SANBI
South African National
Bioinformatics Institute

The South African National Bioinformatics Institute

2015 Annual Report

RESPICE PROSPICE
UNIVERSITY *of the*
WESTERN CAPE

FROM HOPE TO ACTION THROUGH KNOWLEDGE

contents

about sanbi

Bioinformatics is a specialist discipline straddling the fields of biology and computer science and is integral to modern biological research.

The South African National Bioinformatics Institute (SANBI) is situated at the University of the Western Cape in Cape Town. Our primary focus is the development and implementation of computational methodologies which allow biomedical researchers to accelerate their genomics data analyses.

SANBI provides a focus for biological research located in Africa and as such, is dedicated to:

- the development of an online specialised resource for genomics and genome informatics;
- capacity development in genomics and bioinformatics in South Africa; and
- the development and implementation of genome annotation methods.

SANBI aims to heighten awareness of bioinformatics in South Africa and to assist the country in making optimal use of this technology.

As the leading bioinformatics entity on the African continent, we continue to foster local and regional collaborations on health-related topics that cover both communicable and non-communicable diseases.

Policy Mandates

National Strategic Plan for HIV/ AIDS, STIs and TB (2012 – 2016)

The vision and mission of SANBI align with Draft Zero of the National Strategic Plan (NSP) which specifies Research and Innovation as a key enabler of the NSP, and argues that *"relevant research provides information and the impetus for innovation within the implementation of the NSP" and that strategic priorities should include "concrete plans to improve capacity for research" and "a budget for research".*

The Department of Science and Technology's 10-Year Innovation Plan (2008 - 2018)

One of the five Grand Challenge areas specified in this Plan is the *"Farmer to Pharma"* value chain to strengthen the bioeconomy. SANBI's genomics programme, which straddles both communicable and non-communicable diseases, aligns clearly with this Grand Challenge.

The MRC Act (Act 58 of 1991)

As an extramural unit of the MRC, SANBI falls under the legislative mandates of the MRC. At Section 3, this Act states that the Legislative Mandate of the MRC is:

"…through research, development and technology transfer, to promote the improvement of the health and quality of life of the population of the Republic, and to perform such functions as may be assigned to the MRC by or under this Act".

Our Vision

To become a center of excellence, achieving the highest level in biomedical research and education in the global, African and South African context.

Our Mission

To conduct cutting edge bioinformatics and computational biology research relevant to South African, African and global populations.

To develop human resources in bioinformatics and computational biology by educating and mentoring scientists.

To increase awareness of and access to bioinformatics and computational biology resources.

Our Goals

To generate and publish high quality, relevant biomedical research.

To train and graduate competent and productive researchers.

To add value to the academic programme of the University of the Western Cape.

To enhance other research fields through collaborative projects.

To establish sources of renewable funding to pursue the mission of the institute.

As you read the 2015 annual report, you will appreciate the cutting edge work carried out by staff and students at SANBI.

director's message

The funding structure at SANBI has and continues to enjoy external funding from national and international sources. However, staff started the year aware of the completion of a 15-year baseline funding as a South African Medical Research Council Unit. With the result that the year has been characterised by a range of applications for additional baseline funding. These efforts have paid dividends through successful funding proposals to the NIH, Gates Foundation and the European Union that has ensured continuity specifically in the applied health sciences.

SANBI contines to fulfil its human capital development mandate through 7 PhD and 2 MSc graduations in 2015. Our staff have contributed to computational biology training nationally and internationally. Our academic leadership on the African continent has been reinforced through the pivotal role that our staff played in regional bioinformatics projects and awards such as the Human Genome Organisation Africa Prize and the NRF Hamilton Naki Award.

Prof Alan Christoffels PhD, M.ASSAf
Director & DST/NRF Research Chair in Bioinformatics and Health Genomics
SA MRC Unit for Bioinformatics Capacity Development
South African National Bioinformatics Institute, University of the Western Cape

year in review

The research teams at SANBI have once again successfully delivered on their mandate to produce high impact research outputs and to provide international exposure for our postgraduate students in our pursuit of excellence in bioinformatics capacity development.

Research Highlights

2015 was an exciting year for Simon Travers' group with the launch of spin-out company, **Hyrax Biosciences** to commercialise the exatype™ HIV drug resistance analysis pipeline which enables the use of next-generation sequencing for HIV drug resistance testing. In December 2015 Hyrax Biosciences signed an exclusive licensing agreement with UWC for the use of exatype™. It is a proud achievement to use technology that was created at SANBI to positively affect the lives of potentially millions of people affected by the HIV scourge.

Awards

HUGO-Africa Prize: This prize was awarded to Alan Christoffels at the HUGO meeting in Malaysia in March 2015. The Human Genome Organisation recognises researchers in Africa who have contributed leadership and advanced genetics research. The selection panel included an international committee comprising 6 members.

Hamilton Naki Prize: Alan Christoffels received this award in August 2015 at the NRF annual awards evening in Durban. The prize is awarded by the NRF for contribution to changing the scientific landscape in South Africa.

Funding

In May 2015, we secured funding from the **European Union (Horizon2020)** for a biobanking project entitled B3Africa. This consortium comprises 4 African labs and 4 labs in Europe. The primary goals are to develop an informatics framework to support biobanks in resource limited settings. UWC hosted the first meeting of the project "B3Africa – Bridging Biobanking and Biomedical Research across Europe and Africa" from 24 - 25 August 2015.

DST-NRF Centre of Excellence in HIV Prevention. This award will enable SANBI to continue its successful collaboration with CAPRISA. Specifically, SANBI researchers will undertake projects focused on HIV prevention through the development of a HIV vaccine. Simon Travers is the senior scientist representing UWC/SANBI on this Centre of Excellence.

Alan Christoffels with B3Africa project members.

US National Institutes of Health RO1 grant: Identifying sources of HIV infection in adolescent girls in rural South Africa. Simon Travers is a principle investigator on this award. The purpose of this study is to understand the HIV transmission dynamics and sexual networks that adolescents might belong to, potentially increasing their risk for HIV.

Technology Innovation Agency (TIA): Junaid Gamieldien was awarded this funding for a 12 month period to work on the development of a cloud-based exome/genome and RNAseq data analysis services platform.

International Initiatives

As part of an international collaborative project, Alan Christoffels' laboratory was invited to annotate the Asian Seabass genome in 2014. Subsequently in March 2015 he visited the Reproductive Genomics group at the Temasek Life Sciences Laboratory in Singapore to finalise aspects of the project that need to be completed before the data could be published. A manuscript has been submitted. The genome project web portal is found at **http://seabass.sanbi.ac.za/.**

Research Outputs

In 2015, SANBI produced a total of 19 journal publications and 2 book chapters. Six of the manuscripts represent student first author publications.

Capacity Development

The student cohort comprised 8 Postdoctoral fellows, 11 Doctoral and 14 Masters students in 2015. Of the total 33 students, 42% are female and 64% are South African.

SANBI graduated 7 PhD and 2 MSc students in 2015.

staff

As a research institute, the Director of SANBI reports through the faculty of Natural Sciences at the University of the Western Cape. The academic staff comprises of 5 principal investigators who are tasked with research, student training and and capacity development. The computing infrastructure is maintained by a team of 7 technical staff. Administrative support to the institute is provided by 4.5 admin staff.

academic

Alan Christoffels, Prof
Director & DST/NRF Research Chair in Bioinformatics & Health Genomics,
SA MRC Unit for Bioinformatics Capacity Development
DST/NRF Research Chairs Programme

Junaid Gamieldien, Dr
Senior Lecturer
UWC

Gordon Harkins, Dr
Senior Lecturer
UWC

Uljana Hesse, Dr
Senior Lecturer
UWC

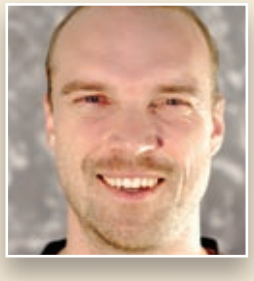
Nicki Tiffin, Dr
Senior Lecturer
SA Medical Research Council

Simon Travers, Prof
Associate Professor
UWC

technical

Hocine Bendou
Software Developer
NIH funding

Mario Jonas
Database Administrator
SA Medical Research Council



Phillip Labuschagne
Senior Software Developer
HVTN funding from the Bill and Melinda Gates Foundation

Baruch Lubinsky
Software Developer
DST-HIV funding

Campbell Rae
Web Developer
DST/NRF Research Chairs Programme

Peter Van Heusden
Senior Systems Developer
UWC

Long Yi
Software Developer
SA Medical Research Council

administrative

Fungiwe Mpithi
Reception
SA Medical Research Council

Ferial Mullins
Finance Administrator
UWC Dean's budget

Maryam Salie
Student Administrator
SA Medical Research Council

Natasha Schoeman
PA and HR Administrator
DST/NRF Research Chairs Programme

Junita Williams
Administrator part-time
DST/NRF Research Chairs Programme

capacity development

SANBI offers graduate education and training courses that are recognised world-wide as being of the highest calibre. Students are encouraged to participate in workshops and conferences relevant to their research projects or visit collaborators from South Africa or abroad. Students are also expected to publish their work in peer-reviewed journals and to present at laboratory meetings or wider audiences. This support provided by SANBI exposes students to internationally competitive research environments.

Undergraduate Training Programme

Bioinformatics Module (BTN 323)

The UWC undergraduate Bioinformatics Module is taught to approximately 107 third year students by the SANBI faculty. Students attend lectures, practicals and tutorials on the following topics:
Databases and their Applications in Disease Genomics;
Introduction to Molecular Evolution and Phylogenetics;
Comparative Genomics; and
Alignments and Database Searching.

Honours Progamme

Although SANBI does not have an Honours programme, students who attain a pass rate of >60% can include a bioinformatics component to their Honours project. Two honours students are currently working with their SANBI PIs:

- Warren Jacobus is currently registered in the Computer Science department and is carrying out his thesis project in collaboration with Alan Christoffels and Peter van Heusden.
- Lynley Abdoll is registered in the Medical Bioscience department and working on a joint TB drug metabolising enzyme project with Alan Christoffels.

Internship Programme

As part of the DST/NRF Research Chairs Programme, SANBI provides a 1-year internship programme to students who graduated with a BSc degree. Students who completed this programme in the past have pursued postgraduate studies either at SANBI or other universities. Darryn Zimire was the intern for 2015.

Postgraduate Training Programme

Masters Programme

SANBI offers a research MSc in Bioinformatics. Candidates with an Honours (BSc) degree or equivalent in a related scientific subject area may apply. The MSc degree is usually completed within two years.

Doctoral Programme

The PhD programme at SANBI is well-established and alumni of this programme are now working all over South Africa and at Bioinformatics research sites around the world. Candidates should be in possession of a Masters degree in Bioinformatics or in a related scientific field subject area such as Computer Science, Mathematics, Biochemistry or Engineering. The PhD degree must be completed within five years although most students aim to complete the degree in three years.

Postdoctoral Programme

Postdoctoral fellows are admitted to the research programme after consultation with a potential supervisor. Fellowships may last for a maximum of three years.

International Training Opportunities for SANBI Students

SANBI students enjoy the benefit of internships at international research labs. These internships usually occur as a result of a collaborative research project.

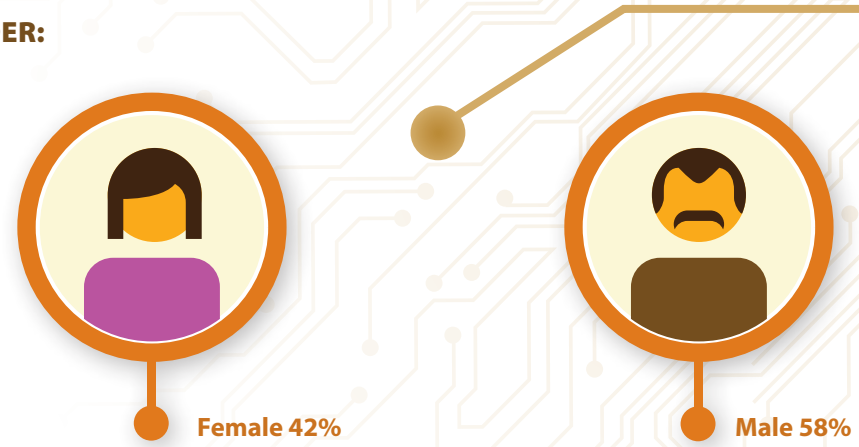
Student Name	Institution	Date	Project details
Ruben Cloete	A-Star Bioinformatics Institute of Singapore	June - July	Protocols for small peptide design: Ruben learnt a computational technique that will be applied to our tuberculosis project for finding new drug molecules.

Ruben Cloete with member's of Prof Chandra Verma's lab group in the Biomolecular Modeling and Design Division at A-star Bioinformatics Institute, Singapore.

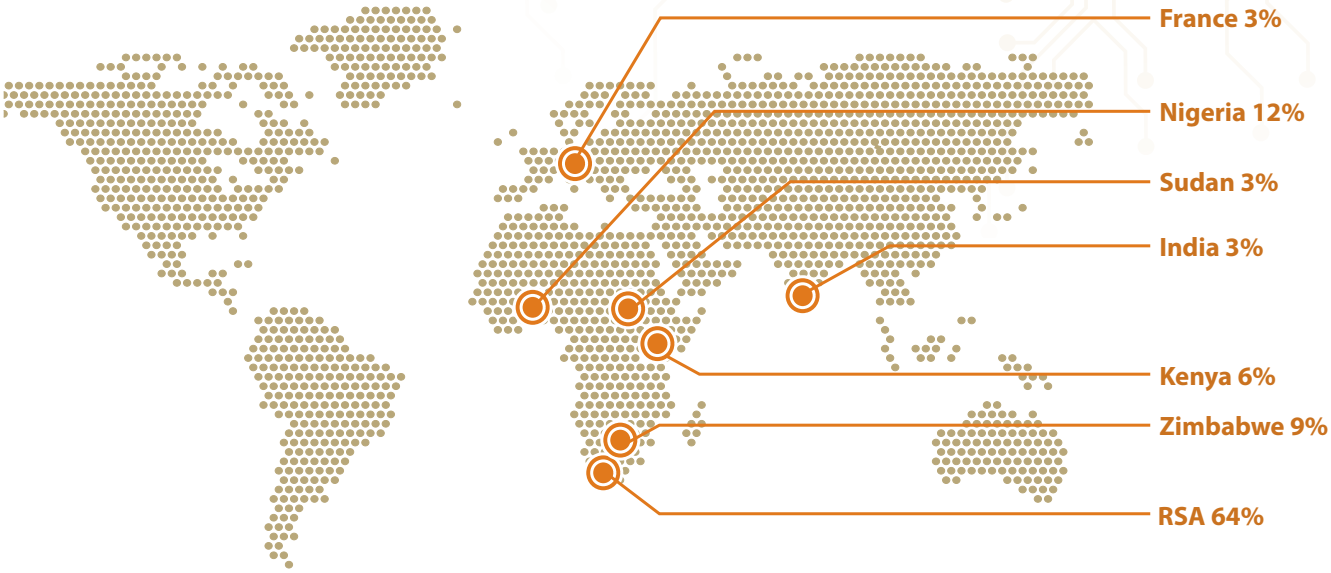
2015 SANBI Postgraduate Registration

In 2015 the student cohort comprised of 8 Postdoctoral fellows, 11 Doctoral and 14 Masters students. The diversity of SANBI students is represented in the following graphs:

