Responding to a parasitic invasion

Research Chair in Bioinformatics and Health Genomics, **Dr Alan Christoffels,** shares some of the details of his laboratory's work on genomics and the role that he plays as an educator of future scientists

You lead the South African National Bioinformatics Institute (SANBI) lab's research on host-pathogen interaction. Could you provide a brief overview of your work in the field?

Different host immune responses are invoked during parasitic invasion. These immune responses can represent evolutionary adaptations specific to certain organisms, and they are clearly illustrated in the three model systems that are used in my laboratory. The three relationships we focus on are between *Anopheles* and *Plasmodium*, humans and *Mycobacterium*, and *Glossina* and *Trypanosoma*. My lab uses a combination of computational and experimental approaches to unravel the molecular interactions at the host-pathogen interface. We also aim to develop high throughput genomics methods, including next generation sequencing data analysis approaches, in order to study communicable diseases such as tuberculosis, malaria and sleeping sickness. As part of an international tsetse genome consortium, my group has been analysing the tsetse genome with a view to understanding the evolution of the immune system and its response to invading parasites.

Why did you decide to pursue a career in computational biology research and, specifically, its application to communicable diseases?

My career started out in the medical sciences, and as time went on, I found myself having to engage with data more and more. A natural affinity for computers made the transition to formal computational biology training an easy career shift for me. The first bioinformatics institute was established at the time when I was considering my PhD options. My choice of research topic is linked to the prevailing disease burden on the African continent and specifically in South Africa.

By what means is SANBI developing throughput genomics methods such as next generation sequencing data analysis for the study of diseases like tuberculosis, malaria and sleeping sickness?

Across all these disease model systems, the issue of data integration is key. We have access to large sequencing datasets, and we often find ourselves having to either alter existing computational approaches such as genome annotation pipelines or detect genetic variation. On the other hand, we are using machine-learning techniques to identify protein interaction networks between humans and mycobacteria.

How is the biomedical research conducted at SANBI relevant in both African and international contexts?

The research portfolio at SANBI spans both communicable and non-communicable diseases. Each project focuses on a disease that is pertinent to South Africa, the African region and the international community; for example, our work on developing a cost effective assay for HIV drug resistance testing has a global impact. Multidrug resistance and extreme drug resistance is also a global problem. We are using a combination of metabolic networks and drug design to look at new drug targets.

Alongside pioneering biomedical research, SANBI aims to foster the highest levels of excellence in education. Could you outline the undergraduate and postgraduate programmes and training courses offered by the Institute?

Currently, we do not offer any complete undergraduate programmes, although we do teach a term module to third year BSc students as part of the biotechnology degree at the University of the Western Cape. At SANBI, we offer an MSc and PhD degree programme. These postgraduate degrees are research-based and supplemented with short courses. Among these short courses is a compulsory seven-week course hosted at SANBI for all South African universities.

What can students expect from SANBI with regard to mentoring, research opportunities and teaching environments?

We have five research laboratories and their research themes cover HIV dynamics, non-communicable diseases, plant viruses, clinical biomarker discovery and host-pathogen interactions. Each of these labs is embedded in multinational collaborative projects, and students will have an opportunity to engage with an international network of researchers. Participation in this global collaboration will also allow students to enjoy overseas visits as part of their graduate experience.

On a personal level, what do you believe to be SANBI's greatest accomplishment to date?

I will not single out any specific research project, but the institute as a whole has achieved a number of important accomplishments. We have delivered on our mandate to develop computational biology in South Africa and the African region, for example, and our graduate students have taken up prominent academic positions at local universities – or have gone abroad, and are contributing significantly on an international stage.

Bioinformatics investigations in Africa

The **South African National Bioinformatics Institute** plays a vital role in computational biology research in Africa, as well as training future leaders in the field. Its importance is set to grow over the coming years

SOUTHERN AFRICA, AND particularly the country of South Africa, is currently an exciting location for those in the field of bioinformatics. 2011 saw the National Institutes of Health (NIH) and Wellcome Trust make a strong commitment to support research on the continent with the Human, Heredity and Health (H3Africa) programme, as well as a Pan African Bioinformatics Network to support genetics projects funded by H3Africa and biobanks across the African continent. 2012, similarly, was the inaugural year of the South African Society for Bioinformatics and Computational Biology and saw the establishment of the Southern Africa Human Genome Programme, designed to increase South Africa's capacity for genomics research. Overall, the region has poised itself to become a major contributor to genomic research in the near future.

The continent of Africa is plagued by a number of major health problems that could benefit hugely from genomic, and subsequently bioinformatics, investigation. This approach may, for example, help to shed a light on the parasite/host interactions that result in malaria and sleeping sickness or to unravel the question of why some people are resistant to currently available drugs for the treatment of HIV and tuberculosis. Even with these specific questions aside, the countries that make up Southern Africa have been found to be among the most genetically diverse in the world; therefore, the region simultaneously reflects a great need for genetic study and a great opportunity for it.

FILLING THE GAP

According to the World Health Organization (WHO), no more than 10 per cent of the total staff required to provide adequate genetic services were available in South Africa in 2008 – and human health is only one of the areas that has suffered because of this deficiency. Greater powers of genomic investigation could help Africa in its production of food crops and animals, for example, and there are also unique opportunities to conduct more fundamental research into plant, animal and human populations. Southern Africa's promising initiatives are the first step on the path towards this bright future, but the next will be to ensure that researchers and scientists in the field of computational biology and bioinformatics are being sufficiently trained.

