



Public Health Alliance for Genomic Epidemiology

PHA4GE Pre -Conference Workshops 2023

Dates: 27 -28 October 2023

Venue: South African National Bioinformatics Institute (SANBI), 5th Floor, Life Sciences Building, University of the Western Cape, Cape Town, South Africa

****Rooms for workshops will be confirmed on the day**

****Please refer to the colour coding below for the description of each workshop**

PRE-CONFERENCE WORKSHOPS PROGRAMME

Day 1 - 27 October 2023	Session 1: 9.30am-12.5 0pm	9.30 -11.00 am	Workshop 1 Part One	Workshop 4 Part One
		11.00 - 11.20am	Tea	Tea
		11.20am - 12.50pm	Workshop 1 Part Two	Workshop 4 Part One
		12.50 -1.35pm	Lunch	Lunch
	Session 2: 1.35-4.55pm	1.35 -3.05pm	Workshop 2 Part One	
		3.05 -3.25pm	Tea	
		3.25 -4.55pm	Workshop 2 Part Two	
Day 2 - 28 October 2023		9.00 - 10.30	Workshop 5 Part One	Workshop 3 Part One
		10.30-10.50	Tea	Tea
	Session 4: 10.50am-1.2 0pm	10.50am- 12.20pm	Workshop 5 Part Two	Workshop 3 Part Two
		12.20 -1.20 pm	Workshop 5 Part Three	
		1.20 -2.05pm	Lunch	Lunch
	Session 5: 2.05- 5.00pm	2.05 -4.35pm	Workshop 5 Part Four	Workshop 3 Part Three
		4.35 -5.00pm	Tea	Tea

Key:**Workshop 1:
Plumbing the Metadata: Implementing PHA4GE
data standards with minimal pain**

Data standards enable interoperability across datasets and systems. Public health laboratories need to be able to implement data standards quickly in order to put them into practice efficiently. This workshop will focus on the tools and practical steps that enable the semi-automated implementation of PHA4GE standards in labs and public repositories, in order to make data harmonization and integration easier.

Topics: hAMRonization, the DataHarmonizer, PHA4GE QC contextual data tags, amplicon sequencing primer scheme nomenclature standards, and more

Format: Lecture and hands-on exercises
Duration: 3 hours, including mid-session networking break

Facilitators: Finlay Maguire and Emma Griffiths

**Workshop 3:
Using metagenomics to survey vector-based
pathogens – Chan Zuckerberg Initiative**

Metagenomic next-generation sequencing (mNGS) is a powerful tool that enables the sequencing of all the genetic material in a sample. The ability to sequence all of the microorganisms in a sample enables researchers to identify pathogens and novel taxa without a priori knowledge. This course will cover the basics of mNGS, from experimental design to data analysis and its importance in global health. The data analysis portion of the course will be hands-on, focused on analyzing an mNGS dataset from mosquitoes collected from California using the tool CZ ID. CZ ID is a free, open-source tool that enables researchers to analyze metagenomic datasets from their laptops without the need for computational skills. Although this course focuses on mNGS data acquired from mosquitoes, it should be noted that mNGS can be used to characterize the microbes in any sample type, so the tools learned in this course can be translated to other sample types of interests. Case studies of how mNGS has contributed to global health policies will also be covered.

Topics: real-world application of mNGS, Wet lab methods, Bioinformatics Intro, CZ ID introduction and hands-on activity

Format: Sessions and hands-on exercises
Duration: 9 hours, including lunch break
Facilitators: TBC

**Workshop 5:
Python for Biologists**

Python is a high-level, general-purpose programming language that can be used by biologists for data analysis

Topics: Getting started with Python, Working with Python and Data Structures, Python Example Codes

Format: In-person
Duration: 1 Day
Facilitator: Olaitan Awe

**Workshop 2:
Exploring and Extending GA4GH Standards for Pathogens and
Public Health**

The Global Alliance for Genomics and Health (GA4GH) is a not-for-profit organization that develops genomics data standards and policies for human genomics applications (e.g. genomic/phenotypic/clinical data representation; security/regulation/ethics tools and policies). Hackathons are social coding/development events that bring computer programmers, bioinformaticians, clinicians, and other domain experts together to evaluate, improve upon, or build new software. In this ½ day hackathon, participants will assess existing GA4GH standards for their suitability, utility and implementability in public health pathogen genomics.

Topics: Phenopackets, Passports, Data Use Ontology, and more

Format: Discussion and coding

Duration: 3 hours, including mid-session networking break
Facilitators: Peter van Heusden, Finlay Maguire and Emma Griffiths

**Workshop 4:
Bioinformatics Solutions for Public Health Pathogen Genomics**

Public health pathogen genomics is a rapidly evolving domain with constantly changing best practices. Despite its significance, there is a scarcity of documentation that effectively communicates these best practices to public health scientists, making it challenging for them to stay updated. The objective of this workshop is to highlight the open-source and open-access bioinformatics resources accessible to the public health community. These resources aim to standardize and ensure reproducibility in pathogen genomics research. Throughout the session, we will delve into the diverse challenges that public health laboratories encounter while implementing bioinformatics solutions. Furthermore, we will explore the solutions provided by the PHA4GE Bioinformatics Pipelines and Visualization working group to address these challenges.

Topics: Public health pathogen genomics, open-source solutions, proposed pipeline standards, pathogen analysis guidance documents, and more

Format: Lecture and hands-on exercises
Duration: 3 hours, including mid-session networking break
Facilitators: Kevin Libuit and Jamie Southgate