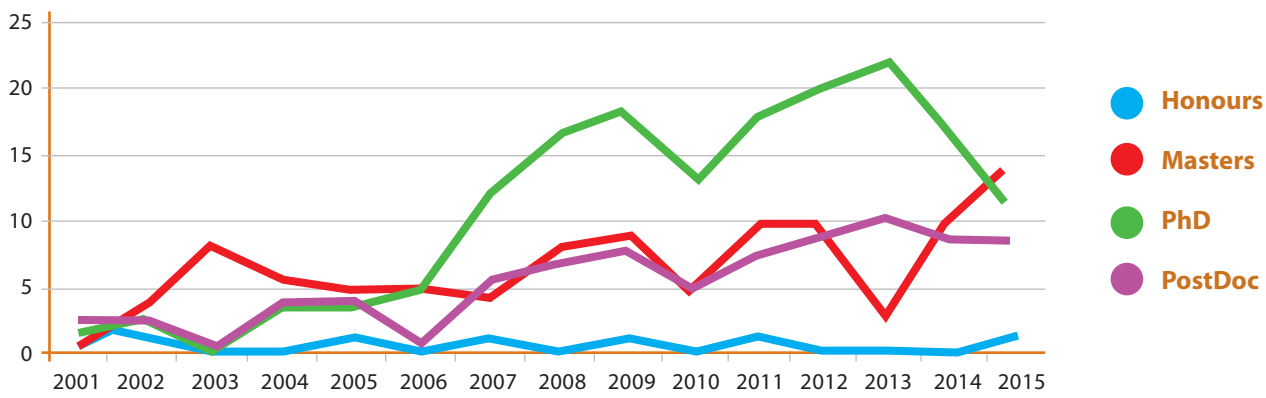
DIVERSITY BY GENDER:



DIVERSITY BY NATIONALITY:



SANBI POSTGRADUATE REGISTRATIONS TREND 2001 – 2015:



2015 SANBI Graduations

SANBI has great pleasure in announcing the graduates for 2015:

Michael Berry

Degree	Thesis Title	Supervisor
PhD	Massively–parallel computational identification of novel broad spectrum antivirals to combat coronavirus infection.	Junaid Gamieldien

Azeez Fatai

Degree	Thesis Title	Supervisor
PhD	Computational analysis of multilevel omics data for the elucidation of molecular mechanism of cancer.	Junaid Gamieldien

Alecia Naidu

Degree	Thesis Title	Supervisor
PhD	Computational characterisation of DNA Methylomes in Mycobacterium tuberculosis Beijing hyper- and hypovirulent strains.	Alan Christoffels

George Obiero

Degree	Thesis Title	Supervisor
PhD	Genome-wide annotation and characterization of odorant receptor genes in genus Glossina.	Alan Christoffels

Kavisha Ramdayal

Degree	Thesis Title	Supervisor
PhD	A characterization of HIV-1 subtype Cgp 120 envelope sequencing diversity in the female genital tract and plasma during acute and chronic infection.	Gordon Harkins

Adugna Woldesemayat

Degree	Thesis Title	Supervisor
PhD	Genomic and proteomic analysis of drought tolerance in sorghum *(Sorghum bicolour (L.) Moench)*.	Alan Christoffels

Imogen Wright

Degree	Thesis Title	Supervisor
PhD	Development of rapid, sensitive sequence alignment tools for application to high-throughput sequencing data.	Simon Travers

Leendert Cloete

Degree	Thesis Title	Supervisor
MSc	The Molecular Evolution and Epidemiology of Rubella Virus.	Gordon Harkins

Batsirai Mabvakure

Degree	Thesis Title	Supervisor
MSc	Global population dynamics and spread of Tomato yellow leaf curl virus.	Gordon Harkins

Workshops and Courses organised or hosted by SANBI

SANBI faculty are involved in numerous capacity development efforts in the form of workshops and courses throughout the year.

National Introduction to Bioinformatics Course

The annual 7-week full time course was held at SANBI from February to April. It aims to give postgraduate students an introduction to a range of bioinformatics topics. Students develop in-depth knowledge of key topics that inform their postgraduate thesis design. Lecturers from SANBI and other national universities presented the course to 27 students from SA institutions.

Principles and Practice of Clinical Research Workshop

During the week of 11 - 16 May 2015, lecturers from the NIH, MRC and SANBI presented techniques in Clinical Practice. Approximately 80 biomedical researchers or practitioners from all over Africa attended. SANBI hosted the last day of this workshop.

Colloquium on databases

In August 2015, SANBI researchers hosted the informal colloquium at SANBI offices. Presentations covered topics on storing, analysing and presenting data using databases. Approximately 15 students from SANBI and UWC Biotechnology Department were in attendance.

Workshop on N-linked glycosylation and vaccine response

On 04 November 2015, Simon Travers presented a workshop which was hosted at the CHIL (Cape Town HVTN immunology laboratory) laboratories in Cape Town. The workshop brought together 16 researchers from the USA and all over South Africa to establish a strategy for combining molecular dynamics simulations with traditional "wet-lab" work to enhance our understanding of the N-linked glycosylation in the susceptibility to (and escape from) broadly cross-neutralising antibodies.

Students who attended the national Bioinformatics Course early in 2015.

Python Workshop

On 5 - 6 November Peter Van Heusden taught Python as part of a 2-day Software Carpentry workshop at North West University. The 25 students were drawn from the university's Potchefstroom and Mafikeng campuses and ranged from undergraduates to academic staff. The workshop led to discussions around NWU's work on the lovebird *(Agapornis)* genome and attempts at discovering biomarkers for drug resistant tuberculosis.

H3ABionet

SANBI continues to be actively involved in providing training through the H3ABioNet workshops around Africa. SANBI node members were particularly active as lecturers and trainers in 2015:

- Peter Van Heusden taught Linux Systems Administration at the H3ABioNet Advanced Systems Administration Workshop at the Bioinformatics and Computational Biology Unit at the University of Pretoria from 2 - 11 February.
- Jean-Baka Domelevo-Entfellner lectured at the Introduction to Biostatistics for Genome Wide Association testing workshop at the Institut Pasteur de Tunis, Tunis, Tunisia from 16 – 26 March.

Invitations to Lecture

Alan Christoffels, assisted by PhD student Mmakamohelo Direko, taught the module "Introduction to Bioinformatics" to a group of approximately 30 students from the Universities of Limpopo, Venda and North West from 6 – 7 July at the University of Venda.

Peter Van Heusden with Python Workshop attendees at North West University.

Seminar Series

SANBI students benefited from two seminars presented by subject experts from other institutions in 2015.

Speaker	Date	Seminar Title
Dr Jeffrey Dorfman (University of Cape Town)	21 Jan	Narrowly neutralizing anti-HIV-1 antibodies directed against targets of broadly neutralizing antibodies: What can they tell us about vaccine priorities?
Dr Oliver Grant (Complex Carbohydrate Research Centre, University of Georgia)	4 Nov	Advances in modelling glycan-protein interactions

Algorithms Circle

Peter Van Heusden and senior students worked to train undergrad Computer Science students at UWC about practically implementing algorithms. Lunch time sessions were held every week where approximately 15 students work towards solving algorithmic problems.

Academics hosted at SANBI

For the month of September 2015, Prof Philip Machanick from the Computer Science Department at Rhodes University visted SANBI. He was exploring bioinformatics research collaborations during his stay.

Alan Christoffels and Mmakamohelo Direko at the University of Venda.

During October – November 2015, Dr Oliver Grant visited from University of Georgia on a NRF funded Knowledge, Interchange and Collaboration award. The purpose of his visit was to provide training in carbohydrate molecular dynamic modelling and to strengthen the collaborations between researchers at University of Georgia, UWC, UCT and the NICD on the analysis of the role of N-linked glycosylation in HIV.

Intern Hosted at SANBI

A bioinformatics student Jonas Sträßer, in his fifth semester at the University of Applied Science in Gießen, Germany spent 6 months (since October 2015) in Simon Travers' lab to complete his final year project working on applying the exatype™ HIV drug resistance genotyping software to other disease organisms.

computational resources

SANBI Compute Cluster

We continued to grow our computing capabilities in 2015, with the addition of 2 new servers that provide an extra 32 CPU cores and 192 GB of RAM to our cluster. We now have 168 CPU cores and 1328 GB of RAM on our Dell cluster.

Virtual Machine Infrastructure

SANBI has consolidated its virtual machine hosting resources and currently hosts 91 virtual machines using 136 virtual CPU cores, 224 GB of RAM and 5 TB of disk space, spread across eight virtual machine hosts. This is used to host key SANBI infrastructure as well as web applications presenting SANBI research to the outside world. We have started migrating our VM storage to use the new Ceph storage cluster.

Networking and Internet

We have upgraded our broadband connection to 30 Mbit/s, doubling our Internet speed. Our Globus Online server, connected to SANReN at 1 Gbit/s, has gone into production and been used for data transfers to and from other academic sites in South Africa and beyond.

Storage and backup

The Ceph storage array we had been planning since 2014 arrived in 2015, providing 72 TB of extra storage. This storage array will make our disk usage as scalable as our computing, and we have already added a third storage server to the initial two, which we expect will provide another 60 TB of disk space. The storage expansion has boosted our ability to do science as we no longer need to worry so much about the disk space an analysis will consume. The new storage has also relieved pressure on our older Dell Equallogic SAN, ensuring that we don't have work-interrupting full disks. We also installed a Backblaze storage server, providing 60 TB of disk-based backup to keep SANBI data safe.

Support and Research

Software development to support research was given a boost with the hiring of Thoba Lose, a dedicated scientific software developer. Thoba assisted in building a database and web portal for the Asian seabass *(L. calcarifer)* genome project and has since moved on to start construction of software to support a comprehensive Tuberculosis research portal.

SANBI software developers participated in the symposium organised on databases for research and they continued to support research on the UWC campus by attending meetings on e-Research and Data Science organised at UWC. We also assisted in training the UWC team for the CHPC Student Cluster Competition and were happy to see our students walk away with 3rd place in the national competition.

Peter van Heusden and Long Yi also participated in the H3ABionet Infrastructure Working Group and assisted colleagues across the continent in configuring their scientific software installations.

awards and honours

Alan Christoffels receiving the HUGO-African prize in Malaysia.

Sarah Mwangi being awarded her certificate of completion by Sheila Casey, Deputy Director, Office of Citizen Exchanges at U.S Department of State, Bureau of Educational and Cultural Affairs.

Zahra Jalali receiving her award at the conference in Canada.

HUGO-African Prize

The Human Genome Organisation recognises researchers on the African continent who have contributed leadership and advanced genetics research. This prize was awarded to Alan Christoffels at the HUGO meeting in Malaysia in March 2015.

Hamilton Naki Award

The annual NRF Awards ceremony recognises South African scientists. In addition to the rating linked awards, special recognition awards honour researchers for career achievements and contributions to knowledge creation and dissemination, as well as capacity development and transformation in the national research system. Alan Christoffels received this award for contribution to changing the scientific landscape in South Africa at the annual awards evening held in August 2015 in Durban.

6th International Conference on Biomarkers and Clinical Research

Postdoctoral fellow, Zahra Jalali's poster entitled "Detection of novel genetic variants associated with Mycobacterium tuberculosis drug resistance", won the best poster presentation award at the Conference on Biomarkers and Clinical Research in Toronto, Canada in September.

TechWomenMentorship Programme

This is an Initiative of the US Department of State's Bureau of Educational and Cultural Affairs. In October, Sarah Mwangi underwent a five week mentorship by Dr Julie Baher at Illumina Inc, where she was exposed to the latest technologies in DNA sequencing and participated in the bioinformatics unit group meetings, talks and webinars. "Overall, the TechWomen programme was transformational, in many ways. I got to experience the Silicon Valleys' rich culture, was exposed to the cutting-edge trends in technology, met several innovators and was really inspired. I want to take lessons learned here for my professional development and to inspire young girls to pursue STEM careers in Africa", said Sarah.

sanbi in the media

Three articles appeared in online media during 2015.

How South Africa can stop HIV drug resistance in its tracks

May 22, 2015

By Imogen Wright

http://theconversation.com/how-south-africa-can-stop-hiv-drug-resistance-in-its-tracks-41779

Health care is entering a new era of personalised medicine, where treatment is tailored to the individual patient. To usher in this era, researchers across the world are trying to create cheaper tests that can find DNA mutations in humans, bacteria and viruses…

UWC team slashes cost of drug test

11 June, 2015

By Tanya Farber

http://www.timeslive.co.za/thetimes/2015/06/11/UWC-team-slashes-cost-of-drug-test

A group of scientists at the University of the Western Cape has developed a more affordable test for drug-resistance to antiretroviral therapy in people with HIV.

Simon Travers, a researcher at the university's South African National Bioinformatics Institute, said: "International best practice says that everyone should be tested prior to starting treatment. "But this is not the case in South Africa."

He said not testing resistance to therapy reduces chances of successful treatment but that routine testing "would be too expensive given current resistance testing approaches"…

NRF awards laud country's top researchers

25 August 2015

By Megan Van Wyngaardt

http://www.engineeringnews.co.za/article/nrf-awards-laud-countrys-top-researchers-2015-08-28

To celebrate South Africa's top researchers' continued pioneering work in advancing knowledge creation and innovation, the yearly National Research Foundation (NRF) Awards were held in Durban on Thursday….

…University of the Western Cape Prof and South African National Bioinformatics Institute director Alan Christoffels received the Hamilton Naki Award, in recognition of an individual's outstanding efforts to advance his/her career in science against all odds, and for achieving world-class research performance, despite considerable challenges…

research outputs

SANBI's profile of excellence is reflected in its ability to publish high impact scientific articles in internationally recognised publications. Details of the institute's outputs are described in the tables that follow.

JOURNAL PUBLICATIONS

19 papers and 2 book chapters were published in 2015. First author papers were achieved by 6 students.

No.	Authors (SANBI contributors)
1.	Daisy Stainton, Darren P. Martin, Brejnev M. Muhire, Samiuela Lolohea, Mana'ia Halafihi, Pascale Lepoint, Guy Blomme, Kathleen S.Crew, Murray Sharman, Simona Kraberger, Anisha Dayaram, Matthew Walters, David A. Collings, **Batsirai Mabvakure**, Philippe Lemey, **Gordon W. Harkins**, John E. Thomas, Arvind Varsan. The global distribution of Banana bunchy top virus reveals little evidence for frequent recent, human-mediated long distance dispersal events. Virus Evolution, 2015, Vol. 1, No. 1; DOI: http://dx.doi.org/10.1093/ve/vev009
2.	Kotze MJ, Lückhoff HK, Peeters AV, Baatjes K, Schoemann M, van der Merwe L, Grant KA, Fisher LR, van der Merwe N, Pretorius J, van Velden DP, Ettienne J Myburgh E J, Pienaar FM, van Rensburg SJ, Yako YY, September AV, Moremi KE, Cronje FJ, **Tiffin N,** Bouwens CSH, Bezuidenhout J, Apffelstaedt JP, Hough FS, Erasmus RT, Schneider JW. Genomic medicine and risk prediction across the disease spectrum. Critical Reviews in Clinical Laboratory Sciences, 2015 Jan 19:1-18.
3.	Budd A, Dinkel H, Corpas M, Fuller JC, Rubinat L, Devos DP, Khoueiry PH, Forstner KU, Georgatos F, Rowland F, Sharan M, Binder JX, Grace T, Traphagen K, Gristwood A, **Wood NT**. Ten simple rules for organizing an unconference. PLoS Computational Biology 2015; 11(1):e1003905.
4.	**Stanley Kimbung Mbandi, Uljana Hesse, Peter van Heusden and Alan Christoffels.** Inferring bona fide transfrags in RNA-Seq derived-transcriptome assemblies of non-model organisms. BMC Bioinformatics 2015, 16:58; doi:10.1186/s12859-015-0492-5.
5.	**Fadhal E, Gamieldien J, Mwambene EC.** Self-similarity of human protein interaction networks: a novel strategy of distinguishing proteins. Scientific Reports 2015; 5: 7628.
6.	**Michael Berry, Junaid Gamieldien**, and Burtram C. Fielding. Identification of New Respiratory Viruses in the New Millennium. Viruses 2015, 7, 996-1019; doi:10.3390/v7030996.
7.	Celia van der Merwe, **Zahra Jalali Sefid Dashti, Alan Christoffels**, Ben Loos, Soraya Bardien. Evidence for a common biological pathway linking three Parkinson's Disease causing genes: Parkin, PINK1 and DJ-1. European Journal of Neurosciences 41(9); DOI: 10.1111/ejn.12872.
8.	Scheepers C, **Shrestha RK**, Lambson BE, Jackson KJ, **Wright IA**, Naicker D, Goosen M, Berrie L, Ismail A, Garrett N, Abdool Karim Q, Abdool Karim SS, Moore PL, **Travers SA**, Morris L. Ability To Develop Broadly Neutralizing HIV-1 Antibodies Is Not Restricted by the Germline Ig Gene Repertoire. Journal of Immunology 2015. pii: 1500118.
9.	**Michael Berry**, Burtram Fielding and **Junaid Gamieldien**. Human coronavirus OC43 3CL protease and the potential of ML188 as a broad-spectrum lead compound: Homology modelling and molecular dynamic studies. BMC Structural Biology (2015) 15:8; DOI 10.1186/s12900-015-0035-3.

