# UFS-NGS Unit 5th Data to Discovery Bioinformatics Workshop for Beginners
# (07-09 July 2025)

### Introduction

NGS data analysis training workshops are essential. This demand has been addressed in the previously organized bioinformatics workshops by the UFS-NGS Unit. However, despite the growing significance of bioinformatics in research and healthcare, there remains a persistent gap in the foundational training available for beginners in this discipline of sequence data handling. In the last months the UFS-NGS Unit has received several notable requests from students and researchers for foundational bioinformatics training tailored towards early beginners and absolute novices. For many of these researchers and students, the barriers to entry in bioinformatics such as unfamiliarity with programming languages, data analysis tools, and computational workflows is overwhelming and complicated with the rapid evolution of bioinformatics tools and methods.

### Aim

To bridge this gap, the UFS-NGS Unit is committed to providing foundational bioinformatics training designed specifically for beginners, equipping them with the essential skills to navigate bioinformatics tools, data analysis workflows, and computational methods effectively. The workshop is designed to equip the researchers with the necessary bioinformatic skills required to analyse NGS data focusing on three aspects: whole genome data, 16S metagenomics and viral metagenomics data. The workshop will combine both theoretical and practical sessions.

### Intended audience

The course is aimed at individuals from a molecular biology background, interested in bioinformatics, and who work (or are planning to work) on genomic datasets.

### Participants will be introduced to

- Basic Linux command lines for NGS data analysis
- 16S/Shotgun metagenomics data analysis aspects
- Whole genome data analysis aspects including phylogenetic analysis
- Proteomics and molecular modelling

### Host

The University of the Free State-Next Generation Sequencing Unit

### Venues

- Metro 4
- Computer labs
- Faculty of Health Sciences Foyer

**Guest Speakers**

- **Prof. Chris Viljoen (Acting Dean and Head of School of Biomedical Sciences, Faculty of Health Sciences, University of the Free State)**

**Prof. Chris Viljoen** is a Full Professor in Human Molecular Biology, School of Biomedical Sciences, Faculty of Health Sciences, University of the Free State, South Africa. He is the Head of the School of Biomedical Sciences, which comprises the Animal Research Centre, Clinical Simulation and Skills Unit, Department of Basic Medical Sciences, Department of Biostatistics, Human Molecular Biology Unit, and Next Generation Sequencing Unit. His research focuses on food authenticity diagnostics, including the detection of genetically modified (GM) food and animal species.

Prof. Viljoen has published 45 papers in accredited journals, one peer-reviewed book chapter, eight popular scientific articles, co-authored one report for the World Health Organization, and an online chapter for the United Nations Convention on Biological Diversity for the Cartagena Protocol on Biosafety. His publications have a citation i10 index of 30 and an H-index of 19 with 1283 citations. Notable publications in Q1 journals: Journal of Food Composition and Analysis, Food Control, Environmental Sciences Europe and Food Chemistry, and Q2 journals: Gene and Food Additives & Contaminants.

He is internationally recognised as an expert in detecting genetically modified organisms and animal species in the United Nations roster of experts in the Biosafety Clearing-House record BCH-CON-SCBD-48313-3 (https://bch.cbd.int). He initiated and developed the GMO Testing Facility, which is one of two laboratories in South Africa that is approved to perform genetically modified status certification by the Department of Agriculture, Land Reform and Rural Development (DALRRD) and is the only laboratory in South Africa recognised by the United Nations under the Convention on Biological Diversity: Biosafety Protocol for the detection and identification of GMOs in the Biosafety Clearing-House (BCH-ORG-SCBD-48314-2 https://bch.cbd.int/en/database/record?documentID=250666).

Prof. Viljoen teaches postgraduate modules in Molecular Biology diagnostics and Research Techniques and has supervised 28 Masters and 10 Doctoral graduates. He has delivered 11 invited keynote addresses at international conferences, made 14 presentations at international and more than 50 at national conferences. He is the editorial board member of the international ISI journal Euphytica, a reviewer for international journals, and an examiner for Masters dissertations and Doctoral theses. He has also served on several international committees, including the European Commission Global GMO Network Forum, and is the coordinator of the Southern African Network for GM detection laboratories (SANGL). He is a recipient of several research grants, including R1.7 million from 2008 to 2010 and R570,133 from 2021 to 2023.

Prof. Viljoen is a registered member of the Health Professions Council of South Africa (HPCSA) and the South African Council for Natural Scientific Professions (SACNASP), a member of the South African Society for Biochemistry and Molecular Biology and the European Hematology Association.

