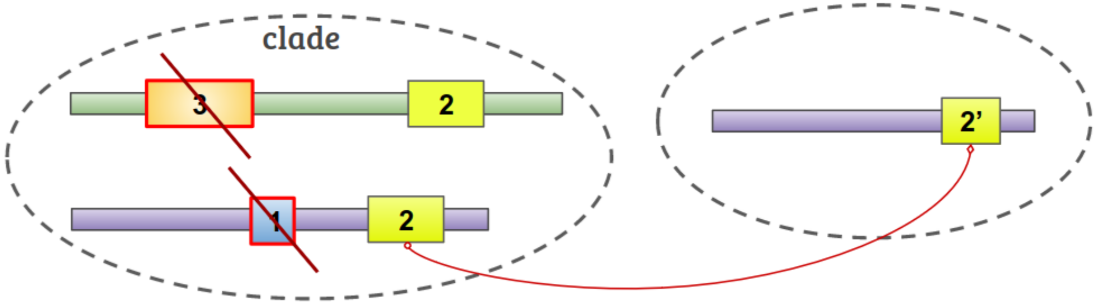
Red fox *(Vulpes vulpes)*

Metaphlan: Reference database

- The general process:
  - Each genome → bag-of-genes representation
  - Only conserved genes in the clade are saved
  - Inter-clade uniqueness index elimination
  - Single-copy genes were preferred of multi-copy genes

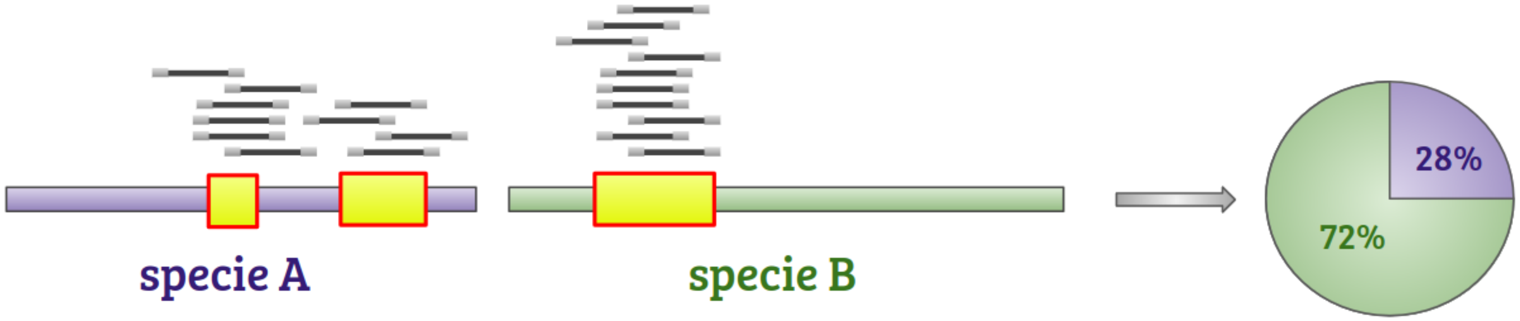
- Properties of the markers
  - Gene level
  - 5.1M filtered genes
  - 27K species-level genome bins
  - Not necessarily continuous (bag-of-genes)
  - ~4% of the total genome length
  - ~260 markers per specie





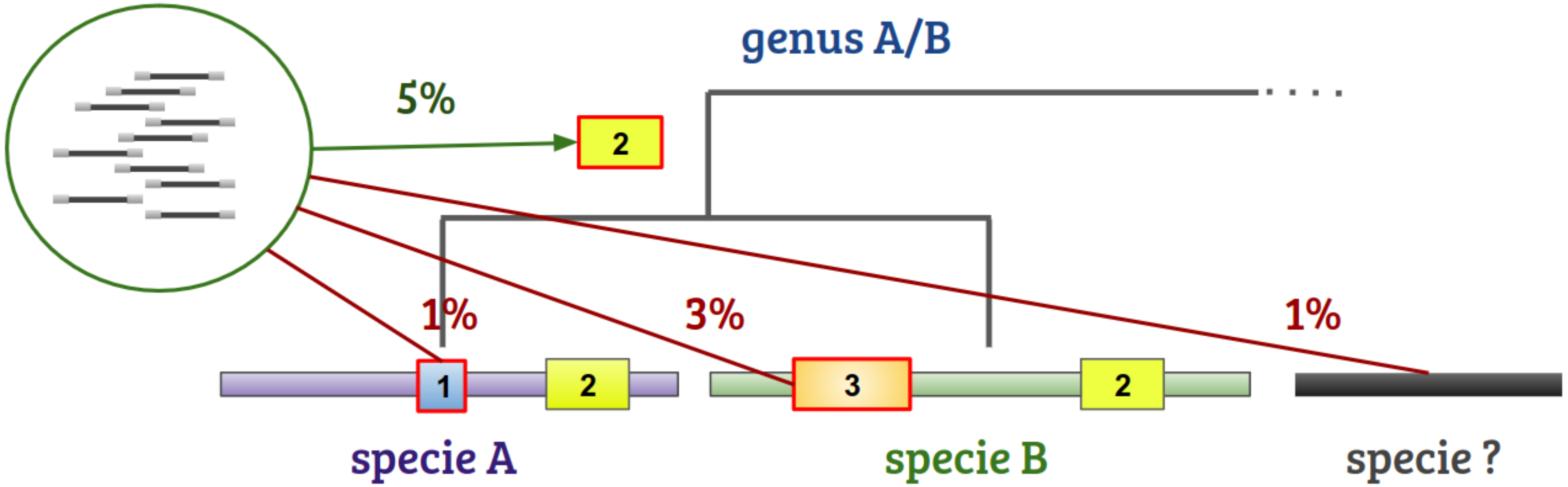
Metaphlan: Taxonomic profiling

- Map reads against reference database of marker genes
- Calculate relative abundance
  - Sum the total reads mapped to clade markers
  - Divide by marker’s total length
  - Abundances in every clade-level sum up to 100%