No.	Authors (SANBI contributors)
10.	Charlotte Osafo, Yemi Raheem Raji, David Burke, Bamidele O Tayo, **Nicki Tiffin**, Marva M Moxey-Mims, Rebekah S Rasooly, Paul L Kimmel, Akinlolu Ojo, Dwomoa Adu, Rulan S Parekh. Human Heredity and Health (H3) in Africa Kidney Disease Research Network: A Focus on Methods in Sub-Saharan Africa. Clinical Journal of the American Society of Nephrology, CJN. 11951214.
11.	Azeez Butali, Peter Anthony Mossey, **Nikki Tiffin**, Wasiu Lanre Adeyemo, Mekonen Abebe Eshete, Chrispinanus Mumena, Rosemary Audu, Chika Onwuamah, Pius Agbenorku, Mobolanle Olugbenga Ogunlewe, Adetokunbo Raphael Adebola, Hecto Oladapo Olasoji, Babatunde Aregbesola, Ramat Oyebunmi Braimah, Abimibola Victoria Oladugba, Ifeanyichukwu Igwilo Onah, Ezekiel Adebiyi, Peter Babatunde Olaitan, Lukman Olajide Abdur-Rahman, Adebowale Adeyemo. Multidisciplinary approach to genomics research in Africa: the AfriCRAN model. The Pan African Medical Journal. 2015;21:229; doi:10.11604/pamj.2015.21.229.7380
12.	Xiaowei Jiang, Felix Feyertag, Conor Meehan, Grace McCormack, **Simon A Travers**, Charles Craig, Mike Westby, Marilyn Lewis, David L Robertson. Characterising the diverse mutational pathways associated with R5-tropic maraviroc resistance: HIV-1 that uses the drug-bound CCR5 coreceptor. J Virol 89:11457–11472; doi:10.1128/JVI.01384-15.
13.	**Sarah Mwangi**, Geoffrey Attardo, Yutaka Suzuki, Serap Aksoy and **Alan Christoffels**. TSS seq based core promoter architecture in blood feeding Tsetse fly *(Glossina morsitans morsitans)* vector of Trypanosomiasis. BMC Genomics (2015) 16:722; DOI 10.1186/s12864-015-1921-6.
14.	B Hodkinson, **D Mapiye**, D Jayne , A Kalla, **N Tiffin** and I Okpechi. The African Lupus Genetics Network (ALUGEN) registry: standardized, prospective follow-up studies in African patients with systemic lupus erythematosus. Lupus (2015) 0, 1–6.
15.	Malan-Müller S, Fairbairn L, Daniels WM, **Dashti MJ**, Oakeley EJ, Altorfer M, Kidd M, Seedat S, **Gamieldien J**, Hemmings SM. Molecular mechanisms of D-cycloserine in facilitating fear extinction: insights from RNAseq. Metabolic Brain Disease Sep 2015. 10.1007/s11011-015-9727-4.
16.	Nicola J Mulder, Ezekiel Adebiyi, Raouf Alami, Alia Benkahla, James Brandful, Seydou Doumbia, Dean Everett, Faisal M Fadlelmola, Fatima Gaboun, Simani Gaseitsiwe, Odile Ouwe Missi Oukem-Boyer, Ellis Owusu-Dabo, Sumir Panji, Hugh Patterton, Fouzia Radouani, Khalid Sadki, Fouad Seghrouchni, Özlem Tastan Bishop, **Nicki Tiffin**, Nzovu Ulenga. H3ABioNet, a sustainable Pan-African Bioinformatics Network for Human Heredity and Health in Africa. Genome Research 12/2015; DOI:10.1101/gr.196295.115.
17.	Lindisa Mbuli, **Darlington Mapiye**, Ikechi Okpechi. Lupus nephritis is associated with poor pregnancy outcomes in pregnant SLE patients in Cape Town: a retrospective analysis. Pan African Medical Journal 01/2015; 22; DOI: 10.11604/pamj.2015.22.365.7897.
18.	Nontobeko Eunice Mvubu, Bala Pillay, **Junaid Gamieldien**, William Bishai, Manormoney Pillay. Canonical Pathways, Networks and Transcriptional Factor Regulation by Clinical Strains of *M. tuberculosis* in Pulmonary Alveolar Epithelial Cells. Tuberculosis 12/2015; DOI: 10.1016/j.tube.2015.12.002.
19.	**Michael Berry**, Burtram Clinton Fielding, **Junaid Gamieldien**. Potential broad spectrum inhibitors of the coronavirus 3CLpro: A virtual screening and structure based drug design study. Viruses 12/2015; 7(12):6642-6660; DOI: 10.3390/v7122963.

BOOK CHAPTER PUBLICATIONS

No.	Authors
1.	**N Tiffin, J Gamieldien**. A Prioritization Analysis of Disease Association by Data-mining of Functional Annotation of Human Genes. Book chapter in Sakharkar, K.R., Sakharkar, M.K., and Chandra, R. (2015). Post-Genomic Approaches in Drug and Vaccine Development. River Publishers ISBN: 9788793102842 http://www.riverpublishers.com/view_details.php?book_id=291

No.	Authors
2.	**Michael Berry**, Burtram Fielding, **Junaid Gamieldien**. Practical Considerations in Virtual Screening and Molecular Docking. Book chapter in: Emerging Trends in Computational Biology, Bioinformatic and Systems Biology (2015) . 1st Edition Chapter: 27 Elsevier (MK imprints) Editors: Q.N. Tran, H.R. Arabnia DOI: 10.1016/B978-0-12-802508-6.00027-2 In book.

Software Resources and Computational Tools Developed

SANBI was actively involved in the H3ABionet NetCapDB functionality completion as well as the semi-automated reporting functionality. **http://netcapdb.sanbi.ac.za/**

Additionally, the RedCAP platform has been shown to function successfully as a pan-African patient database.

Alan Christoffels' lab developed the webportal for the Asian Seabass project to provide access to genome annotation data. **http://seabass.sanbi.ac.za/**

INVITED TALKS/KEYNOTES

Presenter	Date	Title	Details
Alan Christoffels	19 Jan	Management of a research programme and managing a research team.	PLUME Workshop, Lord Charles Hotel, Somerset West, Cape Town.
Alan Christoffels	17 Mar	Comprehensive analytical toolkit for tuberculosis research.	Temasek Lifesciences Lab, Singapore.
Alan Christoffels	19 Mar	Effects of nucleotide variation on the structure and function of TB drug metabolising enzymes.	Bioinformatics Institute of Singapore.
Alan Christoffels	20 Mar	Can Africa cope with a genomics data explosion?	Human Genome Meeting, HUGO, Kuala Lumpur, Malaysia.
Alan Christoffels	6 May	What's the fuss about doing a PhD?	PhD Conference, SA MRC, Parow, Cape Town.
Simon Travers	03 Jun	Harnessing the power of "next-generation" sequencing platforms to enable high-throughput, cost-effective HIV drug resistance testing.	UCT Institute of Infectious Diseases and Molecular Medicine seminar series.
Simon Travers	22 Jun	From PhD project to spinoff company – enabling low-cost HIV drug resistance testing.	1st Massachusetts-South Africa Conference for Technology Transfer and Global Innovation, Stellenbosch.
Simon Travers	25 Jun	Big data in public health.	NRF Consultative Workshop on Community of Practice in Public Health, UWC.
Mahjoubeh Jalali	02 Sep	Application of clinical exome sequencing and semantic knowledge base approach identifies potential causative variants in a family pedigree with atypical autosomal dominant distal muscular dystrophy.	Young Researchers Forum, 6th International Conference on Biomarkers & Clinical Research, Toronto, Canada.
Simon Travers, Imogen Wright, Natasha Wood, Baruch Lubinsky	05 Oct	Enabling cost-effective, high throughput HIV drug resistance genotyping.	Ion World Tour 2015, Pretoria.
Simon Travers	15 Oct	Tools and pipelines for the analysis of NGS data in HIV vaccine research.	HVTN/CHIL NGS Sequencing Workshop, Cape Town.

SCIENTIFIC/RESEARCH MEETINGS

Attended by	Date	Meeting Name	Purpose of meeting
Peter Van Heusden	2 - 3 Feb	Big Data Workshop, UCT	Presentation of data intensive research and Big Data initiatives from IBM Research, researchers from UWC and UCT and the Department of Science and Technology with discussions towards collaboration on Big Data initiatives.
Alan Christoffels	16 - 21 Jun	University of Missouri - UWC partnership	Explore new collaborations between SANBI and UM informatics institute.
Alan Christoffels	24 - 25 Aug	B3Africa first EU consortium meeting	SANBI hosted the first meeting of a new EU funded project on Biobanking.
Colleen Saunders	16 - 19 Aug	16th Southern African Society of Human Genetics Congress, Pretoria	Scientfic conference. Theme: "The Next Generation" – focused on the rapid expansion of new technologies that have enhanced the ability to study Human Health and Disease.
Alan Christoffels, Junaid Gamieldien, Peter Van Heudsen, Hocine Bendou, Nicki Tiffin, Jean-Baka Domelevo-Entfellner	10 - 13 Nov	H3ABioNet and Scientific Advisory Board Annual Meeting, Cape Town	To discuss 3 year progress and update on working groups activities including long term goals.
Alan Christoffels, Hocine Bendou, Lunga Sizani	13 - 16 Dec	B3Africa, Sweden University for Agricultural Studies, Uppsalla, Sweden	Workshop to define the strategies, resources and workflows for the implementation of the eB3Kit.

CONFERENCE PARTICIPATION

Presenter	Conference Details	Month	Title	Type
Mmakamohelo Direko	ISCB Africa ASBCB Conference on Bioinformatics, Dar es Salaam, Tanzania	Mar	Long and small non-coding RNAs in Mycobacterium tuberculosis H37Rv.	Poster
Ruben Cloete	ISCB Africa ASBCB Conference on Bioinformatics, Dar es Salaam, Tanzania	Mar	Molecular dynamics of the *M.tuberculosis* drug target nicotinate-nucleotideadenylyltransferase (Rv2421c) and docking studies to identify novel inhibitors.	Oral
Imogen Wright, Simon Travers	ISCB Africa ASBCB Conference on Bioinformatics, Dar es Salaam, Tanzania	Mar	NOAH: A novel guide tree ordering and progressive alignment heuristic.	Oral
Alan Christoffels	MRC Parow	May	Big data: Privacy, storage, use and resources. Principles and Practice of Clinical Research.	Oral
Natasha Wood, Fadda E, Grant O, Woods R, Travers SA	22nd International HIV Dynamics and Evolution Conference, Budapest, Hungary	May	Subtype-specific structural characteristics and molecular dynamics of glycosylated HIV-1 gp120 proteins.	Poster
Ereshia Gabier, Natasha Wood, Fadda E, Grant O, Woods R, Simon Travers	7th SA AIDS Conference, Durban, South Africa	Jun	The Dynamics of HIV-1 gp160 N-linked Glycans in the context of Broadly Cross Neutralizing Antibodies.	Oral
Mercuur C, Wood N, Fadda E, Grant O, Woods R, Travers SA	7th SA AIDS Conference, Durban, South Africa	Jun	Using molecular dynamics to investigate the effects of N-linked glycosylation on the gp120 envelope trimer of HIV-1.	Poster
Darlington Mapiye	23rd Annual ISMB/ECCB Conference, Dublin	Jul	Computational genomics approaches for Kidney disease in Africa.	Poster
Zahra Jalali	6th International Conference on Biomarkers and Clinical Research in Toronto, Canada	Aug	Detection of novel genetic variants associated with Mycobacterium tuberculosis drug resistance.	Poster

Presenter	Conference Details	Month	Title	Type
Mahjoubeh Jalali	6th International Conference on Biomarkers and Clinical Research in Toronto, Canada	Aug	Clinical exome sequencing and semantic discovery identifies potential causative variants in a family with atypical distal muscular dystrophy having unclear mode of inheritance.	Oral
Ruben Cloete	Biophysics Conference, Stellenbosch, Cape Town	Nov	Structural and functional effects of nucleotide variation on the Tuberculosis drug metabolizing enzyme human arylamine N-acetyltransferase1 protein.	Poster

COURSES/WORKSHOPS ATTENDED

Participant	Topic	Month	Location
Colleen Saunders	Wellcome Trust Advanced Course: Human Genome Analysis: Genetic Analysis of Multifactorial Diseases. Comprehensive overview of the statistical methods currently used to map disease susceptibility genes in human populations.	Jul	Hinxton, Cambridge, United Kingdom
Peter Van Heusden	GOBLET Train the Trainers workshop	Nov	UCT Medical School, Cape Town

THESIS EXAMINATION FOR STUDENTS AT OTHER INSTITITUTIONS

PI Name	Institution	Degree
Alan Christoffels	University of Cape Town	PhD
	University of Stellenbosch	MSc
Junaid Gamieldien	Rhodes University	PhD

EXTERNAL MODERATION

PI Name	Institution	Month
Junaid Gamieldien	University of KwaZulu-Natal: Genetics Research Skills	Jul
Simon Travers	University of Cape Town: Bioinformatics Honours programme	Nov

JOURNAL EDITING AND REVIEWS

PI Name	Journal
Alan Christoffels	Editorial Board Member: PLoS ONE Reviewer: BMC Bioinformatics BMC Biorepository BMC Biopreservation
Junaid Gamieldien	Nature PLoS ONE

EXPERT PANEL OR COMMITTEE MEMBERSHIPS

PI Name	Membership
Alan Christoffels	NRF ex-officio International Council on Science: Regional Committee for Africa Chair of ISMB: Special Interest Group (SIGs), 23rd Annual International Conference on Intelligent Systems for Molecular Biology, 11-15 July 2015, Dublin, Ireland Department of Science and Technology: Bioinformatics Service Platform Conference Review Panel: ISCB Travel Fellowships for ISMB July, 2015

research projects

SANBI's vibrant research environment is reflected in the diversity and number of projects that are currently underway.

