WORKSHOP ON BIOINFORMATICS APPLICATIONS IN GENOMICS SURVEILLANCE OF BACTERIAL ANTIMICROBIAL RESISTANCE

CURRICULUM

Summarised curriculum plan

Item	Notes
1. Training Title	Bioinformatics applications in genomics surveillance of bacterial antimicrobial resistance
2. Goals/aims	**Goal:** To build human capacity for analysis of genomics data to support antibiotic resistance surveillance in Cameroon and Central Africa **Objectives:** - To deliver 5 training courses per year to early career scientists in application of bacterial pathogen genomic surveillance tools. - To equip next-generation researchers with bioinformatics skills to support bacterial genomics research in Cameroon - To establish an African network of scientists conducting surveillance of bacterial AMR
3. Target audience	- Early career researchers with background in microbiology, molecular biology, and related biology fields (PhD student, PhD holder under 3 years) - Good knowledge or interest in antibiotic resistance and bacterial genetics - No previous prior experience in programming is required - Participants willing to learn bacterial bioinformatics and how to analyse NGS (Illumina and Nanopore) bacterial genomes can attend to the course. - Having sequencing data will be an advantage - The course is for 25 learners
4. Learning outcomes	At the end of this course, learners should be able to: - Recognise the main bioinformatics data file formats after sequencing and have an overview on the different types of sequencing technologies - Use the biological databases to perform gene-based analysis - Apply online bioinformatics tools to analyse genomic data starting from the quality control to phylogenetic analyses - Interpret bacterial genomics data - Run a pre-written script to automate bioinformatics tasks
5. Logistic information	- The training will take place in Yaoundé, Cameroon within the premise of the CEDBCAM-RI quarterly as from February 2023 whereby recorded video will be shared and live zoom lecture delivery will be done for international trainers. - The course will be face-to-face to facilitate interaction with and support of the participants. - Trainers will be certified researchers from CEDBCAM-RI, Inqaba Biotech, MboaLab, the University of Yaoundé I and the University of Dschang, Sanger Institute - The course will run for 3 days. - A minimum registration fees will be required to attend the training

6.	Potential challenges	The main anticipated challenges are: - Network issue. It will be addressed through the procurement of bundles from different network providers. We will also use recorded video that will be shared to all, and a set of already run experiments for practical with outputs - Lack of computer room on site which may not allow trainers to fully enjoy the course if they lack a good computer. The web-based tool will require a normal computer - Lack of trainers (We will make effort to invite some experienced trainers from H3ABioNet, Wellcome Sanger Institute and other network to give lectures - Poor energy supply in the country may also disrupt the course. A generator is available on site to cope with this. - Dropouts of some participants (We will try to have little number of trainees about 25 and close monitoring and follow up for the trainee will be enhanced)
7.	Content and Assessment	Inaugural lecture: Module 1. Introduction to genomics and bioinformatics Module 2. Quality control and Basic Local Alignment Tool (BLAST) Module 3. Online-based tools for bacterial genotyping Module 4. Introduction to phylogenetics, phylogenetic analysis and phylogenetic tree visualization Module 5. Introduction to Linux The course does not cover hands-on experience on sequencing platforms The learning outcomes will be assessed with homework, workgroup exercises, discussion, quizzes and a final test
8.	Activities	Activities - Lecture presentations - Quiz + puzzle - Group discussion - Case-based learning - Hands-on computation practical analysis of NGS data - Series of seminars by invited speakers, who will highlight their cutting-edge work harnessing the power of next generation sequencing technologies to address a wide range of biological questions for pathogens.
		Resources - Web-browsers (PathogenWatch, Center for Genomic Epidemiology, BactGen, NCBI) - We will prepare the tailored handouts, slides and notice - Links - We have NGS dataset and those open access data archived in NCBI
		Module 1 & 2: 22 February 2023 Module 3 & 4: 23 February 2023 Module 5 : 24 February 2023
9.	Delivery of training/ session/presentation	I will deliver the training with collaborators within and outside my institution. We are the best to deliver the training as we have all been trained through several bioinformatics workshops and courses. We have previous experience in training and analyses of our genomic data independently. The unique skills and expertise required to be trainers are : - Expertise in bioinformatics and antibiotic resistance - Understanding of adult learning