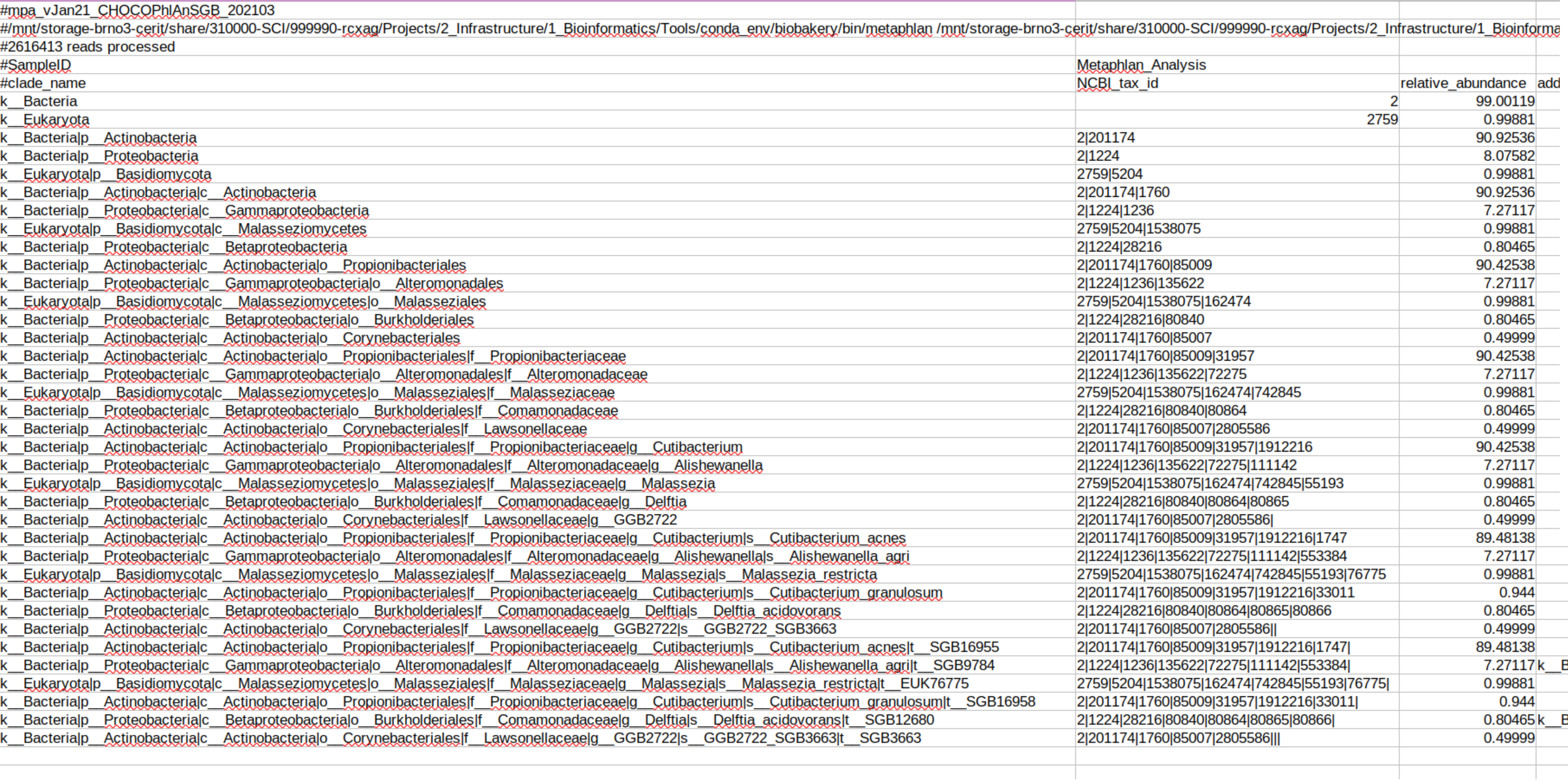
Metaphlan: Taxonomic profiling

- Unclassified case
  - Move up in the taxonomy tree

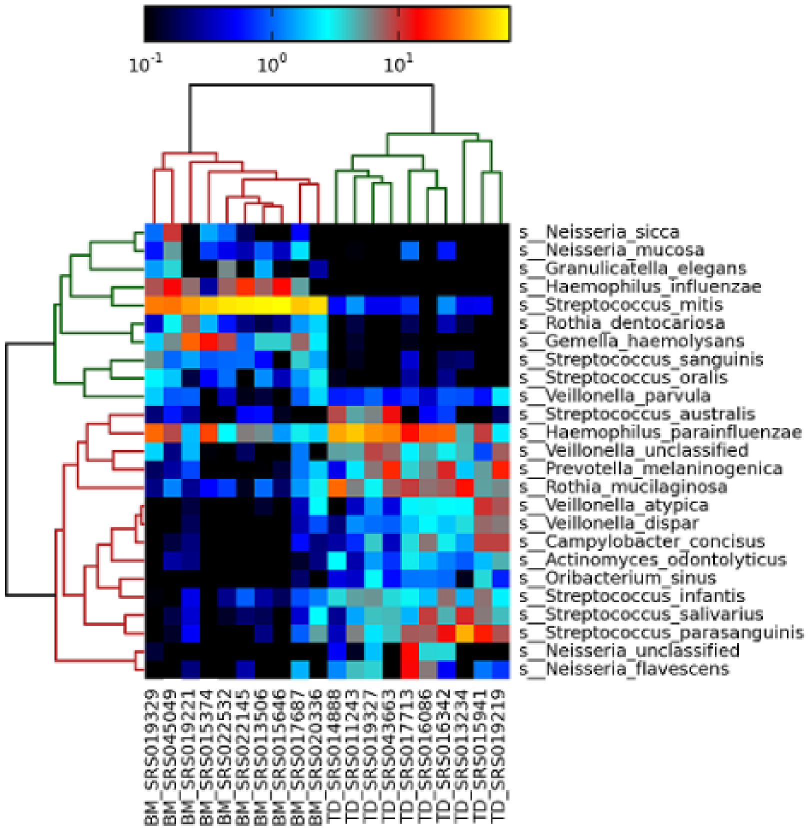


Metaphlan: outputs and visualisations

- bugs_list.tsv
- Utility scripts



#mpa_vJan21_CHOCOPhlAnSGB_202103			
#/mnt/storage-brno3-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/conda_env/biobakery/bin/metaphlan /mnt/storage-brno3-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinforma			
#2616413 reads processed			
#SampleID	Metaphlan_Analysis		
#clade_name	NCBI_tax_id	relative_abundance	add
k__Bacteria	2	99.00119	
k__Eukaryota	2759	0.99881	
k__Bacteria|p__Actinobacteria	2|201174	90.92536	
k__Bacteria|p__Proteobacteria	2|1224	8.07582	
k__Eukaryota|p__Basidiomycota	2759|5204	0.99881	
k__Bacteria|p__Actinobacteria|c__Actinobacteria	2|201174|1760	90.92536	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria	2|1224|1236	7.27117	
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes	2759|5204|1538075	0.99881	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria	2|1224|28216	0.80465	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales	2|201174|1760|85009	90.42538	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales	2|1224|1236|135622	7.27117	
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes|o__Malasseziales	2759|5204|1538075|162474	0.99881	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales	2|1224|28216|80840	0.80465	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales	2|201174|1760|85007	0.49999	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae	2|201174|1760|85009|31957	90.42538	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae	2|1224|1236|135622|72275	7.27117	
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes|o__Malasseziales|f__Malasseziaceae	2759|5204|1538075|162474|742845	0.99881	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae	2|1224|28216|80840|80864	0.80465	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Lawsonellaceae	2|201174|1760|85007|2805586	0.49999	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae|g__Cutibacterium	2|201174|1760|85009|31957|1912216	90.42538	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella	2|1224|1236|135622|72275|111142	7.27117	
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes|o__Malasseziales|f__Malasseziaceae|g__Malassezia	2759|5204|1538075|162474|742845|55193	0.99881	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia	2|1224|28216|80840|80864|80865	0.80465	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Lawsonellaceae|g__GGB2722	2|201174|1760|85007|2805586|	0.49999	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae|g__Cutibacterium|s__Cutibacterium_acnes	2|201174|1760|85009|31957|1912216|1747	89.48138	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella|s__Alishewanella_agri	2|1224|1236|135622|72275|111142|553384	7.27117	
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes|o__Malasseziales|f__Malasseziaceae|g__Malassezia|s__Malassezia_restricta	2759|5204|1538075|162474|742845|55193|76775	0.99881	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae|g__Cutibacterium|s__Cutibacterium_granulosum	2|201174|1760|85009|31957|1912216|33011	0.944	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_acidovorans	2|1224|28216|80840|80864|80865|80866	0.80465	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Lawsonellaceae|g__GGB2722|s__GGB2722_SGB3663	2|201174|1760|85007|2805586||	0.49999	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae|g__Cutibacterium|s__Cutibacterium_acnes|t__SGB16955	2|201174|1760|85009|31957|1912216|1747|	89.48138	
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella|s__Alishewanella_agri|t__SGB9784	2|1224|1236|135622|72275|111142|553384|	7.27117	k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales
k__Eukaryota|p__Basidiomycota|c__Malasseziomycetes|o__Malasseziales|f__Malasseziaceae|g__Malassezia|s__Malassezia_restricta|t__EUK76775	2759|5204|1538075|162474|742845|55193|76775|	0.99881	
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Propionibacteriaceae|g__Cutibacterium|s__Cutibacterium_granulosum|t__SGB16958	2|201174|1760|85009|31957|1912216|33011|	0.944	
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_acidovorans|t__SGB12680	2|1224|28216|80840|80864|80865|80866|	0.80465	k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Lawsonellaceae|g__GGB2722|s__GGB2722_SGB3663|t__SGB3663	2|201174|1760|85007|2805586|||	0.49999	



Metaphlan: application

- Shotgun sequencing
- Microbiome Profiling
- Metagenomics
- Metatranscriptomics

- As input for HUMAnN
  - profiling the abundance of microbial metabolic pathways and other molecular functions