Alan Christoffels

Theme	Projects
Host-pathogen interactions	• Virulence mutations: In collaboration with the Tygerberg MRC Unit, we are developing methods to analyse high throughput sequencing data for microbial genomes. • Identification of novel drug targets in pathways known to contain drug resistant genes. • Predicting the interaction networks between human and mycobacteria • In collaboration with the National Institute for Communicable Diseases (NICD), we are investigating miRNA targets in *Anopheles funestus* to understand regulation of mosquito development. • Identification of iron-regulatory networks in blood-feeding insects including tsetse fly. • Characterisation of promoters in tsetse immunity genes. • Characterising olfactory genes in tsetse.
Biobanking	• Our efforts to develop biobank standards continue through the NIH funded H3Africa project. This NIH-funded project ended in 2015. New funding obtained in partnership with European Union researchers. • Identifying biomarkers in HIV lymphomas.

Junaid Gamieldien

Theme	Projects
Tools and SOPs for clinical sequencing	• Development of a pipeline for identifying likely function impacting variants from raw exome and genome sequencing data. • Development of a semantic database and associated search strategies for prioritizing candidate disease causing/modifying variants.
Clinical application of exome sequencing	• Collaborative projects on: muscular dystrophy, myasthenia gravis, motor neuron disease, Parkinson's, schizophrenia and monogenic diabetes.
RNA sequencing	• Transcriptomic mechanisms of D-cycloserine in fear extinction in a rat model of post traumatic stress disorder using mRNA, miRNA and lncRNA sequencing. • Strain-specific gene expression by epithelial cells infected with Mycobacterium tuberculosis strains.
Cancer genomics	• Integrative analysis of multi-omic data to elucidate the molecular mechanisms of glioblastoma multiforme and breast cancer.

Gordon Harkins

Theme	Projects
Molecular evolution of viruses	• We have received funding from the NRF to investigate how virulence and pathogenicity are evolving in viruses important to human health using a combination of laboratory experimentation and state-of-the-art computational analyses. The spatiotemporal dynamics and evolution of a broad range of viral pathogens are being investigated focusing primarily on viruses that cause important human (HIV, SARS and Hepatitis C) and crop diseases (MSD, CMD and TYLCD). • Around a third of the world's countries still lack rubella vaccination programs, and as a consequence, rubella outbreaks remain frequent in many developing nations. We will employ Bayesian Coalescent Inference methods to estimate both how fast this viral pathogen is evolving, and the historical global dissemination pathways that underpin its current distribution while, simultaneously taking into account the roles of genetic recombination, pervasive secondary structure and natural selection. • In collaboration with scientists from the University of Cape Town, we have received funding from the MRC to sequence gp41 and p24 in 84 HIV-infected heterosexual couples in stable relationships attending the Manyanani clinic in Crossroads into this study to determine whether HIV sharing and super-infection between HIV+ partners in concordant relationships may account for increased viral load set points noted in such couples compared to those in discordant HIV+ relationships. • A longitudinal study investigating the evolution of HIV-1 subtype C viruses in the female genital tract relative to the blood plasma.

Uljana Hesse

Theme	Projects
Rooibos Transcriptome	• Sequencing the transcriptomes of rooibos *(Aspalathus linearis)* andeukaryotic micro-symbionts.

Nicki Tiffin

Theme	Projects
Human genetics underlying diseases in African populations	• Analysis of transcriptome profiles that characterise disease flare in immune cells of lupus patients from Cape Town. • Establishment of a patient registry clinical database for lupus patients in Cape Town. • Meta-analysis of public transcriptomic data to identify key pathways and biomarkers implicated in lupus. • Analysis of genetic factors underlying inherited myoclonic epilepsy in two South African families, using next generation exome sequencing techniques. • Analysis of genetic factors underlying inherited ESRD in a family from Cape Town, using next generation exome sequencing techniques. • Analysis of genetic factors underlying inherited lupus in a family from Cape Town, using next generation exome sequencing techniques. • Researching genetics and environmental factors underlying kidney disease in African patients.

Simon Travers

Theme	Projects
HIV Dynamics	• HIV drug resistance: In collaboration with groups in South Africa, Malawi and Ireland we are using next-generation sequencing approaches to study the emergence of HIV drug resistance in people exposed to ARVs and the presence of drug resistance in individuals who have never been exposed to ARVs. • HIV drug resistance testing: We have established a spin-off company (Hyrax Biosciences) that supports the establishment of routine HIV drug resistance testing in South Africa and other resource limited countries with a high burden of HIV using the Seq2Res computational pipeline developed at SANBI. • We are working with researchers at the South African National Institute for Communicable Diseases (NICD) to characterize the spectrum of diversity of IGHV gene variants in individuals capable of producing broadly neutralising antibodies against HIV. • Research investigating the role of N-linked glycosylation in the escape of HIV from, and susceptibility to, neutralizing antibodies. • Investigating the role of N-linked glycosylation on protein structure and function in relation to host cell recognition and entry. • Studying the role of sexual networks in the spread of HIV in schools in South Africa. • A number of studies as part of the HIV vaccine trials network (HVTN) analysing sequence data generated in multiple projects including sieve analysis for vaccine trials, ontogeny of B cell responses, transcription profiles.

Peter van Heusden

Theme	Projects
Cloud Computing	• We have initiated the testing of various private cloud computing solutions for biomedical research applications and to enhance our training laboratory.

research laboratories

Members of the Christoffels Lab

Research Laboratory of Alan Christoffels

Research projects

My bioinformatics laboratory focuses on host-pathogen interactions where we are developing high throughput genomics methods including next generation sequencing data analysis approaches to study communicable diseases such as tuberculosis, malaria and sleeping sickness.

The DST/NRF Research Chair funding supplemented with the MRC Flagship project (COMBAT-TB) provides the primary funding to support a range of infectious disease projects. These findings are only meaningful if translated into a clinical intervention strategy. To achieve this goal, a comprehensive research programme is underway that includes investigating genetics determinants in both host (Human) and pathogen (Mycobacterium tuberculosis) to understand drug resistance, and protein structure determination to inform patient-centric drug design.

These findings inform the development of a scientific workflow management system to support reproducible high throughput computational experiments.

The medicinal benefits of plant extracts generates a plethora of data. In collaboration with the UWC School of Pharmacy, we are exploring the use of computational tools to identify signatures in common between registered anti-malarial drugs and plant extracts that show anti-plasmodial activity.

The laboratory continues to expand its footprint on the African continent through a tsetse genome annotation project focusing specifically on chemosensory perception genes. The gene prediction and annotation expertise in the laboratory continues to expand our international footprint. In collaboration with reseachers in Kenya, we are describing the repertoire of chemosensory genes in the tsete fly. More recently, we partcipated in an International consortium to annotate the Asian Seabass genome.

Highlights of the lab for 2015:

Publications

4 papers published, 8 submitted.

Mbandi Kimbung's PhD work published in BMC Bioinformatics has been listed as a "highly cited" paper demonstrating the relevance of his research to an international audience.

Graduations

Congratulations to the 3 PhD graduates in my lab: Alecia Naidu, George Orbiero and Adugna Woldesemayat.

Funding

In May 2015, we secured funding from the European Union (Horizon2020) for a biobanking project entitled B3Africa. This consortium comprises 4 labs each in Africa and Europe. The primary goal is to develop an informatics framework to support biobanks in resource limited settings

International collaborations

In March 2015, I visited the Reproductive Genomics group in Singapore which sequenced the Asian Seabass. Our discussions finalised aspects of the project that needed to be completed before we could publish the data. This followed on from the 2014 annotation of the Asian Seabass genome which my lab performed as part of this international collaborative project.

In June 2015 I visited Missouri University, Columbia to explore collaboration opportunities in bioinformatics. We identified a few projects where researchers share common interests with SANBI.

Research Collaborations of Alan Christoffels

1. Prediction of human-tuberculosis interaction networks

Collaborating Parties:

Prof Peter Witbooi - Mathematics Department, UWC;

Prof Eileen Hoal-van Helden - Medical Biochemistry, Stellenbosch University

Nature and Purpose:

Use machine-learning techniques to improve the prediction of host-pathogen interactions between human mycobacteria.

Output in the last 12 months:

A PhD student submitted their thesis.

2. Develop a biobank informatics management system

Collaborating Parties:

B3Africa consortium funded through Horizon2020 and including:

Prof Akin Abayomi - Hematology Department, Tygerberg Hospital;

Prof Erik Bongcam-Rudloff, Swedish Agriscience, Uppsala;

Dr Heimo Muller, Graz University, Austria

Nature and Purpose:

Development of a laboratory management system that will integrate with other biobanking tools through an open-source platform called Biobank in a Box (BiBox).

Output in the last 12 months:

We secured EU funding to support the software development.

3. Chemosensory genes in Tsetse

Collaborating Parties:

Prof Serap Aksoy, Dr Paul Mireji - Yale University

Dr Dan Masiga - ICIPE, Kenya

Nature and Purpose:

To annotate chemosensory genes in 5 tsetse species

Outputs in the last 12 months:

Submitted two papers.

1 PhD graduation and 1 PhD thesis submission.

4. Identification of miRNAs in An.funestus

Collaborating Parties:

Prof Lizette Koekemoer - NICD, Pretoria

Nature and Purpose:

Mosquitoes from the Anopheline species are the principle vectors of malaria in Africa. Limited data exist for miRNA functions in Anopheline mosquitoes and few reports suggest miRNA control of plasmodium survival rate. We identified miRNAs that function in the survival and development of mosquitoes using NGS data and predicted miRNA targets **(www.insectar.sanbi.ac.za)**.

Outputs in the last 12 months:

Experimental quantification of selected miRNAs.

5. Identification of novel drug targets for drug resistant tuberculosis

Collaborating Parties:

Prof Wolf Dieter-Schubert – University of Pretoria

Dr Edwin Murungi, Egerton University, Kenya

Prof Samantha Sampson – University of Stellenbosch,

Prof Rob Warren – University of Stellenbosch,

Dr Cedric Wereley – University of Stellenbosch

Nature and Purpose:

Current TB drugs are more than 30 years old and have unacceptable efficacy and safety profiles, emphasising the need for new drugs. We mapped drug resistance genes, derived from comparative genome analysis of three *M. tuberculosis* strains (susceptible, multidrug, and extensively drug resistant), and from published literature to metabolic pathways and identified nine potential drug target candidate genes. These genes were ranked for further computational analyses in the quest for identifying inhibitors that could be lead compounds.

Output in the last 12 months:

Submitted a paper.

6. Chemo-informatics profiling of plant extracts that show anti-plasmodial activity

Collaborating Parties:

Profs Sarel Malan and James Syce, School of Pharmacy, UWC

Nature and Purpose:

At least 1000 compounds were identified in plant extracts and show antiplasmodial activity. We are using a range of computational tools to compare these compounds with currently registered antimalarial drugs to identify common signatures and ultimately prioritise the list of compounds identified in plant extracts.

Outputs in the last 12 months:

Submitted two papers and 1 PhD thesis.

Members of the Gamieldien Lab.

Research Laboratory of Junaid Gamieldien

Research projects

Semantic technology for disease genomics

The rapid growth in the volume, types and complexity of leads generated from the data has highlighted the need for knowledge-driven approaches for their further prioritization. Our BioOntological Graph Relationship Database (BORG) database assimilates and integrates multiple sources of genomic and biomedical knowledge and metadata and is able to learn rules about diseases and the phenotypes and gene functions associated with disease, which it utilizes to automatically prioritize likely causative candidates. This has proven especially useful in our disease variant and transcriptomic studies, where biologically and biomedically plausible links between identified mutations or differentially expressed genes, etc. and a disease or phenotype are automatically identified. The major benefit of the approach is the ability to make discoveries based on existing knowledge of a disease's molecular mechanisms, pathology and even 'intuitions' that clinical geneticists may have about the disease genetics. This 'guilt-by indirect-association' approach discovers links that would likely have been missed when directly consulting the literature or individual databases, while also providing the researcher with an easy to evaluate report for further vetting leads.

Several improvements additions have been made to knowledge corpus modelled in the semantic database, as well as to the methods that automatically explore the network for plausible 'genotype-to-phenotype' links. In addition to existing models for: demyelinating diseases, anxiety disorders, Parkinsonism, muscular dystrophies, motor neuron disease, soft-tissue injury, atypical diabetes and cancer, we are now able to rapidly develop new models for even disease cases with atypical presentation. We are thus especially able to assist researchers struggling to priorities candidates where standard filtering approaches fail to produce an obvious best candidate(s).

Development of SOPs for whole exome sequencing (WES)

For the clinical application of WES, which is the major applied research focus of our lab, it is crucial that the variant discovery process is carefully developed and rigorously optimized. We are therefore constantly identifying, testing and optimising best-of-breed software components for our WES processing pipelines. Furthermore, since WES produces large numbers of single nucleotide variants (SNVs), we are developing a variant prioritization method that uses an ensemble of tools that predict their likely functional impact, even in non-coding regions. Used in conjunction with our semantic discovery approach, we have been able to identify strong candidate causal variants in several in-house and collaborative disease studies that would otherwise have been overlooked or discarded due to the unobvious links to the phenotype.

Development of an exome based genetic diagnosis framework for monogenic and atypical diabetes, including MODY 'X'

As part of our MRC Strategic Health Innovation Partnerships grant, we are developing an exome sequencing based instrument for the genetic confirmation of Mature Onset Diabetes of the Young (MODY), neonatal diabetes (NDM) and ketosis prone diabetes.

MODY, which has an age of onset of less than 25 years and is a group of clinically heterogeneous, often non-insulin-dependent forms of diabetes mellitus (DM) defined at the molecular genetic level by deleterious mutations in different genes. The disease is estimated to be responsible for 2% of all cases of DM misdiagnosed in up to 80% of MODY cases, which are then sub-optimally treated. Equally important, a confirmed MODY genetic diagnosis often enables individualized treatment depending onthe gene involve and allows prognostication and early intervention in other family members carrying the mutation.

We are confident that we are able to identify known MODY mutations, novel mutations in known MODY genes and deleterious mutations in genes not yet associated with the disease, but which have the correct functional profile.

Clinical exome sequencing projects

We have several ongoing disease/risk variant discovery projects that benefit from customized versions of our WES pipeline and semantic discovery process:

- Distal muscular dystrophy
- Tendinopathy
- Early onset motor neuron disease

Highlights of the lab for 2015:

- 6 publications in international peer reviewed journals, with 2 more accepted for publication
- 2 book chapters
- Successful application of our semantic technology to identify what we believe to be the first ever genetic modifier of dysferlinopathy
- Postdoc Mahjoubeh Jalali presenting her dysferlinopathy exome work in at an international conference in Canada as an invited speaker
- Successful application for Technology Innovation Agency (TIA) seed funding to develop cloud based NGS analytical services platforms

Research Collaborations of Junaid Gamieldien

1. Clinical exome sequencing to identify the genetic cause of apparently atypical form of distal muscular dystrophy in a Cape Town family

Collaborating Parties:

Dr Mahjoubeh Jalali - SANBI, UWC

Prof Jeanine Heckmann and Dr Melissa Nel - University of Cape Town

Nature and Purpose:

To investigate the genetic basis of a seemingly autosomal dominant form of Miyoshi myopathy/dysferlinopathy using exome sequencing and semantic discovery.

Outputs in the last 12 months:

We have identified and confirmed by Sanger sequencing a novel mutation in the dysferlin gene that confirms Myoshi myopathy in the 1st generation and a 2nd novel mutation in a different that may explain the neuromuscular symptoms in the DYSF mutation carriers in the second generation and may also act as a genetic modifier of dysferlinopathy (manuscript in review).

Future Direction:

We will further profile the modifier gene to determine whether the pathways and processes it is involved in may provide leads for clinical intervention in the symptomatic carriers.

2. Exome sequencing of sporadic motor neuron disease: flail arm ALS

Collaborating Parties:

Dr Mahjoubeh Jalali - SANBI, UWC

Prof Jeanine Heckmann - University of Cape Town

Nature and Purpose:

To identify causative variants in two unrelated but highly similar cases of atypical flail arm amyotrophic lateral sclerosis.

