Southern African Human Genome Programme

Workshop 1
21 to 25 July 2014
University of the Western Cape

The NGS data from the 24 genomes have arrived and we are now planning two workshops to build capacity to understand the data, manipulate the data, characterise the genomes and formulate research questions. A policy for the storage of the data and access to the data will be formulated by the Core SAHGP group in accordance with the expectation from the DST.

The workshop participants are limited to 20 individuals. They will be chosen to represent geographical and institutional distribution. One to two people may be nominated from each of the 7 participating Institutions and 6 or more individuals from other institutions will be included to total not more than 20 people.

Please complete the application form at the following URL:
https://docs.google.com/forms/d/1U3ST1Ab2ECj_6VVAI46FJ40GQVfF1IxPH1AC0OulteQ/viewform?c=0&w=1&usp=mail_form_link
For assistance you may contact Freedom Mukomana (freedom.mukomana@wits.ac.za)
Deadline for applications: Monday 7 July at 12 noon.
Accommodation and travel will be funded by the SAHGP and arrangements will be made by Martie Madgwick (martie.madgwick@up.ac.za) from Michael Pepper's Office.

Workshop 1 - **Training in NGS analysis**

This course is designed to introduce participants to manipulating next generation sequencing data from raw reads to analysis. The course will consist of both theoretical lectures and hands on exercises and the emphasis will be on giving the participants practical skills. There will also be invited lectures on current research topics.

Overview of content: Introduction to Linux; Remote data access; Copying data between systems; Introduction to NGS: NGS technologies and data formats; Quality control theory and practice; Mapping reads to genomes; Variant calling; Manipulating variant data; Introduction to PLINK; Population structure and admixture.

Date and venue: 21 to 25 July (5 days) at UWC
Lead organisers: Alan Christoffels & Scott Hazelhurst
Format: Lectures and hands on data analysis

Objectives:
1. To understand the nature and format of data (file types and formats)
2. To manipulate raw sequence reads
3. To align data and perform QC
4. Variant analysis based on Illumina calling (SNPs and indels)
5. Reformatting data for PLINK (and other analysis tools)

Participants:
1. Hands on manipulation of data and analysis will be part of the course
2. Computational or biological background suitable

Outcomes:
1. Capacity development for large data management
2. Reformatting of data (e.g. PLINK)
3. Creating data sets for downstream analysis
4. Data query skills
5. Data storage in easily accessible format