Hands-on course on shotgun metagenomics data analysis

Proposed content

Day 1: Classification of metagenomics data

9:00 am - 9:30 am

Welcome and Introduction

- Overview of the course and learning objectives
- Icebreaker: Participant introductions and their research focus
- Brief introduction to shotgun metagenomics: concepts and applications

9:30 am - 10:30 am

Session 1: Basics of Shotgun Metagenomics Data

- Structure and format of raw sequencing data (FASTQ, quality scores)
- Overview of experimental design: Importance of controls, replicates, and metadata
- Hands-on: Exploring and assessing raw data quality using FastQC

10:30 am - 10:45 am

Break

10:45 am - 12:00 pm

Session 2: Preprocessing Shotgun Metagenomics Data

- Quality control and trimming: Adapters, low-quality reads, and minimum length
- Removing host contamination using tools like Bowtie2 or Centrifuge
- Hands-on: Trimming reads with Trimmomatic and removing host contamination

12:00 pm - 1:00 pm

Lunch Break

1:00 pm - 2:30 pm

Session 3: Taxonomic Profiling

- Assigning taxonomy: Kraken2, Kaiju, MetaPhlAn

- Effect of sequencing depth
- Hands-on: Running Kraken2 on cleaned data
- Interpreting outputs: Taxonomic summary tables

2:30 pm - 3:00 pm

Discussion/Q&A

- Review of Day 1 concepts
- Addressing questions and issues encountered during hands-on sessions

3:00 pm - 3:15 pm

Break

3:15 pm - 4:00 pm

Session 4: Functional Profiling Introduction

- Overview of functional annotation: Understanding pathways and genes
- Hands-on: Introduction to HUMAnN or similar tools

Day 2: Advanced Analysis and Visualization

9:00 am - 10:30 am

Session 5: Functional Profiling (Continued)

- Hands-on: Running HUMAnN on taxonomic profiling outputs
- Interpreting pathway abundance and gene family abundance
- Data normalization and pathway reconstruction

10:30 am - 10:45 am

Break

10:45 am - 12:00 pm

Session 6: Comparative Metagenomics

- Comparing multiple samples: Beta diversity, statistical tools (e.g., ANCOM, LEfSe)
- Hands-on: Comparing groups using tools in R (Microbiome package)

12:00 pm - 1:00 pm

Lunch Break

1:00 pm - 2:30 pm

Session 7: Visualizing Metagenomics Data

- Creating publication-ready plots: Bar plots, heatmaps, PCA
- Hands-on: Visualizing taxonomic and functional data with R packages (e.g., ggplot2, phyloseq)

2:30 pm - 3:00 pm

Discussion/Q&A

- Troubleshooting and advanced tips
- Review of visualization tools and interpretation

3:00 pm - 3:15 pm

Break

3:15 pm - 4:00 pm

Session 8: Workflow Automation and Resources

- Workflow managers: Snakemake, Nextflow
- Discussion: Best practices and available resources for metagenomics
- Final wrap-up and feedback