Metaphlan: pros & cons

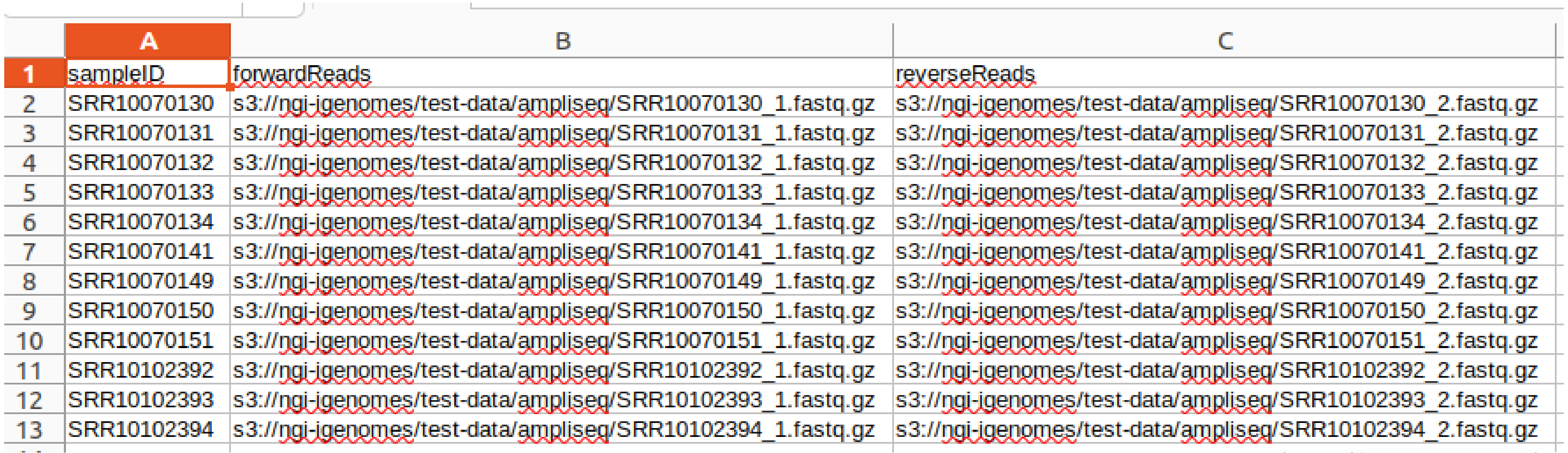
Pros

- Rapid profiling
- Accuracy
- Versatile
- Quantitative output

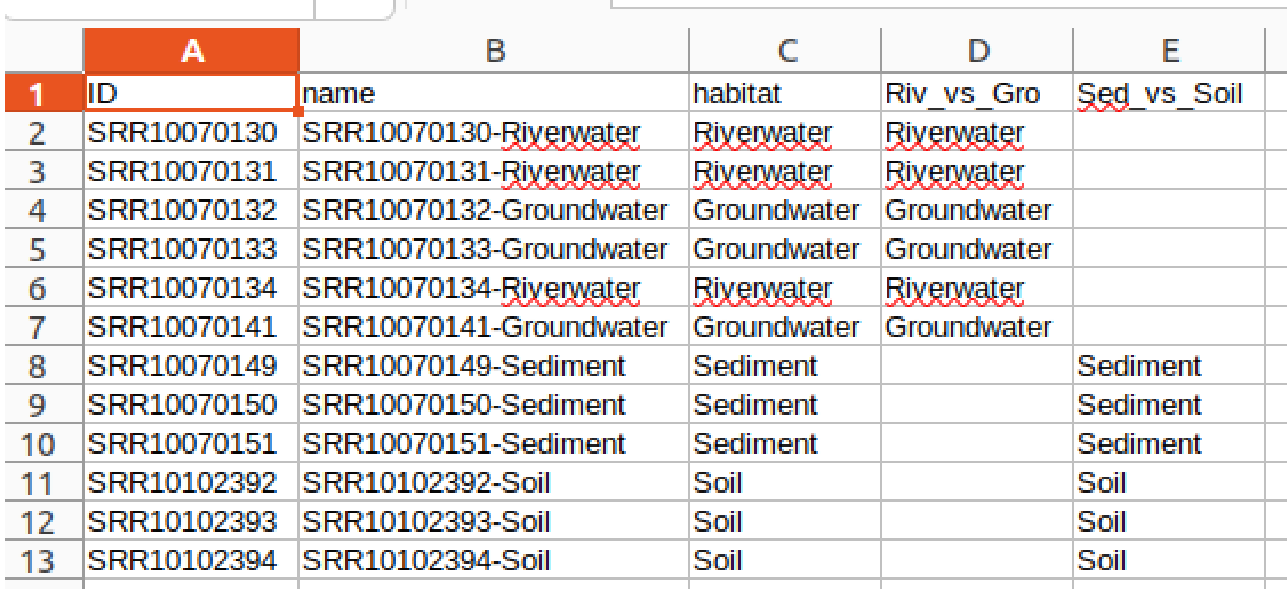
Cons

- Limited functional information
- Reference-dependent
- Computational resources
- Interpreting unknowns

QIIME2

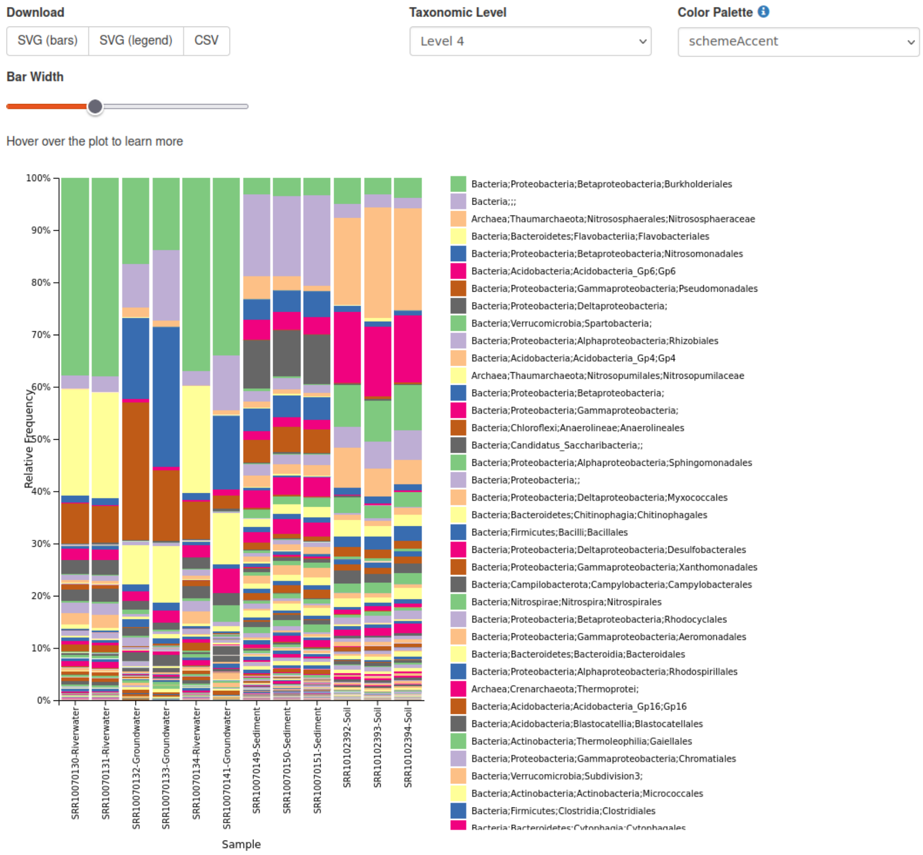


	A	B	C
1	sampleID	forwardReads	reverseReads
2	SRR10070130	s3://ngi-igenomes/test-data/ampliseq/SRR10070130_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070130_2.fastq.gz
3	SRR10070131	s3://ngi-igenomes/test-data/ampliseq/SRR10070131_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070131_2.fastq.gz
4	SRR10070132	s3://ngi-igenomes/test-data/ampliseq/SRR10070132_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070132_2.fastq.gz
5	SRR10070133	s3://ngi-igenomes/test-data/ampliseq/SRR10070133_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070133_2.fastq.gz
6	SRR10070134	s3://ngi-igenomes/test-data/ampliseq/SRR10070134_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070134_2.fastq.gz
7	SRR10070141	s3://ngi-igenomes/test-data/ampliseq/SRR10070141_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070141_2.fastq.gz
8	SRR10070149	s3://ngi-igenomes/test-data/ampliseq/SRR10070149_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070149_2.fastq.gz
9	SRR10070150	s3://ngi-igenomes/test-data/ampliseq/SRR10070150_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070150_2.fastq.gz
10	SRR10070151	s3://ngi-igenomes/test-data/ampliseq/SRR10070151_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10070151_2.fastq.gz
11	SRR10102392	s3://ngi-igenomes/test-data/ampliseq/SRR10102392_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10102392_2.fastq.gz
12	SRR10102393	s3://ngi-igenomes/test-data/ampliseq/SRR10102393_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10102393_2.fastq.gz
13	SRR10102394	s3://ngi-igenomes/test-data/ampliseq/SRR10102394_1.fastq.gz	s3://ngi-igenomes/test-data/ampliseq/SRR10102394_2.fastq.gz