10. Assessment	The learning outcomes will be assessed with homework exercise, group work, discussion/reflection, quiz, test, and a final project
11. Facilities, infrastructure and resources	A large conference room will be used for the training within the CEDBCAM-RI. The centre also has good internet network to facilitate hands-on practicals online. Projectors, printers and papers for assignment and group discussion. All participants will come along with their laptop as the centre does not have a computer lab.
12. Feedback and Evaluation	
13. Additional information, thoughts or ideas	

Detailed curriculum plan

Training: Bioinformatics applications in genomics surveillance of bacterial antimicrobial resistance

Summary

Antimicrobial resistance is a global public health concern. With the advent of whole genome sequencing (WGS) and the collection of sufficiently large datasets of various isolates, it is now plausible to use bioinformatics to gain new insights in the more complex molecular mechanisms of antibiotic resistance. This training aims to equip next-generation researchers with bioinformatics skills to support genomics surveillance of antibiotic resistance in Cameroon. It will introduce participants to data file formats used in bacterial sequencing, web-based browsers to analyse genomic data, how to predict the resistance phenotype of bacteria, and how to share data so that others around the world can benefit. These skills are the building blocks for scaling up genomics to antibiotic resistance surveillance in Cameroon.

Goal

To build capacity for analysis of genomics data to support antibiotic resistance surveillance in Cameroon

Objectives

- To deliver 5 bioinformatics modules applied to bacterial genomic surveillance to early career scientists.
- To equip next-generation researchers with bioinformatics skills to support bacterial genomics research in Cameroon
- To establish a network of scientists conducting genomic surveillance of bacterial AMR

Overall learning outcomes:

At the end of this course, participants should be able to:

- Describe the bioinformatics workflow for bacterial genome analyses and NGS data file formats
- Use the biological databases to perform gene-based analysis
- Apply online bioinformatics tools to analyse genomic data
- Interpret bacterial genomics data
- Run a pre-written script to automate bioinformatics tasks

Topics: Antibiotic resistance, Bacterial genetics, bioinformatics, Sequencing

Target audience

- Early career researchers based in Cameroon with background in microbiology, molecular biology, and related fields
- Participants willing to learn bacterial bioinformatics and how to analyse NGS (Illumina and Nanopore) bacterial genomes can attend to the course.
- The course is for 25 learners

Course pre-requisites:

- Background in microbiology, molecular biology, and related fields
- Good knowledge or interest in antibiotic resistance, bacterial genetics
- No previous prior experience in programming is required

Module developers – Lead/Co-Leads:

Appendix *(include any guideline documents related to content development such as breaking down content for remote classroom format, how to write LO's, designing assignments etc…)*

Timetable/schedule: 3 Days physical course

Contents:

- Module 1. Introduction to genomics and bioinformatics
- Module 2. Quality Control and Basic Local Alignment Tool (BLAST)
- Module 3. Online-based tools for bacterial genotyping
- Module 4. Introduction to phylogenetics and phylogenetic analysis
- Module 5. Introduction to Linux

Module 1: Introduction to genomics and bioinformatics

- **Module Lead/Co-lead:**

- **Summary or objectives**

- The role of genomics and bioinformatics
- Application of bioinformatics to infectious diseases and AMR containment
- To recognise the job opportunity related to this field

- **List of learning outcomes specific for this module**

At the end of this module, participants should be able to:

- Define key theoretical concept of bioinformatics
- Recognise NGS technologies and NGS data file format
- Understand the sequencing workflow
- Use biological databases and resources to search and retrieve file

- **Total number of hours/days to be spent on this content:** 3 hours

- **Tools/software and resources to be used**

  - **Databases: NCBI, EMBL, DDJI, UniProt,**

  - **Description of datasets to be used:** FastA and GenBank files of ST398 LA-MRSA strain PJFA-521M

