

1KSA - Decoding South Africa's Biodiversity

Tuesday, 3 December 2024 14:10 (20 minutes)

South Africa is one of the most biodiverse countries in the world with many institutions researching and documenting local biodiversity. However, South African scientists often conduct biodiversity-related genetic sequencing overseas, to take advantage of competitive prices internationally, and the data follow. This contributes to a drain of skills, knowledge and opportunity out of South Africa.

1KSA (www.1kSA.org.za), a DIPLOMICS initiative, was launched in 2023 to demonstrate that large-scale genetic sequencing and the associated data analysis (bioinformatics) can be done in country, thereby mitigating the brain drain and ultimately benefiting people of South Africa.

Using data generated using Oxford Nanopore Technology (long-read) to conduct whole genome sequencing the 1KSA draft genome assembly pipeline has been successfully implemented on the Centre for High Performance Computing (CHPC) within nextflow and makes use of seriallong and bigmem resources. To date, 1KSA has assembled the genomes (to draft level) of nearly 50 South African species on the CHPC - 29 plants; 4 mammals; 13 fish and 3 anthropods. The expected genome size of these draft genome assemblies ranges from 162.3 Mb to 2.6 Gb. However, there are still some computational challenges that need to be addressed to tackle the sequencing and assembly of larger genomes.

Meanwhile, the successful assembled draft genomes and the raw data are stored using the Data Intensive Research Initiative of South Africa and are made known via the generation of species information cards on the 1KSA website. These sequenced genomes of biodiversity and economically important species are to become tools to enable researchers to investigate species populations, conservation and the impacts of climate change, as well as enable the discovery of novel compounds.

Student or Postdoc?

Email address

Co-Authors

CHPC User

Yes

CHPC Research Programme

CBBI1622 / CBBI1617 / CBBI1691 / CBBI1369

Workshop Duration

Primary authors: Dr SWART, Patricia (DIPLOMICS); 1KSA PROJECT CONSORTIUM; Dr NEWMAN, Tim (DIPLOMICS / CPGR); Dr MURRAY, Shane (DIPLOMICS)

Presenter: Dr SWART, Patricia (DIPLOMICS)

Session Classification: HPC Applications

Track Classification: Bioinformatics and Biological Sciences