- **Prof. Olihile Sebolai (Department of Microbiology and Biochemistry, University of the Free State)**

**Prof. Olihile Sebolai** is a professor of microbiology at the University of the Free State, South Africa. He leads a research laboratory with an established programme on Cryptococcus neoformans, particularly focusing on antifungal resistance and host-pathogen interactions. In recent years, Prof. Sebolai has incorporated molecular docking and in silico screening into his research to elucidate pathogen-to-pathogen interactions. Notably, he co-authored a few studies that explored the structural alignment and docking of fungal Kex2-like proteases with the SARS-CoV-2 spike protein, providing new insights into cross-kingdom protease interactions and potential therapeutic implications.

Prof. Sebolai has published widely in the fields of medical mycology and infectious disease, and his work spans both laboratory-based and computational approaches. He is also a dedicated mentor, having supervised numerous MSc and PhD students, and plays an active role in promoting interdisciplinary collaborations that address fungal disease burden in African settings. His research contributes to expanding the toolkit for antifungal drug discovery through repurposing and enhancing our understanding of fungal virulence and persistence.

- **Mrs. Mandy Jampies (Research Development, University of the Free State)**

**Mrs. Mandy Jampies** is a Senior Officer in the Directorate: Research Development at the University of the Free State (UFS). She holds a National Diploma in Commercial Administration and a B.Tech in Office Management and Technology from the Central University of Technology.

With a career spanning over 15 years in higher education, Mandy coordinates the postdoctoral fellowship programme across UFS campuses and manages internal bursaries for master's and PhD students. She provides vital support in research administration, data reporting and funding processes.

She was also part of the team that led UFS to achieve the prestigious Platinum Status for Good Financial Grant Practice (GFGP) accreditation, reflecting her contribution to strengthening research governance and operational excellence.

## LOCAL ORGANIZING COMMITTEE AND CO-FACILITATORS

### Coordinator of the sub-committees

- Prof. Martin Nyaga (UFS-NGS Unit)

### Planning/content sub-committee

- Dr. Ayodeji Ogunbayo (UFS-NGS Unit)
- Dr. Milton Mogotsi (UFS-NGS Unit)

### Logistics sub-committee

- Ms. Hlengiwe Sondlane (UFS-NGS Unit)
- Ms. Mamello Maku (UFS-NGS Unit)
- Ms. Somila Nazo (UFS-NGS Unit)
- Ms. Nkosazana Shange (UFS-NGS Unit)
- Ms. Suprise Baloyi (UFS-NGS Unit)

### Marketing sub-committee

- Ms. Manyi Eyong (UFS-NGS Unit)
- Ms. Sesiyanda Maseko (UFS-NGS Unit)
- Ms. Angela Mbele (UFS-NGS Unit)
- Ms. Nkosazana Shange (UFS-NGS Unit)
- Ms. Thabisa Mpaxa (UFS-NGS Unit)
- Ms. Mbali Ncube (UFS-NGS Unit)
- Ms. Palesa Tsunke (UFS-NGS Unit)

### ICT sub-committee

- Mr. Stephanus Riekert (ICT, UFS)

## FACILITATORS/TRAINERS/ASSISTANT TRAINERS

- Dr. Morne Du Plessis (Genetics, UFS)
- Dr. Frank Maleka (Genetics, UFS)
- Dr. Setshaba Taukobong (SAMRC, Cape Town)
- Ms. Rehema Mukami (Rhodes University)
- Ms. Tori Williams (Microbiology and Biochemistry, UFS)
- Mr. Adedayo Lanrewaju (Durban University of Technology)
- Dr. Emmanuel Ogunbayo (UFS-NGS Unit)
- Dr. Milton Mogotsi (UFS-NGS Unit)
- Ms. Nkosazana Shange (UFS-NGS Unit)
- Ms. Surprise Baloyi (UFS-NGS Unit)

## FUNDING ORGANIZATIONS

- Distribution Platform in Omics (DIPLOMICS)

## PARTICIPATING ORGANIZATIONS

- University of the Free State, Bloemfontein campus, South Africa
- National Health Laboratory Services, Bloemfontein, South Africa
- Central University of Technology, Bloemfontein, South Africa
- Vaal University of Technology, Vanderbijlpark, South Africa
- Africa Health Research Institute, Durban, South Africa

**NUMBER OF PARTICIPANTS FOR PRACTICAL TRAINING:** 25

## MEET THE TRAINERS

**Dr. Morne Graham Du Plessis** is a lecturer in the Department of Genetics at the University of the Free State, where his teaching and research span molecular biology, biodiversity, infectious disease, and applied bioinformatics. Over the past 15 years, he has worked across a range of systems (from vultures and bushbabies to microbial communities and disease-associated human samples) using next-generation sequencing and marker-based approaches to explore questions in conservation, population genetics, and pathogen dynamics. His work in infectious diseases focuses on generating and interpreting genomic data to better understand microbial diversity, resistance, and host–pathogen interactions. This applied perspective also informs his teaching. In this introductory bioinformatics course, he introduces students to core concepts and tools for analysing sequence data, with an emphasis on making the field approachable and relevant. The course draws on practical examples from relevant research, allowing students to engage with real-world data while building foundational skills in bioinformatics.