Outputs in the last 12 months:

Participants have been exome sequenced and using our semantic model of motor neuron disease, strong candidate causative variants have been identified in an affected-child + unaffected parents trio.

Future Direction:

Candidates will be confirmed using Sanger sequencing.

3. Exome sequencing of myasthenia gravis-related ophthalmoparesis

Collaborating Parties:

Dr Mahjoubeh Jalali - SANBI, UWC

Prof Jeanine Heckmann (PI), Dr Melissa Nel (PhD candidate) - University of Cape Town

Nature and Purpose:

To identify the genetic cause of myasthenia gravis-related ocular muscle complications seen only in African patients.

Outputs in the last 12 months:

Exome data of control (ophthalmoparesis negative) and affected patients have been processed by the SANBI team and the identified candidate variants confirmed by Sanger sequencing and further prioritized by the UCT team.

Future Direction:

Candidate variants will be genotyped in a larger cohort.

4. Exome sequencing of atypical diabetes

Collaborating Parties:

Dr Mahjoubeh Jalali - SANBI, UWC

Prof Alison September (co-PI) - University of Cape Town

Prof Naomi (Dinky) Levitt - University of Cape Town

Nature and Purpose:

To test a clinical exome-sequencing pipeline developed as part of an MRC SHIP funded project by evaluating its ability to identify causative mutations in patients suspected of having mature onset diabetes of the young (MODY) or ketosis prone diabetes.

Outputs in the last 12 months:

We have developed a specialized semantic model of diabetes in our BORG database, which we have shown to be able to identify strong MODY for variants previously published as 'unexplainable'. Cases of possible MODY are being identified for exome sequencing-based genetic diagnoses.

Future Direction:

Selected DNA samples will be sent for whole exome sequencing. Additional cases that may benefit from WES will be sought.

5. Identification of genetic biomarkers for risk for musculoskeletal soft-tissue injuries (MSI)

Collaborating Parties:

Dr Colleen Saunders - SANBI, UWC

Prof Alison September and Prof Malcolm Collins - University of Cape Town

Nature and Purpose:

To identify candidate Achilles tendinopathy susceptibility genes and variants by comparing whole exome sequences of individuals with severe tendinopathy phenotypes and healthy physically active controls.

Outputs in the last 12 months:

We have developed a specialized semantic model of tendinopathy in our BORG database, which has been applied to a reanalyzed microarray expression dataset from the public repositories. Several strong candidate genes and mechanisms have been identified and a manuscript has been published in Scientfic Reports (Nature). A cohort of suitably matched cases and controls at the extreme ends of the phenotypic spectrum has been exome-sequenced and candidate predisposing variants identified.

Future Direction:

Candidates will be genotyped in a larger cohort.

6. Identification of genetic signatures to assess risk and recovery of concussion

Collaborating Parties:

Prof Alison September (PI) and Dr Michael Posthumous - University of Cape Town

Dr Jon Patricios - Pretoria University

Nature and Purpose:

Develop various risk and recovery assessment models for concussion injuries using sports concussion injuries as the model.

Outputs in the last 12 months:

None to date.

Future Direction:

We aim to identify potential genomic signature motifs which may assist in the

(i) risk assessment of concussion,
(ii) prediction of recovery,
(iii) assist in the confirmation of diagnoses of concussion in the absence of clinical symptoms (suspected concussion).

7. Identification of signatures of circulating miRNAs predictive of smoke inhalation injury in burn victims

Collaborating Parties:

Dr Kareemah Gamieldien (PI) - Cape Peninsula University of Technology

Dr Wayne Kleintjies - Head of the Tygerberg Hospital Burns Unit

Nature and Purpose:

To identify microRNAs in blood that may be used in the early diagnosis of smoke inhalation related lung injury in burn victims.

Outputs in the last 12 months:

Ethical clearance has been obtained from both CPUT and Stellenbosch University.

Future Direction:

Sample collection will commence in 2016 and small RNAs will be isolated from blood of patients meeting the clinical criteria.

8. An investigation into the molecular aetiology of Parkinson's disease in South African patients

Collaborating Parties:

Dr Soraya Bardien (PI) - Stellenbosch University

Nature and Purpose:

To investigate the genetic basis of Parkinson's disease in South African patients.

Outputs in the last 12 months:

The PhD thesis of Brigitte Glanzmann was passed.

Future Direction:

Manuscript preparation.

9. Molecular mechanisms of D-cycloserine (DCS) in a fear extinction posttraumatic stress disorder (PTSD) animal model

Collaborating Parties:

Dr Sian Hemmings, Prof Soraya Seedat, Dr Stefanie Malan-Muller - Stellenbosch University

Novartis, Switzerland

Nature and Purpose:

To investigate the molecular mechanism of action of DCS in facilitating fear extinction in an animal model of PTSD by performing gene expression and epigenetic analyses. Our research group is the bioinformatics partner in the project and has performed the RNAseq and microRNAseq data processing and downstream computational analysis to identify candidate drug targets and to elucidate possible mechanism of action of DCS.

Outputs in the last 12 months:

One manuscript has been published in an international peer reviewed journal and another submitted for publication.

Future Direction:

Confirmation of identified differentially expressed long noncoding RNAs and further exploration for possible epigenetic roles in the fear extinction process.

10. Understanding the SHARED ROOTS of Neuropsychiatric Disorders and Modifiable Risk Factors for Cardiovascular Disease

Collaborating Parties:

A South African Medical Research Council Flagship project - multiple collaborating institutions and departments, with Prof Soraya Seedat of Stellenbosch University as PI. Our group is the bioinformatics lead on the project.

Nature and Purpose:

The overarching aim of SHARED ROOTS is to use a 'whole systems' biology approach to identify biological, environmental and behavioural signatures of disease comorbidity in neuropsychiatric disease in order to direct future prevention and treatment. It will combine genomic, transcriptomic, epigenetic, and complementary phenotypic and multimodal neuroimaging data, to disentangle mechanistic pathways that lead to the development of comorbidity of these disorders.

Outputs in the last 12 months:

We have performed an ancillary study, which applied our optimized exome pipeline and variant prioritization protocol to the whole exome sequences of 3 female siblings that all carry the LRRK2 G2019S Parkinson's disease (PD) predisposing mutation but one patient is unaffected, one patient has PD and the third patient was diagnosed with schizophrenia prior to PD. We have thusfar identified a small number of strong candidates that may explain the schizophrenia case and submitted a manuscript for publication.

Future Direction:

Genomics data for the main project will be generated for qualifying participant samples.

11. RNAseq evaluation of gene expression by epithelial cells infected with Mycobacterium tuberculosis

Collaborating Parties:

Dr Manormoney Pillay (PI), Prof Bala Pillay and Dr Nonto Mvubu - University of KwaZulu-Natal

Nature and Purpose:

To elucidate potential strain-specific gene expression by epithelial cells infected with Mycobacterium tuberculosis strains of varying degrees of pathogenecity. Dr Gamieldien is the bioinformatics partner on the project.

Outputs in the last 12 months:

One manuscript has been published in an international peer reviewed journal and another submitted for publication.

Future Direction:

None yet defined.

Research Laboratory of Gordon Harkins

My research focuses on the evolution and molecular epidemiology of ssDNA and RNA viral pathogens of animals and plants. I am a member of a highly productive, plant-virus epidemiology network that seeks to determine the evolutionary underpinnings of the emergence and spread of the numerous viral diseases that seriously threaten the health and food security of Africa and the rest of the developing world. The rapid rate of evolutionary change in viruses means that the epidemiological and ecological processes that shape their genetic diversity act on approximately the same timescale as mutations that become fixed within viral populations. Consequently, the patterns of genetic variation present in viral genomes can be used to infer the processes underlying the dynamics of viral evolution, providing a unique molecular perspective on their ancestry and mechanisms of change.

A summary of some of the research projects that members of my research group have been involved in 2015 is presented below.

Plant-infecting viral pathogens

Virulence Evolution

We have received funding from the National Research Foundation (NRF) to investigate how the traits of virulence and transmission rates are evolving in viral pathogens important to human health. Using a combination of laboratory experimentation and state-of-the-art computational analyses, we are investigating the spatiotemporal dynamics and evolution of a broad range of viral pathogens focusing primarily on viruses that cause important crop diseases (Maize Streak Disease, Cassava Mosaic Disease and Tomato Yellow Leaf Curl Disease). Collectively, these studies are aimed at improving our understanding of the different life history strategies of pathogens, including reproduction, migration between and within hosts and virulence. By explicitly connecting our pathogenicity and virulence measurements to the natural history of the pathogen, we are confident this will provide a broader perspective on the evolution of pathogens for which current findings are mainly limited to theoretical or experimental settings.

Viral Emergence

In collaboration with researchers from CIRAD France, the University of Canterbury and the University of Cape Town, we have adopted a "geo-metagenomics" approach to investigate whether ecological disturbances such as intensive agriculture, cause measureable changes in the spatial and temporal diversity, demographics and evolutionary dynamics of viral communities inhabiting natural ecosystems that are linked to the emergence of socially-relevant crop-infecting geminiviral diseases. Our study population is the fynbos ecosystem situated in the Western Cape region of South Africa – a biodiversity hot-spot which is extremely threatened by increasing human population densities, urbanisation and agriculture. To complement the data on contemporary viral diversity derived from field studies, we are sequencing "ancient" viral genomes present in fynbos plant specimens stored within French and South African herbaria to more accurately estimate when and where, key ancestral plant viruses existed. By producing infectious clones of both these ancient genomes and re-synthesized computationally inferred ancestral genomes, it will be possible to pinpoint

when and where over the past few centuries, major changes in virus virulence have occurred. It is hoped that this project will complement the results obtained from our previous work on virulence evolution in Maize streak virus and increase our understanding of the evolutionary and epidemiological processes by which members of this important group of crop pathogens (i) emerged during the colonization era from their indigenous African hosts, (ii) began causing diseases in crop species introduced by Europeans and (iii) disseminated across the continent from sites of initial emergence to become what are today among the most important biotic threats to African food security.

Viral dynamics

The ongoing global spread of pathogenetic crop-infecting (chickpea, tomatoes, maize, bananas and cassava) viruses within the families *Geminiviridae* and *Nanoviridae* represent a serious looming threat to food production in many regions of the world. We applied Bayesian phylogeographic inference and recombination analyses to all available viral sequences and reconstructed a plausible history ongoing diversification and movements throughout the world and identified a range of predictor variables that influence the source-sink dynamics of these pathogens using Generalized Linear Models (GLM).

Human–infecting viral pathogens

RNA viruses

Our researchers are investigating the degree to which RNA viral evolution is constrained by secondary structure within the genomes of positive sense single-stranded RNA viruses in the order *Picornavirales* (one of the most genetically diverse of the positive-sense single-stranded RNA viral orders and the most common cause of infections in humans in developed countries). Among the most notable serious health threats include clinical syndromes such as HIV/AIDS, aseptic meningitis, paralytic poliomyelitis, SARS and hepatitis.

From left to right: Jean-Paul Toutain, Attaché for science and technology at the French Embassy in Pretoria, Mr. Xavier d'Argoeuves, Consul of France in Cape Town, Jean Albergel, IRD - CNRS - CIRAD Representative for Southern Africa, Brejnev Muhire (PhD student), Michel Eddi, CIRAD President Managing Director, Darren Martin, Associate Prof at University of Cape Town, Philippe Roumagnac, CIRAD Montpellier, France, Jacques Lançon, East and Southern Africa CIRAD Director, Gordon Harkins, senior lecturer at SANBI.

Human immunodeficiency virus

In collaboration with researchers from the University of Cape Town, we have received funding from the SA MRC to sequence gp41 and p24 in 84 HIV-infected heterosexual couples in stable relationships attending the Manyanani clinic in Crossroads to investigate why HIV-infected individuals in stable heterosexual relationships with HIV-infected partners (concordant HIV+) experience significantly higher plasma viral loads during chronic infection (higher set point) than HIV-infected individuals in relationships with uninfected partners (discordant HIV+). The molecular basis for this difference in plasma viraemia in concordant HIV+ couples is uncertain, although it is clear that viral load set point is a predictor of clinical disease course suggesting that concordant HIV+ couples will be at risk for more rapid disease progression. We aim therefore, to elucidate the molecular mechanism for increased viral loads in concordant HIV+ couples by using 454 sequencing to identify cases of dual, co- or super-infection.

Highlights of the lab for 2015

- 2 journal articles published in 2015 with an average impact factor of 3.015:
  - Molecular diversity of Chickpea chlorotic dwarf virus in Sudan: high rates of intra-species recombination a driving force in the emergence of new strains.
    Infection, Genetics and Evolution: (29) 2013-2015, 2015
  - The global distribution of banana bunchy-top virus reveals little evidence of frequent recent, human-mediated long-distance dispersal events.
    Virus Evolution, 1(1): vev009. doi: 10.1093/ve/vev009. 2015
- 2 MSc (cum laude) students and 1 PhD graduated in 2015.
- The MSc student Batsirai Mabvakure obtained his first coauthored publication in 2015 and has submitted a second for review in 2016.
- In recognition of their productivity in terms of scientific publications (34 in five years), Gordon Harkins along with his collaborators from the University of Cape Town (Associate Prof Darren Martin and Dr. Brejnev Muhire) were invited to present their research on important crop infecting viral pathogens to members of the Centre de coopération internationale en recherche agronomique pour le dévelopement (CIRAD). Visiting members included Jean-Paul Toutain, Attaché for science and technology at the French Embassy in Pretoria, Mr. Xavier d'Argoeuves, Consul of France in Cape Town, Jean Albergel, IRD - CNRS - CIRAD Representative for Southern Africa, Michel Eddi, CIRAD President Managing Director, Jacques Lançon, East and Southern Africa CIRAD Director and Philippe Roumagnac, CIRAD Montpellier, France.
- Philippe Roumagnac is currently spending two years in South Africa where he is funded by a Marie Curie fellowship to conduct research on the ecological and molecular mechanisms of viral emergence. The purpose of this visit was to showcase the research that has emanated from this highly productive international collaboration and to determine how to ensure that future bioinfomatics research initiatives between the respective French and South African scientific research institutes may be sustained when Philippe Roumagnac returns to France at the end of 2016. To this end, a further meeting will be held in April 2016 in Cape Town with Daniel Barthélémy (Director of Cirad System Biology Department) to determine the best way to proceed in this respect with the specific aim of identifying funding opportunities for students grants and to enable researchers to travel between the respective countries.

Research Collaborations of Gordon Harkins

1. Characterization of HIV-1 subtype C envelope sequences in the female genital tract and blood plasma during acute and chronic infection

Collaborating Parties:

Bronwen Lambson, Nthabeleng Ranchobe, Penny Moore and Lynn Morris - National Institute of Communicable Disease, South Africa.

Carolyn Williamson, Gama Bandawe, Joanne Passmore and Darren Martin - University of Cape Town, South Africa.

Salim Abdool Karim - CAPRISA.

Nigel Garrett, Koleka Mlisana - University of KwaZulu Natal, South Africa.

Gerhard Walzl - University of Stellenbosch, South Africa.

Nature and purpose:

To determine whether distinct viral populations exist within the female genital tract and blood plasma during acute and chronic infection. We are investigating the use of recently developed Bayesian spatial diffusion models to estimate the timing, direction and relative rates of movement between the cervix and the systemic vascular system inferred from HIV-1 subtype C sequences from several of the envelope genes from 4 CAPRISA patients sampled up to 4 years post infection.

Outputs in the last 12 months:

This work has resulted in one PhD graduate (Miss Kavisha Ramdayal) at SANBI in 2015.

Future Direction:

Two manuscripts reporting the findings of this study will be submitted for publication in 2016.