	A	B	C	D	E
1	ID	name	habitat	Riv_vs_Gro	Sed_vs_Soil
2	SRR10070130	SRR10070130-Riverwater	Riverwater	Riverwater	
3	SRR10070131	SRR10070131-Riverwater	Riverwater	Riverwater	
4	SRR10070132	SRR10070132-Groundwater	Groundwater	Groundwater	
5	SRR10070133	SRR10070133-Groundwater	Groundwater	Groundwater	
6	SRR10070134	SRR10070134-Riverwater	Riverwater	Riverwater	
7	SRR10070141	SRR10070141-Groundwater	Groundwater	Groundwater	
8	SRR10070149	SRR10070149-Sediment	Sediment		Sediment
9	SRR10070150	SRR10070150-Sediment	Sediment		Sediment
10	SRR10070151	SRR10070151-Sediment	Sediment		Sediment
11	SRR10102392	SRR10102392-Soil	Soil		Soil
12	SRR10102393	SRR10102393-Soil	Soil		Soil
13	SRR10102394	SRR10102394-Soil	Soil		Soil

QIIME2



Download

SVG (bars) | SVG (legend) | CSV

Taxonomic Level

Level 4

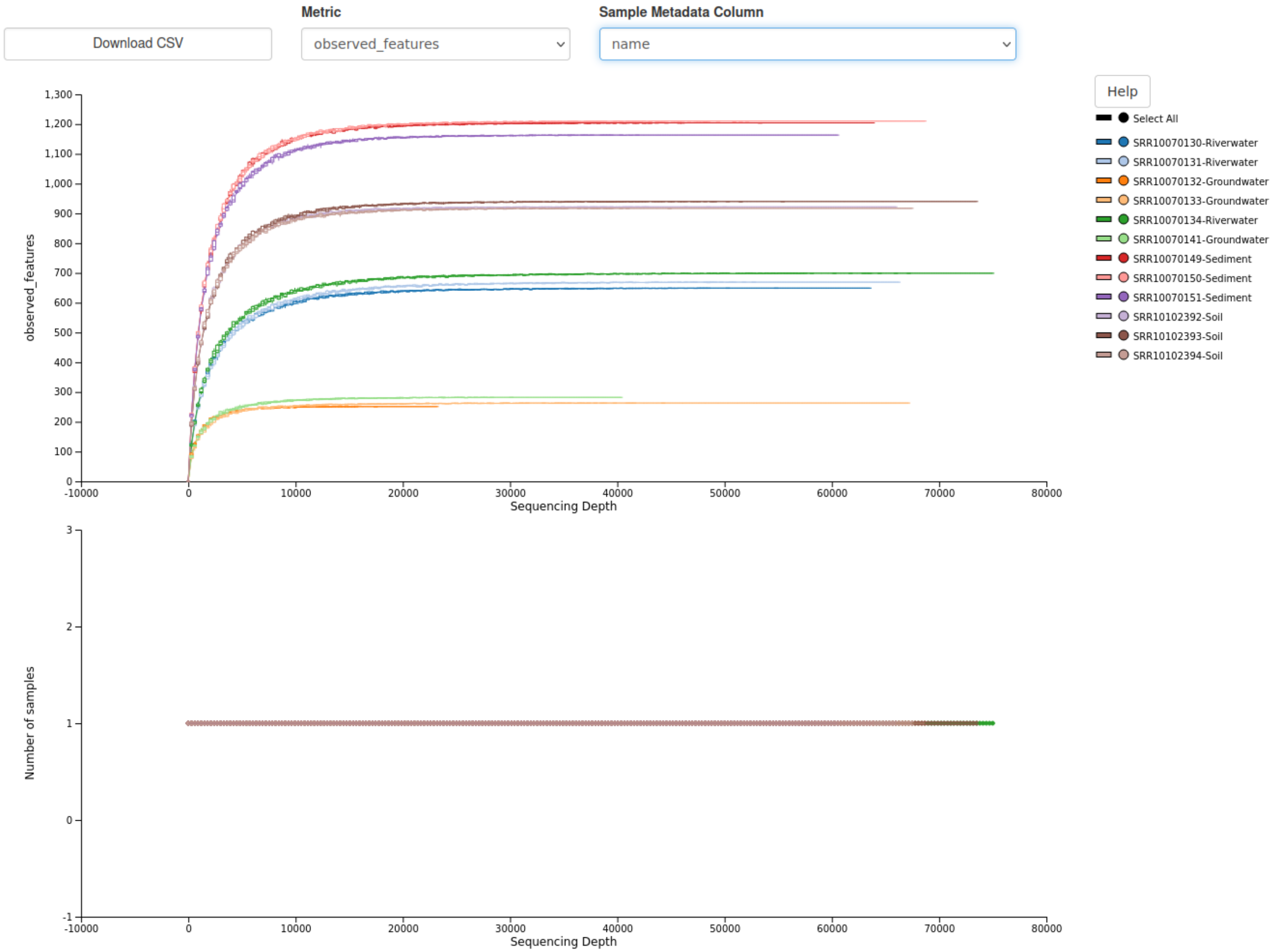
Color Palette

schemeAccent

Bar Width

Hover over the plot to learn more

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Relative Frequency
SRR10070130-Riverwater
SRR10070131-Riverwater
SRR10070132-Groundwater
SRR10070133-Groundwater
SRR10070134-Riverwater
SRR10070141-Groundwater
SRR10070149-Sediment
SRR10070150-Sediment
SRR10070151-Sediment
SRR10102392-Soil
SRR10102393-Soil
SRR10102394-Soil
Sample
Bacteria;Proteobacteria;Betaproteobacteria;Burkholderiales
Bacteria;;;
Archaea;Thaumarchaeota;Nitrososphaerales;Nitrososphaeraceae
Bacteria;Bacteroidetes;Flavobacteriia;Flavobacteriales
Bacteria;Proteobacteria;Betaproteobacteria;Nitrosomonadales
Bacteria;Acidobacteria;Acidobacteria_Gp6;Gp6
Bacteria;Proteobacteria;Gammaproteobacteria;Pseudomonadales
Bacteria;Proteobacteria;Deltaproteobacteria;
Bacteria;Verrucomicrobia;Spartobacteria;
Bacteria;Proteobacteria;Alphaproteobacteria;Rhizobiales
Bacteria;Acidobacteria;Acidobacteria_Gp4;Gp4
Archaea;Thaumarchaeota;Nitrosopumilales;Nitrosopumilaceae
Bacteria;Proteobacteria;Betaproteobacteria;
Bacteria;Proteobacteria;Gammaproteobacteria;
Bacteria;Chloroflexi;Anaerolineae;Anaerolineales
Bacteria;Candidatus_Saccharibacteria;;
Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales
Bacteria;Proteobacteria;;
Bacteria;Proteobacteria;Deltaproteobacteria;Myxococcales
Bacteria;Bacteroidetes;Chitinophagia;Chitinophagales
Bacteria;Firmicutes;Bacilli;Bacillales
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales
Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales
Bacteria;Campilobacterota;Campylobacteria;Campylobacterales
Bacteria;Nitrospirae;Nitrospira;Nitrospirales
Bacteria;Proteobacteria;Betaproteobacteria;Rhodocyclales
Bacteria;Proteobacteria;Gammaproteobacteria;Aeromonadales
Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales
Bacteria;Proteobacteria;Alphaproteobacteria;Rhodospirillales
Archaea;Crenarchaeota;Thermoprotei;
Bacteria;Acidobacteria;Acidobacteria_Gp16;Gp16
Bacteria;Acidobacteria;Blastocatellia;Blastocatellales
Bacteria;Actinobacteria;Thermoleophilia;Gaiellales
Bacteria;Proteobacteria;Gammaproteobacteria;Chromatiales
Bacteria;Verrucomicrobia;Subdivision3;
Bacteria;Actinobacteria;Actinobacteria;Micrococcales
Bacteria;Firmicutes;Clostridia;Clostridiales
Bacteria;Bacteroidetes;Cytophagia;Cytophagales



