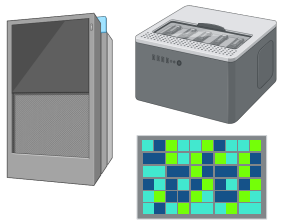
Bioinformatics pipeline for analysis of long read metagenome data

Generation metagenomic data and QC



Sequencing

RAW
Seq
Qual

Base call

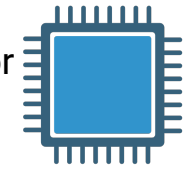
transform

RAW
FastQ

Data QC

Pass
FastQ

Simulator

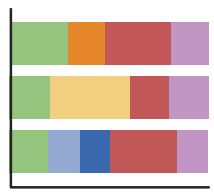


Remove host
sequence

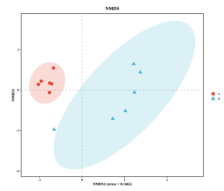
Analyses with taxonomy

Clean
FastQ

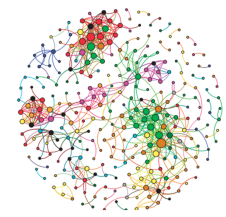
Classify



Diversity



Correlation



Analysis with MAGs

Assembly
Long-read only
Hybrid

Primary
Contig

Polish

Long read
Short read

Polished
Contig

Binning

Long-read
MetaHI-C
Methylation

MAGs

Quality evaluation

Taxa classificaion

Phylogeny inference

SV and popluation

Genome catalogue

Analysis with Gene

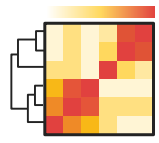
Gene prediction

Gene

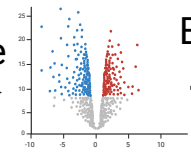
Cluster

Non-
redundant
gene

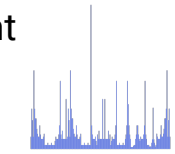
Quantification



Diffenence



Enrichment



Annotation

Functional
Taxonomy
MGEs
BGCs et...

Genic or genomic
catalogue

Databases