2. The impact of HIV sharing in HIV concordant heterosexual South African couples on viral load and clinical disease progression

Collaborating Parties:

Darren P. Martin, Alistair Grey, Carolyn Williamson and Anna-Lise Williamson - UCT, South Africa.

Aderito Monjane - University of Uppsala, Sweden.

Nature and Purpose:

HIV-infected individuals in stable heterosexual relationships with HIV-infected partners (concordant HIV+) experience significantly higher plasma viral loads during chronic infection (higher set point) than HIV-infected individuals in relationships with uninfected partners (discordant HIV+). As viral load set point is a predictor of clinical disease course concordant HIV+ couples will be at risk for more rapid disease progression. We aim therefore, to elucidate the molecular mechanism for increased viral loads in concordant HIV+ couples by using 454 sequencing to identify cases of dual, co- or super-infection.

Outputs in the last 12 months:

The first of publication is expected in 2016.

Future Direction:

This will be dependent on the findings from the current investigation.

3. Virulence Evolution

Collaborating Parties:

Philippe Roumangac - CIRAD, Montpelier, France.

Darren P. Martin, Adérito Luis Monjane, Dionne Natalie Shepherd, Ed Rybiscki - UCT, South Africa.

P. Lemey and Simon Dellicour - Department of Microbiology and Immunology, Katholieke Universiteit Leuven, Belgium.

Pierre Lefeuvre and Jean-Michel Lett - CIRAD, Reunion Island.

Arvind Varsani - University of Canterbury, Christchurch, New Zealand.

Mark Wamalwa - Bioscience Eastern and Central Africa, Nairobi, Kenya.

Nature and purpose:

We have received funding from the National Research Foundation (NRF) to investigate how the traits of virulence and transmission rates are evolving in viral pathogens important to human health. Using a combination of laboratory experimentation and state of-the-art computational analyses, we are investigating the spatiotemporal dynamics and evolution of a broad range of viral pathogens focusing primarily on viruses that cause important crop diseases (MSD, CMD and TYLCD). In July 2014 Prof Remy Froissant from the University of Montpelier, France spent six months in South Africa conducting experiments on the transmission and mortality rates of plants infected with Maize streak virus (MSV).

Outputs in the last 12 months:

None

Future Direction:

Besides formally testing the trade-off hypothesis using Maize streak virus (MSV) as a model plant-infecting viral pathogen, there are plans to extend the scope of the investigation to include pathogenic cassava and tomato-infecting viruses.

4. Molecular Mechanisms of Viral emergence

Collaborating Parties:

Pauline Bernardo, Sarah François, Maëlle Deshoux, Denis Filloux, Emmanuel Fernandez, Serge Galzi, Romain Ferdinand, Martine

Granier, Michel Peterschmitt, and Philippe Roumagnac - CIRAD-INRA-SupAgro, Montpellier, France.

Brejnev Muhire, Darren P. Martin - Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa.

Simona Kraberger, Arvind Varsani - University of Canterbury, Christchurch, New Zealand.

Anna-Liisa Laine, Mikko J. Frilander - University of Helsinki, Finland.

Armelle Marais, Thierry Candresse - INRA, UMR Villenave d'Ornon Cedex, France.

Pablo Monge, Fernando Escriu - Unidad de Sanidad Vegetal Zaragoza, Spain.

Nature and purpose:

To investigate whether ecological disturbances such as intensive agriculture, cause measureable changes in the spatial and temporal diversity, demographics and evolutionary dynamics of viral communities inhabiting natural ecosystems that are linked to the emergence of socially- relevant crop-infecting geminiviral diseases such as TYLCD, MSV and CMD.

Outputs in the last 12 months:

This work has resulted in one PhD graduate and a paper submitted in 2015.

Future Direction:

The rate of discovery of "emergent" viruses has increased over the past two decades. However, it remains difficult to determine whether these viruses are truly emergent, or, if they have simply remained undetected until now. We have shown that one such plant-infecting single-stranded DNA geminivirus species named Euphorbia caput-medusae latent virus displays a high prevalence in samples of the wild spurge c aput-medusae collected between Cape Town and Paternoster on the west coast of South Africa is related to other newly discovered species that cause severe infections in French beans in India and alfalfa in France. It is known that the latter virus is aphid transmitted and, if the other related viruses are also transmissible by aphids, it would suggest these viruses have the potential to emerge as serious crop pathogens. Therefore, in 2016 we will conduct experiments to definitively identify the insect vector responsible for the transmission of Euphorbia caput-medusae latent virus in South Africa to complement the work conducted elsewhere, and submit a manuscript describing our results for publication.

5. Endogenous Viral Elements (EVE's)

Collaborators:

Philippe Roumangnac, Denis Filloux and Emmanuel Fernandez - CIRAD, Montpelier, France.

Darren Martin - University of Cape Town.

Véronique Jamilloux, Florian Maumus - INRA, URGI, Versailles, France.

Pierre-Yves Teycheney - CIRAD, UMR AGAP, Guadeloupe, France.

Pierre Lefeuvre, Jean-Michel Lett - CIRAD, UMR PVBMT, Réunion Island, France.

Thierry Candresse, Armelle Marais, Sébastien Theil and Chantal Faure - INRA, UMR BFP, Villenave d'Ornon France.

Andrew Geering - the University of Queensland, Australia.

Nature and purpose:

To study endogenous viral elements (EVE's) and their role in virus evolution and their functions in plants to address fundamental biological issues that are critical to the understanding of plant virus evolution and the contribution of plant viruses to both the host genome evolution and biology. This involves approaches in virology, bio-informatics, genomics, proteomics, evolution and synthetic biology, and aims at disentangling key mechanisms of adaptive plasticity of plant and viral genomes such as gene domestication, genetic and epigenetic regulation of plant gene expression and deciphering mechanisms that are central to molecular plant-virus interactions such as recombination, transcriptional or post transcriptional gene silencing.

Outputs in the last 12 months:

None

Future Direction:

A funding proposal to support this research has been re-submitted to the French National Research Foundation in November 2015.

Uljana Hesse on a visit to a rooibos farm in Nieuwoudtville.

Research Laboratory of Uljana Hesse

Since joining SANBI in April 2015, I have

- successfully applied for funding for the rooibos transcriptome project
- lectured "comparative genomics" in the UWC undergraduate Bioinformatics course (BTN323)
- trained two undergraduate students in bio-computational sequence assembly (these are interns at IMBM).

Research Collaboration of Uljana Hesse

1. Sequencing the transcriptomes of rooibos *(Aspalathus linearis)* andeukaryotic micro-symbionts

Collaborators:

Prof Jeanine L Marnewick - Cape Peninsula University of Technology (Oxidative Stress Research Centre), Cape Town, South Africa

Dr Marilize Le Roes-Hill - Cape Peninsula University of Technology (Biocatalysis and Technical Research Group), Cape Town, South Africa

Nature and Purpose:

Rooibos *(Aspalathus linearis),* widely used for anti-colic, anti-allergic, and anti-aging properties, is one of the most important cash-crops developed from the endemic flora of South Africa. Production currently amounts to ~18.000t, most of which is exported overseas. It provides work for ~5000 people in the Cederberg area where it is a key cash-crop, earning an approximate R500 Million per year. Health promoting qualities are mainly attributed to distinctive polyphenolic antioxidants found in this plant. Preliminary investigations have shown that rooibos is home to a plethora of micro organisms, which may affect the production of polyphenols and other secondary metabolites or even produce some themselves. So far, neither the biosynthetic pathways of secondary metabolites in rooibos nor their origin (plant or symbiont) have been resolved. We have just received funding from the Research and Technology Fund (NRF) to investigate the transcriptomes of rooibos and potential eukaryotic symbionts using a meta-transcriptomics approach.

Future Direcion:

Our study will focus on

1) the development of interactive databases for fully annotated rooibos and symbiont transcriptomes
2) identification of genes potentially involved in polyphenolics production,
3) study the genetic divergence of rooibos ecotypes,
4) develop molecular biomarkers for rooibos ecotype diversity studies and targeted plant propagation,
5) initiate a culture-bank of characterized symbiont isolates.

The project is supported by the leading rooibos producer (Rooibos LTD) and by the rooibos farmers from the three main rooibos producing areas (Clanwilliam, Wupperthal and Nieuwoudtville).

Members of Nicki Tiffin's lab.

Research Laboratory of Nicki Tiffin

I work on human genetics underlying disease, specifically in African populations, aiming to characterise genetic diversity in South Africa patient populations within the disease context. Ongoing projects include a collaborative project establishing a registry of patients from Cape Town who have systemic lupus erythematosus (SLE).

We are building a database for effective storage and data-mining of extensive clinical and biochemical patient data for these patients, and will use this data to design and implement –omics studies to further elucidate the genetic and environmental contributors to this disease, as well as to characterise the clinical characteristics of lupus in South African patients. We are actively involved in establishing the African Lupus Genetics Network (ALUGEN), which is an informal network of African researchers and clinicians who see patients with SLE and would like to use the registry database structure established at SANBI to collect standardised data about their patients.

We are also undertaking research to understand the molecular drivers of disease flare in lupus patients in Cape Town. In parallel, we are undertaking bioinformatics approaches to harness the public data available from studies on lupus genetics, and to understand gene regulatory pathways and processes that are implicated in lupus. Ongoing exome sequencing projects are investigating genetic factors underlying disease in Cape Town families with inherited early-onset, autosomal dominant end stage renal disease; with autosomal dominant myoclonic epilepsy, and with familial systemic lupus erythmatosus. We have also contributed to a collaborative project investigating variants potentially associated with non-syndromic sensorineural deafness in Africa.

We continue to investigate issues for genomics research in Africa, including the genetic diversity found in Africa, ethical and practical challenges.

Highlights of the lab for 2015:

- Highlights for the year include the implementation of both the ALUGEN database and the NetCapDB database.
- This year sees the submission of degree theses by Ms Tracey Calvert-Joshua (MSc), Ms Stephanie Pitts (MSc) and Mr Darlington Mapiye (PhD).

Research Collaborations of Nicki Tiffin

1. Genetics underlying systemic lupus erythematosus (SLE/lupus) in South African Patients; the ALUGEN network.

Collaborators:

Dr Ikechi Okpechi - Division of Nephrology and Hypertension, University of Cape Town/Groote Schuur Hospital

Dr Bridget Hodkinson - Division of Rheumatology, University of Cape Town/Groote Schuur Hospital

Nature/purpose:

Historically, there has been a belief that lupus does not occur in Africans on the African continent. In our experience, however, there are many lupus patients in South Africa and other African countries and we would like to gain insights into the mechanisms and presentation of lupus in Africans. We have established the lupus registry to capture clinical information about our patient population, and we biobank DNA from consenting participants. Further to this, we are making the purpose-built database framework accessible to African partners who may wish to start their own lupus registries. We have an ongoing study using RNAseq next generation sequencing technologies to investigate gene expression changes in patients in disease flare.

Outputs in last 12 months:

Publication: B Hodkinson, D Mapiye, D Jayne , A Kalla, N Tiffin and I Okpechi. The African Lupus Genetics Network (ALUGEN) registry: standardized, prospective follow-up studies in African patients with systemic lupus erythematosus. Lupus (2015) 0, 1–6

Future direction:

The database is live and in 2015 we will be encouraging African partners to join the ALUGEN consortium.

2. Establishing a bioinformatics network in Africa for bioinformatics capacity and skills development.

Collaborators:

The H3Africa Bioinformatics Network (H3ABioNet).

P.I. Dr Nicky Mulder - Department of Computational Biology, University of Cape Town.

Nature/purpose:

H3ABioNet is a pan-African network of bioinformatics groups, and the aim of the network is to develop bioinformatics skills and capacity to support both ongoing H3Africa research projects and the African health genomics research community. The SANBI node has developed a database to capture bioinformatics capacity within the network, and to capture changes in capacity and research output across the five funded years of the project, and beyond. The database has a browser-based front end for data entry by H3Africa nodes, and can also be queried through a web-based query page which ouputs graphs and charts of the collected metrics. Reporting is semi-automated to assist with data collation and presentation to the NIH. The SANBI node is also supporting the RedCap platform for clinical databasing.

Outputs in last 12 months:

Publication: Mulder NJ, Adebiyi E, Alami R, Benkahla A, Brandful J, Doumbia S, Everett D, Fadlelmola FM, Gaboun F, Gaseitsiwe S, Ghazal H, Hazelhurst S, Ibrahimi A, Hide W, Jaufeerally Fakim Y, Jongeneel V, Joubert F, Kassim S, Kayondo J, Kumuthini J, Lyantagaye S, Makani J, Alzohairy AM, Masiga D, Moussa A, Nash O, Ouwe Missi Oukem-Boyer O, Owusu-Dabo E, Panji S, Patterton H, Radouani F, Sadki K, Seghrouchni F, Tastan Bishop Ö, Tiffin N, Ulenga N; H3ABioNet Consortium. H3ABioNet, a sustainable Pan-African Bioinformatics Network for Human Heredity and Health in Africa. Genome Res. 2015 Dec 1. pii: gr.196295.115. [Epub ahead of print]

NetCapDB database and reporting interface: **http://netcapdb.sanbi.ac.za/** in use.

Future direction:

H3ABioNet is currently in its third year of funding. SANBI will be involved in the reporting of change in capacity of the network to date, as well as the development of Standard Operating Procedures for next generation sequencing methodologies and variant analysis.

3. Researching genetics and environmental factors underlying kidney disease in African patients.

Collaborators:

The H3Africa Kidney Disease Research Network

PIs Prof. Dwomoa Adu - University of Ghana Medical School,

Prof Akinololu Ojo - University of Michigan.

Nature/purpose:

The H3A KDRN is a multi-centre pan-African research network that investigates kidney disease in African patients at the clinical and molecular level. SANBI is a bioinformatics partner in the NIH-funded programme.

Future direction:

Currently the focus of the project is on collecting samples and data; but the focus will shift to data analysis, going forward.

Outputs in last 12 months:

Publication: Azeez Butali, Peter Anthony Mossey, Nicki Tiffin, Wasiu Lanre Adeyemo, Mekonen Abebe Eshete, Chrispinanus Mumena, Rosemary Audu, Chika Onwuamah, Pius Agbenorku, Mobolanle Olugbenga Ogunlewe, Adetokunbo Raphael Adebola, Hecto Oladapo Olasoji, Babatunde Aregbesola, Ramat Oyebunmi Braimah, Abimibola Victoria

Oladugba, Ifeanyichukwu Igwilo Onah, Ezekiel Adebiyi, Peter Babatunde Olaitan, Lukman Olajide Abdur-Rahman, Adebowale Adeyemo.

Multidisciplinary approach to genomics research in Africa: the AfriCRAN model.

The Pan African Medical Journal. 2015;21:229 doi:10.11604/pamj.2015.21.229.7380

4. Genetics underlying Autosomal Dominant Kidney Disease in a South African family.

Collaborators:

Dr Ikechi Okpechi - Department of Nephrology and Hypertension, University of Cape Town/Groote Schuur Hospital

Prof Iqbal Parker - ICGEB Cape Town.

Nature/purpose:

End stage renal disease is a debilitating illness that affects many Africans. In order to understand the aetiology of the disease, we are performing exome sequencing of six individuals from a Cape Town family that suffers from an autosomal dominant form of severe, early onset ESRD.

Future direction:

The sequencing has been completed and the analysis of identified variants is complete. Manuscript in preparation.