Download CSV

Metric

observed_features

Sample Metadata Column

name

Help

QIIME2: Output artifacts and visualisations

- *.qza - zip folder, containing data and metadata
- *.qzv - zip folder, containind data, metadata and visualisations

QIIME2: pros & cons

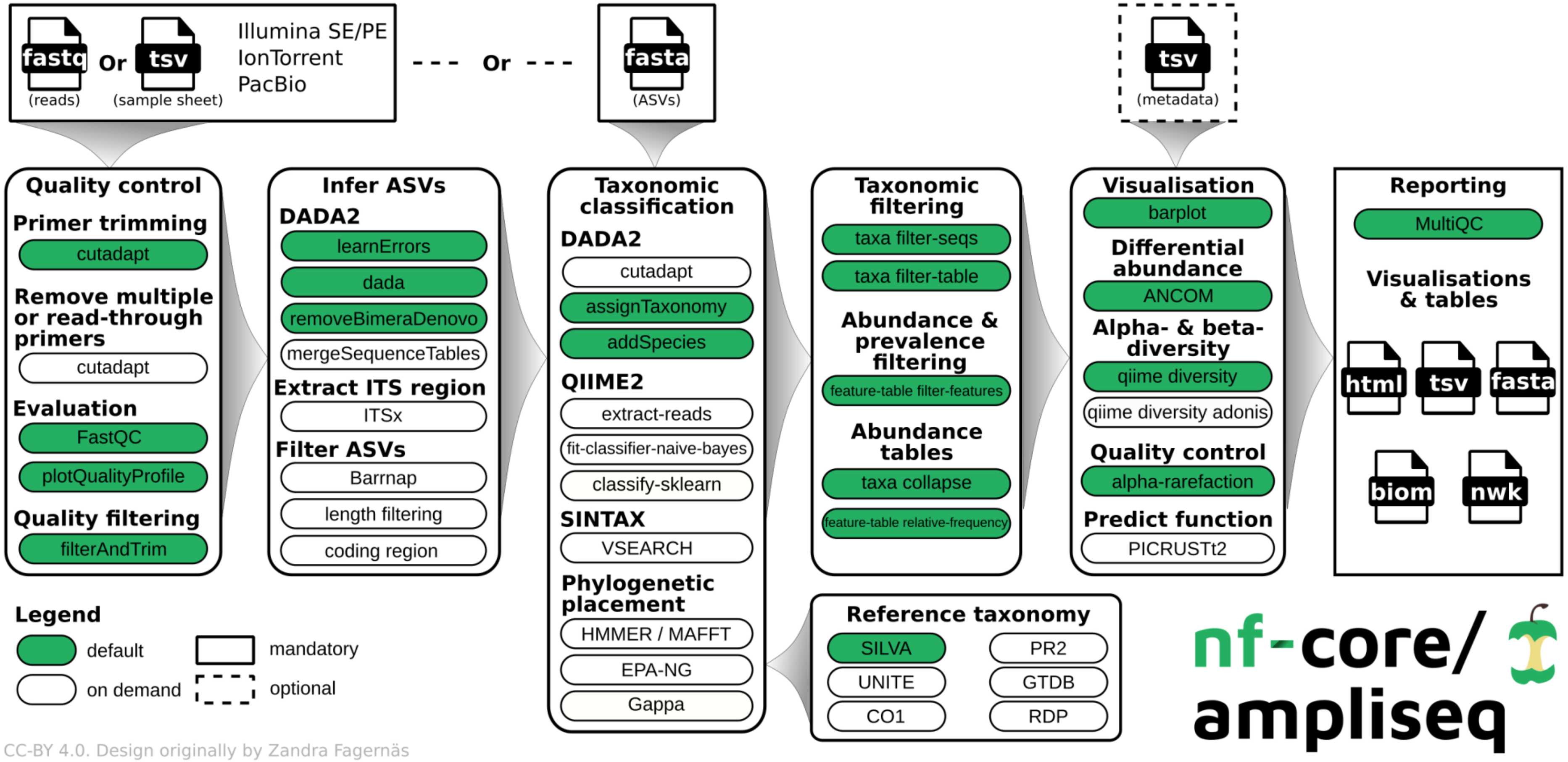
Pros

- Comprehensive pipeline
- Plugins
- GUI
- Modularity

Cons

- Own data types
- Learning curve
- Reference dependent

nf-core/Ampliseq pipeline



fastq (reads) Or tsv (sample sheet)
Illumina SE/PE
IonTorrent
PacBio

--- Or ---

fasta (ASVs)

tsv (metadata)

Quality control

Primer trimming
- cutadapt

Remove multiple or read-through primers
- cutadapt

Evaluation
- FastQC
- plotQualityProfile

Quality filtering
- filterAndTrim

Infer ASVs

DADA2
- learnErrors
- dada
- removeBimeraDenovo
- mergeSequenceTables

Extract ITS region
- ITSx

Filter ASVs
- Barrnap
- length filtering
- coding region

Taxonomic classification

DADA2
- cutadapt
- assignTaxonomy
- addSpecies

QIIME2
- extract-reads
- fit-classifier-naive-bayes
- classify-sklearn

SINTAX
- VSEARCH

Phylogenetic placement
- HMMER / MAFFT
- EPA-NG
- Gappa

Taxonomic filtering
- taxa filter-seqs
- taxa filter-table

Abundance & prevalence filtering
- feature-table filter-features

Abundance tables
- taxa collapse
- feature-table relative-frequency

Visualisation
- barplot

Differential abundance
- ANCOM

Alpha- & beta-diversity
- qiime diversity
- qiime diversity adonis

Quality control
- alpha-rarefaction

Predict function
- PICRUSTt2

Reporting
- MultiQC

Visualisations & tables

html tsv fasta
biom nwk

Reference taxonomy

SILVA	PR2
UNITE	GTDB
CO1	RDP

Legend

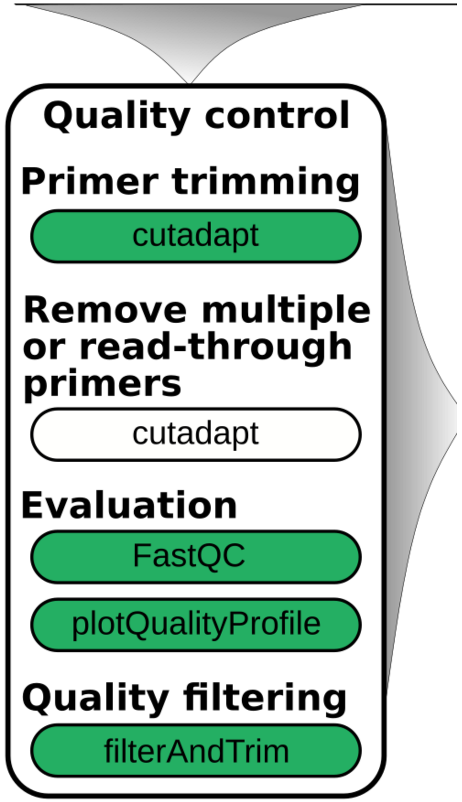
🟩 default ▭ mandatory
○ on demand ⬚ optional