**Dr. Setshaba Taukobong** is a Bioinformatician with an academic background and hands-on experience in genomics data analysis. She holds a Master of Science in Bioinformatics, along with an honours degree in Bioinformatics and a Bachelor of Science in Genetics and Microbiology. Her research background spans to whole genome sequencing, transcriptomics and epigenetic profiling, particularly in AMR, Bacteriology and Virology.

She previously worked at the National Institute for Communicable Diseases, where she developed pipelines for SARS-CoV-2 variant surveillance in wastewater. She also contributed to genome assembly and bioinformatics support initiatives at the Centre for Proteomic and Genomic Research (CPGR - DIPLOMICS).

She is currently a Senior Scientist at the SAMRC, working on leveraging machine learning models for SARS-CoV-2 variant surveillance in wastewater as well projects involving somatic mutation detection and differential methylation analysis in cancer research. She is now seeking to broaden her expertise into marine microbiomes and metagenomics through a PhD program.

**Mr. Adedayo Ayodeji Lanrewaju** is a Biotechnology PhD candidate at Durban University of Technology (DUT) whose research bridges computational biology, viral genomics and environmental surveillance. Using virtual screening, molecular dynamics simulations, density functional theory, and next-generation sequencing, he evaluates promising antiviral candidates and maps viral diversity in wastewater treatment plants. His publications report small molecule inhibitors of Rotavirus A, SARS-CoV-2, and metabolic disease druggable targets.

A strong science communicator, he won the first-place oral presentation and people's choice award at the 2024 Computational & Systems Biology Conference as well as DUT's Three-Minute Flash Fact competition. As a regular facilitator of DUT's Computational & Systems Biology Workshop, he guides participants through bioinformatics fundamentals, in silico drug discovery pipelines and data visualisation, grounding each lesson in real world antiviral and environmental virology case studies.

**Ms. Rehema Mukami** is a PhD student of Bioinformatics at Rhodes University with the Research Unit in Bioinformatics (RUBi). My research is based on structural pharmacogenomics. I did my MSc in Bioinformatics and Computational Molecular Biology at Rhodes University with the same unit. I have a BSc in Biomedical Science and Technology from Chuka University-Kenya.

**Ms. Tori Shay Williams** is an early-career researcher specializing in extreme microbiology at the University of the Free State. She graduated with best honours in Microbiology and is pursuing her masters degree, which is a temporal study investigating the composition and assembly of a bacterial biofilm at the 1,200-meter level within Moab Mine. Her work on deep subsurface microbial communities has been presented by her at respected international and bilateral prestigious platforms, an oral presentation at SAJU6 and a poster presentation at ISME19. Tori's passion for uncovering the resilience of life in extreme environments fuels her research, while in her free time she enjoys reading and spending time in nature.

## PROGRAMME SCHEDULE AND OUTLINE

### Day 1: Monday, 07 July 2025

### Introduction to NGS data analysis

| TIME | DESCRIPTION | VENUE | RESPONSIBLE PERSON(S) |
|---|---|---|---|
| **Rapporteur: Dr. Emmanuel Ogunbayo and Ms. Sesiyanda Maseko** | | | |
| **08:00-09:00** | Registration | Foyer | **All facilitators** |
| **09:00-09:15** | Welcome Address (Overview of the UFS-NGS Unit) | Metro 4 | **Prof. Martin Nyaga** |
| **09:15-09:45** | Guest Speaker: Kex2 proteases as potential biological bridges: A molecular docking study linking fungal pathogenesis and SARS-CoV-2 | Metro 4 | **Prof. Olihile Sebolai** |
| **09:45-10:00** | Institutional Talk: Empowering research excellence: Enhancing impact through strategic support | Metro 4 | **Mrs. Mandy Jampies** |
| **10:00-10:30** | **Tea/Coffee break** | | |
| **10:30-13:00** | Theoretical talk: Introduction to NGS and Bioinformatics data analysis.<br>Introduction to Linux environment and basic commands for file management: ls, cd, mkdir, cp, mv, rm, grep, head, tail, man, pwd, etc | Lab A | **Dr. Morne Du Plessis and Dr. Maleka,** supported by the rest of the trainers |
| **13:00-14:00** | **Lunch break** | | |
| **14:00-16:30** | Hands-on training: Quality assessment of NGS data using FastQC and MultiQC and Trimming of genomes using Trimmomatic, Geneious, and genome assembly using SPAdes | Lab A | **Dr. Morne Du Plessis and Dr. Maleka,** supported by the rest of the trainers |
| **18:00 -20:00** | Welcoming dinner<br>Welcome Speech by Acting Dean and Head of School of Biomedical Sciences, Faculty of Health Sciences: Prof. Chris Viljoen | Health Sciences Foyer | **All** |