5. Genetics underlying Myoclonic Epilepsy in a South African family.

Collaborators:

Prof Jonathan Carr - Division of Neurology, University of Stellenbosch

Prof Charles Rotimi - CRGGH, USA

Nature/purpose:

Two large South African families have a rare phenotype inherited autosomal dominant myoclonic epilepsy. In order to better understand the disease aetiology, we have exome sequenced six individuals from the family and are analysing the variants. In order to gain some information about the location of aetiological variants, our collaborators in the group of Prof Charles Rotimi (NIH) are genotyping 21 individuals from the family in order to identify genomic regions associated with the disease.

Outputs in last 12 months:

Publication: Wright GEB, Bardien S, Carr JA, Tiffin N. Poster: Exome sequencing of a familial form of adult myoclonic epilepsy in two South African pedigrees. Human Genome Meeting 27 – 30 April 2014, Geneva.

Future direction:

The GWAS analysis is ongoing and will assist in identifying the pathogenic variant.

6. Genetics underlying Familial Lupus in a South African family.

Collaborators:

Dr Mike Urban - Department of Genetics, University of Stellenbosch

Dr Monika Esser - Department of Rheumatology, University of Stellenbosch/Tygerberg Hospital

Dr Ikechi Okpechi - Department of Nephrology, University of Cape Town/Groote Schuur Hospital

Nature/purpose:

In rare instances, families may be affected by an inherited form of lupus, and the familial nature of the disease may offer insights into the pathology of lupus in these individuals. We are performing exome sequencing of five individuals from a South African family with familial lupus, in order to identify variants that may contribute to the disease phenotype.

Future direction:

The exomes have been sequenced and data analysis is currently underway.

7. (Dr Jean-Baka Domelevo-Entfellner) Bootstrapping large phylogenies for HIV and mammalian phylogenies with a more robust confidence measure on splits.

Collaborators:

Prof Olivier Gascuel (LIRMM, CNRS/Université de Montpellier), France

Dr Emmanuel Douzery (ISEM, CNRS/Université deMontpellier), France

Dr Tulio de Oliveira (UKZN/Africa Centre for Health and Population Studies), South Africa

Dr Eduan Wilkinson (UKZN Africa Centre for Health and Population Studies), South Africa

Nature/purpose:

The bootstrap method is a classical tool in statistics to assess the robustness of a given model, from a resampling of the data. While present-day genomic datasets are orders of magnitude bigger than those in use decades ago, it is observed that the bootstrap method, when used on large trees made of several thousand taxa, performs poorly: well-established nodes of the phylogeny suffer low bootstrap values with no apparent reason. In this collaboration we try to clearly establish this phenomenon on large HIV phylogenies and also mammalian genomic datasets; and to provide the community with a more robust confidence measure on splits that is free from this systematic issue.

Future direction:

Manuscript in preparation.

Research laboratory of Simon Travers

2015 was an exciting year for my group and saw us spin-out a company, Hyrax Biosciences, to commercialise the exatype™ HIV drug resistance analysis pipeline to enable the use of next-generation sequencing for HIV drug resistance testing. In December 2015 we (Hyrax Biosciences) signed an exclusive licensing agreement with the University of the Western Cape for the use of exatype™. This has been the result of many years of intensive research by a dedicated and brilliant team and the credit for this exciting move must go to them. I'm very excited and proud that we can use technology that was created at SANBI to positively affect the lives of potentially millions of people affected by the scourge on society that is HIV.

Aside from the HIV drug resistance work we have strengthened our partnership with the HIV vaccine trials network through the recruitment of Phillip Labuschagne who is responsible for the development of computational pipelines for the routine analysis of viral sequence data generated as part of HIV vaccine trials being undertaken in South Africa.

My other major research programme focuses on using molecular modelling to explore the diversity of HIV's "glycan-shield" and its susceptibility to, and escape from, neutralising antibodies. This programme is highly novel and we have spent the last number of years establishing the pipelines, protocols and capacity to be able to do this work. 2015 saw the completion of this establishment meaning that 2016 will be an exciting year with the results from this work being released.

In 2015 we recruited a PhD student to continue this work and two MSc students completed their projects and are anticipated to graduate in 2016.

The award of the DST/NRF Centre of Excellence in HIV Prevention to CAPRISA (of which we are a member) is an exciting prospect as it means we will be able to significantly expand our work on glycosylation in HIV and the role it has in the development, and eventual success, of a HIV vaccine.

Research projects

The development of computational approaches for highly sensitive analysis of next-generation sequencing (NGS) data.

One of the biggest challenges with NGS is the complexity involved in the analysis of the data. Our exatype™ platform is an easy to use, cloud-based platform that facilitates the use of NGS for HIV drug resistance testing.

It supports the analysis of data generated by the Illumina, Roche/454 and Ion Torrent sequencing platforms and is accessed through an intuitive web site that enables people with no bioinformatics training to analyse their data.

Current research within this programme is focused on the addition of support for resistance genotyping of other infectious organisms within the exatype™ platform and the establishment of partnerships/research to facilitate the future rollout of point of care resistance genotyping.

The use of high-throughput sequencing approaches to explore viral and host diversity.

While NGS approaches are immensely powerful for sequencing the entire breadth of a whole genome (e.g. human genome) in a cost-effective and efficient manner, they also provide considerable power for generating 'ultra-deep' data enabling us to identify low abundance viral variants infecting individuals. In 2015 we published a paper that described the use of NGS approaches to characterize the potential of individuals to produce broadly cross-neutralising antibodies against HIV. We also are heavily involved in the HIV vaccine trials network (HVTN) where we are focused on the development of pipelines to analyse NGS data produced using the primer ID approach, a method that allows the accurate identification of the prevalence of viral variants within the viral quasispecies of a HIV infected individual.

The application of molecular dynamics approaches to characterise the viral epitopes of broadly cross-clade neutralising (BCN) antibodies with a view to identifying novel targets for HIV vaccine design.

Recent studies have identified antibodies that are capable of neutralising a wide range of HIV strains, prompting interest in eliciting them as part of a vaccine response. These studies have shown that carbohydrates bound to the surface of HIV's gp120 surface protein form all, or part, of the epitope for many of these antibodies. Very little is known, however, about the structural dynamics of these carbohydrates and how they may affect susceptibility to, and escape from, neutralisation by these carbohydrate binding antibodies.

We work with researchers at the University of Georgia (US), Maynooth University (Ireland), NICD and UCT (SA) to undertake molecular dynamic modeling work to enhance our knowledge of N-linked glycans bound to the surface of HIV virions.

The use of molecular epidemiology to explore the modes of transmission among school learners in a rural setting in KwaZulu-Natal (South Africa).

Recent work has shown that the ratio of HIV prevalence between learners in rural KwaZulu-Natal is disproportionate with low levels of HIV infection in young men while the prevalence in young female learners is greater than 20%. The reasons for this are not fully understood and, together with collaborators in CAPRISA and the University of Cape Town, we are undertaking a study exploring the modes of transmission in order to facilitate the effective implementation of HIV prevention strategies. 2015 saw the awarding of a RO1 collaborative grant from the United States National Institutes of Health (NIH) to undertake a large study identify the sources of HIV infection in adolescent girls in rural South Africa. Our role is to undertake a study using molecular phylogenetics to understand the transmission dynamics within the community and to identify transmission clusters within the community.

Research Collaborations of Simon Travers

1. The role of N-linked glycosylation in the escape of HIV from, and susceptibility to, neutralizing antibodies.

Collaborating Parties:

Dr Oliver Grant - University of Georgia

Prof Robert Woods - University of Georgia

Dr Elisa Fadda - Maynooth University, Ireland

Prof Penny Moore – NICD, South Africa

Dr Jeffrey Dorfman – University of Cape Town, South Africa

Dr Natasha Wood – University of Cape Town, South Africa

Nature and purpose:

As part of the post-translational processing of a HIV virion carbohydrates are added to the surface of the virion by the hosts glycosylation mechanism. The binding of such N-linked glycans conveys protection to a virion's surface proteins by acting as a shield to avoid detection by the host's immune system. These carbohydrates, however, may comprise a novel target for HIV therapeutics and we are using molecular dynamic simulations to further understand the conservation and dynamics of these carbohydrates and their effect on the accessibility of the underlying protein.

Outputs in the last 12 months:

Gabier E, Wood N, Fadda E, Grant O, Woods R, Travers SA (2015) The Dynamics of HIV-1 gp160 N-linked Glycans in the context of Broadly Cross Neutralizing Antibodies. 7th SAAIDS Conference.

Wood N, Fadda E, Grant O, Woods R, Travers SA (2015) Subtype-specific structural characteristics and molecular dynamics of glycosylated HIV-1 gp120 proteins 22th International HIV Dynamics and Evolution Conference.

Mercuur C, Wood N, Fadda E, Grant O, Woods R, Travers SA (2015) Using molecular dynamics to investigate the effects of N-linked glycosylation on the gp120 envelope trimer of HIV-1. 7th SAAIDS Conference.

Future Direction:

We are part of the DST-NRF centre of excellence in HIV prevention awarded to CAPRISA and this award will facilitate the expansion of the research programme in 2016.

2. Development and application of the exatype™ HIV drug resistance genotyping pipeline.

Collaborating Parties:

Prof Maria Papathanasopoulos - Wits Medical School

Dr Miguel Lacerda - University of Cape Town

Nature and purpose:

In 2015 we undertook a study in collaboration with the NHLS at the University of Witwatersrand exploring the ability of various NGS sequencing platforms for sensitive HIV drug resistance genotyping. We have worked with a number of research and pathology laboratories throughout South Africa assisting in the analysis of their HIV resistance data generated using various NGS platforms. Version 1 of the exatype™ pipeline was released in December 2015 by Hyrax Biosciences (spun out from SANBI) following the signing of an exclusive licensing agreement with the University of the Western Cape.

Output in the last 12 months:

Travers SA. Tools and pipelines for the analysis of NGS data in HIV vaccine research. HVTN/CHIL NGS Sequencing Workshop, 15-16th October 2015, Cape Town.

Travers SA, Wright IA, Wood N, Lubinsky B (2015) Enabling cost-effective, high throughput HIV drug resistance genotyping. Ion World Tour 2015, Pretoria (Invited speaker).

Wright IA, Travers SA (2015) NOAH: A novel guide tree ordering and progressive alignment heuristic. ISCB Africa ASBCB 2015.

Travers SA (2015) Big data in public health. NRF Consultative Workshop on Community of Practice in Public Health, UWC (Invited speaker).

Travers SA (2015) From PhD project to spinoff company – enabling low-cost HIV drug resistance testing. 1st Massachusetts-South Africa Conference for Technology Transfer and Global Innovation, Stellenbosch (Invited speaker).

Travers SA (2015) Harnessing the power of "next-generation" sequencing platforms to enable high-throughput, cost-effective HIV drug resistance testing. UCT Institute of Infectious Diseases and Molecular Medicine seminar series (Invited speaker).

Future Direction

Future work will see the marketing of the exatype™ pipeline as a commercial product. Further, we will expand the ability of the analysis pipeline to facilitate genotypic resistance testing of various infectious organisms.

3. Using next-generation sequencing to explore the diversity of the HIV viral quasispecies.

Collaborating Parties:

Prof Lynn Morris - NICD, South Africa

Prof Carolyn Williamson - UCT, South Africa

The HIV vaccine trials network (HVTN)

Nature and Purpose:

We are involved in a number of research projects that are harnessing the power of NGS to enable us to fully understand the diversity of the viral quasispecies in HIV infected individuals. The vast majority of this work is focused on the identification of the potential of individuals to produce broadly cross-neutralising antibodies as well as exploring the viral diversity of breakthrough infections following vaccination with HIV vaccine candidates.

Outputs in the last 12 months:

Cathrine Scheepers, Ram K. Shrestha, Bronwen E. Lambson, Katherine J. L. Jackson, Imogen A. Wright, Dshanta Naicker, Mark Goosen, Leigh Berrie, Arshad Ismail, Nigel Garrett, Quarraisha Abdool Karim, Salim S. Abdool Karim, Penny L. Moore, Simon A. Travers and Lynn Morris (2015) Ability To Develop Broadly Neutralizing HIV-1 Antibodies Is Not Restricted by the Germline Ig Gene Repertoire. The Journal of Immunology. 194(9):4371-8

Travers SA. Tools and pipelines for the analysis of NGS data in HIV vaccine research. Invited speaker at HVTN/CHIL NGS Sequencing Workshop, 15-16th October 2015, Cape Town

Future Direction:

The development of pipelines and analysis of data will continue over 2016 as we move towards the establishment of a number of HIV vaccine trials that will start in South Africa in the next number of years.

4. The use of molecular epidemiology to explore the modes of transmission among school learners in a rural setting in KwaZulu-Natal (South Africa)

Collaborating Parties:

Prof Ayesha Kharsany - CAPRISA

Prof Carolyn Williamson - University of Cape Town

Nature and Purpose:

Recent work has shown that the ratio of HIV prevalence between learners in rural KwaZulu-Natal is disproportionate with low levels of HIV infection in young men while the prevalence in young female learners is greater than 20%. The reasons for this are not fully understood and we are undertaking a study exploring the modes of transmission in order to facilitate the effective implementation of HIV prevention strategies.

Future Direction:

The research team lead by Prof Ayesha Kharsany at CAPRISA has been awarded a US National Institutes of Health RO1 grant that will build upon the preliminary work studying transmission dynamics within and between school learners in Vulindlela. This project will involve cross-sectional sampling from a diverse range of more than 23,000 individuals across the entire area. Viral fragments will be sequenced in individuals identified as HIV positive and a molecular epidemiological study will be undertaken at SANBI to identify the presence of transmission clusters within the population that contain sequences sampled from learners. This clustering will be correlated with epidemiological data collected during sampling and will be used to identify the modes of transmission.

research collaborations

SANBI researchers have established a vast network of partnerships and collaborations all over the world.

2015 financials

SANBI's income and expenditure trends for 2015 are shown in this section. 68% of the grant income was expensed in 2015. Factors that contributed to the unspent funds are the late receival of Letters of Awards and UWC shutdown periods during protest action.

FIGURE 1. DISTRIBUTION OF INCOME RECEIVED FROM ALL SOURCES.

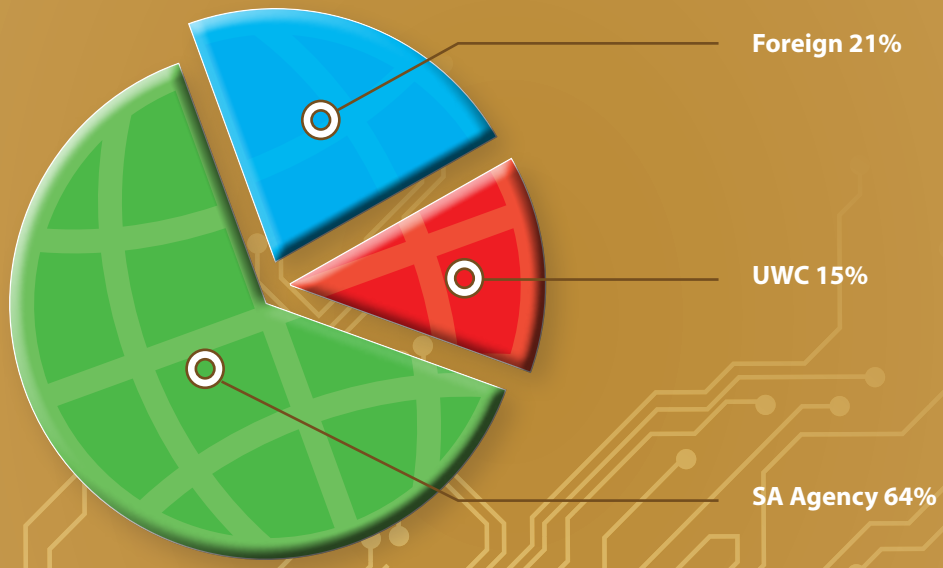


FIGURE 2. INCOME RECEIVED FROM SA SOURCES.