nf-core/ampliseq

CC-BY 4.0. Design originally by Zandra Fagernäs

Ampliseq: quality control

- Data preprocessing
- Check reads quality
- Perform filt&trim



Quality control

Primer trimming

cutadapt

Remove multiple or read-through primers

cutadapt

Evaluation

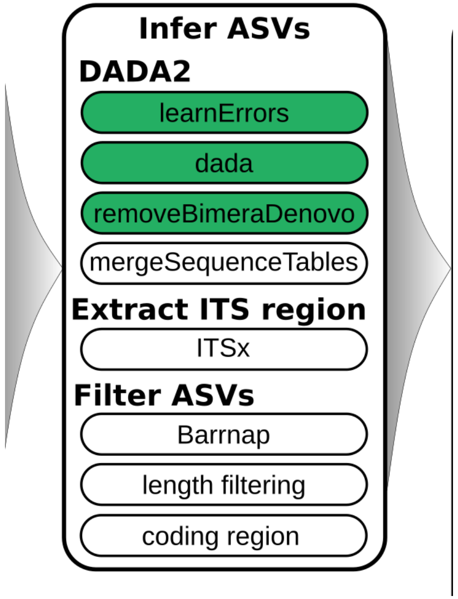
FastQC

plotQualityProfile

Quality filtering

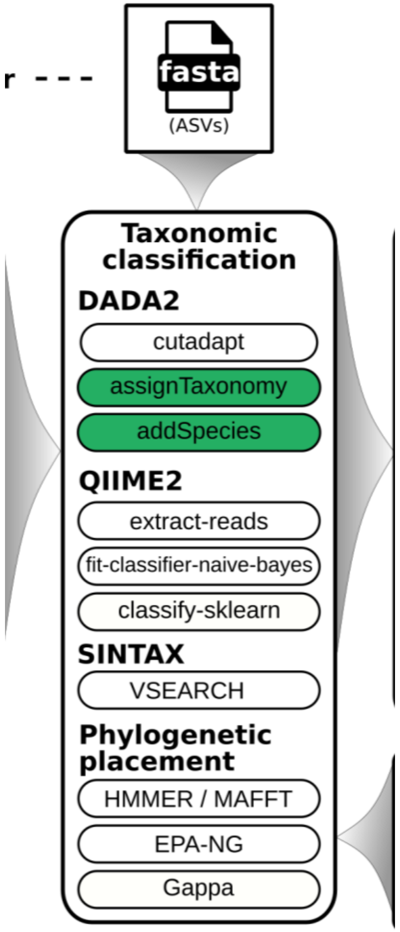
filterAndTrim

Ampliseq: ASVs calculation



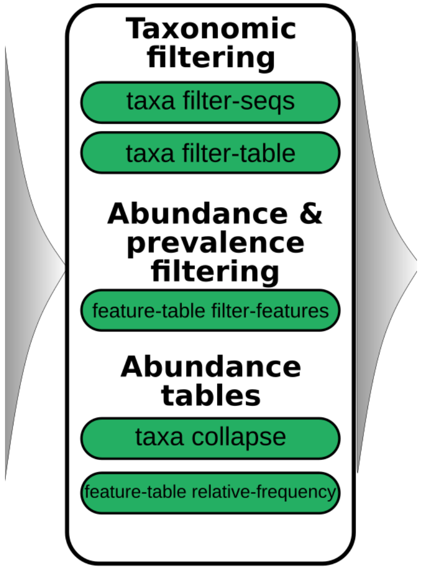
- DADA2
- Error estimation
- Chimera removal
- Contamination removal
- Filtering

Ampliseq: Taxonomic classification



- Database dependent
- Infer species
- Confidence intervals
- Multiple assignment

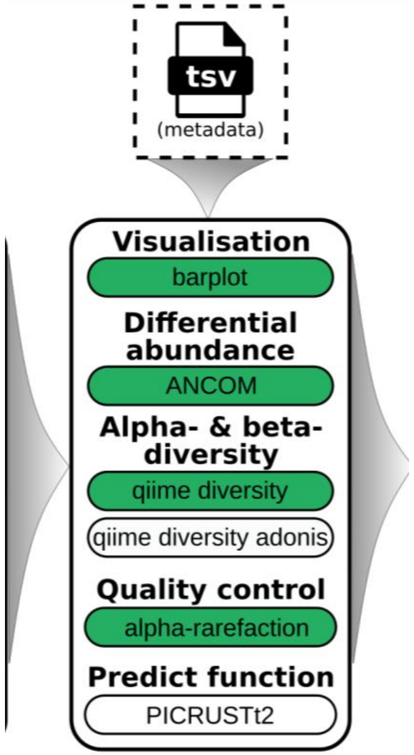
Ampliseq: Taxonomic filtering


Taxonomic filtering
taxa filter-seqs
taxa filter-table
Abundance & prevalence filtering
feature-table filter-features
Abundance tables
taxa collapse
feature-table relative-frequency

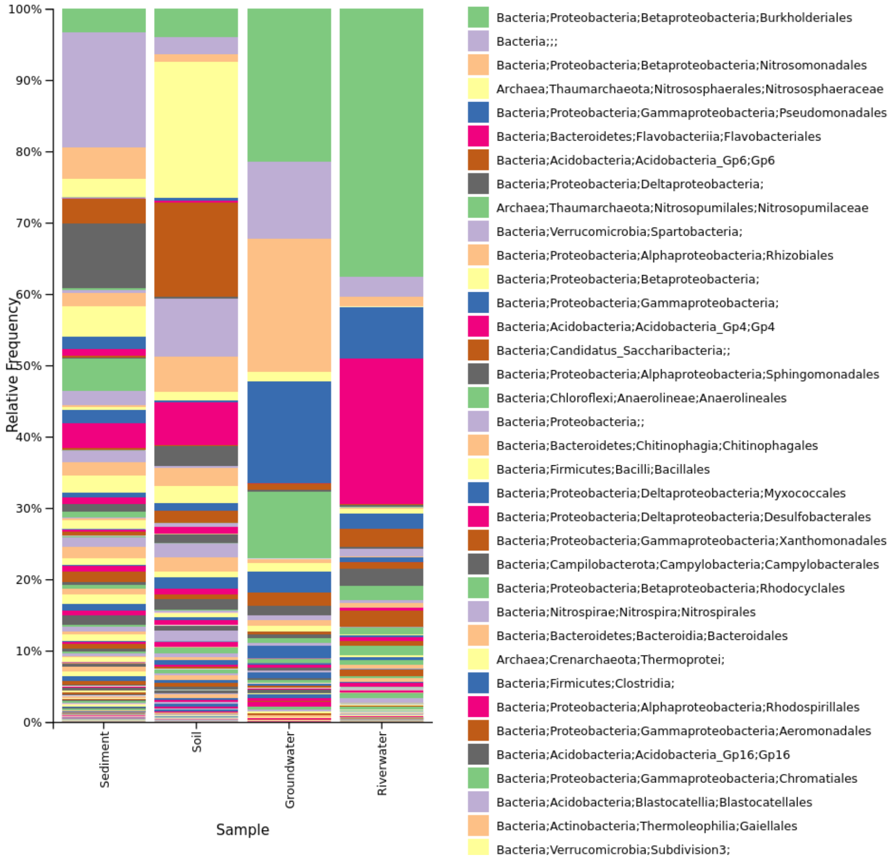
- Filter specific taxa
- Abundance filtering

