

The Goal: to develop bioinformatics capacity in Africa through training and support for data analysis, while also facilitating access to informatics infrastructure in the form of pipelines and tools for human, microbiome, and pathogen genomic data analysis.

The Problem

Researchers working in the life sciences are generating data of increasing size and complexity due to modern research techniques. Interpreting this data is essential for understanding health and disease and developing solutions to improve African and global health. Considering this and also the lack of modern research infrastructure in Africa, there is an ever-increasing need to develop the human and physical capacity for bioinformatic research in Africa to enable the interpretation of valuable African genomics data.

Project Strategy

1. Implement a pan-African informatics infrastructure
2. Develop an H3Africa Data Coordinating Center
3. Provide high quality informatics support to H3Africa
4. Enable and enhance innovative translational research
5. Establish outreach and sustainability plans

Successes to Date

Training and Development

- 40 Courses/workshops run, 29 face-to face, 2,256 people trained
- Annual introductory and intermediate online courses
- New bioinformatics degree programs developed in 4 African countries
- 15 accreditations in analysis of genomic data awarded

Enabling Genomic Medicine in Africa

- Designed new genotyping chip for African populations
- African genome reference panel for imputation
- Genomic medicine curriculum designed and 225 nurses trained
- Clinical data collection and phenotype harmonization support

Computing Infrastructure

- 17 computing facilities developed across 11 countries
- 3432 cores for data processing
- 1.2 petabytes available for storage and processing of African genomic data
- 4 genomic data analysis workflows developed and openly accessible

Data Management, Structure, and Transfer

- 14 African genomics datasets (135TB) stored in H3Africa archive
- 13 datasets QC'd and submitted to public repositories
- Created the H3Africa Data Catalogue, which allows users to search for datasets and biospecimens of interest
- 77.3 terabytes of African genomic data transferred

Project Lead



Prof. Nicola Mulder
University of Cape Town
nicola.mulder@uct.ac.za

Project Sites



H3ABioNet is a pan-African network consisting of 28 nodes across 16 African countries.

*** = Lead Institution:**

South Africa: University of Cape Town

Non-African Collaborators:

USA: University of Illinois

This work is supported by the U.S. National Institutes of Health (NIH) and the National Human Genome Research Institute (NHGRI) grant number U24HG006941.