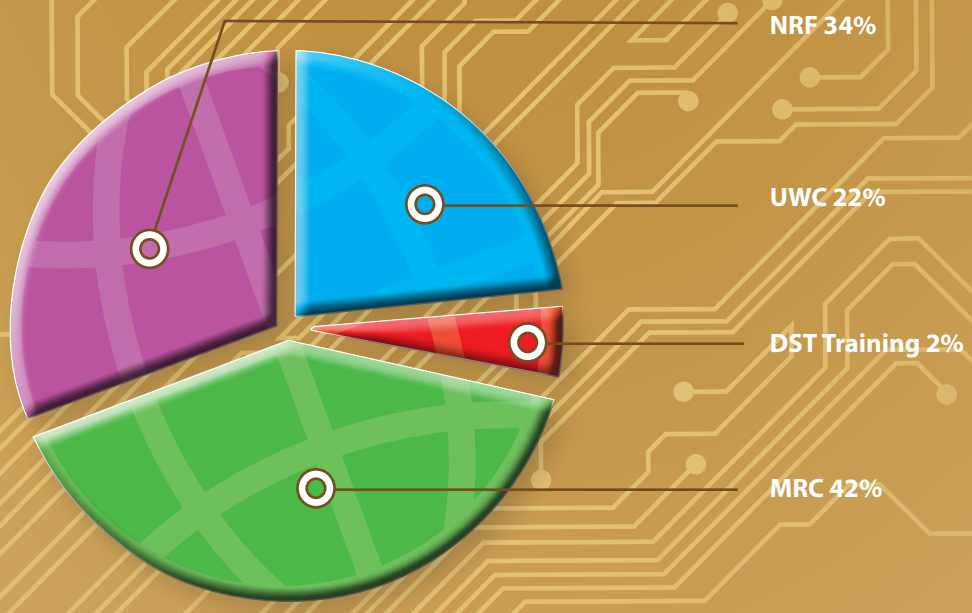


FIGURE 3. INCOME VS. EXPENDITURE 2010 – 2015.

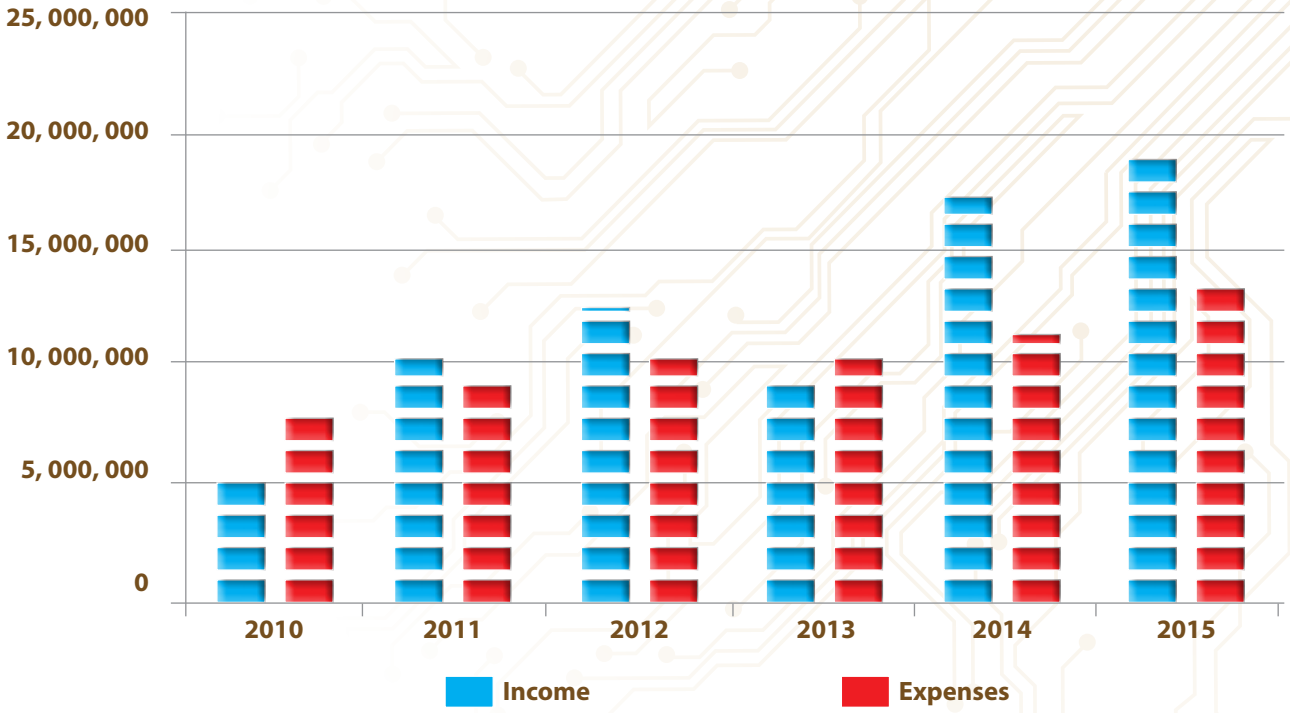
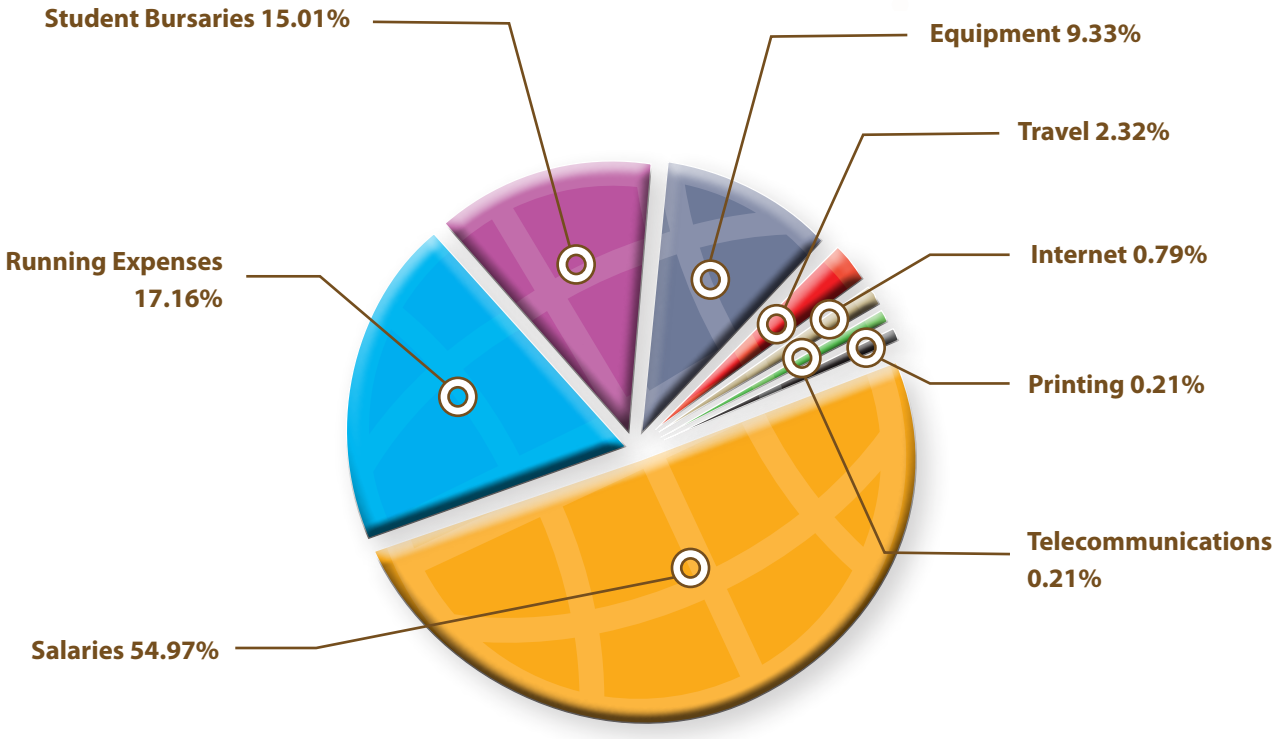


FIGURE 4. EXPENDITURE IN 2015.



students 2015

POSTDOCTORAL FELLOWS


Ruben Cloete


Jean-Baka Domelevo Entfellner


Mahjoubeh Jalali


Zahra Jalali


Sarah Mwangi


Pavan Rallabandi


Colleen Saunders


Imogen Wright

DOCTORAL


Ibrahim Ahmed


Mmakamohelo Direko


Samuel Egieyeh


Azeez Fatai


Roux-Cil Ferreira


Rosaline Macharia


Darlington Mapiye


Alecia Naidu


Anati Nkaule


Catherine Rossouw

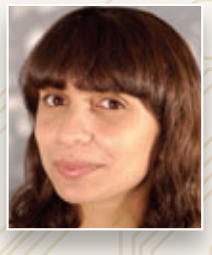

Emil Tanov

MASTERS

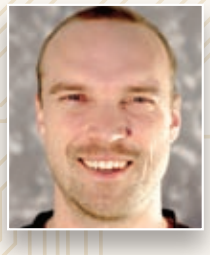
Toluwaleka Ademuyiwa

Olabode Ajayi

Tracey Calvert-Joshua



Ereshia Gabier



Phillip Labuschagne

Batsirai Mabvakure

Eugene Madzokere

Clint Mercuur

Stephanie Pitts

Lunga Sizani

Peter Van Heusden

Larry Van Vuuren

Werner Veldsman

Luvuyo Morris

alumni

STAFF

Name	Currently
Winston Hide	Chair of Computational Biology Sheffield Institute for Translational Neuroscience, Department of Neuroscience, University of Sheffield Adjunct Prof, Harvard School of Public Health Principal Faculty, Harvard Stem Cell Institute.
Vladimir Bajic	Director & Professor Computational Bioscience Research Center, King Abdullah University of Science and Technology
Heikki Lehvaslaiho	Senior Research Scientist Computational Bioscience Research Centre, King Abdullah University of Science and Technology
Tulio de Oliviera	Senior Bioinformatics Researcher Africa Centre for Health and Population Studies, University of KwaZulu-Natal
Nicky Mulder	Group Head Computational Biology Group, University of Cape Town
Cathal Seoighe	Stokes Prof of Bioinformatics School of Mathematics, Statistics and Applied Mathematics, National University of Ireland, Galway
Dale Gibbs	Self-employed

POSTDOCTORAL FELLOWS

Name	Date completed	Currently
Soraya Bardien-Kruger	2002	Associate Prof, University of Stellenbosch
Vladimir Babenko	2002	Senior Staff Scientist, IC&G
Janet Kelso	2004	Max Planck Institute for Evolutionary Anthropology
Raphael Isokpehi	2004	Director of the Center for Bioinformatics & Computational Biology at Jackson State University
Konrad Scheffler	2005	Theodore Gildred Research Facility, University of California, San Diego
Nicki Tiffin	2005	Senior Lecturer, SANBI, UWC
Gwen Koning	2006	Global Seed Core Manager – Syngenta Crop Protein AG, Basel, Switzerland
Chris Maher	2007	Assistant Prof, Washington University School of Medicine
James Patterson	2009	
Adam Dawe	2009	SANBI Staff, 2012
Sunil Sagar	2009	Research Scientist, KAUST
Mandeep Kaur	2009	Wits School of Molecular and Cell Biology (MCB)
Stuart Meier	2009	Research Scientist, KAUST
Adele Kruger	2010	Wayne State University, Detroit, Michigan
Oliver Hofmann	2010	Affiliated Faculty, Harvard Stemcell Institute, Associate Director at Harvard School of Public Health
Sundarajan Seshadri	2010	Nanyang Technology University, Singapore
Ashley Pretorius	2010	Senior Lecturer, Biotechnology, UWC
Sumir Panji	2012	H3A BioNet, Network Manager, UCT
Samson Muyanga	2012	
John Pool	2012	Academic post, UKZN
Uljana Hesse	2013	IMBM, UWC
Barbara Picone	2013	University Stellenbosch

Monique Maqungo	2013	
Edwin Murungi	2013	Egerton University, Kenya
Hannah Ajogee	2013	
Natasha Wood	2014	UCT Computational Biology Department
Halimit Ebrahim	2009	
Katlego Motlhatlego	2012	
Siyanda Tsaba	2012	
Stacey Moses	2012	MSc, UWC Biotechnology Department

PHD

Name	Date completed	Currently
Alan Christoffels	2001	Director, SANBI; DST/NRF Research Chair
Ekow Oppon	2002	Zambia
Junaid Gamieldien	2002	Senior Lecturer, SANBI, UWC
Zhuo Zhang	2007	Research Scientist, University of Singapore
Alan Chong	2009	Research Fellow, Beth Israel Deaconess Medical Centre, Harvard Medical School
Magbubah Essack	2009	Research Scientist, KAUST
Sebastian Schmeier	2009	New Zealand
Ulf Schaefer	2009	Research Scientist, KAUST
Mark Wamalwa	2011	International Livestock Research Institute, Kenya
Musa Gabere	2011	USA
Samuel Kwofie	2011	Ghana
Aleksander Radovanovic	2010	Research Scientist, KAUST
Mushal Ali	2013	National Institute of Communicable Diseases, Gauteng
Kavisha Ramdayal	2014	SAP ERP Analyst, City of Cape Town
Michael Berry	2015	KappaBiosystems
Azeez Fatai	2015	Senior Lecturer, Lagos State University

HONOURS

Name	Date completed
Clifford Omorogie	2001
Grant Carelse	2002
Thurayah Davids	2005

MSC

Name	Date completed	Currently
Bukiwe Lupindo	2005	SA Government Administration
Tzu-Ming Chern	2003	PhD, KAUST
Elana Ernstoff	2003	
Etienne Swart	2003	Graduate Student, Princeton University
Victoria Nembaware	2003	Postdoctoral Fellow, UCT
Zayed Albertyn	2003	Bioinformatics Director, Malaysia
Anelda Boardman	2004	Stellenbosch University, Sequencing Facility Manager
Faisel Mosoval	2005	Senior Professional Officer, Information Systems and Technology, City of Cape Town
Nothemba Gwija-Kula	2005	Health Consultant, World Bank
Farahnaz Ketwaroo	2005	PhD, UCT
Mario Jonas	2006	Web Administrator, SANBI
Oliver Bezuidt	2007	PhD, University of Pretoria
Eugene Duvenhage	2009	Software Developer, Corporate
Frederick Kamanu	2009	PhD, KAUST
Feziwe Mpondo	2009	South African MRC Research Scientist
Saleem Adam	2011	
Firdous Khan	2012	PhD UWC Biotechnology Department
Wisdom Akurugu	2013	Bioinformatician, H3AbioNet, Ghana
Fred Nindo	2013	Computational Biology Department, UCT
Ram Shrestha	2013	London, UK
Leendert Cloete	2014	Bioinformatics scientist, KappaBiosystems, Cape Town

funders

SANBI
South African National Bioinformatics Institute

UNIVERSITY of the WESTERN CAPE

saMRC advancinglife

science & technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

NRF National Research Foundation

NIH National Institutes of Health

H3ABioNet

SANBI
South African National
Bioinformatics Institute

Postal Address:
South African National Bioinformatics Institute
University of the Western Cape | Private Bag X17 | Bellville | 7535

Physical Address:
South African National Bioinformatics Institute
5th Floor | Life Sciences Building | University of the Western Cape
Robert Sobukwe Road | Bellville | 7530 | South Africa

Telephone: +27 (0)21 959-3645
Facsimile: +27 (0)21 959-2512
Email: info@sanbi.ac.za

www.sanbi.ac.za

RESPICE PROSPICE
UNIVERSITY *of the*
WESTERN CAPE