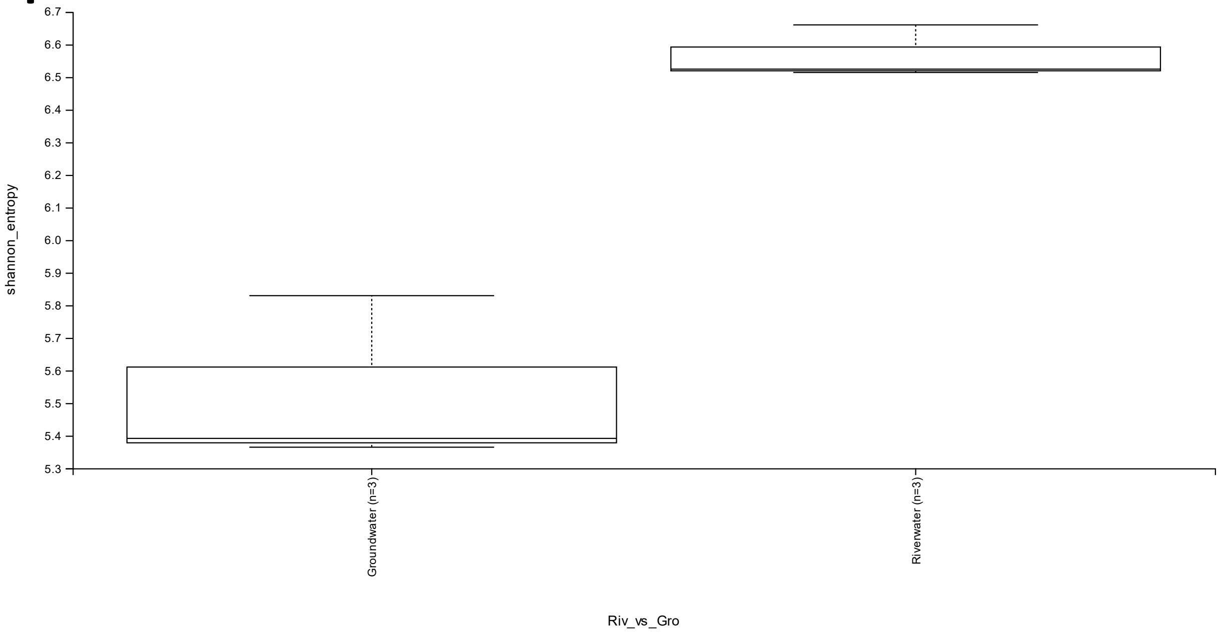
Ampliseq: Post processing

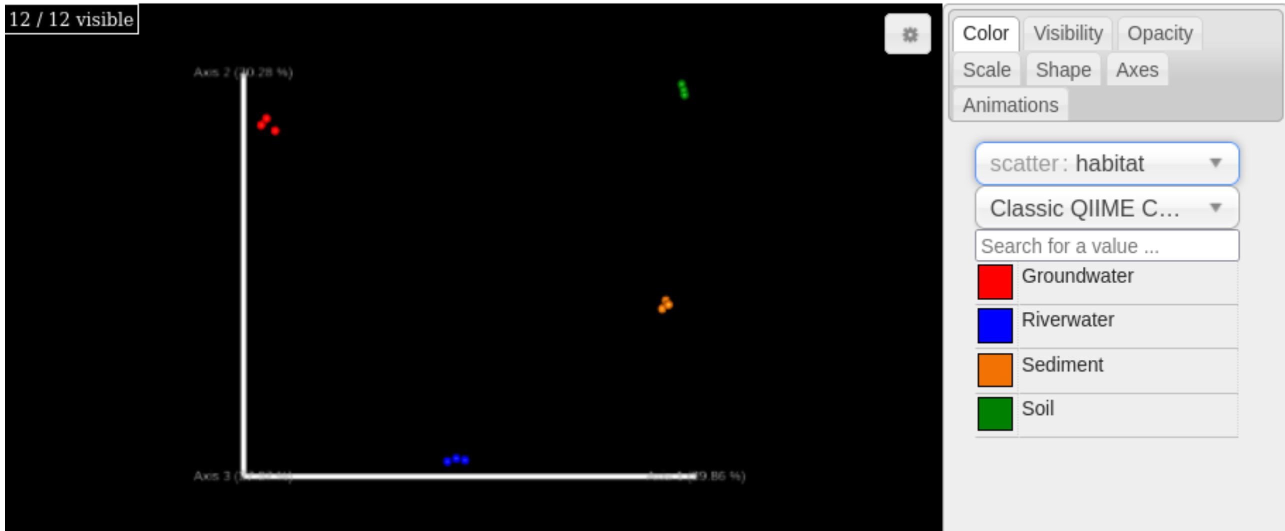
- Visualisation
- Diversity computation
- Functional analysis



Ampliseq: Visualisation





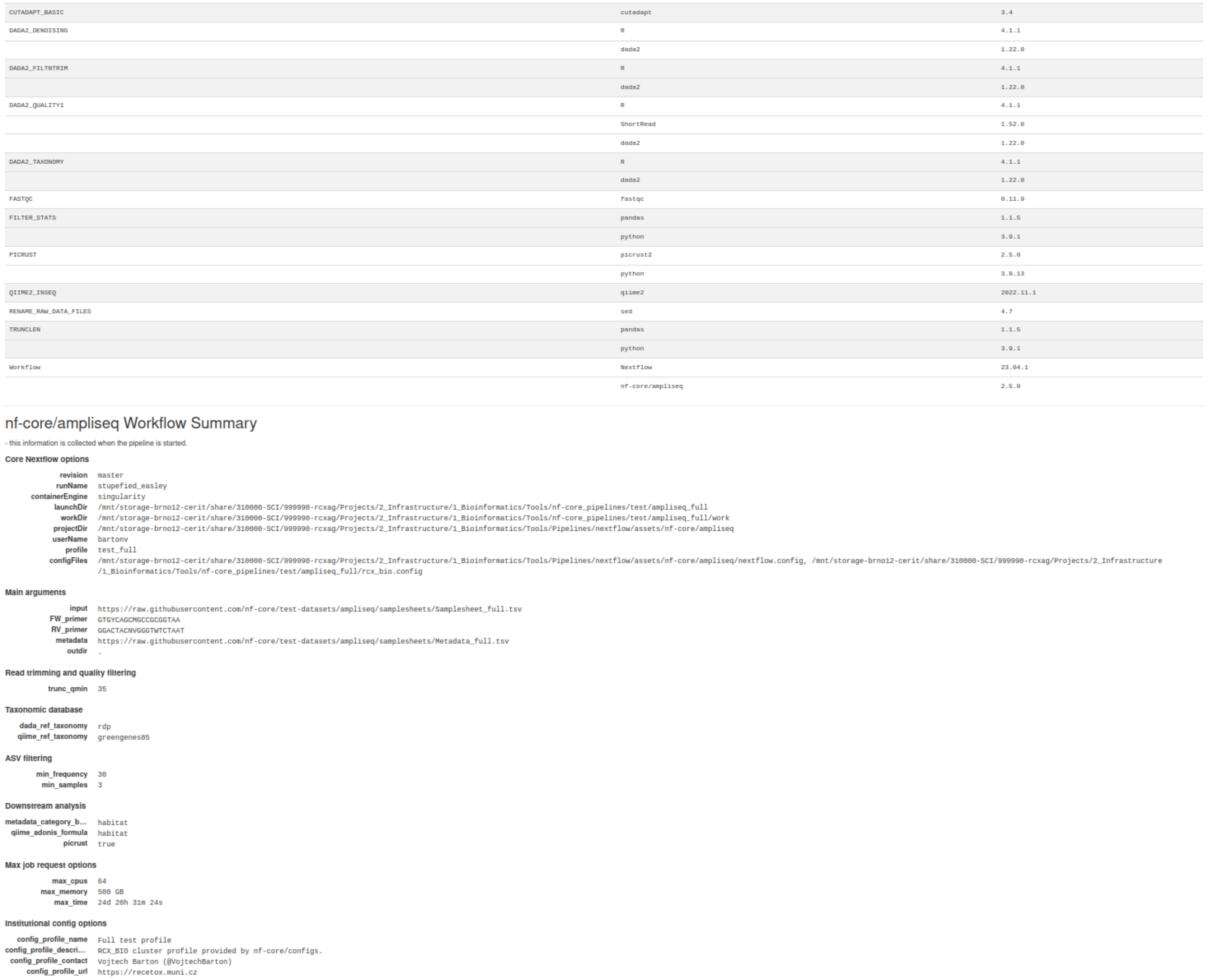


Ampliseq: functional profiling

- Picrust2
- Phylogenetic Investigation of Communities by Reconstruction of Unobserved States

- KEGG and COG database
- Based on phylogeny
- Genes present in microbial genomes are similar amongst relatives
- When sufficient genome sequences are available, it is possible to predict which gene families are present in a given microbial OTU from phylogeny alone.

Ampliseq: Report



CUTADAPT_BASIC	cutadapt	3.4
DADA2_DENOISING	R	4.1.1
	dada2	1.22.0
DADA2_FILTNTRIM	R	4.1.1
	dada2	1.22.0
DADA2_QUALITY1	R	4.1.1
	ShortRead	1.52.0
	dada2	1.22.0
DADA2_TAXONOMY	R	4.1.1
	dada2	1.22.0
FASTQC	fastqc	0.11.9
FILTER_STATS	pandas	1.1.5
	python	3.9.1
PICRUST	picrust2	2.5.0
	python	3.8.13
QIIME2_INSEQ	qiime2	2022.11.1
RENAME_RAW_DATA_FILES	sed	4.7
TRUNCLEN	pandas	1.1.5
	python	3.9.1
Workflow	Nextflow	23.04.1
	nf-core/ampliseq	2.5.0

nf-core/ampliseq Workflow Summary

- this information is collected when the pipeline is started.