A major player in this research landscape, and a vital provider of new scientists and researchers, has been the South African National Bioinformatics Institute (SANBI), located at the University of the Western Cape in South Africa. SANBI was founded in 1996, and in 2000, it was responsible for hosting South Africa's first workshop on genomics. Comprising four research teams headed by Drs Gamieldien, Travers, Tiffin and Harkins, under the directorship of Professor Alan Christoffels, the institute has continued to achieve its important goals: to conduct nationally and internationally relevant, cutting-edge research and to educate and mentor biological and computational scientists. The institution produces around six graduates per year, half of these being Masters students and the other half PhDs. Many of these students are drawn from historically disadvantaged backgrounds, and all of them go on to further study or to take positions amongst the faculty of other national or international institutions.

PROFILING PARASITES

Concerning human health, one of SANBI's most important projects has been investigating the genetic adaptations hosts and parasites have developed towards one another, with particular regard to three pairings which are responsible for tuberculosis, malaria and sleeping sickness. In the case of sleeping sickness, a disease caused by the protozoan *Trypanosome brucei*, its work focuses on analysing the genome of the tsetse fly. The tsetse fly is a common vector for parasitic infection, but is itself immune to the protozoan's attacks; its genes may reveal why that is. At the same time, the team is using comparative phylogenetic analyses to identity SNARE proteins in the parasites themselves.

The research work extends far beyond HIV and even parasitic infections, into plant viruses, and genetic determinants of non-communicable diseases. Underlying all of these studies is a coordinated approach to integrate diverse datasets

In terms of tuberculosis, the abundance of mycobacterial samples in South Africa has led to the rapid sequencing of hundreds of mycobacterial genomes. Christoffels' team is in the process of building workflows to automate the assembly of sequencing data from Pac Bio and illumina. Ultimately, the goal is to map the genetic variation of the bacteria against their phenotypic traits to inform their drug discovery platform.

EVOLUTION AND IMMUNITY

The SANBI also devotes much of its research to the problems presented by HIV, looking at the genomics of both the virus and the host from a number of angles. In collaboration with groups in Malawi and Ireland, they are studying the use of the gene CXCR4 – a gene thought to have an impact on drug resistance – during disease progression. Another important group of genes under study are the immunoglobulin heavy chain variable region genes, which the group is attempting to sequence in its different variations across a number of HIV patients. A third project aims to use their next-generation sequencing platform for research, where the SANBI team is incorporating it into the development of a software solution for user-friendly, cost effective HIV drug resistance testing.

But all of these studies represent only a fraction of the diverse, ongoing HIV research projects at SANBI, and their medical work extends far beyond HIV and even parasitic infections, covering studies into plant viruses and genetic determinants of non-communicable diseases. Underlying all of these studies is a coordinated approach to integrate diverse datasets. The institute's committed approach to collaboration and training also ensures that the information, practices and resources cultivated as part of these projects are widely shared – the SANBI participates in the International Glossina Genome Initiative, for example, which is concerned with assembling the tsetse fly genome. As part of this, the SANBI decided to implement the Glossina Functional Genomics Network, an initiative to provide African researchers with bioinformatics training and, more recently, internship opportunities in American and European laboratories.

AGRICULTURAL APPLICATIONS

The institute's research even extends to the genomic investigation of fungal pathogens affecting plants. *Venturia inaequalis*, or apple scab, is a fungal disease that can ruin fruit crops, and so SANBI has been working in collaboration with the Agricultural Research Council for the last three years to assemble and annotate the *V. inaequalis* genome. The team has identified a few classes of genes that are expanded specifically in this organism and not in other fungi, and this discovery provides examples of genes that could be key to the pathogenicity of the fungus. The findings provide a starting point for wet-lab experiments. Moreover, the collaborators have developed a web portal to share the genomic data with other users in the agricultural research community.

THE AGE OF BIOINFORMATICS

The SANBI has served the South African community well over the last 18 years, but it seems that its best years are still ahead; as bioinformatics becomes increasingly important in the African research environment, and worldwide, so too the SANBI will continue its ascension, providing new scientists, cutting edge research and practical solutions for Africans along the way.

INTELLIGENCE

SOUTH AFRICAN NATIONAL BIOINFORMATICS INSTITUTE

OBJECTIVES

The South African National Bioinformatics Institute delivers biomedical discovery appropriate to an international and African context. The institute's mission is:

- 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.

KEY COLLABORATORS

National Institute for Communicable Diseases • University of Cape Town • University of Kwa-Zulu Natal • University of Witwatersrand

FUNDING PARTNERS

Atlantic Philanthropies • South African Medical Research Council • South African National Research Foundation • Wellcome Trust • World Health Organization (WHO) • EU Sixth Framework Programme • International Business Machines Inc. South Africa

CONTACT

Dr Alan Christoffels
Director

South African National Bioinformatics Institute
University of the Western Cape
Private Bag X17, Bellville 7535
South Africa

T +27 21 959 2969
E alan@sanbi.ac.za

www.sanbi.ac.za

DR ALAN CHRISTOFFELS is the current Director & DST/NRF Research Chair in Bioinformatics and Health Genomics, at the South African National Bioinformatics Institute. His bioinformatics lab focuses on host-pathogen interactions and is developing high throughput genomics methods, including next generation sequencing data analysis approaches, to study communicable diseases such as tuberculosis, malaria and sleeping sickness.

UNIVERSITY of the WESTERN CAPE

SANBI South African National Bioinformatics Institute