  - **Other resources, or readings**

- **Overview of activities and exercises** (list specific practical activities which participants will do)
  Introduction lecture: Introduction to genomics for bacterial antimicrobial resistance
  Exercise/Activity 1: Discussion on the selection of sequencing platforms based on application
      Exercise 2: Discussion on sequencing workflow
      Exercise 3: Search, Retrieve and export files on NCBI

- **Assessment**
  Assignment 1: (i) Place the sequencing steps in order, (ii) select the NGS platform based on various applications
  Assignment 2: Search, Retrieve and export files on NCBI
  Quiz 1: List the biological databases and their use
  Quiz 2 : What are the different NGS data file format and their specificities

- **Competencies/outcomes**
  If you are comfortable with competencies, please list the competency/ies this module will address: data retrieval and searching,

Module 2: Basic Local Alignment Tool (BLAST)

- **Module Lead/Co-lead**

- **Summary or objectives**

- Sequence alignment theory and applications
- BLAST theory and practice
- Overview of primer design

- **List of learning outcomes specific for this module**

At the end of this module, participants should be able to:

- Query a variety of sequences on NCBI BLAST
- Interpret the information conveyed in NCBI BLAST search outputs and infer their significance
- Examine the annotations of reported matches and their provenance
- Export BLAST results

- **Total number of hours/days to be spent on this content:** 3 hours

- **Tools/software and resources to be used**

  o **Software: none**

  o **Description of datasets to be used**: NCBI BLAST, UniProt

- **Overview of activities and exercises**
  Introduction lecture:
  Exercise/Activity 1: Workgroup on sequence alignment
  Exercise/Activity 1: Case-based learning Fastqc and Multiqc
  Exercise 2: Practical BLAST with selected nucleotide query sequences
  Exercise 3: Practical BLAST with selected protein query sequences

- **Assessment**
  Assignment 1: Run BLASTn to retrieve the nucleotide sequence related to a resistance gene
  Assignment 2: Run BLASTx to retrieve the protein expressed by a resistant bacteria
  Quiz 1 : What are the common BLAST programs
  Quiz 2 : Name two different web-based BLAST programs. What type of input sequence & target database do they each require?

- **Competencies/outcomes**
  If you are comfortable with competencies, please list the competency/ies this module will address:

Module 3: Online-based tools for bacterial genotyping

- **Module Lead/Co-lead**

- **Summary or objectives**

- Analyse the quality of genomic data with fastqc
- Explore web-based tools as a method to access publicly available genomes and analyse genomic data

- **List of learning outcomes specific for this module**

- Perform quality check with fastqc
- Browse publicly available genomes and associated metadata.
- Use online tools to analyse genome sequence data
- Interpret output of *in silico* prediction

- **Total number of hours/days to be spent on this content:**

- **Tools/software and resources to be used**

   o **Software: Fastqc, CARD, RGI of the Center of Genomic Epidemiology, AMRFinderPlus, ARG-ANNOT, Virulence Finder, Plasmid Finder, MLST**

- **Overview of activities and exercises** (list specific practical activities which participants will do)
  Introduction lecture:
  Exercise/Activity 1: Discussion on fastqc and multiqc report
  Exercise 2: Practical session on CARD, RGI, AMRFinderPlus, and PathogenWatch
  Exercise 3: Bacteria subtyping approach (MLST from CGE, PathogenWatch)

- **Assessment**
  Assignment 1: Use CARD and RGI to predict the AMR genes in the five fasta files of *S. aureus* provided. Interpret the results
  Assignment 2: Use PathogenWatch and MLST CGE to subtype the *E. coli* isolates provided. Interpret the results
  Quiz 1 : Observe the following fastqc report, how many reads are in *K. pneumoniae* forward read file?
  Quiz 2: Is there any Ns and adaptor in the *E. coli* reverse read file?