Core Nextflow options

revision	master
runName	stupefied_easley
containerEngine	singularity
launchDir	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/nf-core_pipelines/test/ampliseq_full
workDir	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/nf-core_pipelines/test/ampliseq_full/work
projectDir	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/Pipelines/nextflow/assets/nf-core/ampliseq
userName	bartonv
profile	test_full
configFiles	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/Pipelines/nextflow/assets/nf-core/ampliseq/nextflow.config, /mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/nf-core_pipelines/test/ampliseq_full/rcx_bio.config

Main arguments

input	https://raw.githubusercontent.com/nf-core/test-datasets/ampliseq/samplesheets/Samplesheet_full.tsv
FW_primer	GTGYCAGCMGCCGCGGTAA
RV_primer	GGACTACNVGGGTWTCTAAT
metadata	https://raw.githubusercontent.com/nf-core/test-datasets/ampliseq/samplesheets/Metadata_full.tsv
outdir	.

Read trimming and quality filtering

trunc_qmin	35

Taxonomic database

dada_ref_taxonomy	rdp
qiime_ref_taxonomy	greengenes85

ASV filtering

min_frequency	30
min_samples	3

Downstream analysis

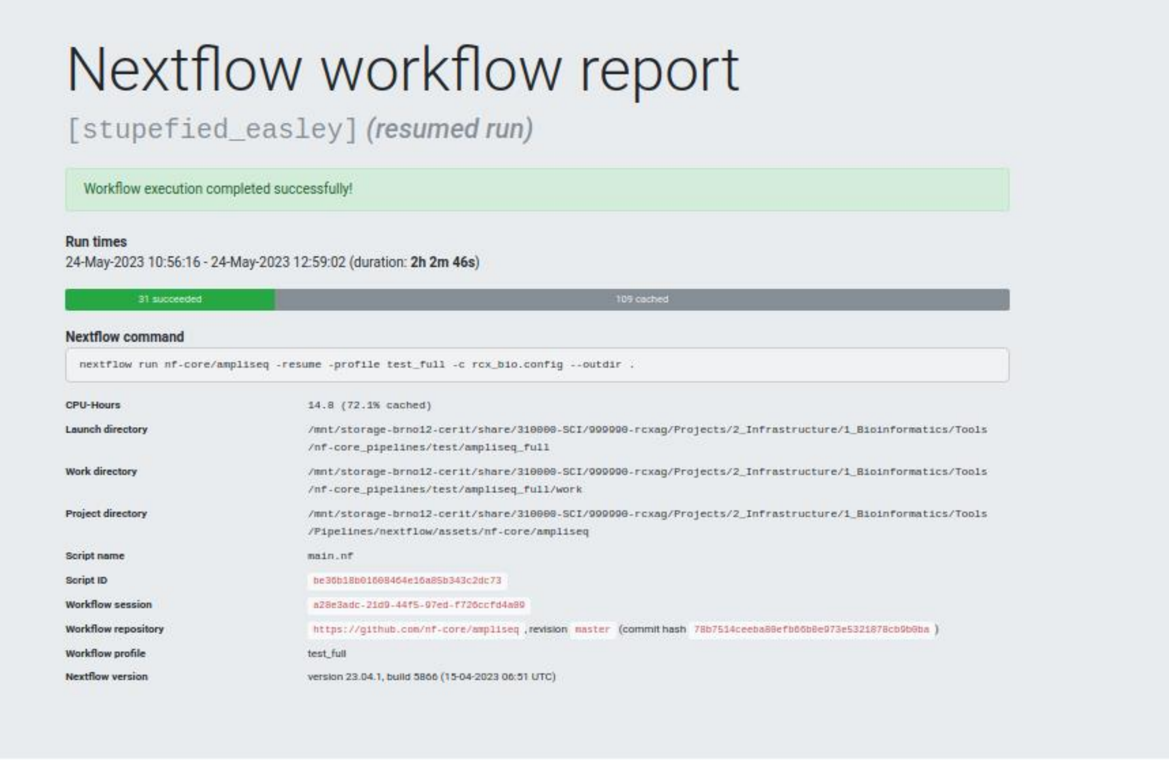
metadata_category_b...	habitat
qiime_adonis_formula	habitat
picrust	true

Max job request options

max_cpus	64
max_memory	500 GB
max_time	24d 20h 31m 24s

Institutional config options

config_profile_name	Full test profile
config_profile_descri...	RCX_BIO cluster profile provided by nf-core/configs.
config_profile_contact	Vojtech Barton (@VojtechBarton)
config_profile_url	https://recetox.muni.cz



Nextflow workflow report

[stupefied_easley] *(resumed run)*

Workflow execution completed successfully!

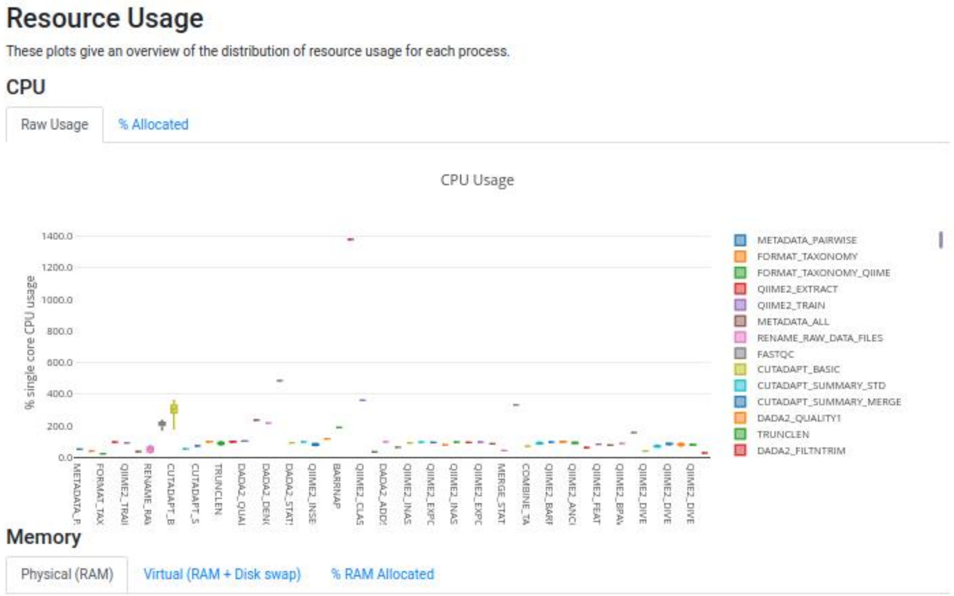
Run times

24-May-2023 10:56:16 - 24-May-2023 12:59:02 (duration: **2h 2m 46s**)

Nextflow command

```
nextflow run nf-core/ampliseq -resume -profile test_full -c rcx_bio.config --outdir .
```

CPU-Hours	14.8 (72.1% cached)
Launch directory	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/nf-core_pipelines/test/ampliseq_full
Work directory	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/nf-core_pipelines/test/ampliseq_full/work
Project directory	/mnt/storage-brno12-cerit/share/310000-SCI/999990-rcxag/Projects/2_Infrastructure/1_Bioinformatics/Tools/Pipelines/nextflow/assets/nf-core/ampliseq
Script name	main.nf
Script ID	be36b18b01608464e16a85b343c2dc73
Workflow session	a28e3adc-21d9-44f5-97ed-f726ccfd4a89
Workflow repository	https://github.com/nf-core/ampliseq , revision master (commit hash 78b7514ceeba89efb66b0e973e5321878cb9b0ba)
Workflow profile	test_full
Nextflow version	version 23.04.1, build 5866 (15-04-2023 06:51 UTC)



Resource Usage

These plots give an overview of the distribution of resource usage for each process.

CPU

Memory

Ampliseq: pros & cons

Pros

- Standardized
- Easy to run
- Comprehensive analysis
- Community-driven

Cons

- Learning curve
- Resource intensive
- Needs setup
- Software versions dependent