MAG Dereplication, Pangenomes & Population Genomics — Hands-on

Build non redundant genome sets and explore strain level diversity from metagenome assembled genomes (MAGs). This module covers MAG dereplication using ANI based clustering, construction of pangenomes, gene presence absence matrices and core population genomics summaries that connect MAGs back to microbiome ecology and phenotypes.

MAG Dereplication, Pangenomes & Population Genomics

Help Desk · WhatsApp

Session Index

Session 1 — MAGs, Quality & Redundancy Concepts
Session 2 — Dereplication, ANI Clustering & dRep Style Workflows
Session 3 — Pangenomes, Gene Presence–Absence & Population Structure
Session 4 — Mini Capstone: Non Redundant MAG Set & Pangenome Story

Session 1

Fee: Rs 8800 Apply Now

MAGs, Quality & Redundancy Concepts

Recap of MAG reconstruction and QC

bins to MAGs overview **completeness and contamination thresholds** **basic quality tiers (high, medium, low)**

Why dereplicate MAG collections

inflation of near duplicate genomes **downstream bias**

in diversity estimates | need for representative non redundant sets

Distance measures and ANI basics

average nucleotide identity (ANI) concepts | species level ANI cut offs | fast ANI approximations overview

Session 2

Fee: Rs 11800 Apply Now

Dereplication, ANI Clustering & dRep Style Workflows

Setting up dereplication workflows

input requirements and genome QC filters | pairwise similarity matrices overview | clustering strategy and thresholds

dRep style dereplication concepts

primary and secondary clustering | choosing representatives per cluster | scoring by completeness and contamination

Outputs: non redundant sets and mapping back to samples

cluster reports and representative lists | sample coverage of each representative | simple abundance matrices for MAG representatives

Session 3

Fee: Rs 14800 Apply Now

Pangenomes, Gene Presence–Absence & Population Structure

Pangenome concepts for MAG clusters

core, accessory and unique gene sets | gene presence absence matrices | functional annotation of gene families overview

Population genomics summaries from MAG sets

simple SNP and allele frequency ideas **strain level diversity metrics overview** **linking gene content to phenotypes**

Visualising pangenomes and population structure

gene presence absence heatmaps **PCA and clustering on gene content** **simple trees or networks for strain groups**

Session 4

Fee: Rs 18800 Apply Now

Mini Capstone: Non Redundant MAG Set & Pangenome Story

Designing a dereplication and pangenome analysis

guided theory plus practical

From MAG collection to non redundant set and pangenome outputs

run one dereplication workflow **build gene presence absence matrix** **summarise core and accessory content**

Deliverables: figures, tables and short narrative on strain level structure

pangenome heatmaps and cluster plots **summary of non redundant MAG representatives** **brief report relating strain patterns to ecology or metadata**