### Day 1 topics

- Introduction to Linux and command lines for NGS data analysis
- Basic commands: ls, cd, pwd, mkdir, rm, cp, mv, cat, less, head, tail
- File manipulation: touch, nano, chmod, chown
- Searching and retrieving data: grep, find, curl, wget
- Retrieving biological databases: NCBI, UniProt, Ensembl, etc.
- Quality Control and Pre-processing: Hands-on exercises in quality control of whole-genome sequencing data using different tools

## Day 2: Tuesday, 08 July 2025

## 16S/Shotgun Metagenomic Analysis

| TIME | DESCRIPTION | VENUE | RESPONSIBLE PERSON(S) |
|---|---|---|---|
| **Rapporteur: Ms. Sesiyanda Maseko/Ms. Thabisa Mpaxa** | | | |
| **08:00-10:00** | Hands on training on 16S/shotgun metagenomic analysis: Overview of tools and databases, QC, Pre-processing, Assembly | Lab A | **Ms. Tori and Dr. Setshaba,** supported by the rest of the trainers |
| **10:00-10:30** | **Tea/Coffee break** | | |
| **10:30-12:30** | Hands on training on 16S/shotgun metagenomic analysis: Taxonomic classification, diversity analysis | Lab A | **Ms. Tori and Dr. Setshaba,** supported by the rest of the trainers |
| **12:30-13:30** | **Lunch break** | | |
| **13:30 -15:00** | Hands on training on 16S/shotgun metagenomic analysis continues | Lab A | **Ms. Tori and Dr. Setshaba,** supported by the rest of the trainers |
| **15:00 – 17:00** | Excursion to the Planetarium | Naval Hill | **All** |

### Day 2 topics

- Tools and Databases: Overview of commonly used bioinformatics tools and databases for shotgun Quality Control and Pre-processing of Shotgun Metagenomics Data: Hands-on exercises in quality control of shotgun metagenomics sequencing data using different tools.
- Sequence assembly of shotgun metagenomics data using applicable tools.
- metagenomics analysis and hands-on exercises in accessing and using these resources.
- Hands-on exercises in taxonomic classification of shotgun metagenomics data
- Hands-on exercises in diversity and community analysis of metagenomics data

**Day 3: Wednesday, 09 July 2025**

**Whole Genome Analysis**

| TIME | DESCRIPTION | VENUE | RESPONSIBLE PERSON(S) |
|---|---|---|---|
| **Rapporteur: Dr. Mhlekazi Molatoli/Ms Angela Mbele** | | | |
| **08:00 -10:00** | Hands-on training on whole genome analysis: Multiple sequence alignments and phylogenetic analysis | Lab A | **Dr. Maleka and Dr. Setshaba,** supported by the rest of the trainers |
| **10:00 -10:30** | **Tea/Coffee break** | | |
| **10:30: -13:00** | Hands-on training on whole genome analysis: Multiple sequence alignments and phylogenetic analysis | Lab A | **Dr. Maleka and Dr. Setshaba,** supported by the rest of the trainers |
| **13:00 -14:00** | **Lunch break** | | |
| **14:00 -15:30** | Hands-on training: Proteomic analysis and Molecular Modelling | Lab A | **Mr. Adedayo and Ms. Rehema,** supported by the rest of the trainers |
| **15:30-16:30** | Hands-on training: Proteomic analysis Proteomic analysis and Molecular Modelling continues | | **Mr. Adedayo and Ms. Rehema,** supported by the rest of the trainers |
| **16:30 -16:40** | Closing remarks and vote of thanks | Lab A | **Prof. Martin Nyaga** |
| **18:00 -20:00** | Closing dinner for the guests and trainers | Longhorn Grill | **Trainers/facilitators/guests** |

**Day 3 topics**

- Sequence Alignment for Phylogenetics: Hands-on exercises in sequence alignment for phylogenetic analysis, including multiple sequence alignment and the use of software such as MUSCLE and MAFFT.
- Model Selection: Overview of model selection for phylogenetic analysis, including the use of likelihood-based methods and model comparison tools such as jModelTest and ModelTest.
- Phylogenetic Tree Construction: Hands-on exercises in phylogenetic tree construction using software such as IQTree and MEGA
- Proteomic analysis and molecular modelling