



Whole Genome sequencing and Bioinformatics Training

4-7 March 2025

Lilongwe, Malawi

Day 1 – Introduction to WGS and Online Bioinformatics Basics

Focus: Setting the stage, understanding sequencing technologies, performing BLAST and exploring different online tools.

Morning Session 9:00-12:00: Welcome and Overview

- **Introductions & Goals:**
 - o Welcome and participant introductions.
 - o Outline training goals.
- **WGS Overview:**
 - o Overview of whole genome sequencing (WGS) technologies and their laboratory roles.
 - o Review basic bioinformatics concepts: file formats and sequence data types.
- **Hands-On Exercise:**
 - BLAST Mystery Species Identification: Use BLAST to identify a mystery species.
 - Emphasize effective communication between the lab and bioinformatics teams.

Lunch 12:00-13:00

Afternoon Session 13:00-17:00: Introduction to Online Tools

- **Demonstrations**
 - o Overview of browser-based platforms such as BLAST, Galaxy, Epi2me, PathogenWatch, CGE, “Assembly Server”, Illumina Dragen, Terra Bio, etc.
 - o Walk-through of each platform’s interface and navigation tips.
 - o Brief demo illustrating a command-line workflow (e.g., running an Epi2me workflow or assembly).
 - o Showcase Marco’s bacterial assembly pipeline.
- **Hands-On Exercise:**
 - o Explore various online platforms through guided exercises.

Day 2 – Quality Control (QC) and Assembly

Focus: Understanding data quality, performing QC on raw reads, and exploring assembly theory and practice.

Morning Session 9:00-12:00: Quality Control (QC) of Raw Reads

- **Presentation:**
 - o Concepts of data quality and its impact on downstream analysis.
 - o Common issues: adapter contamination and low-quality reads.
- **Demonstration:**
 - o Review examples of good and poor run reports from Oxford Nanopore Technologies (ONT).
- **Hands-On Exercises:**
 - o Use online QC tools (e.g., FastQC and ONT QC tool) to assess sample sequencing data.
 - o Review pre-loaded sample datasets and interpret QC reports.

Lunch 12:00-13:00

Afternoon Session 13:00-17:00: Assembly Theory, Exercises, and QC

- **Presentation:**
 - o Assembly Approaches:
 - Reference-based vs. de-novo assembly (including de Bruijn graphs).
 - Issues such as contamination and high contig counts.
 - o Assembly Types:
 - Long, short, and hybrid assemblies.

- o Tools like BV-BRC for assembly.
 - o Quality Assessment:
 - QC tools such as QUAST, MultiQC, CheckM, and BUCO.
- **Hands-On Exercises:**
 - o Generate assemblies (ideally using participants' own data) via the BV-BRC Assembly Service.
 - o Evaluate assembly quality using QC reports (reference QC reports from NICD as examples).
 - o Run assemblies through BLAST to further assess their quality.

Day 3 – Typing and AMR Analysis Using Online Platforms

Focus: Understanding antimicrobial resistance (AMR) fundamentals and using online tools for bacterial typing and AMR gene detection.

Morning Session 9:00-12:00: Typing & AMR Fundamentals

- **Presentation:**
 - o Compare traditional lab-based typing methods with bioinformatics approaches.
 - o **Typing Methods:**
 - KmerFinder (including kmer theory), MLST (using tools like KMA), serotyping, PlasmidFinder, VirulenceFinder.
 - Additional tools: VirulenceFinder and PlasmidFinder.
- **Hands-On Exercise:**
 - Review Marco's slides on building a ResFinder/PlasmidFinder workflow.
 - Build your own Resfinder/PlasmidFinder with MyDBFinder (upload section from a 16S reference database, download from NCBI).
 - AMR package for R - determining MDR

Lunch 12:00-13:00

Afternoon Session 13:00-17:00: Hands-On AMR Exercise

- **Review & Discussion:**
 - Revisit CGE results and discuss findings.
- **AMR Overview Presentation:**
 - Clinical significance and mechanisms of resistance (briefly).
 - Importance of surveillance and the genotype-to-phenotype correlation in AMR reporting (include Marco's story on cholera).
- **Hands-On Exercises:**
 - Use web-based tools for AMR gene detection (e.g., ResFinder, PathogenWatch, MEFinder, PlasmidFinder).
 - Discuss interpretation of results and the importance of consensus among tools (e.g., comparing ResFinder, AMRFinderPlus, CARD; demo Harmonize if possible).
 - Examine HTML reports generated from assemblies.

Day 4 – Bacterial Typing and Phylogenetic Analysis via Online Tools

Focus: Leveraging online platforms for bacterial typing and exploring phylogenetic relationships.

Morning Session 9:00-12:00: Bacterial Typing Overview

- **Presentation - Overview of Typing Methods:**
 - Introduction to methods such as MLST and cgMLST and their applications in tracking AMR epidemiologically.
 - Discussion on SNP and cgMLST analysis (kmer distance vs. SNPs vs genes/alles).
- **Additional Topics:**
 - Use of multiple sequence alignment (MSA) and tree-building tools.
 - Overview of tools such as CSI Phylogeny, MinTyper, and BEAST (including rooted trees, bootstrapping, and clock/time analysis).
- **Demonstration:**
 - Live demo using Enterobase.

Lunch 12:00-13:00

Afternoon Session 13:00-17:00: Hands-On Phylogenetic Analysis & Simulated Outbreak Scenario

- **Hands-On Exercise:**
 - Participants work with sample datasets to generate and visualize phylogenetic trees using online tools such as GrapeTree, Microreact, and iTOL.
- **Theory & Discussion:**
 - Interpreting tree topology and understanding outbreak clusters.
 - Discuss practical implications for AMR surveillance.
- **Simulated Outbreak Scenario**
 - **Tasks:**
 - Identify AMR genes using a web-based tool.
 - Perform bacterial typing and generate a phylogenetic tree.
 - Interpret the integrated results to propose potential clinical and epidemiological actions.
- **Wrap-Up:**
 - Q&A session and final discussion.