- **Competencies/outcomes**
  If you are comfortable with competencies, please list the competency/ies this module will address:
  Online-based analysis, interpretation of data quality and in silico outputs

Module 4: Introduction to phylogenetics and phylogenetic analysis

Module Lead/Co-lead

- **Summary or objectives** – what you plan to cover for this module

- Principles of phylogenetics
- Phylogenetic tree construction, interpretation and visualisation

- **List of learning outcomes specific for this module**

- Understand the key principles of phylogeny
- Use Patric for retrieval of available genome
- Construct a tree with snp-tree of RGI
- Visualise a phylogenetic tree with MicroReact
- Interpret a phylogenetic tree
    - **Total number of hours/days to be spent on this content:** 4 hours

    - **Tools/software and resources to be used**

        o **Software: MicroReact, Patric, Figtree**

        o **Description of datasets to be used**

        o **Other resources, or readings**

- **Overview of activities and exercises**
  Introduction lecture:
  Exercise/Activity 1: Discussion on MicroReact phylogenetic tree
  Exercise 2: Practical session on RGI
  Exercise 3: Practical session on Figtree and MicroReact

    - **Assessment**
      Assignment 1: Use Patric to retrieve genome and related metadata.
      Assignment 2: Use snp-tree to construct a phylogenetic tree
      Assignment 3: Use figtree and PathogenWatch to visualize and annotate the tree. Interpret the results
      Quiz 1 : What can you say about the constructed tree?
      Quiz 2: Is there any linked strains?

- **Competencies/outcomes**
  If you are comfortable with competencies, please list the competency/ies this module will address:
  Online-based analysis, interpretation of data quality and in silico outputs

Manual steps
Provide detailed manual steps or exercises which participants will follow for this session or module (see examples of various courses online e.g. https://github.com/WCSCourses)

Module 5: Introduction to Linux

- **Module Lead/Co-lead**

- **Summary or objectives**

- Introduction to Unix, file structure and navigation
- Basic Linux commands
- Navigation in Linux directories
- Basic manipulating file commands

- **List of learning outcomes specific for this module**

- Understand the Linux file structure
- Understand the command line structure and learn basic commands
- Learn how to create, access files and directories and navigate through them
- Learn how to read files content and extract information from them
- Run a pre-written Linux script

- **Total number of hours/days to be spent on this content:** 3 hours

- **Tools/software and resources to be used**

  o **Software: Virtual Machine, MacOS terminal, Linux**

  o **Other resources, or readings: lecture slides**

- **Overview of activities and exercises**
  Introduction lecture: Introduction to Linux
  Exercise/Activity 1: Using the terminal
  Exercise 2: Create directories and navigate through the file structure
  Exercise 3: Basic manipulating file commands
- **Assessment**
  Assignment 1: Apply Linux command to create directory and move file in it
  Assignment 2: Use text editor to modify a script and run a bioinformatics script
  Quiz 1 : What is the difference between a relative and absolute path in Linux
  Quiz 2: Which command is used to:
    o Know the working directory?
    o List the files in a directory?
    o To remove a directory?
    o To create a directory?

- **Competencies/outcomes**
  If you are comfortable with competencies, please list the competency/ies this module will address: **s, navigation in Linux, run pre-written script**

Sponsorship Package for Bioinformatics Workshop
(Cohort 1 – February 2023)

No	Item	Quantity	Unit Cost (USD)	Total Cost (USD)
1	Participant registration	25	110	2750
2	Facilitators' compensation (USD $60/day $180 for 3 days	5	180	900
3	Teaching assistants' compensation (USD $30/day $90 for 3 days)	3	90	270
4	Zoom licence (1 year)	1	200	200
5	Certificates production and signature with all sponsors logos	40	12.5	500
6	Gifts for guest speakers (Branded Conference Mugs, T-shirt, pen, and agenda)	3	50	150
7	Internet modem and data bundles	1	200	200
8	Branded stationaries for participants (notebook, pen, USB flash drive and workshop files)	40	12.5	500
9	Office supply (chair & table)	40	12.5	500
10	Coffe-break & Lunch & Drink	3	200	600
11	**Communication outlets** Roll-up, Event Banners Workshop Badges for participants and facilitators	 2 2 40	 65 70 2	**350** 130 140 80